

# Chapter 9

## Compression Performance Analysis in HEVC

Ali Tabatabai, Teruhiko Suzuki, Philippe Hanhart, Pavel Korshunov,  
Touradj Ebrahimi, Michael Horowitz, Faouzi Kossentini, and Hassene Tmar

**Abstract** In this chapter, performance analysis of HEVC (Recommendation ITU-T H.265 | ISO/IEC 23008-2) in comparison with AVC (Recommendation ITU-T H.264 | ISO/IEC 14996-10) in terms of both objective as well as subjective quality assessments are given. Because of the increased flexibility offered by HEVC, methods to select the best coding parameters, in a rate-distortion sense, are also described. Special care has been taken to apply a unified approach when conducting subjective and objective quality evaluation between HEVC and AVC. Our overall evaluation study results show the coding efficiency of HEVC to be about twice higher than that of AVC.

### 9.1 Performance Analysis

Performance analysis of HEVC is in general a complex undertaking since it can be conducted in number of different ways based on, for example, compression efficiency, complexity, visual quality, application of rate distortion optimization (RDO), delay, robustness, etc. The goal of this chapter is to present HEVC compression efficiency in comparison with AVC both in terms of objective and subjective quality assessments while taking into account some aspects of complexity, RDO,

---

A. Tabatabai • T. Suzuki (✉)  
Sony Corporation, Tokyo, Japan  
e-mail: [ali.tabatabai@am.sony.com](mailto:ali.tabatabai@am.sony.com); [teruhikos@jp.sony.com](mailto:teruhikos@jp.sony.com)

P. Hanhart • P. Korshunov • T. Ebrahimi  
Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

M. Horowitz • F. Kossentini • H. Tmar  
eBrisk Video, Inc., Vancouver, BC, Canada

and delay. Note that it is important to include both quality measures; relying solely on objective quality evaluations could, in cases, underestimate the amount of bit rate reduction and hence affect our analysis of compression efficiency. Subjective quality evaluations on the other hand, although difficult to conduct, correlate directly with perceptual experience of the viewers. This chapter is organized as follows. Section 9.1 provides the background information and forms the basis for the sections that will follow. In Sect. 9.2 encoder settings and testing conditions are described by considering various encoder configurations according to complexity and delay requirements; moreover, a list of test sequences used, test cases and the description of non-normative R–D optimization tools that contribute significantly to coding efficiency improvement are also covered in this section. In Sect. 9.3, objective quality evaluations of HEVC and AVC reference implementations are investigated. In Sect. 9.4, we present the results of HEVC subjective quality testing and visual assessments of HEVC and AVC. Section 9.5 describes an informal subjective video quality comparison of production–quality HEVC and AVC encoders in the context of 4K streaming applications. Conclusions appear in Sect. 9.6.

## 9.2 Encoder Setting

To conduct HEVC and AVC performance evaluations, a well-defined encoder setting and testing environment need to be established. In this section, we will describe, the HEVC and AVC reference encoder software (SW) used in our investigations. In addition, we will also describe various encoder configurations and prediction structures that are appropriate for different application requirements in terms of coding efficiency, complexity and delay.

### 9.2.1 Encoder Software

In the standardization of HEVC, the reference software, which is called HM (HEVC Test Model, reference software) [15] has been developed as a common SW platform for further improvement and study. Using SVN servers, the HM reference software is maintained at two sites [16]: HHI (Heinrich Hertz Institute) maintains the main SVN server and BBC (British Broadcasting Corporation) maintains the mirroring repository site.

The reference software for AVC, which is called JM (Joint Test Model), has been developed, as a common test platform, for AVC performance evaluations. The JM reference software is maintained at SVN server [6]. In this chapter, in order to compare the coding performance of HEVC with AVC, HM12.1 and JM18.5 SW are used for HEVC and AVC encoders, respectively.

## 9.2.2 Test Conditions

During the development of the HEVC specification, establishment of Common Test Conditions (CTC) provided a well-defined platform on which experiments for coding tool evaluations are performed [3]. Since HEVC coding performance evaluations are carried out according to the CTC, a detailed description of CTC key elements will follow.

## 9.2.3 Prediction Structure

For performance evaluation, CTC defines the following prediction structures.

1. All Intra (AI)
2. Random Access (RA)
3. Low Delay P picture (LDP)
4. Low Delay B picture (LDB)

In these configurations, QP (Quantization Parameter) value can be modified by adding to it a “QP offset” value. That is, CTC defines QP of the first picture (QP of an I picture, QPI, with I picture defined below) and the QP of the following pictures are derived as  $QP = (QPI + QP \text{ offset})$ , with QP offset being determined according to the picture type (e.g., P & B pictures, defined below) or a picture temporal ID. An I (intra) picture refers to a picture that can be decoded independently without requiring prediction data from other decoded pictures. A P (predicted) picture, in general, requires picture sample data from one other I, P or B picture to generate each predicted sample block. A B (bi-predicted), in general, requires picture sample data from two other I, P or B pictures to generate each predicted sample block.

### 9.2.3.1 All Intra (AI)

In this configuration, each picture is encoded as an I picture. Because no inter picture prediction is used, it is thus suitable for low delay and higher bit rate applications. QP offset in this configuration is 0 since QP is kept constant over the whole sequence. Figure 9.1 shows an example of this prediction structure.

### 9.2.3.2 Random Access (RA)

In this configuration, a hierarchical B structure is used [21]. Figure 9.2 shows an example of this prediction structure. The coding efficiency achieved by the bi-directional hierarchical prediction structure is higher than the other configurations. It has however a larger delay due to the reordering of the pictures. To control possible error propagation and ease of random access, I pictures are inserted periodically. QP offset values for each picture are summarized in Fig. 9.2.

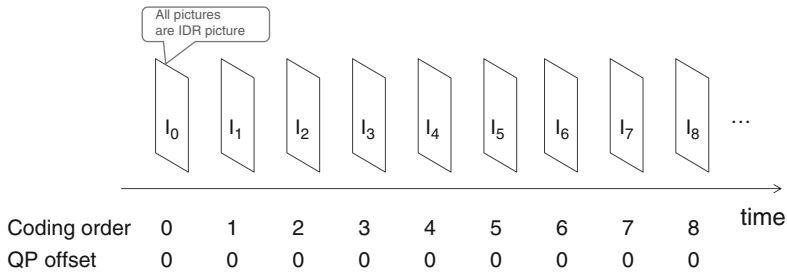


Fig. 9.1 The prediction structure of the intra-only configuration

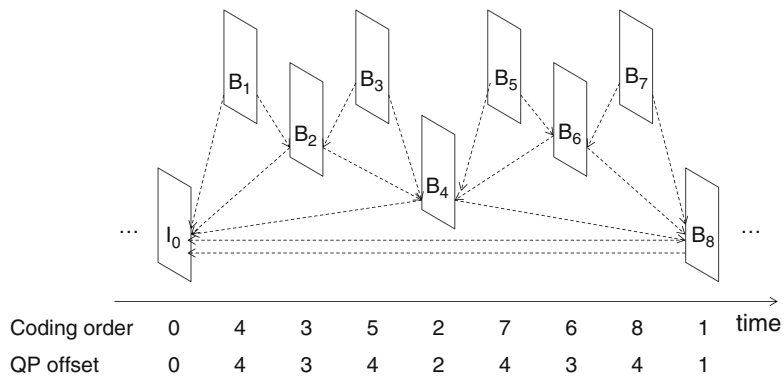


Fig. 9.2 The prediction structure of the random access configuration

### 9.2.3.3 Low-Delay P (LDP)

In this configuration, the first picture is encoded as an I picture and the subsequent pictures are encoded as P pictures. Since reordering of pictures is not allowed and only past pictures are used for prediction, the coding delay, in this configuration, may be made small. Figure 9.3 shows an example of this prediction structure. QP offset values are summarized for each picture in Fig. 9.3.

### 9.2.3.4 Low-Delay B (LDB)

In this configuration, similar to the previous configuration, reordering of pictures is not allowed. The first picture is encoded as an I picture and subsequent pictures are encoded as B pictures. Moreover, since past B pictures are used for prediction, a low coding delay, similar to LDP, but with higher coding efficiency (because of bi-prediction) is achieved. QP offset values, for each picture, are summarized in Fig. 9.3.

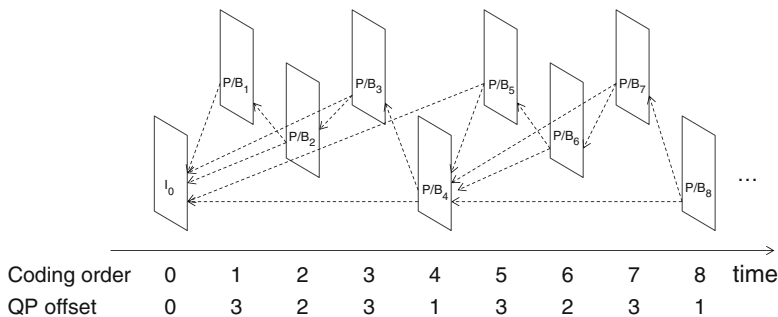


Fig. 9.3 The prediction structure of low-delay P and B configurations

### 9.2.4 Test Sequences

Test sequences are defined according to the picture size and applications and they are classified into six classes (class A to class F). Class A is the set of sequences with higher resolution than 1080p HDTV. The sequences are used to evaluate the coding performance of 4K/8K video. To reduce computation time, picture sizes are cropped to 2,560 × 1,600 pixels. Class B is for coding performance evaluation of 1080p HDTV and the set contains HDTV sequences, with a picture size of 1,920 × 1,080 pixels. Classes C and D are the set of test sequences with picture sizes of 832 × 480 pixels and 416 × 240 pixels, respectively. Test sequences in these two classes are for coding performance evaluation of mobile applications. Class E is the set of test sequences with a picture size of 1,280 × 720 pixels. It is used to evaluate coding performance of low-latency applications such as visual communications. CTC, in addition, defines class F sequences for coding performance evaluation of non-camera captured content such as video screen content, containing, for example, text and computer graphic. The test sequences are listed in Table 9.1.

In addition to the test sequences defined in CTC, 4K test sequences listed in Table 9.2 are used for both objective and subjective quality performance analysis in this chapter.

### 9.2.5 Test Cases and Bit Depth

Two test cases Main and Main 10 are defined to evaluate coding performance of 8-bit and 10-bit video. All test cases are summarized in Table 9.3.

In Main10 configuration, an 8-bit video is converted first to a 10-bit video by a 2-bit left shift, and it is then encoded as 10-bit video. Likewise, in Main configuration, a 10-bit video is first converted to an 8-bit video by a 2-bit right shift,

**Table 9.1** Test sequences

Class	Sequence name	Frame count	Frame rate (fps)	Bit depth
A	Traffic	150	30	8
A	PeopleOnStreet	150	30	8
A	Nebuta	300	60	10
A	SteamLocomotive	300	60	10
B	Kimono	240	24	8
B	ParkScene	240	24	8
B	Cactus	500	50	8
B	BQTerrace	600	60	8
B	BasketballDrive	500	50	8
C	RaceHorses	300	30	8
C	BQMall	600	60	8
C	PartyScene	500	50	8
C	BasketballDrill	500	50	8
D	RaceHorses	300	30	8
D	BQSquare	600	60	8
D	BlowingBubbles	500	50	8
D	BasketballPass	500	50	8
E	FourPeople	600	60	8
E	Johnny	600	60	8
E	KristenAndSara	600	60	8
F	BaskeballDrillText	500	50	8
F	ChinaSpeed	500	30	8
F	SlideEditing	300	30	8
F	SlideShow	500	20	8

**Table 9.2** 4K Test sequences

Resolution	Sequence name	Frame count	Frame rate (fps)	Bit depth
3,840 × 2,160	Book <sup>a</sup>	500	50	10
3,840 × 2,160	BT709Birthday <sup>b</sup>	500	50	10
3,840 × 2,160	HomelessSleeping <sup>c</sup>	600	60	10
3,840 × 2,160	Manege <sup>d</sup>	600	60	8
4,096 × 2,048	Traffic	300	30	8

<sup>a</sup>Book sequence was created by BBC (British Broadcasting Corporation), Research and Development Department

<sup>b</sup>BT709Birthday sequence was created by Technicolor

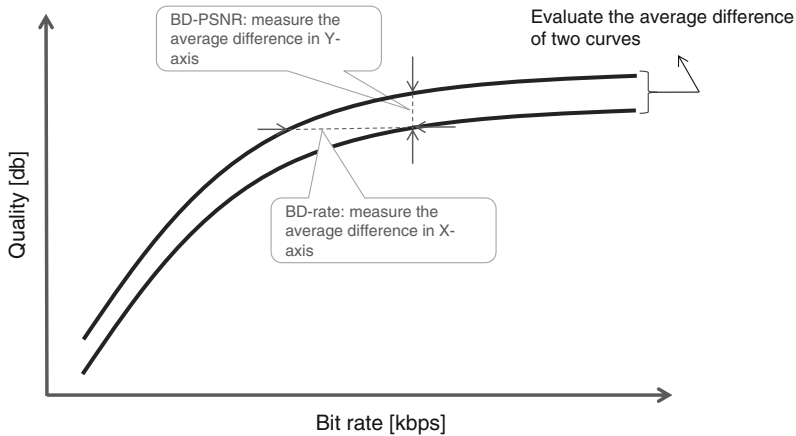
<sup>c</sup>HomelessSleeping sequence was created by Kamerawerk GmbH, Switzerland. The sequence is an excerpt from the film entitled “No Sleep 4K”

<sup>d</sup>Manege sequence was created by 4EVER consortium

and it is then encoded as an 8-bit video. The word “optional” in Table 9.3 means that using certain class of sequences (e.g., class F) or certain prediction structures (e.g., LDP) were not required but recommended, instead. In this chapter, Main configuration is used to evaluate “optional” cases in Table 9.3.

**Table 9.3** Summary of test cases in the common test conditions

Prediction structure	Class					
	A	B	C	D	E	F
AI	Main/Main10	Main/Main10	Main/Main10	Main/Main10	Main/Main10	Optional
RA	Main/Main10	Main/Main10	Main/Main10	Main/Main10	N/A	Optional
LDB	N/A	Main/Main10	Main/Main10	Main/Main10	Main/Main10	Optional
LDP	N/A	Optional	Optional	Optional	Optional	Optional



**Fig. 9.4** An example of R–D curve

### 9.2.6 Rate Distortion Curves

When evaluating the coding performance of a video codec, a graph of R–D curve (Rate–Distortion Curve) is used. R–D curve is generated by plotting the encoded results, in terms of bit rate versus the resulting quality, in a graph. The horizontal axis denotes the bit rate and the vertical axis denotes a measure of distortion or quality of encoded video. In general, a higher compression ratio results in a lower bit rate; however, picture quality is generally reduced. Low compression ratio, on the other hand, improves picture quality but at the cost of an increase in bit rate. Since a high coding efficiency codec can achieve higher quality at lower bit rates, the R–D curve moves toward upper left, as shown in Fig. 9.4.

As an objective measurement of picture quality, PSNR (Peak Signal to Noise Ratio) is widely used. PSNR can be calculated by the following equation.

$$PSNR = 10 \log_{10} \frac{(2^{bitdepth} - 1)^2 * W * H}{\sum_i \{O_i - D_i\}^2}$$

where

bitdepth: Bit depth of each pixel

W: Number of horizontal pixels

H: Number of vertical pixels

O<sub>i</sub>: Pixel value of the reference picture

D<sub>i</sub>: Pixel value of the decoded picture

i: Pixel address

PSNR is calculated for each YCbCr component. In YCbCr domain, human visual system is more sensitive to luminance (Y) than to chrominance (Cb or Cr); accordingly, and in practice, PSNR for luminance (PSNR Y) is a more important metric for objective quality measurements.

In order to compare the coding efficiency of a reference codec vs. the one being evaluated, the average difference of the two R–D curves is calculated. The average bit rate difference (difference in horizontal direction) is referred to as BD (Bjøntegaard's Delta) Rate and the average PSNR difference (difference in vertical direction) is referred to as BD PSNR [1].

In order to calculate BD Rate and BD PSNR, the two R–D curves (corresponding to reference and tested codecs) are approximated by the following cubic polynomial.

$$PSNR = a + b * (bit\ rate) + c * (bit\ rate)^2 + d * (bit\ rate)^3$$

Parameters a–d in the above equation can be derived by using four data points (PSNR and bit rate points). This polynomial approximation will then allow us to derive the BD Rate by integrating the difference of two curves in horizontal direction and BD PSNR by integrating the difference of two curves in vertical direction (see Fig. 9.4).

BD Rate and BD PSNR have been widely used to evaluate coding tools, in the HEVC standardization work. It is however known that such approximation could sometimes lead to large errors, especially for large pictures (e.g. class A sequences). To further improve the approximation accuracy, a piece-wise cubic interpolation is proposed as an alternative [2].

### 9.2.7 R–D Optimization

HEVC encoder flexibility stems from the fact that it contains an increased number of coding tools, beyond those provided by earlier video coding standards e.g., AVC. This added flexibility allows an encoder to adaptively determine block dependent coding parameters in terms of:

1. Coding unit (CU) quadtree structure, prediction unit (PU) partition modes and transform unit (TU) quadtree structure;
2. Intra PU prediction mode;



3. Inter PU motion parameters and reference list index or indices, for motion estimation;
4. Rate–distortion optimized quantization (RDOQ), for quantization process.

The key function and differentiation point of a “good” encoder is the selection of the “best” coding parameters (or so-called syntax element values), for improved coding efficiency. Finding the “best” coding parameters is traditionally performed in a rate–distortion (R–D) sense: it enables tradeoffs between the numbers of bits used to encode a block of the picture vs. the resulting distortion that is produced by using that number of bits. An R–D optimization problem can in general be formulated as:

$$\min_{(\text{coding parameters})} (D) \text{ subject to } R \leq R_T \quad (9.1)$$

where

$D = \text{Distortion}$ ,

$R = \text{Rate (number of bits required to signal coding parameters)}$

$R_T = \text{Target Rate}$

The above minimization is over a combined set of coding parameters and the distortion term is used to quantify the fidelity between original and reconstructed block. In principal, distortion can be measured either by relying on a mathematical distance or by taking into account perception mechanisms. Perceptual metrics correlate well with viewers’ perceptual experience but defining them is challenging because of the complexity of modeling various physiological components involved in human visual system. Objective quality measures based on mathematical distances, on the other hand, are easier to derive and under many circumstances they can still provide good tradeoffs between subjective quality and rate used. They are, moreover, “content-agnostic”. That is, the same error distribution on different content could yield similar objective quality metrics. Examples of distance based objective quality metrics include mean-squared error (MSE), peak-signal-to-noise (PSNR), and sum of absolute differences (SAD).

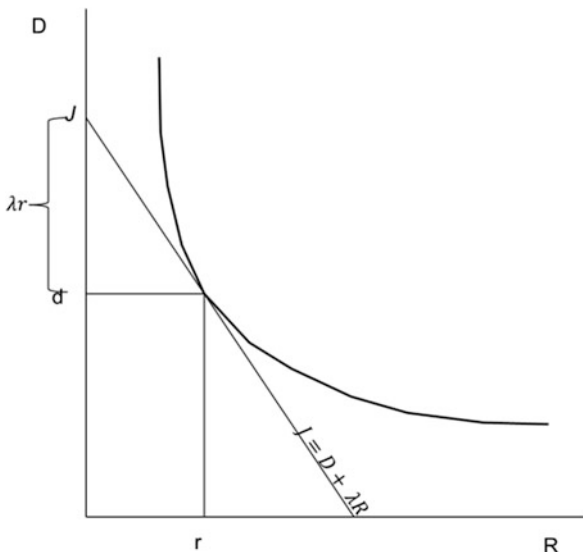
Constrained optimization problem in (9.1) can be turned into an unconstrained optimization problem by the introduction of non-negative Lagrangian multiplier  $\lambda$  which combines R and D into a so-called Lagrangian cost function [20, 22], namely:

$$\min_{(\text{coding parameters})} J = (D + \lambda * R) \quad (9.2)$$

Note that  $\lambda$  acts, in a sense, as a “knob”: changing the value of  $\lambda$  enables tradeoffs between rate decreases vs. distortion increases. For example,  $\lambda = 0$ , in (9.2), corresponds to minimizing distortion; conversely, choosing a large value for  $\lambda$  corresponds to rate minimization. A natural question that arises is what value to choose for  $\lambda$ ? Sullivan and Wiegand [22] and Ohm et al. [19] address this question by establishing a relationship between  $\lambda$  and quantization step size  $Q$ .

$$\lambda = c * Q^2 \quad (9.3)$$

**Fig. 9.5** Typical R–D curve and cost function  $J$  with slope  $-\lambda$



In AVC and HEVC, the quantization step size  $Q$  is controlled by a quantization parameter ( $QP$ ) such that  $Q$  is proportional to  $2^{(QP-12)/6}$  and the constant of proportionality,  $c$ , depends on coding mode decisions.

An example based on a graphical minimization of (9.2) is shown in Fig. 9.5, where a line denoting Lagrangian cost function is plotted against a typical rate–distortion curve that is a non-increasing convex function of  $R$  [4]. Minimum  $J$  can be achieved by finding the point on the rate–distortion curve which is “hit” first by the plane wave of slope  $-\lambda$  [20].

There are many alternative methods to performing R–D cost optimization. One, for example, can minimize a frame level distortion or minimize an average frame distortion, taken over many video frames. These aforementioned methods are not computationally practical as they will incur significant amount of complexity and delay. Instead, and as described in [15, 19], minimization of (9.2) is performed for each block of samples (e.g., CUs) independently and in four stages: (1) mode decision; (2) intra prediction mode estimation; (3) motion estimation; and (4) quantization. Accordingly, for each block an exhaustive pre-calculation of cost function, associated with each combination of coding parameters, is performed: the optimal R–D solution for the block is the combination that minimizes the R–D cost function. Making block independent assumption despite spatial/temporal dependencies that could exist between blocks (e.g., current block predictor is based on the past reconstructed block samples) is generally ignored for practical applicability [19]. We now describe briefly the four R–D optimization stages:

We let  $S_A(i, j)$  and  $S_B(i, j)$  denote the  $(i, j)^{th}$  sample in blocks A and B, of the same size, respectively. For measuring distortion, we use the following metrics as specified in [15]:

$$\text{Sum of Square Error (SSE)} = \sum_{i,j} (s_A(i, j) - s_B(i, j))^2 \quad (9.4)$$

$$\text{Sum of Absolute Difference} = \sum_{i,j} |s_A(i, j) - s_B(i, j)| \quad (9.5)$$

$$\text{Hadamard Transformed SAD (SATD)} = \sum_{i,j} |HT(i, j)| \quad (9.6)$$

$HT(i, j)$  in (9.6) is the  $(i, j)^{th}$  coefficient of a block that is obtained by applying Hadamard transform to the block difference between blocks A and B.

JCT-VC [15] specifies also the following  $\lambda$  values:

$$\lambda_{mode} = \alpha * W_k * 2^{((QP-12)/3.0)} \quad (9.7)$$

$$\lambda_{pred} = \sqrt{\lambda_{mode}} \quad (9.8)$$

$$\omega_{chroma} = 2^{((QP-QP_{chroma})/3.0)} \quad (9.9)$$

$\alpha = 1.0 - \text{Clip3}(0.0, 0.5, 0.05 * \text{number\_of\_B\_frames})$  for referenced pictures

$$\alpha = 1.0 \quad \text{for non-referenced pictures} \quad (9.10)$$

where

$$\text{Clip3}(x, y, z) = \begin{cases} x; & z < x \\ y; & z > y \\ z; & \text{otherwise} \end{cases}$$

Interested readers are referred to [15] for derivation of  $W_k$  as well as  $\lambda$  values for chroma.

CU level mode decision (intra vs. inter) coding is based on finding coding parameters that minimize cost function  $J_{mode}$  in (9.11).

$$J_{mode} = (SSE_{luma} + \omega_{chroma} * SSE_{chroma}) + \lambda_{mode} * R_{mode} \quad (9.11)$$

Distortion terms  $SSE_{luma}$  and  $SSE_{chroma}$  correspond to the SSE between the original and reconstructed luma and chroma CU blocks respectively. Similarly,  $R_{mode}$  represents the total number of bits used for CU level intra or inter mode signaling, PU partition(s) within the CU, PU prediction mode(s) in case of intra mode or PU motion parameters in case of inter mode, TU quadtree partition(s), and finally number of bits required for representing quantized residual transform coefficient levels.

For finding the best inter CU coding cost,  $J_{mode}$  is evaluated for all possible PU partition modes (e.g.,  $2N \times 2N$ ,  $N \times N$ ,  $2N \times N$ ,  $N \times 2N$ ,  $n1 \times 2N$ ,  $nR \times 2N$ ) and a partition that gives the minimum coding cost is chosen.

Motion estimation for each inter PU partition is done based on the minimization of inter prediction cost shown in (4.12).

$$mp^* = \arg \min_{mp \in MP} D_{mp} + \lambda_{pred} * R_{mp} \quad (9.12)$$

For a given reference picture list, set  $MP$ , over which the minimization is carried out, consists of all possible motion parameters, namely motion vectors and associated reference indices. The minimization task in (9.12) is broken into two parts: integer-sample precision and sub-sample precision. For integer-sample precision, distortion term  $D_{mp}$  corresponds to the SAD between original PU block and its motion compensated reference block. For sub-sample motion search however distortion term  $D_{mp}$  represents SATD of the block difference between the original and sub-sample motion compensated reference block.  $R_{mp}$  term represents an estimate of the number of coded bits required to transmit  $mp$ .

For bi-prediction, cost function minimization in (9.12) becomes a joint optimization problem and is solved by the application of an iterative algorithm [5]. The algorithm is initialized first with the two best motion parameters that are obtained independently, for each reference list (L0 and L1). Iteration for further refinement and combined cost minimization is performed by keeping motion parameter of L0 list constant while performing sub-pixel motion search on the complementary list (L1). Once minimum cost is achieved, motion parameter associated with L1 list is held constant and motion parameter of the L0 list is adjusted for computing minimum combined cost. This “ping-pong” like iteration process is continued until convergence is reached.

For intra PU prediction, a two-stage minimization process is performed:

At the first stage, a fix number<sup>1</sup> of candidate intra prediction modes with lowest prediction cost are chosen according to the minimization of the prediction cost function in (9.13).

$$p^* = \arg \min_{p \in P} D_p + \lambda_{pred} * R_p \quad (9.13)$$

Distortion term  $D_p$  in (9.13) represents the SATD between the original block and its prediction block using intra prediction mode  $p$  and  $R_p$  represents number of coded bits required for signaling mode  $p$ . Set  $P$ , over which minimization is carried out, consists of planar, DC and all the 33 angular prediction directions.

In the second stage, the list containing the candidate intra prediction modes from the first stage is augmented with the three most probable modes if not already present in the list. The best intra prediction mode is the one that gives the minimum  $J_{mode}$  among candidate intra prediction modes in this augmented list.

Note that HEVC allows PCM coding of a CU block if the block size is greater or equal to a signaled minimum PCM coding block size. For PCM  $J_{mode}$  evaluation,

---

<sup>1</sup>These fix values are pre-determined and they depend on PU size.

distortion terms  $SSE_{luma}$  and  $SSE_{chroma}$  are set to zero when both input and PCM coded samples have the same bit depth. Term  $R_{mode}$  includes all the bits required for signaling PCM mode and PCM coded samples.

Finally, by applying this CU level mode decision at each level of CU recursion tree a coding tree unit (CTU) level coding mode decision can be obtained.

The goal of Rate distortion optimized quantization (RDOQ), in quantization process stage, is the adjustment of transform coefficient levels, in R–D sense [17]. For an insight to the general concept of minimization process, assume  $c_k$  to be the last non-zero coefficient in a transformed block for a given position  $k$ ; then, for each transform coefficient level  $l_i$ , at position  $i = k - 1, \dots, 0$ , RDOQ tries to find the optimal transform coefficient level,  $l_i^*$  that minimizes the cost function,  $J_k(l_i)$ , below:

$$J_k(l_i) = D_k(l_i) + \lambda * R_k(l_i) \quad (9.14)$$

For computational simplicity, possible values of  $l_i$  are limited to be either zero, or truncated  $l_i$ , or rounded-up  $l_i$  (i.e.,  $l_{floor}$  and  $l_{ceiling}$ ). Distortion term  $D_k(l_i)$  is due to the quantization error and is calculated as normalized SSE in transform domain and  $R_k(l_i)$  denotes number of bits used for transmitting level  $l_i$ . The optimal solution is the vector of re-quantized transform levels at position  $k^*$  with minimum  $J_k$  over all possible positions,  $k$ .

### 9.3 Objective Performance Analysis

This section summarizes the comparison of coding efficiency of HEVC and AVC. The test conditions are summarized in Table 9.4 and all encoders settings described in Table 9.3 are used for the comparisons.

The results for the test sequences in Table 9.1 are summarized in Tables 9.5, 9.6, 9.7 and 9.8. In case of Random Access Main, coding efficiency of HEVC is 42.7 % higher than that of AVC. In case of All Intra Main however the improvement is 21.9 % which indicates that the improvement in Intra picture is lower than that in predictive pictures (P or B picture).

As an example, R–D curves of the sequence, Four People (Class E, RA Main) are shown in Figs. 9.6, 9.7, and 9.8.

**Table 9.4** Test conditions

Test conditions	
Encoder	HM12.1 (HEVC) and JM18.5 (AVC)
Test sequences	All sequences defined in Table 9.1
Bit depth	8-bit (Main configuration)
Prediction structure	AI, RA, LDP and LDB
QPI	22, 27, 32 and 37

**Table 9.5** Comparison of coding performance of HEVC and AVC (All Intra Main)

	All Intra Main		
	Y (%)	U (%)	V (%)
Class A	-23.6	-21.1	-19.9
Class B	-22.7	-22.1	-21.7
Class C	-19.7	-20.8	-21.1
Class D	-16.4	-17.0	-17.7
Class E	-28.8	-27.1	-27.1
<b>Overall</b>	<b>-21.9</b>	<b>-21.4</b>	<b>-21.2</b>
Class F	-28.6	-25.3	-26.2

**Table 9.6** Comparison of coding performance of HEVC and AVC (Random Access Main)

	Random Access Main		
	Y (%)	U (%)	V (%)
Class A	-42.6	-42.5	-44.5
Class B	-47.7	-43.3	-42.5
Class C	-35.5	-34.7	-35.4
Class D	-33.9	-36.6	-37.9
Class E	-56.0	-54.4	-55.8
<b>Overall</b>	<b>-42.7</b>	<b>-41.7</b>	<b>-42.5</b>
Class F	-53.1	-52.2	-53.9

**Table 9.7** Comparison of coding performance of HEVC and AVC (Low Delay B Main)

	Low Delay B Main		
	Y (%)	U (%)	V (%)
Class A	-38.6	-24.9	-26.7
Class B	-42.1	-34.4	-35.3
Class C	-32.7	-32.2	-32.9
Class D	-29.9	-31.2	-33.0
Class E	-44.1	-39.2	-38.8
<b>Overall</b>	<b>-36.6</b>	<b>-32.6</b>	<b>-33.6</b>
Class F	-33.9	-35.3	-37.8

**Table 9.8** Comparison of coding performance of HEVC and AVC (Low Delay P Main)

	Low Delay P Main		
	Y (%)	U (%)	V (%)
Class A	-29.3	-31.4	-33.4
Class B	-38.1	-37.3	-39.6
Class C	-32.4	-38.2	-38.5
Class D	-30.1	-38.2	-38.6
Class E	-44.4	-44.9	-44.7
<b>Overall</b>	<b>-35.3</b>	<b>-38.0</b>	<b>-38.9</b>
Class F	-35.3	-39.5	-40.5

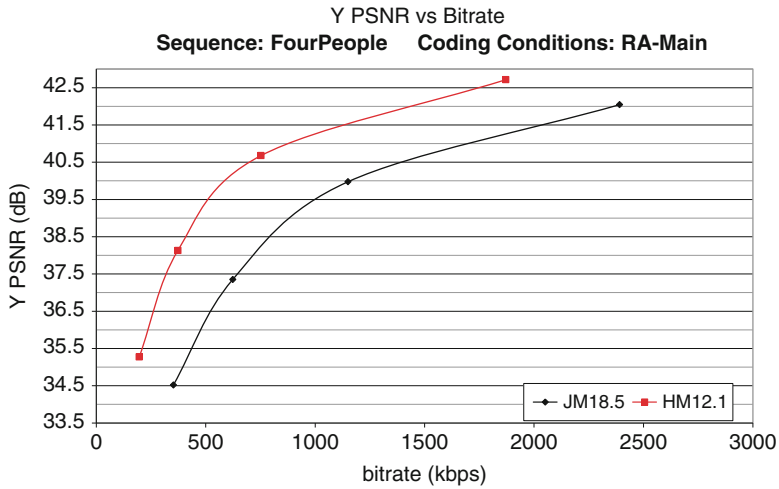


Fig. 9.6 R-D curve of Y (Four People, RA-Main)

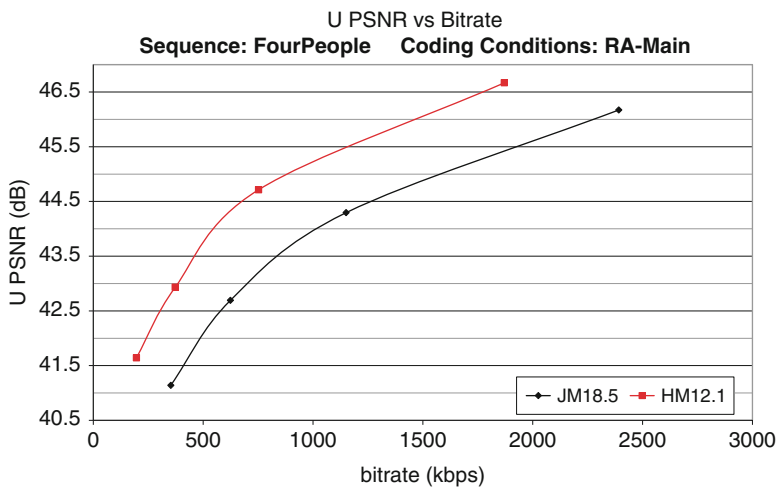


Fig. 9.7 R-D curve of U (Four People, RA-Main)

The results for 4K test sequences in Table 9.2 are summarized in Table 9.9 (Random Access Main only). We observe a coding efficiency improvement of up to 76 % for HEVC.

In addition, still picture coding performance of HEVC based intra coding, relative to JPEG and AVC intra coding is reported in [18]. The results show that bit rate reductions due to HEVC intra coding are about 44 and 32 %, respectively. Comparisons of HEVC intra coding to JPEG and JPEG2000 by means of objective and subjective evaluations are also reported in [9]. The evaluation

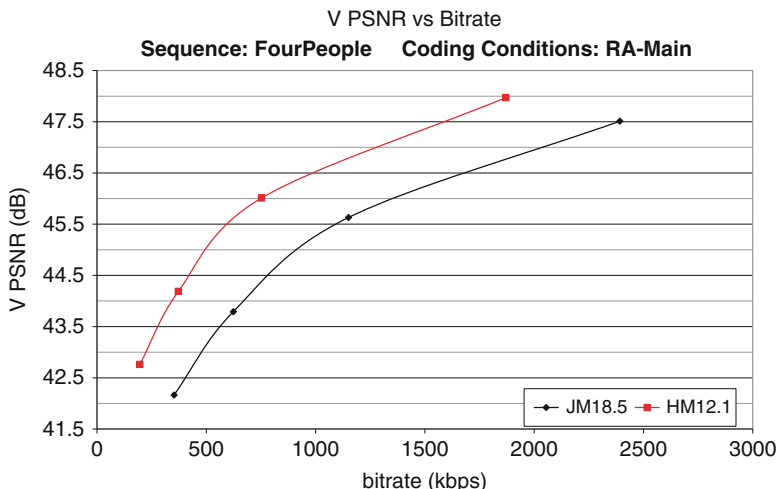


Fig. 9.8 R–D curve of V (Four People, RA-Main)

Table 9.9 Comparison of coding performance of HEVC and AVC (Random Access Main)

	Random Access Main		
	Y (%)	U (%)	V (%)
Book	−58.8	−55.2	−56.7
BT709Birthday	−60.4	−54.3	−57.6
HomelessSleeping	−75.9	−79.6	−83.5
Menage	−33.3	−34.1	−36.2
Traffic	−41.4	−43.7	−43.6

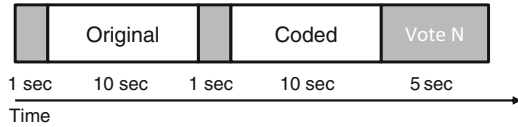
results demonstrate that HEVC intra coding outperforms encoders for still images with an average bit rate reduction ranging from 16 % (compared to JPEG 2000 4:4:4) up to 43 % (compared to JPEG).

### 9.4 Subjective Performance Analysis

Because, subjective evaluation of video content correlates directly with the viewer perceptual experience, it could very well be considered as a more reliable performance measure of a codec. It is therefore important that for conducting subjective evaluation test, the testing methodology be defined in accordance with the universally accepted guidelines and practices, such as those described in Recommendation ITU-R BT.500 [11–14]. In the following sub-sections, we will further elaborate on the testing methodology and environment together with references to subjective evaluation test results.



**Fig. 9.9** DSIS basic test cell (BTC)



### 9.4.1 Test Methodology

First, we provide a brief tutorial about some frequently used subjective quality assessment methods. In general, there are two broad methods to carry visual evaluation tests: double stimulus and single stimulus. In double stimulus test subjects rate either the quality or change in the quality between two video clips reference (original) vs. impaired (coded). In single stimulus test, subjects rate the quality of the impaired (coded) video clip, only. We will now describe two examples of the former, namely, double stimulus impairment scale (DSIS), and double stimulus continuous quality scale (DSCQS).

#### 9.4.1.1 DSIS (Double Stimulus Impairment Scale)

This method is used when the material to be evaluated shows a wider range of visual quality covering all quality scales (and not of the impairments). There are two variants of DSIS: Variant I and Variant II. The structure of the Basic Test Cell (BTC) of Variant I, is shown in Fig. 9.9. It consists of two consecutive presentations of video clips. Original (reference) video clip is presented first followed by the presentation of the impaired (coded) version of the video clip. A message is then displayed for 5 s requesting viewers to vote.

Viewers are expected to mark their visual quality score on an answer sheet with quality rating over a defined scale e.g., scale that is made of 5 levels—ranging from “1” (very annoying) to “5” (imperceptible). In Variant II of DSIS, the pairs of original (reference) video clip and impaired (coded) version of the video clip are presented twice before voting. For visual test evaluations conducted in Sect. 9.4.2, Variant I of DSIS methodology, as described earlier was chosen.

#### 9.4.1.2 DSCQS (Double Stimulus Continuous Quality Scale)

Double Stimulus Continuous Quality Scale (DSCQS) is used in cases when it is not possible to present the full range of quality scales. In this method, the original (reference) and the coded (impaired) samples of a video clip are presented twice and, in random order, for each BTC. At the end of the second presentation, the viewers are asked to grade each of the two original and the two coded video clips, separately. It should be noted that because of the random presentation order, viewers do not have an *a priori* knowledge of whether a video clip shown belongs to the original or to the impaired one.



**Fig. 9.10** DSCQS basic test cell (BTC)

As shown in Fig. 9.10, the BTC structure of the DSCQS method contains two consecutive pairs of presentations. At first, a mid-grey screen with the letter “A”, in the middle, is displayed for a second followed by a 10-s presentation of a video clip—either original or impaired. Then, a mid-grey screen with the letter “B” appears followed by a 10-s presentation of the second video clip. Similar process is repeated during the second round of presentation by changing letters A and B to A\* and B\*, instead. Finally, a message is displayed for 5 s instructing the viewers to vote.

#### 9.4.1.3 Training Session

The outcome of the visual tests could be highly dependent on the proper training of the participants. In order to allow viewers to get familiarized with the testing procedures, it is important that viewers are briefed about the testing procedures and participate in a training session before starting subjective evaluation tests. Also, the video clips shown for the training need to be different from those used during the actual tests. Coding impairments should resemble those that appear on the tested materials, though. In the training session, three BTCs (the worst quality, medium quality and the best quality) should be included allowing viewers to know the quality range of the test.

#### 9.4.1.4 Viewing Environment

In the laboratory where the viewing session is being held, general internal light has to be low and a uniform light has to be placed behind the monitor. The intensity of the light is specified in the ITU-R BT.500 [11, 14]. No light source has to be directed to the screen or cause reflections. Ceiling, floor and walls of the laboratory have to be made of non-reflecting material (e.g. carpet or velvet) and should have a color tuned as close as possible to CIE Standard Illuminant D65 (daylight illuminant, 6500K). The viewing room must be protected from external visual or audio pollution.

## 9.4.2 Subjective Quality Evaluation Test

This section reports the results of subjective quality evaluation conducted at EPFL's MMSPG test laboratory, which fulfills the recommendations for the subjective evaluation of visual data issued by ITU-R BT.500 [11, 14]. It is also worth noting that the testing methodology performed in this section has benefited significantly from the experience gained while conducting the subjective evaluation tests described in [8].

### 9.4.2.1 Test Environment

The test room is equipped with a controlled lighting system with a 6,500 K color temperature and an ambient luminance at 15 % of the maximum screen luminance, whereas the color of all the background walls and curtains present in the test area are in mid grey. The laboratory setup is intended to ensure the reproducibility of the subjective tests results by avoiding unintended influence of external factors.

To display the test stimuli, two Eizo CG301W LCD monitors with a native resolution of  $2,560 \times 1,600$  pixels were used. The monitors were calibrated using an X-Rite i1Display Pro color calibration device according to the following profile: sRGB gamut, D65 white point,  $120 \text{ cd/m}^2$  brightness, and minimum black level.

The experiment involved two subjects per monitor assessing the test material. The subjects were seated in a row perpendicular to the center of the monitor, at a distance of 2.2 times the picture height, roughly corresponding to a visual angle of 1 arc-minute between two adjacent pixels, as suggested in [13].

### 9.4.2.2 Test Methodology

The double stimulus impairment scale (DSIS Variant I) methodology as described earlier was chosen for the testing. A five-grade impairment scale (5: Imperceptible, 4: Perceptible but not annoying, 3: Slightly annoying, 2: Annoying, 1: Very annoying) was used. The subjects were presented with pairs of video sequences (i.e., stimuli), where the first sequence was always a reference video (stimulus A) and the second, the video to be evaluated (stimulus B). After the presentation of each pair of sequences, a 5-s voting time followed. Subjects were asked to rate the impairments of the second stimulus in relation to the first stimulus, and to express these judgments in terms of the wordings used to define the rating scale.

### 9.4.2.3 Dataset

Five video sequences in Table 9.2 were used in the experiments, with different visual characteristics, resolutions, and frame rates. All sequences were stored as

raw video files, progressively scanned, and with YCbCr 4:2:0 color sampling. The sequences were compressed with HEVC and AVC. For each sequence and codec, four quantization parameters were selected, resulting in a total of 40 test stimuli.

Five training samples were generated using the Sintel39 sequence (its resolution is  $3,840 \times 1,744$ ) and manually selected by expert viewers so that the quality of samples were representative of all grades of the rating scale.

The original sequences were cropped to the resolution of the monitor, keeping only the central part, and the 10-bit sequences were clipped to 8-bit.

#### **9.4.2.4 Training Session**

Before the experiment, a consent form was handed to subjects for signature, and oral instructions were provided to explain their tasks. Additionally, a training session was organized to allow subjects to familiarize with the assessment procedure.

#### **9.4.2.5 Test Session**

Since the total number of test samples was too large for a single test session, the overall experiment was split into two sessions of approximately 13 min each. Between the sessions, the subjects took a 10 min break. The test material was randomly distributed over the two test sessions.

Three dummy pairs (one with high quality, one with low quality, and one of mid quality), whose scores were not included in the results, were included at the beginning of each test session to stabilize the subjects' ratings. To reduce contextual effects, the stimuli orders of display were randomized applying different permutation for each group of subjects, whereas the same content was never shown consecutively.

A total of 18 naive subjects (6 females and 12 males) took part in the experiments. They were between 18 and 27 years old with an average of 23.4 years of age. All subjects were screened for correct visual acuity and color vision using Snellen and Ishihara charts, respectively.

#### **9.4.2.6 Analysis of the Results**

The subjective results were processed by first detecting and removing subjects whose scores appeared to deviate strongly from others. The outlier detection was performed according to the guidelines described in Section 2.3.1 of Annex 2 of [14]. In this study, one outlier was detected. Then, the mean opinion score (MOS) was computed for each test stimulus as the mean across the rates of the valid subjects, as well as associated 95 % confidence interval (CI), assuming a Student's *t*-distribution of the scores.

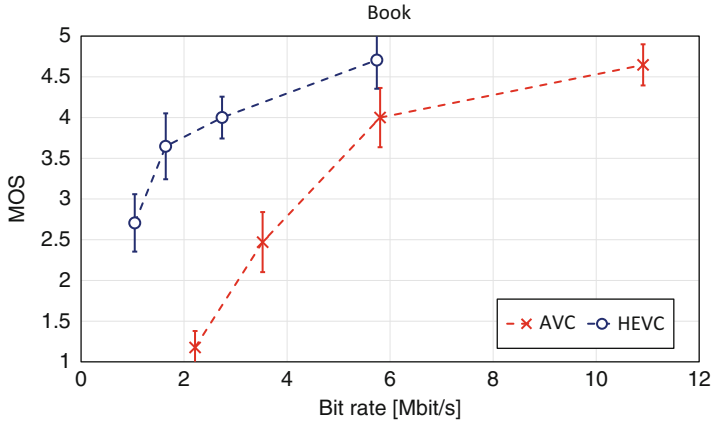


Fig. 9.11 R-D curve (Book)

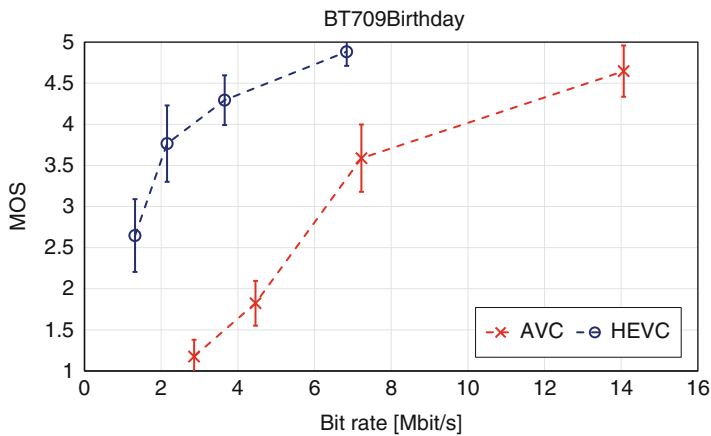


Fig. 9.12 R-D curve (BT709Birthday)

### 9.4.2.7 Rate Distortion Curves Results

The R-D curves obtained by the subjective quality evaluation are shown in Figs. 9.11, 9.12, 9.13, 9.14 and 9.15.

From these figures, it can be seen that HEVC shows substantial visual quality improvements over AVC, especially at lower bit rates.

### 9.4.2.8 Average Bit Rate Difference

The average bit rate difference for HEVC over AVC was computed using the model proposed in [7]. This model is an extension of the Bjøntegaard model [1]

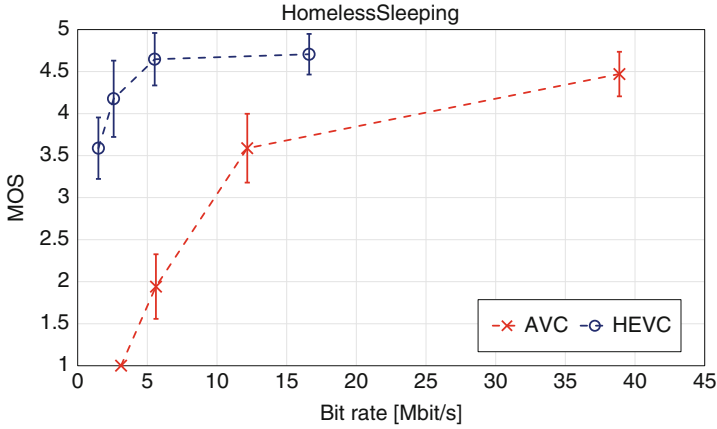


Fig. 9.13 R–D curve (HomelessSleeping)

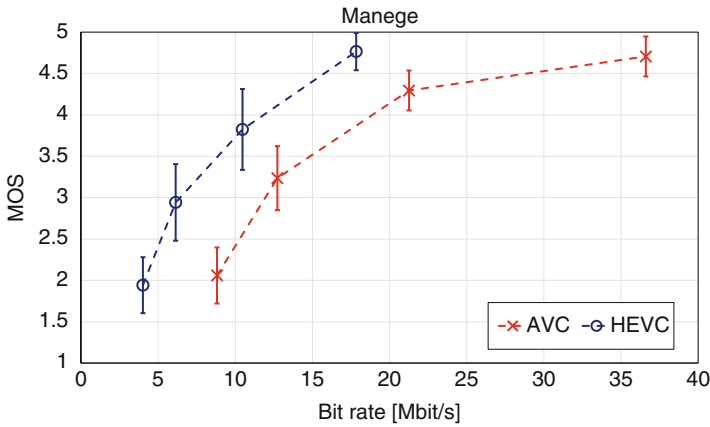


Fig. 9.14 R–D curve (Manege)

for subjective scores:  $\Delta R$  is computed from the MOS;  $[\Delta R_{\min}, \Delta R_{\max}]$  provide a confidence interval on  $\Delta R$  and is determined considering the confidence index (CI) computed on the subjective scores; the confidence index takes into account the spreading of the MOS over the rating scale and the goodness of the fit of the values (Table 9.10).

For visual quality evaluation of CTC test sequences, interested readers are referred to [19], in which results of subjective tests are reported. The reported results indicate that a bit rate reduction of 50 % can be achieved for the example video test set.

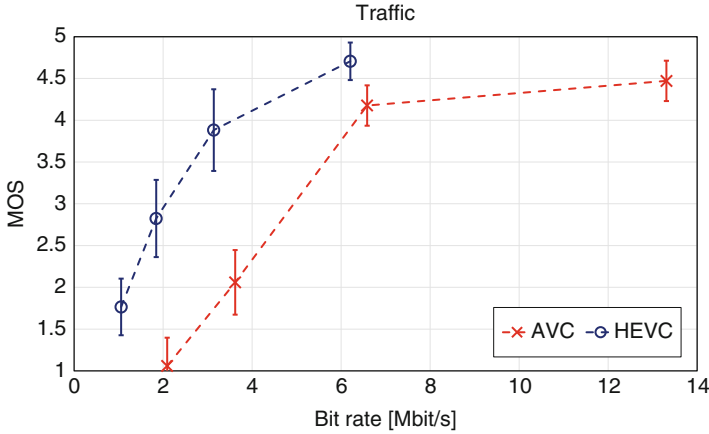


Fig. 9.15 R–D curve (Traffic)

Table 9.10 Bit rate differences of tested bitstreams

Sequence	$\Delta R$ (%)	$[\Delta R_{\min}, \Delta R_{\max}]$	Confidence index (%)
Book	-62	[-75 %, -51 %]	100
BT709Birthday	-71	[-81 %, -60 %]	100
HomelessSleeping	-87	[-94 %, -71 %]	100
Manege	-43	[-60 %, -17 %]	88
Traffic	-55	[-68 %, -39 %]	100
Average	-64	[-76 %, -48 %]	98

## 9.5 Production–Quality Encoder Performance Analysis

This section presents the results of an informal subjective quality comparison between the eBrisk-UHD and x264 [23] production–quality encoders, which were configured to be conformant with HEVC Main Profile and AVC High Profile, respectively. The encoder comparison presented in this section is intended to complement the subjective quality comparisons discussed earlier in this chapter in which HEVC and AVC encoder reference software were used.

### 9.5.1 Test Conditions

In this section, the test conditions, including the encoder configuration and evaluation conditions (e.g., sequence presentation details, viewing equipment, lighting conditions), and tested video sequences are described.

**Table 9.11** Video sequences used for the subjective quality comparison in order of presentation (top to bottom)

Resolution	Sequence name	Frame count	Frame rate (fps)	Bit depth
3,840 × 2,160	IntoTrees	400	50	8
3,840 × 2,160	OldTownCross	400	50	8
3,840 × 2,160	ParkJoy	400	50	8
3,840 × 2,160	DucksTakeOff	400	50	8

### 9.5.1.1 Encoder Settings

The encoders were configured for high coding efficiency operation. More specifically, the HEVC encoder was configured for Main profile and to use a prediction structure similar to that described in Sect. 9.2.3.2 with the period of the intra pictures set to 48 frames. The AVC encoder was configured to use default parameter values except for those parameters that required an explicit setting (e.g., the keyint parameter, used to specify the intra frame period, was set to 48). Several AVC encodings were performed for each video sequence, each with a different quantization parameter (QP) value. In each case, the encoding that yielded an average bit rate closest to 2.5 times that of the corresponding HEVC encoding was selected (i.e., the bit rate of the HEVC encoded sequence was approximately 60 % lower than that of the AVC encoded sequence).

### 9.5.1.2 Subjective Evaluation Conditions

Twenty-seven (27) volunteer viewing subjects were used for the subjective experiments. Fifteen (15) of the viewers had little or no previous experience in evaluating video sequences. The viewing was conducted in a somewhat-darkened room using a XBR55X900A 55" UHD Sony Bravia LED monitor. Additional key elements of the subjective evaluation and video presentation methodology are listed below:

1. Untrained viewers participated one at a time with each viewer seated in a chair that was positioned approximately 1.5 meters from the monitor and centered.
2. For each of the four test sequences listed in Table 9.11, the HEVC and AVC encoded video sequences were cropped in the horizontal dimension and spliced together side-by-side.<sup>2</sup> The relative position of the compared encodings was randomized and the experiments were executed in a double-blind manner.
3. The sequences were displayed at their native spatial resolutions; however, the display rate was set to 30 frames per second (fps) (i.e., the sequences were displayed at 60 % of their native 50 fps frame rate).

<sup>2</sup>The horizontal cropping removed one half of the pixels so as to enable both the HEVC and AVC encoded sequences to be displayed side-by-side on the 3840 pixel-wide 4K monitor.



4. The video sequences were presented to each viewer between one and three times and the viewers were asked to assess the relative quality of the side-by-side encodings according to a 5° scale: *left better*, *left slightly better*, *no preference*, *right slightly better* or *right better*.<sup>3</sup> The viewers were provided approximately 30 seconds between the presentation of each of the spliced video sequences to record their preferences. A total of seven (7) video pairs were presented and each viewing session had a duration of approximately seven (7) minutes.

### 9.5.1.3 Test Sequences

The four video sequences described in Table 9.11 were used in the subjective experiments. The YCbCr 4:2:0 8-bit 4K video sequences have a variety of characteristics typical of what might be encountered in the context of a streaming application, and they were selected from those that are generally available and used for video coding test purposes. All the sequences have a native frame rate of 50 fps. The 4K sequences were displayed at 30 frames per second; the highest frame rate at which the monitor is capable of displaying 4K content. The order of presentation followed that shown in Table 9.11, from top to bottom.

## 9.5.2 Subjective Quality Assessment Results

Table 9.12 shows the results of subjective video quality assessments. For each of the video sequences, the average bit rates of the HEVC and AVC encoders are shown along with the viewers' assessments of the video sequences, according to the 5° scale described in Sect. 9.5.1.2. Each video sequence was encoded using a fixed QP. The QP values for the encodings were selected to yield good, but not especially high video quality, in order to avoid viewing scenarios where either both encodings would yield indistinguishably excellent quality or both encodings would yield substantial coding artifacts.

## 9.5.3 Results

The subjective results presented in Table 9.12 show that the viewers either had no preference or favored the HEVC encoded video at a bit rate that was approximately

---

<sup>3</sup>A 5° scale was selected for this study in lieu of more commonly used 7° and 3° scales, because it is known to work well to prioritize video quality improvement work during commercial encoder development.

**Table 9.12** Subjective viewing comparison results for sequences encoded using the HEVC and AVC encoders (*B* better, *SB* slightly better, *NP* no preference)

Sequence name	HEVC QP	HEVC bit rate	AVC QP	AVC bit rate	HEVC B	HEVC SB	NP	AVC SB	AVC B
InToTrees	34	4.5 Mbps	35	11.9 Mbps	21	6	0	0	0
OldTownCross	32	3.1 Mbps	34	7.8 Mbps	14	3	3	4	3
ParkJoy	38	11.9 Mbps	38	29.0 Mbps	0	2	7	10	8
DucksTakeOff	37	15.4 Mbps	39	37.6 Mbps	12	4	3	6	2
Selection count totals					47	15	13	20	13
Percent totals (%)					43.5	13.9	12.0	18.5	12.0

60 % lower than that of AVC in 69.4 % of the trials.<sup>4</sup> These results are consistent with the subjective results reported in [8]. In addition, comparing the coding efficiency gains of HEVC relative to AVC for certain 4K sequences in this study with those gains reported in earlier studies (e.g., [10]), in which high-quality resampled (lower-resolution) versions of the same video sequences were used, it can be seen that the coding efficiency gains of HEVC relative to AVC are larger for the 4K sequences.<sup>5</sup> This comparison suggests that the increased coding efficiency gains for HEVC compared with AVC observed for 4K sequences cannot be explained solely by differences in content.

## 9.6 Conclusions

In this chapter, performance analysis of HEVC in comparison with AVC in terms of objective as well as subjective quality assessments are given. Because of the increased flexibility offered by HEVC, methods to select the best coding parameters, in a rate–distortion sense, are also described. Special care has been taken to apply a unified approach when conducting subjective and objective quality evaluations between HEVC and AVC. Both objective and subjective tests results indicate significant gains in compression efficiency of HEVC over AVC. More specifically, the bit rate reduction, based on objective evaluation of CTC test sequences, indicates an overall performance improvement of about 22 % for AI, 43 % for RA, 37 % for LDB and 35 % for LDP over AVC. Furthermore, by using non-CTC test sequences, we observe up to 76 % improvement in coding efficiency, as indicated in Table 9.9. Results of subjective evaluation tests indicate that an even higher bit rate saving in the ranges of 55–87 % can be achieved. The informal visual quality evaluation test results also confirm that HEVC yields a substantial improvement in compression capability beyond that of AVC for video streaming applications. It is also suggested that the coding performance gains of HEVC over AVC generally increase with increasing video resolution up to at least 4K resolutions.

**Acknowledgments** Regarding Sect. 9.4, the authors wish to thank 4EVER consortium, British Broadcasting Corporation (BBC), Kamerawerk GmbH and Technicolor for providing the original video test sequences used for visual testing. Our special thanks and gratitude go also to T. K. Tan, of NTT Docomo, for his help and support in the preparation of coded video test sequences used for subjective performance evaluations in Sect. 9.4. Subjective evaluations performed at EPFL were possible thanks to the Swiss National Foundation for Scientific Research (FN 200021-143696-1), EC funded Network of Excellence VideoSense, and COST Action IC1003 European Network on Quality of Experience in Multimedia Systems and Services QUALINET. Efforts by Hiromi Nemoto for set up and conducting subjective evaluations at EPFL are also acknowledged.

---

<sup>4</sup>The 69.4 % result was derived by summing the selection counts in the HEVC B, HEVC SB and NP columns of Table 9.12 and dividing by the total number of selection counts.

<sup>5</sup>The earlier studies also used earlier versions of the HEVC draft standard and a slightly different evaluation methodology.

## References

1. Bjøntegaard G (2001) Calculation of average PSNR differences between RD-curves, ITU-T SG16 Q6 Video Coding Experts Group (VCEG), Document VCEG-M33, Austin, Apr. 2001
2. Bossen F (2011) Excel template for BD-rate calculation based on Piece-wise Cubic Interpolation, JCT-VC Reflector
3. Bossen F (2013) Common test conditions and software reference configurations, Joint Collaborative Team on Video Coding (JCT-VC), Document JCTVC-L1110, Geneva, Jan. 2014
4. Cover TM, Thomas JA (1991) Elements of information theory, Chapter 13. Wiley, New York
5. Flierl M, Girod B (2003) Generalized B pictures and the draft H.264/AVC video compression standard. *IEEE Trans Circuits Syst Video Technol* 13(7):587–597
6. H.264/MPEG-4 AVC Reference Software, Joint Model 18.5: <http://iphome.hhi.de/suehring/tml/download/jm18.5.zip>
7. Hanhart P, Ebrahimi T (2014) Calculation of average coding efficiency based on subjective quality scores. *J Visual Commun Image Represent* 25(3):555–564
8. Hanhart P, Rerabek M, De Simone F, Ebrahimi T (2012) Subjective quality evaluation of the upcoming HEVC video compression standard. In: *Proc. SPIE*. 8499, Applications of Digital Image Processing XXXV, no. 84990V, Oct. 2012
9. Hanhart P, Rerabek M, Korshunov P, Ebrahimi T (2013) Subjective evaluation of HEVC intra coding for still image compression. In: Seventh international workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM), Scottsdale, Arizona
10. Horowitz M, Kossentini F, Mahdi N, Xu S, Guermazi H, Tmar H, Li B, Sullivan GJ, Xu J (2012) Informal subjective quality comparison of video compression performance of the HEVC and H.264/MPEG-4 AVC standards for low-delay applications. In: *Proc. SPIE*. 8499, Applications of Digital Image Processing XXXV, no. 84990W, Oct. 2012
11. ITU-R Rec. BT.500-11 (2006) Methodology for the subjective assessment of the quality of television pictures
12. ITU-T Rec. P.910 (2008) Subjective video quality assessment methods for multimedia applications
13. ITU-R BT.2022 (2012) General viewing conditions for subjective assessment of quality of SDTV and HDTV television pictures on flat panel displays
14. ITU-R BT.500-13 (2012) Methodology for the subjective assessment of the quality of television pictures
15. McCann K, Bross B, Han WJ, Kim IK, Sugimoto K, Sullivan GJ (2013), High Efficiency Video Coding (HEVC) Test Model 13 (HM 13) Encoder Description, Joint Collaborative Team on Video Coding (JCT-VC), Document JCTVC-O1002, Geneva, Oct. 2013
16. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11. HM12.1 Reference Software: [https://hevc.hhi.fraunhofer.de/svn/svn-HEVCSoftware/\(mainsite\)](https://hevc.hhi.fraunhofer.de/svn/svn-HEVCSoftware/(mainsite)) and [http://hevc.kw.bbc.co.uk/svn/jctvc-a124/\(mirrorsite\)](http://hevc.kw.bbc.co.uk/svn/jctvc-a124/(mirrorsite))
17. Karczewicz M, Ye Y, Chong I (2008) Rate distortion optimized quantization, ITU-T SG16 Q6 Video Coding Experts Group (VCEG), Document VCEG-AH21, Antalya, Jan. 2008
18. Nguyen T, Marpe D (2012) Performance analysis of HEVC based Intra coding for still image compression. In: *PCS2012*, May 2012, pp 233–236
19. Ohm J-R, Sullivan GJ, Schwarz H, Tan TK, Wiegand T (2012) Comparison of the coding efficiency of video coding standards –including High Efficiency Video Coding (HEVC). *IEEE Trans Circuits Syst Video Technol* 22(12):1669–1684
20. Ortega A, Ramchandran K (1999) Rate-distortion methods for image and video compression: an overview. *IEEE Signal Process J* 23–50
21. Schwarz H, Marpe D, Wiegand T (2005) Hierarchical B pictures, Joint Video Team (JVT), Document JVT-P014, Poznan, July 2005
22. Sullivan GJ, Wiegand T (1999) Rate-distortion optimization for video compression. *IEEE Signal Process J* 74 -90
23. VideoLANx264 SW library (2013) <http://www.videolan.org/developers/x264.html>. Version core 135 r2345, 30 July 2013