Article

A Novel Consistent Quality Driven for JEM Based Distributed Video Coding

Dinh Trieu Duong [1], Huy Phi Cong [1,2,3,4] and Xiem Hoang Van [1,*]

[1]   VNU—University of Engineering and Technology, Hanoi 100000, Vietnam
[2]   Posts and Telecommunications Institute of Technology, Hanoi 100000, Vietnam
[3]   JTIRC, VNU University of Engineering and Technology, Hanoi 100000, Vietnam
[4]   School of Electrical and Data Engineering, University of Technology Sydney, Sydney 2006, Australia
[*]   Correspondence: xiemhoang@vnu.edu.vn

check for updates

Abstract: Distributed video coding (DVC) is an attractive and promising solution for low complexity constrained video applications, such as wireless sensor networks or wireless surveillance systems. In DVC, visual quality consistency is one of the most important issues to evaluate the performance of a DVC codec. However, it is the fact that the quality of the decoded frames that is achieved in most recent DVC codecs is not consistent and it is varied with high quality fluctuation. In this paper, we propose a novel DVC solution named Joint exploration model based DVC (JEM-DVC) to solve the problem, which can provide not only higher performance as compared to the traditional DVC solutions, but also an effective scheme for the quality consistency control. We first employ several advanced techniques that are provided in the Joint exploration model (JEM) of the future video coding standard (FVC) in the proposed JEM-DVC solution to effectively improve the performance of JEM-DVC codec. Subsequently, for consistent quality control, we propose two novel methods, named key frame quantization (KF-Q) and Wyner-Zip frame quantization (WZF-Q), which determine the optimal values of the quantization parameter (QP) and quantization matrix (QM) applied for the key and WZ frame coding, respectively. The optimal values of QP and QM are adaptively controlled and updated for every key and WZ frames to guarantee the consistent video quality for the proposed codec unlike the conventional approaches. Our proposed JEM-DVC is the first DVC codec in literature that employs the JEM coding technique, and then all of the results that are presented in this paper are new. The experimental results show that the proposed JEM-DVC significantly outperforms the relevant DVC benchmarks, notably the DISCOVER DVC and the recent H.265/HEVC based DVC, in terms of both Peak signal-to-noise ratio (PSNR) performance and consistent visual quality.

Keywords: distributed video coding; joint exploration model (JEM); future video coding (FVC); visual quality consistency; DISCOVER DVC

1. Introduction

Video coding technologies have been playing an important role in the context of audiovisual services, such as digital TV, mobile video, and internet streaming, to cope with the high compression requirements. Most of the available video coding standards, notably the ITU-T H.26x and ISO/IEC MPEG-x standards [1], adopted the so-called predictive video coding paradigm, where the temporal and spatial correlations are exploited at the encoder by using motion estimation/motion compensation and spatial transforms, respectively. As a result, these coding standards typically lead to rather complex encoders and a much simpler decoder.

However, a very low complexity at the encoder becomes a more essential feature for video coding schemes with the explosion of emerging applications, such as low-power video surveillance,

wireless visual sensor networks, and wireless PC/Mobile cameras [2]. At the light of an information theory resulted in the 1970s: The Slepian-Wolf theorem [3] distributed video coding (DVC) can be considered as a new and efficient paradigm to address the requirement of these emerging applications.

In [3], Slepian-Wolf et al. proved that two correlated sources could be independently encoded without any loss in the coding efficiency if they are jointly decoded by exploiting source statistics. In 1976, Wyner-Ziv theorem further proved a similar source coding method while using side information for lossy compression [4]. Video coding schemes in DVC can effectively exploit the temporal correlation at the decoder rather than at the encoder without rate-distortion performance penalty based on Slepian-Wolf and Wyner-Ziv theorems; this allows moving motion estimation and its complexity to the decoder, thus achieving simpler encoders.

Generally, there are two main practical approaches to the DVC design: the DVC Stanford [5] and the DVC Berkeley [6] solutions. As of today, the most popular DVC codec design in the literature is the DVC Stanford architecture, which works at the frame level and it is characterized by a feedback channel based decoder rate control [7].

In the DVC Stanford architecture, the input video is separated into two parts: the key and WZ frames. The key frames are encoded while using the conventional Intra coding and the WZ frames are coded while using the WZ coding scheme. In the case of WZ coding, the outputs of the DVC encoder are parity bits; however, only a part of these parity bits are sent to the decoder to improve the coding efficiency of the DVC codec.

Several approaches have been introduced to improve the performance of the DVC Stanford architecture. In [8], Brites et al. proposed a realistic WZ coding approach by estimating the correlation noise model (CNM) at the decoder for efficient pixel and transform domain WZ coding. In [9], the authors et al. presented a novel H.265/HEVC based DVC, in which the low complexity HEVC Intra profile is employed to encode the key frames. Among the different DVC codecs following the Stanford architecture, DISCOVER DVC [10], as developed by the European project, is one of the best DVC codec reported in the literature. A thorough performance benchmark of this codec is available in [11], where the coding efficiency of the DISCOVER DVC codec is compared to two variants of H.264/AVC with low encoding complexity: H.264/AVC Intra and H.264/AVC No motion [12].

However, we have not compared our approach to a recent video coding that is based on second generation wavelets [13]—so it is not DCT-based—even if it is very efficient in low bit rates, but it may compromise our goal of developing a coder for low complexity constrained video applications.

The problem of providing consistent video quality for the DVC decoded frames is still limited so far, although many approaches have been introduced in the literature to effectively improve the performance of DVC codecs. In [14], Girod et al. proposed a fixed set of seven quantization matrices (QMs), which were defined for effectively encoding the WZ frames while keeping a smooth quality transition between the decoded key and WZ frames. This work is extended in [15] to obtain an additional QM to be added to the fixed set and then applied for verifying the performance of the IST-Transform domain Wyner-Ziv codec (IST-TDWZ). However, regarding to the consistent video quality problem, these approaches are only effective for a limited number of test sequences and specific video coding standards that are applied for DVC codecs, such as H.263+ or H.264/AVC [12] and not so effective for the H.265/HEVC [16] and the new future video coding standard (FVC) [17]. As reported in the experimental results that were obtained for the HEVC based DVC [9], the performance of DVC codec in this case is much better than that of the H.264/AVC based DVC; however, the fluctuation of video quality on the decoded key and WZ frames are very strong and it thus degrades the user's quality of experience.

We propose a novel DVC solution named Joint exploration model based DVC (JEM-DVC) in this paper to solve the problem, which can provide not only higher performance as compared to the traditional DVC solutions, but also an effective scheme for the quality consistency control. In the proposed JEM-DVC solution, several advanced techniques that were provided in the Joint exploration model (JEM) of FVC standard have been employed to effectively improve the performance of

JEM-DVC codec. Subsequently, we propose two novel methods, named key frame quantization (KF-Q) and Wyner-Zip frame quantization (WZF-Q), which determine the optimal values of quantization parameter (QP) and quantization matrix (QM) applied for the key and WZ frame coding for consistent quality control, respectively. Unlike the conventional approaches, the optimal values of QP and QM are adaptively controlled and updated for every key and WZ frames to guarantee the consistent video quality for the proposed codec. Our proposed JEM-DVC is the first DVC codec in literature employing the JEM coding technique, and then all results that are presented in this paper are new. The experimental results show that the proposed JEM-DVC significantly outperforms the relevant DVC benchmarks, notably the DISCOVER DVC and the recent H.265/HEVC based DVC, in terms of both Peak signal-to-noise ratio (PSNR) performance and consistent visual quality.

The rest of the paper is organized, as follows. Section 2 summarizes the background works on DVC. Section 3 describes the proposed JEM-DVC solution in detail. The experimental results are discussed in Section 4. Finally, Section 5 concludes the paper.

2. Relevant Background Works

DVC has attracted significant attention in the last decade due to its specific coding features, notably a flexible distribution of the codec complexity, inherent error resilience, and codec independent scalability [18]. In this section, we briefly describe the state-of-the-art DVC codec, namely DISCOVER DVC, and recent works on DVC.

The DISCOVER DVC codec [10] is one of the most efficient codec in literature and the most widely used as a reference benchmark for evaluating the performance of a DVC codec [7] among the different DVC codecs following the Stanford architecture. In DISCOVER DVC, the input video is separated into two parts: the key and WZ frames. The key frames are encoded while using conventional H.264/AVC Intra coding, due to its low-encoding complexity, and the WZ frames are coded while using WZ coding scheme, as shown in Figure 1. Typically, the key frames are periodically inserted with a certain GOP size. Most of the results available in the literature use a GOP size of 2, which means that odd and even frames are the key and WZ frames, respectively. For each WZ frame coding, an integer 4×4 block-based Discrete cosine transform (DCT) is applied prior to quantization. The quantized values are then split into bitplanes, which go through a low density parity check code (LDPC) encoder, as shown in Figure 1.
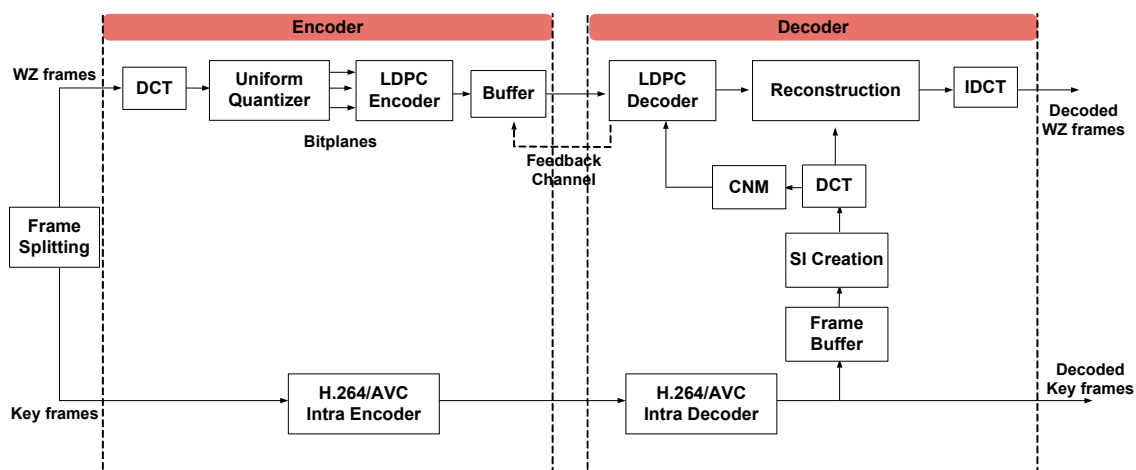

Figure 1. DISCOVER distributed video coding (DVC) architecture.

At the decoder, the use of motion-compensated interpolation or extrapolation schemes that are based on the previously decoded frames generates the side information estimated for the WZ frames. Along with the parity bits generated in the WZ coding, this side information is utilized in the LDPC

decoder to reconstruct the encoded bitplanes, and then to decode the WZ frames. More details regarding the description of DISCOVER DVC architecture can be referred to in [10].

Recently, there are several researches that have been introduced to improve the performance of DVC codecs, including the correlation noise modelling (CNM) [19], the side information improvement [20,21], and distributed scalable video coding (DSVC) [22], etc. Among them, the solutions for high performance key frame coding and consistent video quality are essential issues, since these issues play an important role not only as the most meaningful benchmark for low coding complexity, but also as the essential factors for improving the user's QoE. In the following sections, we describe the proposed methods for these issues in detail.

3. Proposed JEM Based DVC (JEM-DVC)

3.1. Proposed JEM-DVC Architecture

Figure 2 shows the architecture of our proposed JEM-DVC codec, where the highlighted blocks present the new coding methods that are proposed in this paper. As shown in Figure 2, the proposed JEM-DVC architecture also follows the Stanford DVC approach, specifically the DISCOVER DVC codec, which is described in the previous section. However, in the proposed architecture, instead of using the conventional H.264/AVC Intra coding, the proposed JEM-DVC codec employs several advanced coding techniques provided by the JEM/FVC [17]. Though the codec itself is not the core novelty of this paper, our proposed JEM-DVC codec is the first DVC codec in literature that employs JEM Intra coding, and thus all of the results presented in this paper are new. In the next subsections, we describe the proposed JEM based key frame coding and the consistent video quality control for the proposed JEM-DVC codec in detail.
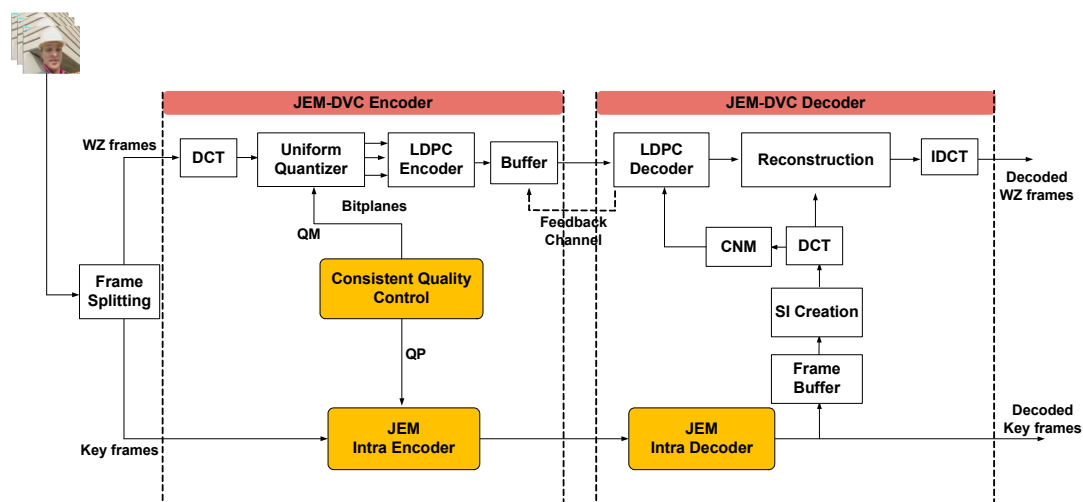

Figure 2. Proposed Joint exploration model based distributed video coding (JEM-DVC) architecture.

3.2. JEM Based Key Frame Coding

As mentioned in Section 1, the key frame coding plays an important role, not only as the most meaningful benchmark of low coding complexity, but also as the essential factor for improving the performance of DVC codecs. Generally, the key frame is encoded while using the Intra coding schemes that were provided by the traditional predictive coding standards, such as H.263, H.264/AVC, or H. 265/HEVC. In [9], the authors et al. showed that the DVC codec utilizing H.265/HEVC for the key frame coding can outperform that utilizing H.264/AVC by up to 2.5~3.0 dB. Therefore, it is expected that the more powerful and effective Intra coding technique applied for key frame coding, the better Rate-distortion performance can be provided for a DVC codec.

Recently, it is worth noticing that the best available video coding standard is no longer H.264/AVC or H.265/HEVC, but rather the FVC standard with JEM [17,23]. As reported in [24], the main goal of JEM is to explore the future video coding technology with a compression capability that significantly exceeds the current HEVC standard. In [25], Sidaty et al. has shown that the FVC/JEM codec can provide up to 40% coding efficiency gains for the Intra/Inter coding as compared to the H.265/HEVC. Additionally, for the Intra coding mode only, the gain that can be achieved is about 20%.

In this work, we propose a novel JEM based key frame coding scheme, being inspired from the recent achievements of the FVC/JEM research. To the best of our knowledge, this is the most up-to-date compression solution for the key frame coding. In addition, we also adaptively configure JEM Intra coding tools to effectively apply to the proposed JEM-DVC codec to meet the requirement of low encoding complexity.

Specifically, in JEM, the number of Intra prediction modes is extended to 67 modes, which include 65 angular modes plus DC and planar modes [16]. All of these prediction modes are available to favor all of the block sizes of the quad tree plus binary tree (QTBT) block structure. QTBT is a novel coding structure block that JEM introduces, which includes coding tree units (CTUs) and each CTU contains smaller coding units (CUs). In JEM software, it is possible to configure a large block size for a CTU (up to 256×256).

Similar to the prior video coding standards, like H.264/AVC and H.265/HEVC, the principle of defining an optimal Intra prediction mode in JEM is to recursively search to find the optimal coding tree structure. On every possible CTU size, a complex rate distortion optimization (RDO) is performed among all of the 67 prediction candidate modes, in order to screen out the best prediction mode [17]. These lead to a very complicated and time-consuming calculation, especially for such a high number of Intra prediction modes that are provided in JEM, and thus strongly against the low coding complexity that is required for a DVC encoder. In addition, the targets of our proposed JEM-DVC coder are usually the low complexity constrained video applications, e.g., wireless sensor networks, wireless surveillance, or remote sensing systems. Thus, the input videos of these emerging applications tend to have low resolutions with low motion activities. In this context, a large number of Intra prediction modes, as introduced in JEM, may not necessary.

Therefore, in this work, many experiments have been performed on several test sequences to empirically define the most suitable prediction modes and CTU sizes to effectively reduce the time that is consumed to find the best prediction mode. From experiments, it is confirmed that, in cases of low resolution applications, we can empirically define the number of Intra prediction modes and the size of CTU setup for all directional prediction modes as $N_{mode} = 35$ and $CTU_{size} = 64 \times 64$, respectively.

Figure 3 illustrates all of the prediction modes utilized for the proposed JEM based key frame coding in detail. In Figure 3, 33 angular modes are numbered from 2 to 34, and each represents a distinct direction. Among them, special modes 10 and 26 are referred to as Horizontal and Vertical, respectively. The non-angular modes 0 and 1 are Planar and DC, respectively, which are defined as the same as in the H.265/HEVC Intra coding [15]. Additionally, in the case of CU coding with several candidates of sizes, such as $CU_{size} = 64 \times 64$, 32×32, 16×16, and 8×8, it is reported from our experimental results that a larger CU_{size} is suitable for flat and homogeneous regions. Additionally, for these regions, DC, Planar, and simple directional modes (e.g., Vertical and Horizontal modes) are the most likely to be chosen as the best mode candidate. In contrast, a smaller CU_{size} is suitable for high activities and significantly detailed regions. Therefore, the exhaustive mode searching for all candidates of CU sizes is unnecessary, since it is very time consuming and complex.

As a good example to verify the above assessment, Figure 4 shows the rate of optimal mode selections that were obtained from our experiments (in the case of large CU size, $CU_{size} = 64 \times 64$ and 32×32). As shown in Figure 4, among a total number of 35 directional predictions, most of the decisions for the best prediction in this case are Planar, DC, Horizontal, and Vertical, which account for up to 91% of total mode candidates. Therefore, we can determine these prediction modes for large CU sizes ($CU_{size} = 64 \times 64$ and 32×32), while the remaining modes are decided for smaller CU sizes

($CU_{size} = 16 \times 16$ and 8×8). These can be referred to as the fast Intra mode decision scheme that is proposed for the JEM based key frame coding.
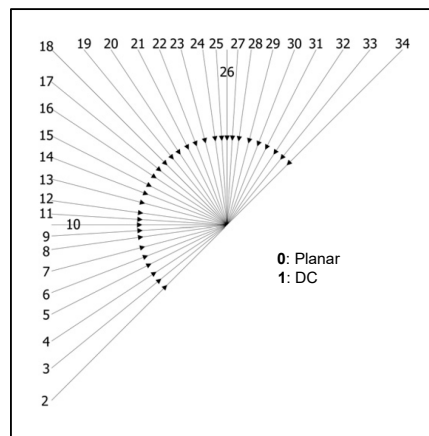

Figure 3. Directional prediction modes of Intra coding scheme proposed for JEM-DVC.
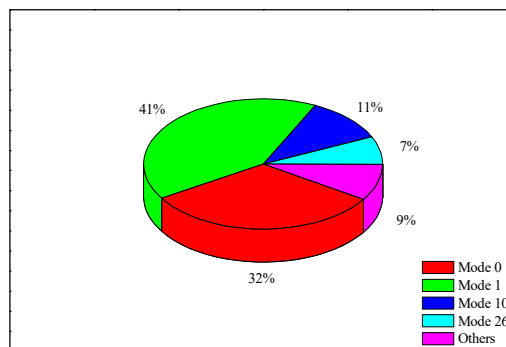

Figure 4. The rate of optimal mode selections in cases of large coding unit (CU) size ($CU_{size} = 64 \times 64$ and 32×32).

3.3. Proposed Consistent Quality Driven (CQD) for JEM-DVC

The scheme that can provide consistent video quality plays an important role for most advanced viewing systems, since viewers are usually sensitive to the fluctuation in the quality of decoded video frames. However, it is the fact that the quality of the decoded key frames is generally much better than that of the decoded WZ frames. This leads to the fluctuant or inconsistent video quality playing between the key and WZ frames, and thus reduces the user's QoE.

Figure 5 shows an example for the problem of quality fluctuation. In Figure 5, the subjective quality of the ith ($i = 14$, 15) decoded key and WZ frames of *Foreman* sequence are shown. These frames are encoded using the DISCOVER DVC coder. The error or distortion images in Figure 5e,g show the differences between the ith ($i = 14$, 15) original key and decoded key frames, respectively, and Figure 5f,h show the differences between the ith ($i = 14$, 15) original WZ and the decoded WZ frames, respectively. The decoded WZ frames generally contain much higher distortion areas as compared to the decoded key frames do, as illustrated in Figure 5. In other words, the image quality of the decoded key frame is generally much better than that of the decoded WZ frames.

The problem becomes more serious when the key frame is encoded while using modern video coding standards, such as H.265/HEVC and FVC/JEM, since the image quality of the decoded key frames in these cases is even much better than that of the decoded key frames that were obtained by the DISCOVER DVC codec.
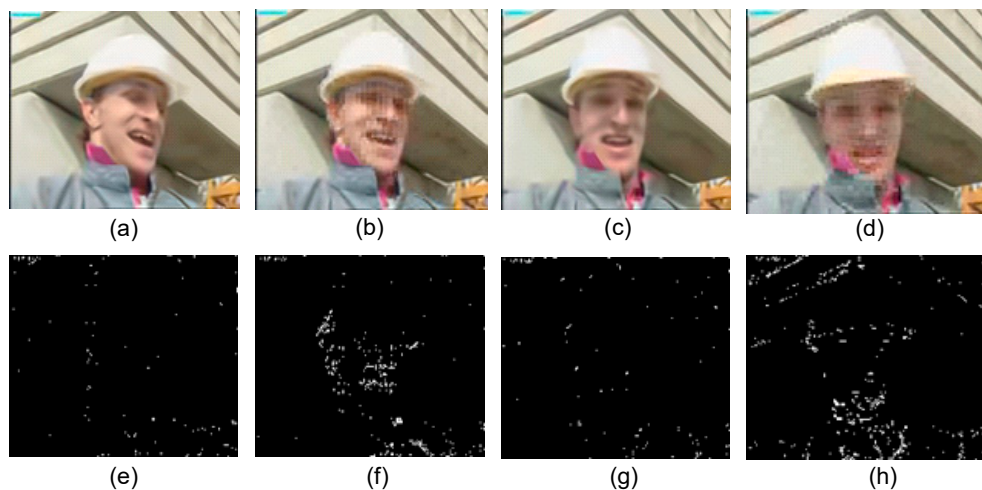
Figure 5. Inconsistent image quality introduced by DISCOVER DVC for the *Foreman* sequence: (**a**) 14th decoded key frame, (**b**) 14th decoded WZ frame, (**c**) 15th decoded key frame, (**d**) 15th decoded WZ frame. The bottom row shows the distortion images between the decoded frames and the original frames, respectively: (**e**) 14th key frame, (**f**) 14th WZ frame, (**g**) 15th key frame, and and (**h**) 15th WZ frame.

It is necessary to propose a new scheme that can control more effectively the quality of both key and WZ frames than the prior traditional schemes in order to solve the problem, especially for the new architecture employed in JEM-DVC. Generally, the coding rate and the quality of the decoded key and WZ frames are controlled while using QPs and QMs, respectively, where the key frame is encoded while using a set of QPs ranging from 0 to 51, while the WZ frame is encoded using the fixed set of eight QMs, $\{QM_i, i = 1, 2, .., 8\}$ [12]. Figure 6 shows the fixed set of QMs, $\{QM_i, i = 1, 2, .., 8\}$. In Figure 6, matrix QM_1 represents the lowest bitrate (and the highest distortion) situation, while matrix QM_8 corresponds to the highest bitrate (and thus lowest distortion). Within a 4×4 quantization matrix, the value at position k in Figure 6 indicates the number of quantization levels that are associated to the DCT coefficients band b_k. These quantization matrices are utilized to determine the rate-distortion performance of a DVC coder and they are assumed to be known by both the encoder and decoder. More details on the description of QMs can be referred to in [11].

In this work, we propose a novel scheme, named consistent quality driven (CQD), which includes two quantization methods to minimize the fluctuations in the quality of decoded frames: key frame quantization (KF-Q) and WZ frame quantization (WZF-Q). The KF-Q and WZF-Q methods can effectively determine the appropriate quantization values that were assigned for QP and QM, respectively. The proposed CQD scheme is described, as follows:

■ Key frame quantization (KF-Q): In the JEM-DVC architecture described in Section 3.1, key frames are encoded while using JEM Intra coding and the quality of the decoded key frame is typically controlled using a constant QP. Let QP_i be the QP value that is defined at the frame level to encode the ith key frame, K_i. Subsequently, in order to effectively reduce the quality fluctuation between the decoded key and WZ frames, QP_i is chosen, so that the average PSNR of the decoded key frames is similar to that of the decoded WZ frames. Therefore, in the case of JEM-DVC codec, a new set of quantization values for the JEM based key frame coding also need to be determined for QP_i, so that QP_i can adapt to the advanced coding tools that are provided in JEM. This is because JEM Intra coding is the most up to date coding scheme and it has not been employed for any DVC architecture before. Thus, several experiments have been performed on a lot of test video sequences to empirically define the appropriate set of quantization values that QP_i can be assigned to in the proposed KF-Q method. These test video sequences are chosen in the media collection [26], which contain different kinds of motions and texture activities, and encoded while using a different values setup for QP. Table 1 shows the most appropriate QP

values that were obtained from our experiments. These values can be referred to as the reference values, QP_{ref}, setup for QP_i to encode the key frame K_i in the JEM Intra coding scheme of the proposed JEM-DVC. It is confirmed from our experimental results that, when QP_i is assigned by QP_{ref}, $(QP_i = QP_{ref})$, the performance of the JEM-DVC codec is not only significantly improved for the key frame encoding, but it can also provide almost smooth video quality for the full set of key and WZ frames. That means that the quality fluctuation between the decoded key and WZ frames in this case can be effectively reduced for the proposed JEM-DVC codec.

Table 1. Reference quantization parameters (QPs) applied for the key frame coding in the JEM-DVC codec.

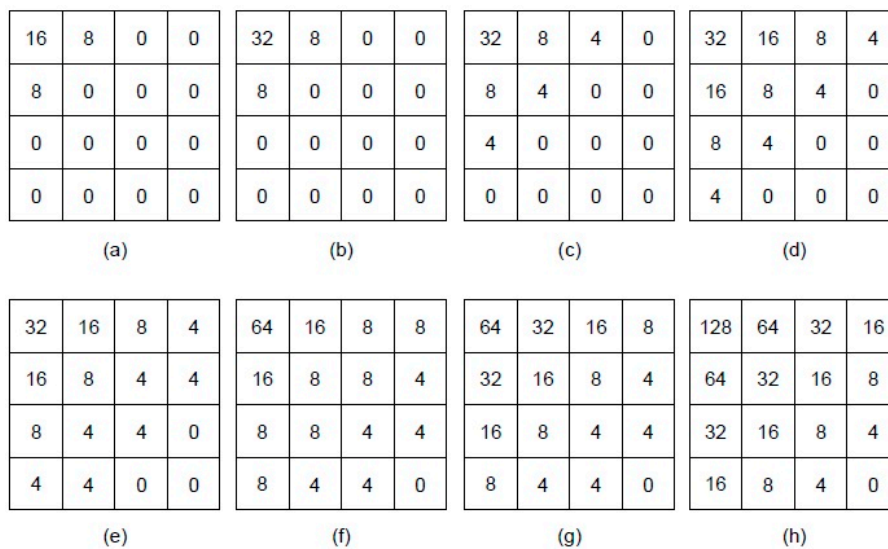|	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8
Hall Monitor	38	36	36	33	33	31	29	24
Coastguard	39	37	37	34	33	31	30	26
Foreman	42	40	38	34	34	32	29	25
Soccer	44	43	40	36	36	34	31	25

Figure 6. Fixed set of eight quantization matrices (QMs).

■　WZ frame quantization (WZF-Q): In this work, for further improving the performance of the proposed JEM-DVC codec and minimizing the PSNR fluctuations that were obtained for the key and WZ frames, we propose a WZF-Q method to define an optimal quantization matrix QM_{opt} applied for the WZ frame coding. Let QM_j be the quantization matrix that is defined to encode the jth WZ frame, WZ_j, in a group of picture (GOP). Without a loss of generality, the size of a GOP, N_{GOP}, is set to 2 ($N_{GOP} = 2$). Afterwards, in each GOP, the PSNR evaluation for the key frame K_i is measured as

$$P(K_i,\ QP_i) = 10log\frac{255^2}{\sigma_e^2} \tag{1}$$

where $P(K_i,\ QP_i)$ denotes the PSNR evaluation for the key frame K_i; QP_i is the QP value defined in the proposed KF-Q method to encode K_i; and, σ_e^2 denotes the Mean-square error between K_i and its decoded frame. Similarly, we can obtain the PSNR value, $P(WZ_j,\ QM_j)$, as measured for the WZ frame, WZ_j. Generally, for the WZ frame coding scheme, QM_j can be chosen among the eight reference matrices of QMs as indicated in Figure 6, $QM_j = QM_{ref}$ ($ref = 1,\ 2, \ldots,\ 8$). The chosen QM_j is the quantization matrix, so that the average PSNR measured for the decoded WZ frame is almost similar to that measured for the decoded key frame, and thus the quality fluctuation between these decoded frames is minimized. However, in the case of JEM-DVC architecture,

the QM matrix that was determined by the traditional approaches is generally not suitable for reducing the quality fluctuation problem. The reason lies in the fact that, when compared to the traditional DVC codecs based on H.263+ or H.264/AVC [10], the PSNR performance of the decoded key frame in the JEM-DVC codec is much higher and, thus, much different than that of the decoded WZ frames. Therefore, unlike the traditional approaches, in our proposed WZF-Q method, the high performance of PSNRs that was achieved for the key frame coding in JEM-DVC is frequently measured and updated to encode the WZ frames. Specifically, according to each PSNR evaluated for the key frame K_i, $P(K_i, QP_i)$, the PSNRs evaluated for the WZ frame WZ_j, $P(WZ_j, QM_j)$, are also measured and frequently updated for WZ_j with all the candidates of QM_j, $QM_j = QM_{ref}$ $(ref = 1, 2, \ldots, 8)$. Among different QM_j candidates defined for WZ_j coding, we choose the optimal QM_{opt} that can provide not only high coding efficiency for the WZ frame coding, but also minimize the quality fluctuation between the key and WZ frames. QM_{opt} is determined by

$$QM_{opt} = \underset{\{QM_j, \; j=1,2,\ldots,8\}}{\text{argmin}} \Delta PSNR(WZ_j, QM_j),$$

$$\text{where } \Delta PSNR(WZ_j, QM_j) = \left| P(K_i, \; QP_i) - P(WZ_j, \; QM_j) \right| \tag{2}$$

4. Performance Evaluation

Several experiments have been performed to illustrate the effectiveness of the proposed JEM-DVC architecture. In this work, the performance of the JEM-DVC codec in terms of objective, subjective, and consistent quality evaluations is compared with that of the relevant benchmarks, including the H.265/HEVC based DVC (H.265-DVC) [9] with H.265-DVC reference code in [27] and H.264/AVC based DVC or DISCOVER DVC [10] codecs. The DISCOVER DVC is one of the most efficient DVC solutions in literature and it is also the most widely used as the reference benchmark for evaluating the performance of a DVC codec.

For evaluation, the proposed JEM-DVC codec is tested on several video sequences, including *Foreman*, *Soccer*, *Coastguard*, and *Hall Monitor*. These test sequences are in 4:2:0 YUV format with CIF (352×288) resolution. The size of a GOP is set to 2 ($N_{GOP} = 2$). In our experiments, several values of QPs and QMs determined in the KF-Q and WZF-Q methods are employed to verify the PSNR performance of the key and WZ frame coding, respectively.

The performance of the proposed codec, including: (1) Key frame coding evaluation; (2) Consistent video quality evaluation; and, (3) Overall rate-distortion performance evaluation, are described in detail in the next subsections, as follows:

4.1. Key Frame Coding Evaluation

Firstly, we compare the performance of the key frame coding scheme in the proposed JEM-DVC with that in the H.265-DVC and DISCOVER DVC, because the performance of the key frame coding has a strong effect on the performance of these DVC codecs, as explained above.

In the proposed JEM-DVC, the key frames are encoded while using the JEM Intra coding in the H.265-DVC and DISCOVER DVC; the key frames are encoded using the H.265/HEVC and H.264/AVC Intra coding, respectively.

For comparison, Figure 7 shows the PSNR performance of the key frames that were obtained for the test video sequences with four values of QPs. As shown in Figure 7, the performance of the key frames that were encoded by the JEM Intra coding significantly outperforms that encoded by the H.265/HEVC and H.264/AVC for all of the test video sequences. For example, as shown in Figure 7, the JEM Intra coding provides up to 2.6 dB and 0.9 dB gains as compared with the H.264/AVC and H.265/HEVC Intra coding, respectively, for the *Foreman* sequence when QP = 27.
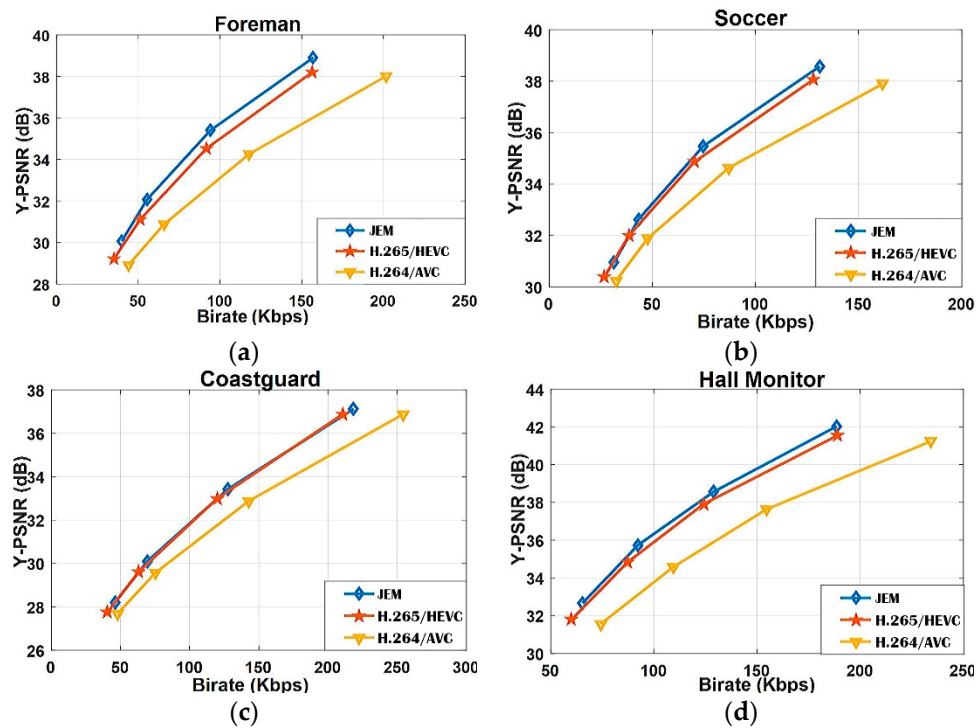
Figure 7. Key frame coding evaluation with different codecs: H.265-DVC, DISCOVER DVC, and the proposed JEM-DVC performed on different test sequences: (**a**) *Foreman*, (**b**) *Soccer*, (**c**) *Coastguard*, and (**d**) *Hall Monitor*.

4.2. Consistent Video Quality Evaluation

Several experiments have been performed with different values of QPs and QMs to show the effectiveness of our proposed consistent quality driven (CQD) method.

Figure 8 shows the results of the decoded key frame at the 14th position in the *Foreman* sequence that was obtained by DISCOVER DVC and the proposed JEM-DVC in terms of subjective quality evaluations.

Figure 8. Evaluation of visual quality consistency for the 14th decoded key frame of *Foremen* sequence obtained by (**a**) Original, (**b**) DISCOVER DVC, (**c**) H.265-DVC, and (**d**) proposed JEM-DVC. The bottom row shows the distortion images between the decoded key frames and the original frame, obtained by (**e**) DISCOVER DVC, (**f**) H.265-DVC, and (**g**) proposed JEM-DVC.

The error or distortion images in Figure 8e–g show the differences between the original key frame and the decoded frames. As shown in Figure 8, both the DISCOVER DVC, H.265, and JEM-DVC codecs can provide a good image quality with a small distortion for the decoded key frame as compared to the original frame. However, for the decoded WZ frame, the performance of DISCOVER DVC codec is critically degraded due to annoying artefacts that were seen around the face of the *Foreman*, as shown in Figure 9b. Correspondingly, the WZ coding scheme of the DISCOVER DVC generally introduces much higher distortion areas in the decoded WZ frames as compared to the key frame coding scheme. This leads to the fluctuation in the video quality playing over the decoded key and WZ frames.
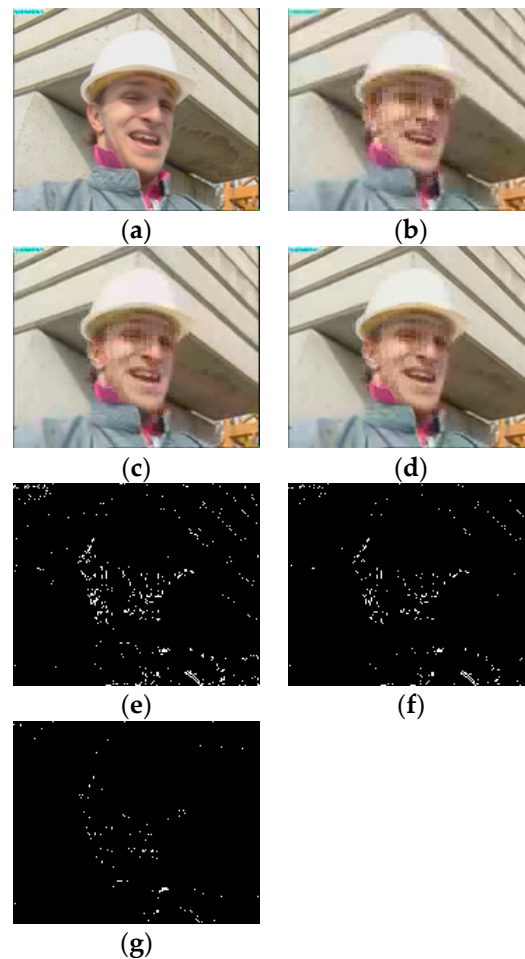

Figure 9. Evaluation of visual quality consistency for the 14th decoded WZ frame of *Foremen* sequence obtained by (**a**) Original, (**b**) DISCOVER DVC, (**c**) H.265-DVC, and (**d**) proposed JEM-DVC. The bottom row shows the distortion images between the decoded WZ frames and the original frame, obtained by (**e**) DISCOVER DVC, (**f**) H.265-DVC, and (**g**) proposed JEM-DVC.

In contrast to the DISCOVER DVC, for the proposed JEM-DVC, as can be seen in Figure 9d, the artifacts seen around the face of the *Foreman* have been effectively reduced, and then the distortion errors yielded in the decoded WZ frame are very small and it is quite similar to that yielded in the decoded key frame. This allows for a smooth and consistent video quality playing for both key and WZ frames for the proposed JEM-DVC codec.

Figure 10 shows more details on the fluctuation in the PSNR performance between the key and WZ frames. In Figure 10, the PSNR fluctuation is evaluated by measuring the different PSNR, $\Delta PSNR = |PSNR_K - PSNR_{WZ}|$, where $PSNR_K$ and $PSNR_{WZ}$ are the average PSNR obtained for the decoded key and WZ frames, respectively, at the different values of QPs and QMs setup for the DISCOVER DVC and JEM-DVC codecs. In Figure 10, $\Delta PSNR$ is measured based on four pairs

of different QPs and QMs, $\{(QP_i, QM_i), i = 1, 2, 3, 4\}$, which are configured for the key and WZ coding schemes of the DISCOVER DVC and JEM-DVC codecs. As expected, for all of the test video sequences, the proposed JEM-DVC can solve the problem of quality fluctuation more effectively than the DISCOVER DVC. For example, in the case of the *Soccer* sequence (Figure 10b), the fluctuation in the PSNR performance is reduced up to 2.1dB for the proposed JEM-DVC as compared to the prior DISCOVER DVC. This feature is very important for the proposed JEM-DVC codec, since it is confirmed that the JEM-DVC codec can provide an effective solution to improve the user's QoEs.
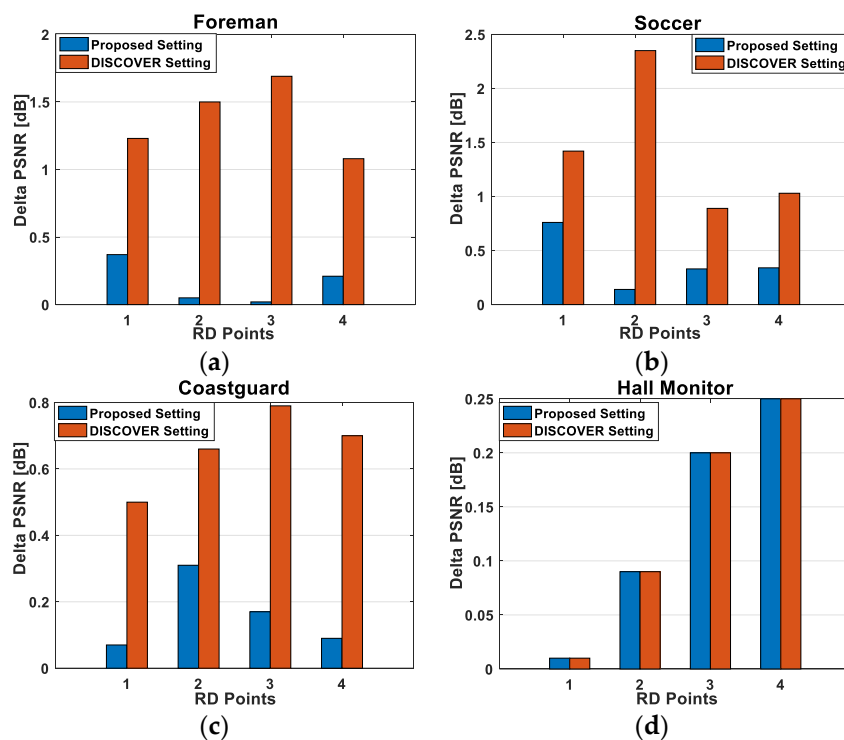

Figure 10. Fluctuation in the Peak signal-to-noise ratio (PSNR) performance between the key and WZ frames obtained by DISCOVER DVC and the proposed JEM-DVC with different test sequences: (**a**) *Foreman*, (**b**) *Soccer*, (**c**) *Coastguard*, and (**d**) *Hall Monitor*.

4.3. Overall Rate-Distortion Performance Evaluation

The key frame coding has a strong effect on the overall performance of a DVC codec, since the image quality of the reconstructed key frame directly affects the accuracy of side information (SI) creation [7], the reliable of a correlation noise model (CNM) [8], and finally, the quality of the decoded WZ frames. As can be seen in Figure 7, the coding efficiency of the key frame coding in the proposed JEM-DVC consistently outperforms the H.265-DVC and DISCOVERY DVC. Subsequently, it is expected that the overall rate-distortion or PSNR performance of the proposed JEM-DVC, including the PSNR performance of both the key and WZ frames is also higher than that of other conventional codecs.

Figure 11 shows the overall PSNR performance of the H.265-DVC, DISCOVERY DVC, and JEM-DVC codecs obtained for *Foreman*, *Soccer*, *Coastguard*, and *Hall Monitor* test sequences. As can be seen in Figure 11, for all of the test video sequences, the proposed JEM-DVC consistently provides better performance than the H.265-DVC and DISCOVERY DVC codecs. For example, as indicated in Figure 11c, for the *Hall Monitor* sequences, the proposed JEM-DVC provides up to 0.4 dB and 2.1 dB gains as compared to the H.265-DVC, DISCOVER DVC, respectively.
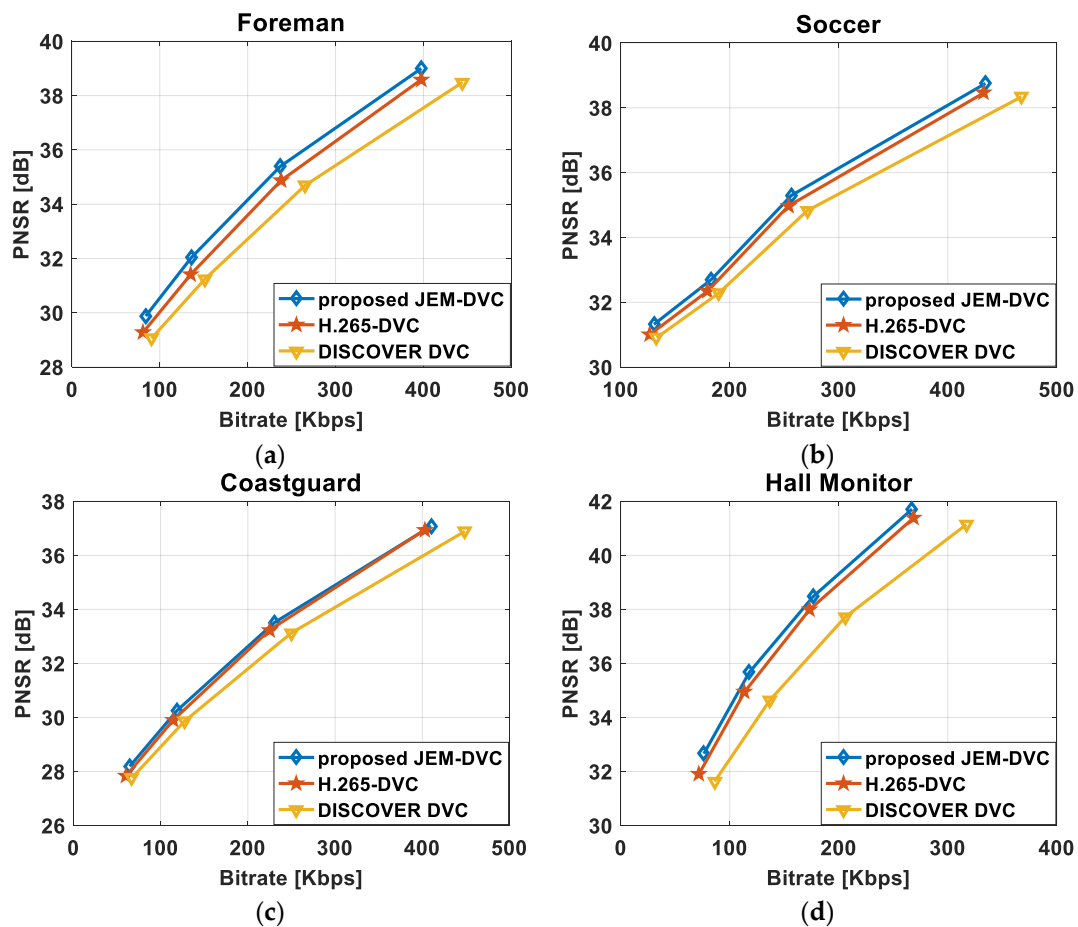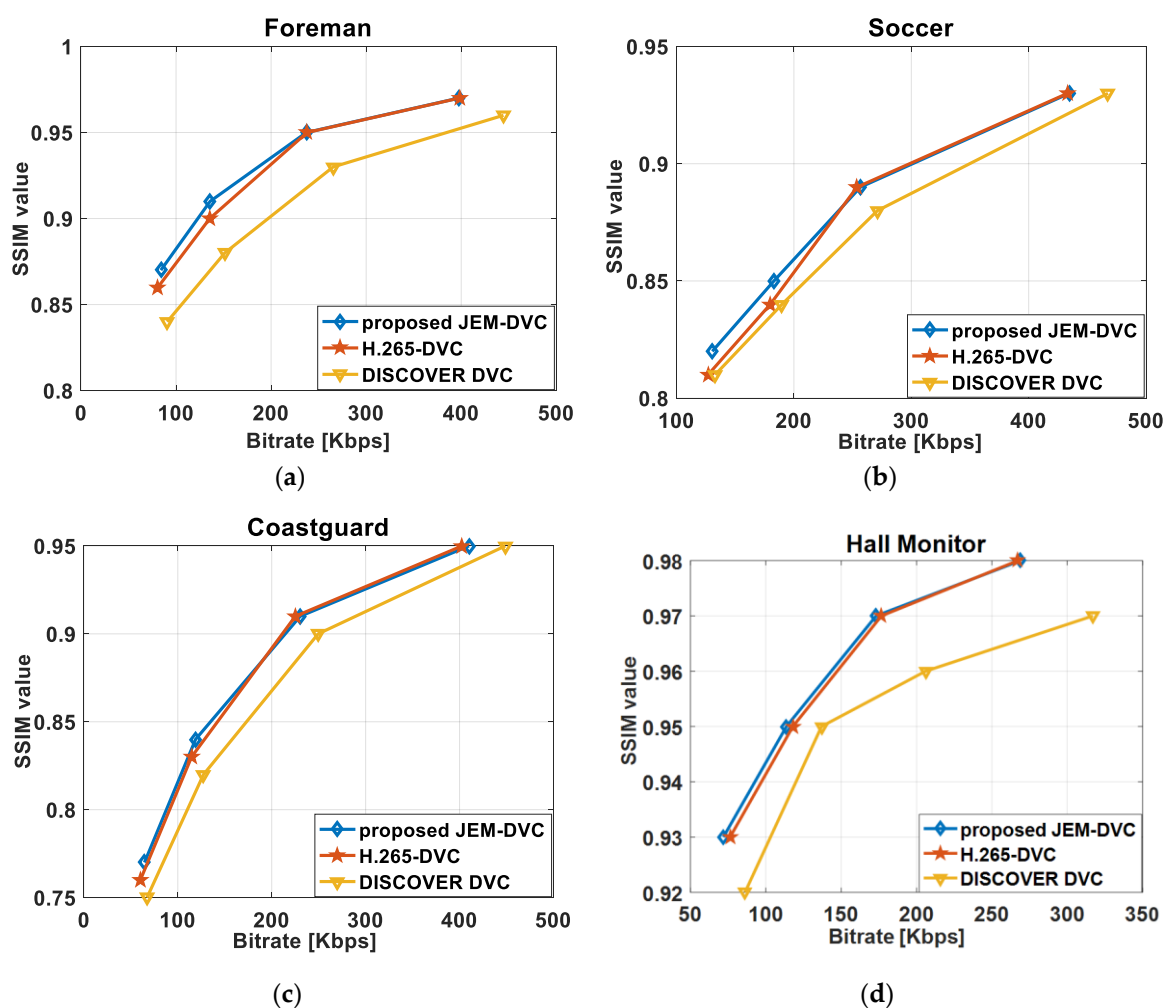
Figure 11. Comparisons of overall rate-distortion (RD) performance obtained by H.265-DVC, DISCOVER DVC, and the proposed JEM-DVC with different test sequences: (**a**) *Foreman*, (**b**) *Soccer*, (**c**) *Coastguard*, and (**d**) *Hall Monitor*.

In terms of structural similarity (SSIM) index, which is an efficient metric for visual quality measurement [28], the proposed JEM-DVC can also provide higher SSIM performance than the conventional DVCs. As shown in Figure 12, for all the test sequences, the proposed JEM-DVC always achieve higher SSIM values for the reconstructed key and WZ frames as compared to the H.265-DVC and DISCOVERY DVC.

In terms of the Bjontegaard metric [29], which can more precisely present the bitrate and PSNR gains; Table 2 shows the BD-Rate and BD-PSNR of the proposed JEM-DVC, H.265-DVC, and DISCOVER DVC codecs. As reported in Table 2, for the BD-Rate, the proposed JEM-DVC can gain about 5% and 17% bitrates saving as compared to the H.265-DVC and DISCOVER DVC, respectively. Additionally, for the BD-PSNR, the proposed JEM-DVC can achieve up to 1.21dB higher coding efficiency when compared to the DISCOVER DVC codec.

Table 2. Comparisons of BD-Rate and BD-PSNR [29] between the proposed JEM-DVC with H.265-DVC and DISCOVER DVC.

Sequences	BD-Rate		BD-PSNR	
	JEM-DVC vs. H.265-DVC (anchor)	JEM-DVC vs. DISCOVER DVC (anchor)	JEM-DVC vs. H.265-DVC (anchor)	JEM-DVC vs. DISCOVER DVC (anchor)
Foreman	−8.56	−20.38	0.54	1.37
Soccer	−3.76	−11.39	0.25	0.76
Coastguard	−3.00	−14.10	0.15	0.74
Hall	−5.29	−23.68	0.40	1.98
Average	**−5.15**	**−17.38**	**0.33**	**1.21**

Figure 12. Comparisons of overall visual quality measurement using SSIM metric obtained by H.265-DVC, DISCOVER DVC, and the proposed JEM-DVC with different test sequences: (**a**) *Foreman*, (**b**) *Soccer*, (**c**) *Coastguard*, and (**d**) *Hall Monitor*.

4.4. Complexity Performance Evaluation

In this work, we also evaluate the encoding complexity of the proposed JEM-DVC and other DVC solutions, since the encoding complexity is an important factor for most low-complexity constrained video applications, such as wireless sensor networks or wireless surveillance systems. Based on the processing time that is consumed for encoding video frames at the DVC encoder, the encoding

complexity is evaluated in terms of encoding times for the proposed JEM-DVC and compared with that of the relevant benchmarks, notably: JEM-Intra and JEM-No motion. For detailed references, the complexities of DISCOVER DVC and H.265-DVC are also calculated and compared with that of H.264-Intra, and H.264-No motion; H.265-Intra, and H.265-No motion, respectively.

In our experiments, the input video sequences are encoded while using a single personal computer (PC), which utilizes an Intel core i7-7700HQ (2.8 GHz) processor, 16GB RAM, and Windows 10- Home OS. Each video sequence has a long of 150 frames and the processing time of all the encoded frames in a sequence is calculated for complexity performance evaluations.

Figure 13 reports the complexity performance of the proposed JEM-DVC, H.265-DVC, and DISCOVER DVC. As shown in Figure 13, the proposed JEM-DVC can save up to 38% encoding time as compared to that of the JEM-Intra and JEM-No motion modes. The proposed JEM-DVC introduces relatively higher encoding times as compared to the DISCOVER DVC and H.265-DVC; however, with the breakthrough on high power IC chip designs nowadays and the high coding efficiency gained by the JEM-DVC, it is worth inspiring and considering the proposed JEM-DVC as an effective solution for a new DVC based video coding codec.
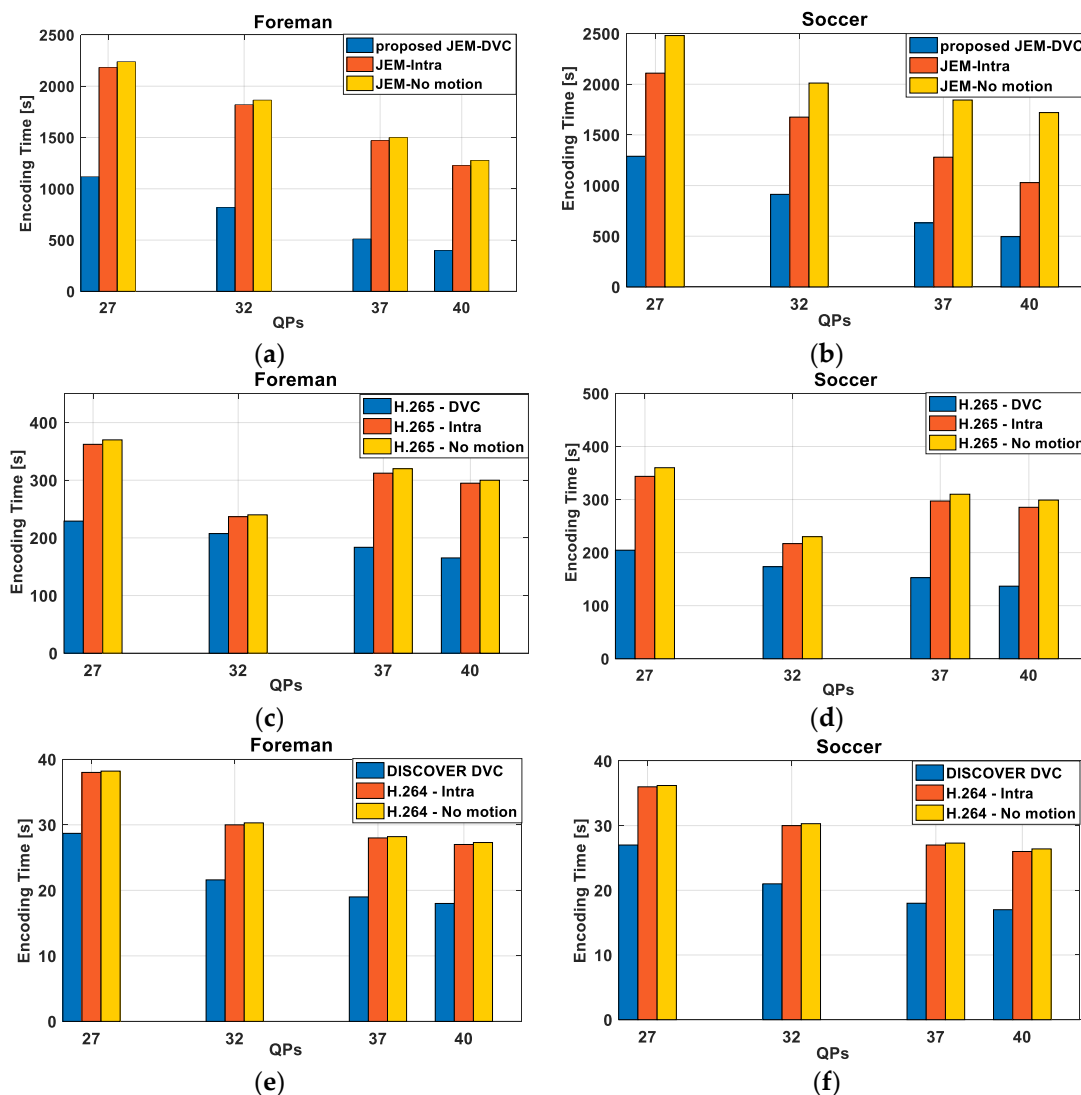

Figure 13. Comparisons of complexity for encoding time (s) obtained by each DVC scheme and video coding standards (Intra and No motion) with different test sequences: (**a**–**f**) *Foreman* and *Soccer*.

5. Conclusions

In this paper, a novel JEM-DVC solution has been proposed to take into account the recent achievements of JEM/FVC upgrades. In the proposed JEM-DVC solution, several advanced techniques of JEM Intra coding have been employed to effectively improve the performance of the JEM-DVC codec. For efficient quality consistency control, two novel methods, named KF-Q and WZF-Q, are introduced and applied for JEM-DVC, which can determine the optimal values of QP and QM that are assigned for the key and WZ frame coding, respectively. The proposed JEM-DVC codec is the first DVC codec in the literature, including JEM coding, and thus all of the performance results that are presented in this paper are new. The experimental results show that the proposed JEM-DVC codec significantly outperforms the relevant DVC benchmarks, notably the recent H.265-DVC and DISCOVER DVC codecs. In future works, several issues, such as adaptive GOP control for JEM Intra coding, complexity reduction for JEM-DVC coder, and scalable coding for JEM-DVC can be considered to further improve the performance of the JEM-DVC codec and adapt this codec to practical video applications.

Author Contributions: Conceptualization, D.T.D. and X.H.V.; Data curation, H.P.C.; Investigation, D.T.D., H.P.C. and X.H.V.; Methodology, D.T.D. and X.H.V.; Software, H.P.C.; Writing—original draft, D.T.D. and X.H.V.; Writing—review & editing, D.T.D., H.P.C. and X.H.V.

References

1. ITU-T. *H.264*: Advanced Video Coding for Generic Audiovisual Services. Available online: https://www.itu.int/rec/T-REC-H.264-200305-S (accessed on 26 June 2019).
2. Distributed Video Coding: Bringing New Applications to Life. Available online: https://pdfs.semanticscholar.org/c518/0c802fa8957b50769e7adc15fc151a561a40.pdf (accessed on 15 March 2019).
3. Slepian, D.; Wolf, J. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory* **1973**, *19*, 471–480. [CrossRef]
4. Wyner, A.D.; Ziv, J. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inf. Theory* **1976**, *22*, 1–10. [CrossRef]
5. Girod, B.; Aaron, A.M.; Rane, S.; Rebollo-Monedero, D. Distributed Video Coding. *Proc. IEEE* **2005**, *93*, 71–83. [CrossRef]
6. PRISM: A New Robust Video Coding Architecture Based on Distributed Compression Principles. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.5913&rep=rep1&type=pdf (accessed on 5 May 2019).
7. Dragotti, P.L.; Gastpar, M. *Distributed Source Coding: Theory, Algorithms and Applications*; Academic Press: Cambridge, MA, USA, 2009.
8. Brites, C.; Pereira, F. Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding. *IEEE Trans. Circuits Syst. Video Technol.* **2008**, *18*, 1177–1190. [CrossRef]
9. Brites, C.; Pereira, F. Distributed Video Coding: Assessing the HEVC upgrade. *Signal Process. Image Commun.* **2015**, *32*, 81–105. [CrossRef]
10. Artigas, X.; Ascenso, J.; Dalai, M.; Klomp, S.; Kubasov, D.; Ouaret, M. The discover codec: Architecture, techniques and evaluation. In Proceedings of the Picture Coding Symposium (PCS '07), Lisbon, Portugal, 7–9 November 2007.
11. Discover DVC Final Results. Available online: http://www.img.lx.it.pt/~{}discover/home.html (accessed on 15 April 2019).
12. Wiegand, T.; Sullivan, G.J.; Bjontegaard, G.; Luthra, A. Overview of the H.264/AVC video coding standard. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, *13*, 560–576. [CrossRef]
13. Ferroukhi, M.; Ouahabi, A.; Attari, M.; Yassine, H.; Abdelmalik, T.A. Medical video coding based on 2nd-generation wavelets: Performance evaluation. *Electronics* **2019**, *8*, 88. [CrossRef]
14. Aaron, A.; Rane, S.D.; Setton, E.; Girod, B. Transform-domain Wyner-Ziv Codec for Video. *Vis. Commun. Image Process.* **2004**, *5308*, 520–528.

15. Brites, C.; Ascenso, J.; Pereira, F. Improving Transform Domain Wyner-Ziv Video Coding Performance. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing, Toulouse, France, 14–19 May 2006; pp. 525–528.

16. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1649–1668. [CrossRef]

17. Marzuki, I.; Sim, D. Overview of potential technologies for future video coding standard (FVC) in JEM software: Status and Review. *IEIE Trans. Smart Process. Comput.* **2018**, *7*, 22–35. [CrossRef]

18. Pereira, F.; Torres, L.; Guillemot, C.; Ebrahimi, T.; Leonardi, R.; Klomp, S. Distributed video coding: Selecting the most promising application scenarios. *Signal Process. Image Commun.* **2008**, *23*, 339–352. [CrossRef]

19. Taheri, Y.M.; Ahmad, M.O.; Swamy, M.N.S. A joint correlation noise estimation and decoding algorithm for distributed video coding. *Multimed. Tools Appl.* **2018**, *77*, 7327–7355. [CrossRef]

20. Wei, L.; Zhao, Y.; Wang, A. Improved side-information in distributed video coding. In Proceedings of the International Conference on Innovative Computing, Information and Control, Beijing, China, 30 August–1 September 2006.

21. Ascenso, J.; Brites, C.; Pereira, F. A flexible side information generation framework for distributed video coding. *Multimed. Tools Appl.* **2010**, *48*, 381–409. [CrossRef]

22. Xiem, H.V.; Ascenso, J.; Pereira, F. Adaptive Scalable Video Coding: An HEVC-Based Framework Combining the Predictive and Distributed Paradigms. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 8.

23. Algorithm Description of Joint Exploration Test 7 (JEM 7). Available online: https://mpeg.chiariglione.org/standards/mpeg-i/versatilevideo-coding (accessed on 12 April 2019).

24. Schwarz, H.; Rudat, C.; Siekmann, M.; Bross, B.; Marpe, D.; Wiegand, T. Coding Efficiency/Complexity Analysis of JEM 1.0 coding tools for the Random Access Configuration. In Proceedings of the Document JVET-B0044 3rd JVET Meeting, San Diego, CA, USA, 20–16 February 2016.

25. Sidaty, N.; Hamidouche, W.; Deforges, O.; Philippe, P. Compression efficiency of the emerging video coding tools. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 2996–3000.

26. Xiph.org Video Test Media. Available online: https://media.xiph.org/video/derf/ (accessed on 10 June 2019).

27. Huy, P.C.; Stuart, P.; Xiem, H.V. A low complexity Wyner-Ziv coding solution for light field image transmission and storage. In Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Jeju, Korea, 5–7 June 2019. (In Press).

28. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Imagequality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]

29. Bjontegaard, G. Calculation of average PSNR differences between RD-curves. In Proceedings of the VCEG-M33, Thirteenth Meeting of the Video Coding Experts Group (VCEG), Austin, TX, USA, 2–4 April 2001.