**RESEARCH**                                                                 **Open Access**

# Two-description distributed video coding for robust transmission

Anhong Wang[1*], Zhihong Li[1], Yao Zhao[2], Wubin Wang[3] and Huihui Bai[2]

## Abstract

In this article, a two-description distributed video coding (2D-DVC) is proposed to address the robust video transmission of low-power capturers. The odd/even frame-splitting partitions a video into two sub-sequences to produce two descriptions. Each description consists of two parts, where part 1 is a zero-motion based H.264-coded bitstream of a sub-sequence and part 2 is a Wyner-Ziv (WZ)-coded bitstream of the other sub-sequence. As the redundant part, the WZ-coded bitstream guarantees that the lost sub-sequence is recovered when one description is lost. On the other hand, the redundancy degrades the rate-distortion performance as no loss occurs. A residual 2D-DVC is employed to further improve the rate-distortion performance, where the difference of two sub-sequences is WZ encoded to generate part 2 in each description. Furthermore, an optimization method is applied to control an appropriate amount of redundancy and therefore facilitate the tuning of central/side distortion tradeoff. The experimental results show that the proposed schemes achieve better performance than the referenced one especially for low-motion videos. Moreover, our schemes still maintain low-complexity encoding property.

**Keywords:** distributed video coding, robust video transmission, multiple description coding

## 1. Introduction

The increasing demand for friendly up-link communication of low-power video captures has generated a lot of research interests in developing video codec of low-complexity encoding. As a new video coding framework, distributed video coding (DVC) [1-3] also called Wyner-Ziv (WZ) video coding makes the low-complexity video encoding a reality, in that DVC shifts the most time-consuming motion estimation to the decoder from the encoder side.

On the other hand, robust DVC methods are desired especially when the video is transmitted over wireless networks. DVC itself takes on inherent robustness because of the error-correcting channel decoding algorithm adopted. However, this robustness is achieved at the cost of compression efficiency. Typically, DVC assumes a correlation existing between the source to be encoded and its side information (SI) available at the decoder. The compression comes from the correlation and the stronger correlation the higher compression efficiency, or vice versa. However, in case of high packet loss rate, such correlation may be destroyed due to poor reconstruction of SI at the decoder, which in turn degrades the coding performance. In some other related studies, WZ coding is used as forward error correction to protect the video transmission. For example, Girod et al. [1,4] provided a systematic lossy error protection (SLEP) method based on WZ coding, which is two-layer scalable in the sense of having one base layer with MPEG encoder and the corresponding WZ bits as the enhancement layer. MPEG stream is firstly decoded and the corrupt data are reconstructed using error concealment, and then the reconstructed signal is used to generate the SI to decode the WZ-encoded data. WZ bits refine the reconstruction thus protecting MPEG stream against the channel packet loss to some degree. However, SLEP scheme still applies motion estimation in its MPEG encoder which sacrifices the desired property of low-complexity. Also, error propagation in the MPEG-encoded stream may negatively impact the quality of SI in WZ coding, which degrades the robustness of the system especially when the packet loss rate is high [5].

* Correspondence: wah_ty@yahoo.com.cn
[1]School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China
Full list of author information is available at the end of the article

Springer

To improve the robustness of SLEP, Crave et al. [5] proposed a distributed multiple-description coding (DMDC), which can be seen as a two-description adaption of SLEP. Nevertheless, the encoding is still of high complexity due to the motion compensation temporal filtering involved in the encoder. In addition, Rane et al. [6] proposed multiple embedded WZ descriptions coding as an extension of SLEP.

MDC has emerged as an attractive framework for robust transmission over unreliable channels. MDC encodes the source message into several bit streams (called descriptions) carrying different but correlated information which then can be transmitted over different channels. When description loss occurs, the decoder can get acceptable reconstruction from the received descriptions. It is the path diversity that makes MDC successful in robust transmission. In view of their desired and complementary features, the robustness problem of DVC is addressed here by combining DVC with MDC. In this article, we attempt to design a robust two description DVC (2D-DVC) under the constraint of low-complexity encoding. It is just the emphasizing on both low-complexity encoding and better robustness that makes our scheme different from those in [5,6]. In our scheme, the input video is first split into two sub-sequences to create two descriptions with each consisting of two parts, i.e., part 1 is a low-complexity encoding bitstream of the corresponding sub-sequence, and part 2 is a WZ bitstream of the other sub-sequence. This is just the so-called 2D-DVC where the WZ bitstream controls the amount of redundancy. However, in [7], it is shown that the residual WZ coding can achieve better rate-distortion performance than the non-residual schemes. This is because that residual WZ technique exploits a second SI accessible at both encoder and decoder. Then, we extend this idea to our 2D-DVC and propose a residual 2D-DVC, where the difference of two sub-sequences is WZ encoded and replaces part 2 of each description. Furthermore, an optimization scheme is employed to introduce the appropriate amount of redundancy. The experimental results show that the proposed schemes achieve better or comparatively rate-distortion performance compared

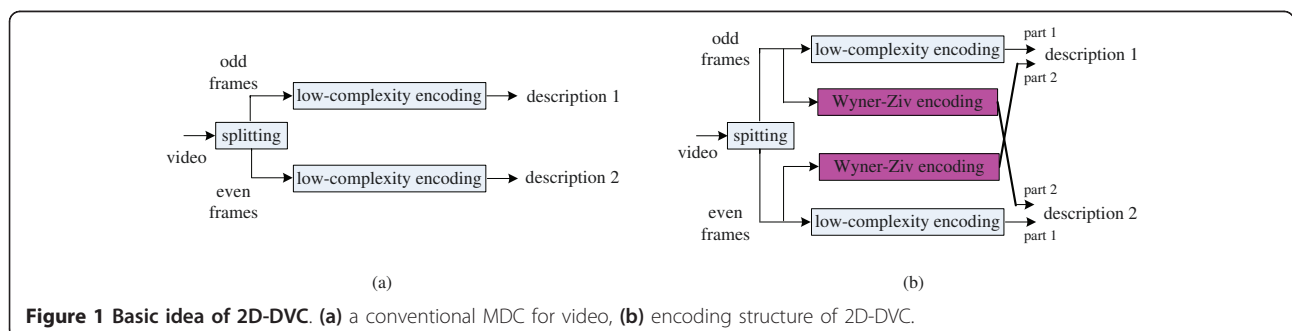with the referenced scheme; while, they maintain low-complexity encoding property.

This article is organized as follows. Section 2 introduces the basic idea and the related techniques, and Section 3 presents the proposed residual 2D-DVC and the optimization scheme in detail. Section 4 provides some experimental results, and Section 5 concludes the article.

## 2. Basic idea and related techniques

2D-DVC is designed to generate and encode two descriptions with correlation exploited only at the decoder, which can support low-complexity encoding and robustness against packet loss.

### 2.1. 2D-DVC scheme

Figure 1 shows the encoding structure of 2D-DVC. Considering a video sequence, its odd and even frames are first split into two sub-sequences. In a conventional MDC scheme of Figure 1a, each sub-sequence produces a description sent to separate channels. When one description is lost, the lost sub-sequence will be estimated by the received one. However, the estimation is normally coarse due to the lack of the other part of information, especially for video sequence with large motion. Figure 1b shows the encoding structure of 2D-DVC, where a WZ stream from the other sub-sequence generates part 2 of each description. The coarse estimation results can act as the SI for WZ decoding. In case of description loss, the SI is refined by WZ stream to recover a better version of the lost sub-sequence. This encoding framework is a two-description adaption of the SLEP like that in [5] but with low-complexity encoding property. We consider a simple encoding scheme for part 1. WZ encoding supports low-complexity encoding because it exploits the correlation of the two sub-sequences only at the decoder. In case of no loss, only part 1 is used to recover the original video, where WZ bitstream is redundant. To further improve the rate-distortion performance, the residual coding and an iterative optimization method are adopted in the scheme.



**Figure 1 Basic idea of 2D-DVC**. **(a)** a conventional MDC for video, **(b)** encoding structure of 2D-DVC.

## 2.2. SW-SPIHT coding

WZ coding [8] refers to lossy compression with SI at the decoder. WZ coding aims to achieve almost the same coding performance by exploiting the correlation only at the decoder as at both the encoder and the decoder. As shown in Figure 2, WZ coding generally consists of the following steps, namely, transform coding, quantization, Slepian-Wolf (SW) coding, as well as the generation and reconstruction of SI at the decoder side. Transform coding and quantization first compress the source to generate the binary sequence, which is then compressed by SW coding. SW coding [9] is generally realized using channel coding, where the binary sequence of the source is first encoded with channel coding and only the syndrome bits are sent to the decoder. In general, the sent syndrome bits are less than that of the source so compression is achieved. At the decoder, SI is also transformed and quantized to generate the SI. The received syndrome bits and the SI are to recover the original source bits by error-correcting channel decoding.

Among some WZ coding schemes, SW set partitioning in hierarchical tree (SW-SPIHT) coding approach [10-12] performs very well in term of its scalability and rate-distortion performance. The process of SW-SPIHT is shown in Figure 3, assuming the source $X$ and the correlated SI $Y$. For $X$, after discrete wavelet transformation (DWT), the traditional SPIHT coding is implemented and we get its binary tree distribution information $SD$, significant information $SP$, sign information $SS$, and the refinement information $SR$. The tree information $SD$ is sent to the decoder after arithmetic coding. $SP$, $SS$, and $SR$ are encoded by channel coding where the syndrome bits are sent to the decoder. At the decoder, the received $SD$ is first decompressed. Then according to $SD$, the side binary sequences $SP_y$, $SS_y$, and $SR_y$ of $Y$ are

obtained by SPIHT encoding. Next, $SP_y$, $SS_y$, $SR_y$, and the received syndrome bits are used to recover the main binary sequence $SP$, $SS$, and $SR$ by channel decoding. Finally, the wavelet coefficients of $X$ will be reconstructed according to

$$W'' = \begin{cases} V_{\max}, & \text{if } W_y > V_{\max} \\ W_y, & \text{if } V_{\min} \leq W_y \leq V_{\max} \\ V_{\min}, & \text{if } W_y < V_{\min} \end{cases} \quad (1)$$

where $W''$ is the final wavelet coefficient, $V_{\max}$ and $V_{\min}$ are the possible maximal and minimal value of $W''$ if SPIHT decoding is implemented to all bit-planes.

## 3. Proposed residual 2D-DVC and optimization

The proposed residual 2D-DVC scheme is sketched in Figure 4, including encoding and decoding processes. Part 1 of each description is generated by zero-motion based H.264 encoding. Part 2, for example, part 2 of description 1, is from the SW-SPIHT stream of description 2. The details are explained as follows.

### 3.1. Zero-motion-based H.264 coding

Zero-motion-based H.264 denoted as H.264 0-mv is employed in our scheme to meet the demands of low-complexity encoding. Zero-motion-based H.264 means that only the previous frame is used as the referenced frame for the inter-coding with motion searching region set to zero, which is therefore similar to the differential pulse coding modulation (DPCM). Because DPCM exploits the temporal correlation between adjacent frames, zero-motion-based H.264 normally outperforms the intra-frame coding in term of rate-distortion performance. With no motion estimation at the encoder, its encoding process is greatly simplified. Typically, in our experiments, the encoding time of zero-motion-based
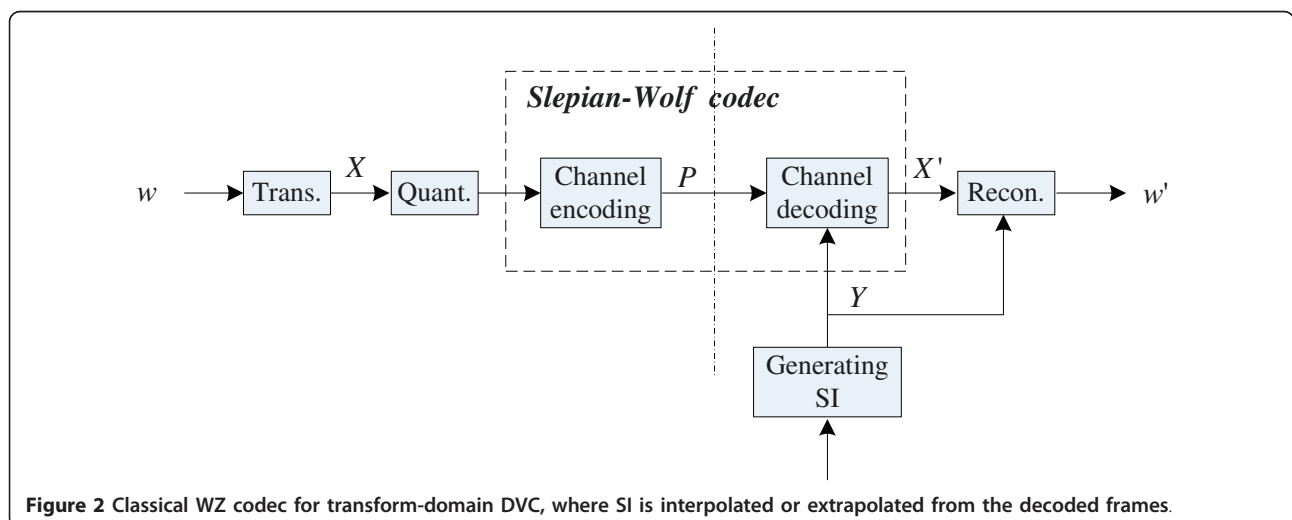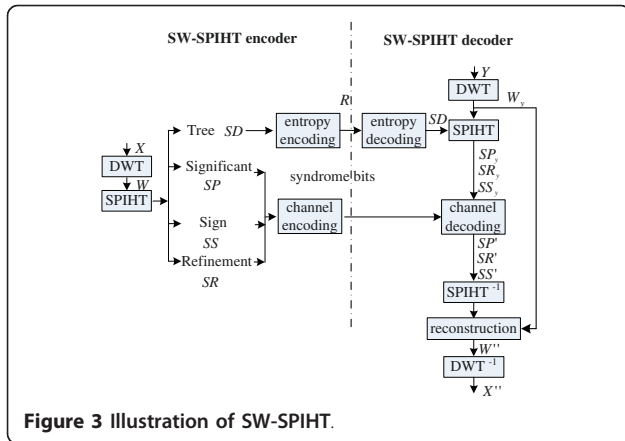


**Figure 2 Classical WZ codec for transform-domain DVC, where SI is interpolated or extrapolated from the decoded frames**.

**Figure 3 Illustration of SW-SPIHT.**

H.264 inter-coding is always shorter than that of the intra-frame in H.264 JM 9.0 program.

### 3.2. Residual-based encoding

In the single-description-based DVC, it has been shown in [7] that the pixel-domain residual WZ coding achieves better rate-distortion performance than non-residual scheme. Here, we extend this idea to our two-description DVC to further improve the rate-distortion performance efficiently. In residual 2D-DVC encoding, SW-SPIHT encodes the difference $D = X - X_{re}$ to generate part 2, where $X_{re}$ is a simple estimation to $X$. In non-residual 2D-DVC encoding, $X$ is directly input to SW-SPIHT to produce part 2. Besides, it is $D_y = Y - X_{re}$ that acts as SI in the residual 2D-DVC, while $Y$ is SI in non-residual 2D-DVC. Residual scheme achieves better performance than the non-residual one mainly due to the use of $X_{re}$. In the residual case, $X_{re}$ can be seen as a second SI accessible to both encoder and decoder [7].

Since $X_{re}$ and $X$ are correlated given $Y$, using $X_{re}$ at the encoder and the decoder amounts to adding an excess condition in encoding $X$. The rate of $X$ will ideally approach the conditional entropy $H(X|(Y, X_{re}))$ which is not greater than the condition entropy $H(X|Y)$ like that in non-residual 2D-DVC, $H(X|(Y, X_{re})) \leq H(X|Y)$.

In the scheme shown in Figure 4, for description 1, $D_1 = X_2 - X_{re2}$, $D_{y1} = Y_2 - X_{re2}$; for description 2, $D_2 = X_1 - X_{re1}$, $D_{y2} = Y_1 - X_{re1}$. $X_{re1}$, $X_{re2}$, as well as $Y_1$ and $Y_2$ are generated in the interpolating process. There are two interpolating methods used, one is the simple average interpolation used both at the encoder and the decoder to generate $X_{re1}$ and $X_{re2}$, while the other is the complex motion-compensated interpolation used only at the decoder to generate $Y_1$ and $Y_2$ when only one description is received. For example, $X_{re2,i}$ and $Y_{2,i}$ for $i$th frame in description 1 are generated according to the following formulae respectively,

$$X_{re2,i} = \frac{X'_{1,i-1} + X'_{1,i}}{2} \tag{2}$$

$$\begin{aligned}
Y_{2,i}(x, y) &= 0.25X'_{1,i-1}(x + 0.5dx_f, y + 0.5dy_f) \\
&+ 0.25X'_{1,i}(x - 0.5dx_f, y - 0.5dy_f) \\
&+ 0.25X'_{1,i-1}(x - 0.5dx_b, y - 0.5dy_b) \\
&+ 0.25X'_{1,i}(x + 0.5dx_b, y + 0.5dy_b)
\end{aligned} \tag{3}$$

where $X'_{1,i-1}$ and $X'_{1,i}$ are the adjacently decoded frames in $X_1$; $(x, y)$ is the coordinates of the interpolated frame; $[dx_b, dy_b]$ and $[dx_f, dy_f]$ are the backward and forward motion vectors between $X'_{1,i-1}$ and $X'_{1,i}$, respectively, which may be obtained by the half-pixel motion estimation similar to literature [3].



**Figure 4 The proposed residual 2D-DVC framework.**
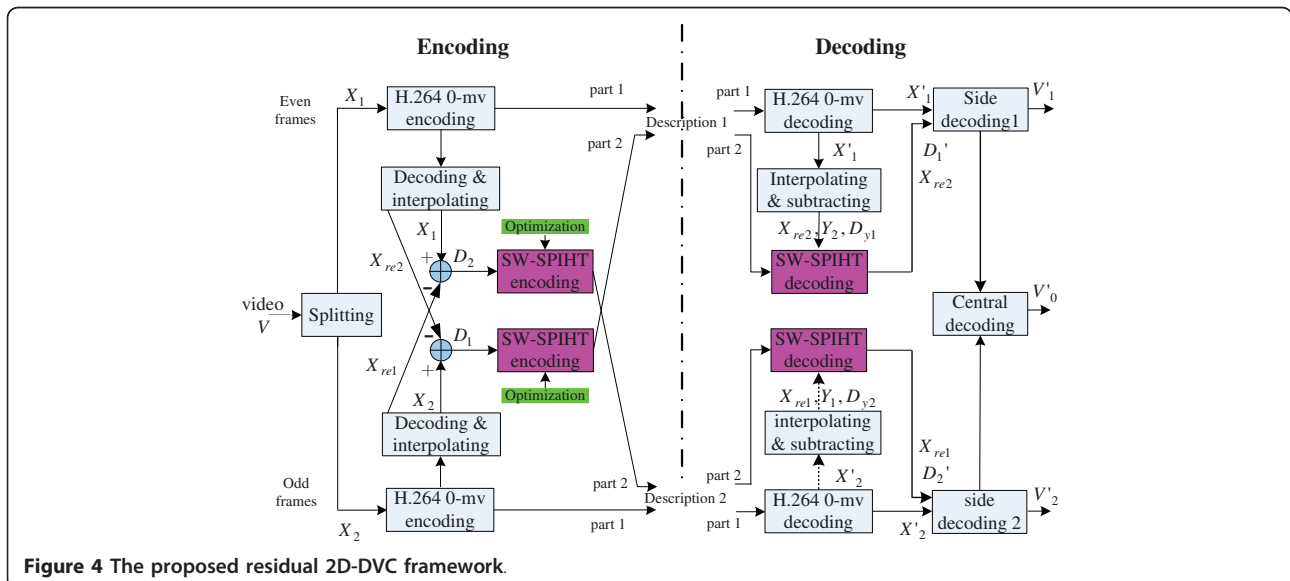
Due to the use of $X_{re}$, some correlation between two descriptions is exploited at the encoder in residual 2D-DVC. However, for the excess encoding complexity over non-residual 2D-DVC, the computation of subtracting and average interpolation operation is very low, which still preserves the low-complexity encoding in the residual 2D-DVC.

In this study, a channel code of low-density parity check with accumulation (LDPCA) [13] is used in SW-SPIHT coding with a feedback. At each bit-plane of SPIHT, the encoder sends a certain amount of syndrome information stage-by-stage on the demand of feedback. If the receiver cannot decode correctly, the encoder will send additional syndrome information.

### 3.3. Decoding

If only one description, for example, description 1 is received, its part 1 is first reconstructed by zero-motion-based H.264 decoding and the interpolation will generate $X_{re2}$ and $Y_2$ based on Equations 2 and 3. Then, SW-SPIHT decoding reconstructs the difference $D'_1$ using the received syndrome and SI $D_{y1}$, and we can obtain $X''_2$ as $X''_2 = D'_1 + X_{re2}$. Finally, $X'_1$ and $X''_2$ are merged to obtain the video $V'_1$ in side decoding 1.

When both descriptions are received, the central decoding works without motion estimation. Part 1 of each description is first decoded by the zeromotion-based H.264 decoding, and then the resulting $X'_1$ and $X'_2$ are refined by WZ bits received. Concretely, $X_{re1}$ and $X_{re2}$ are interpolated according to (2) by $X'_1$ and $X'_2$. The difference $D_{y1}$ and $D_{y2}$ are generated as follows, $D_{y2} = X'_2 - X_{re1}$, $D_{y2} = X'_2 - X_{re1}$. Then, SW-SPIHT decoding recovers $D'_1$ (or $D'_2$) using $D_{y1}$ (or $D_{y2}$) and the received syndrome bits. The refined version of $X'_1$ (or $X'_2$) are obtained, i.e., $X_1 = D'_1 + X_{re2}$, $X_2 = D'_2 + X_{re1}$. Finally, $X_1$ and $X_2$ are merged to recover the video $V'$.

### 3.4. Redundancy optimization

We know that the redundancy will affect the correlation between two descriptions as well as the consequent rate-distortion of central and side coding. In general, when redundancy is more, correlation between two descriptions will be higher thus producing better quality from side decoder, while the central quality drops with the increasing of redundancy. Moreover, too much redundancy may even degrade the side quality. Therefore, optimization is desired to maximize the rate-distortion performance of non-residual and residual 2D-DVC proposed.

Let $d_0(v, N)$ and $d_1(v, N)$ (or $d_2(v, N)$) denote the mean squared errors (MSE) from the central and side decoder for the input video $v$, respectively, given the amount of WZ bits is $N$. Let $R(v, N)$ be the rate for two

descriptions, while $R_1(v, N)$ and $R_2(v, N)$ be the rates for the two balanced descriptions 1 and 2, respectively. Our goal is to find the optimal parameter $N$ in solving the following optimization problem:

$$\min_N d_1(v, N) \tag{4}$$

subject to
condition 1:

$$R(v, N) = 2R_1(v, N) = 2R_2(v, N) \le R_{budget} \tag{5}$$

condition 2:

$$d_0(v, N) \le d_{budget} \tag{6}$$

where $R_{budget}$ is the available total bit rate to encode two descriptions and $d_{budget}$ is the maximum distortion acceptable for central decoder reconstruction. The encoding optimization module is based on the above function. With the constrain on the total bit rate and the central distortion, $N$ is adjusted accordingly to minimize the side distortion.

The optimization for the problem is carried out in an iterative way. The basic algorithm shown in Figure 5 is to make use of the monotonicity of $R$ and $d$ as the function of $N$. After initialization, a smallest $N$ is searched to minimize $d_1$ subject to conditions 1 and 2. Specially, in this study, SW-SPIHT coding generates the redundancy and its rate will affect the performance of 2D-DVC and residual 2D-DVC. For easy realization, the optimization to the bits of SW-SPIHT is just implemented iteratively based on the number of BP (bit plane), $n_{BP}$, i.e., we adjust $n_{BP}$ based on the above function given the QP (quantization parameter) of H.264 0-mv for part 1. Finally, an optimized combining of $n_{BP}$ and QP is chosen, as shown in the following section.

## 4. Experimental results

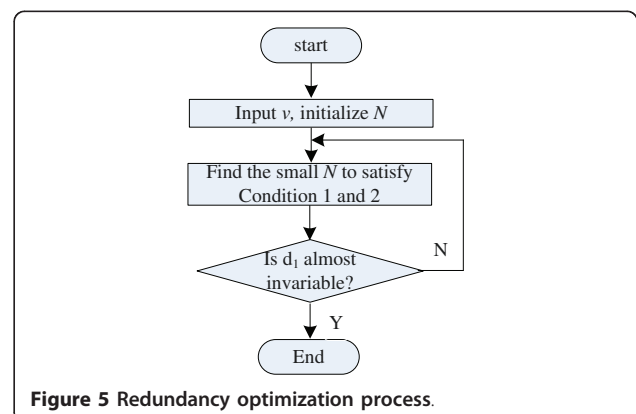For fair comparison, we use four standard video sequences to test the performance. They are *Foreman*,



**Figure 5 Redundancy optimization process**.

**Table 1 Optimized combinations of $n_{BP}$ and QP**

|  | Foreman | | | | Hall | | | | Carphone | | | | Mother | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 |
| QP | 38 | 26 | 20 | 14 | 28 | 25 | 20 | 18 | 34 | 28 | 22 | 16 | 34 | 28 | 22 | 16 |
| $n_{BP}$(non 2D-DVC) | 4 | 4 | 5 | 5 | 4 | 5 | 6 | 6 | 4 | 4 | 5 | 5 | 3 | 3 | 4 | 4 |
| $n_{BP}$(re. 2D-DVC) | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 3 | 1 | 1 | 2 | 3 | 1 | 2 | 2 | 3 |

*Hall*, *Carphone*, and *Mother-daughter* QCIF@15 Hz. There are totally four MDC methods included, the DMDC in [5], zero-motion-based H.264 MDC without any extra WZ bits as shown in Figure 1a, the proposed non-residual 2D-DVC and residual 2D-DVC. The bit rate denotes the total of two descriptions. In the proposed schemes, we obtain four points, Q1, Q2, Q3, and Q4 according to the above optimization process. The optimized combinations of QP and $n_{BP}$ are shown in Table 1. We use the LDPCA codes proposed in [13] for all the simulations. A small amount of feedbacks for additional syndrome bits of the LDPCA code are allowed to achieve the successful LDPCA decoding.

### 4.1. Performance comparison
Figure 6 shows the rate-distortion curves, where the referenced DMDC curves for *Foreman* and *Hall* sequence are the best results of [5]. Experimental results show that the residual 2D-DVC outperforms the non-residual 2D-DVC confidently because the efficient residual WZ coding reduces the encoding rate. Specifically, for low motion *Hall* and *Mother-daughter* sequences, residual 2D-DVC achieves 0.5-1.8 dB side-quality and 0.2-1.7 dB central-quality improvement. For high motion *Foreman* sequence, residual schemes obtain 0.2-1.7 dB side-quality improvement with the comparable central quality. Besides, compared with the best results in [5], residual 2D-DVC achieves about 2-3 dB side-quality and 0.5-2 dB central-quality improvement for *Hall* sequence. The non-residual 2D-DVC achieves about 1-2 dB improvement in side quality with a central-quality decreasing of 0.2-0.7 dB for *Hall* sequence. Residual 2D-DVC has comparable efficiency at high rate for high-motion *Foreman* sequence. However, non-residual 2D-DVC is not efficient for *Foreman* due to the incapability of zero-motion based H.264 in the high-motion cases.
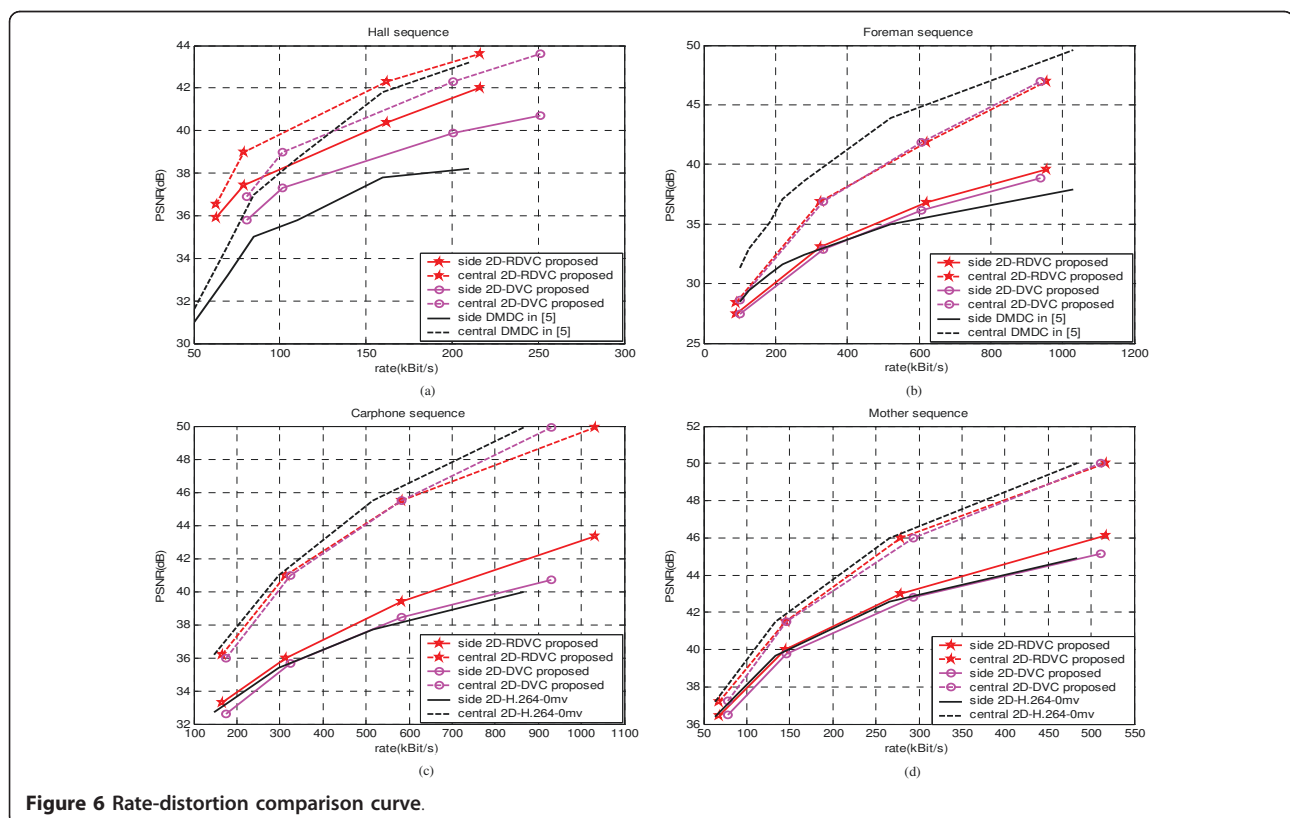


**Figure 6 Rate-distortion comparison curve**.

For *Carphone* and *Mother-daughter* sequences, we compare our scheme with the MDC with zero-motion-based H.264. It is evident that the proposed scheme has comparable rate-distortion performance with MDC zero-motion based H.264. However, the advantage of 2D-DVC over zero-motion-based H.264 lies in the quality consistency when loss occurs, which is shown by the following experimental results.

Figure 7 shows the frame side-PSNR comparison at Q1 point of three MDC schemes, MDC zero-motion-based H.264, residual 2D-DVC and non-residual 2D-DVC.

Here, in Figure 7a for *Hall* sequence, MDC H.264 0-mv is with 56kbps, non-residual 2D-DVC with extra 29kbps and residual 2D-DVC with extra 7 kbps; in Figure 7b for *Foreman* sequence, MDC H.2640-mv with 297 kbps, 2D-DVC with extra 35 kbps and residual 2D-DVC with extra 29 kbps; in Figure 7(c) for *Carphone* sequence, MDC H.264 0-mv with 148 kbps, non-residual 2D-DVC with extra 28 kbps and residual 2D-DVC with extra 17 kbps; Figure 7(d) for *Mother-daughter* sequence, MDC H.264 0-mv with 63 kbps, 2D-DVC with extra 15 kbps and residual 2D-DVC with extra 5 kbps. It is evident that residual
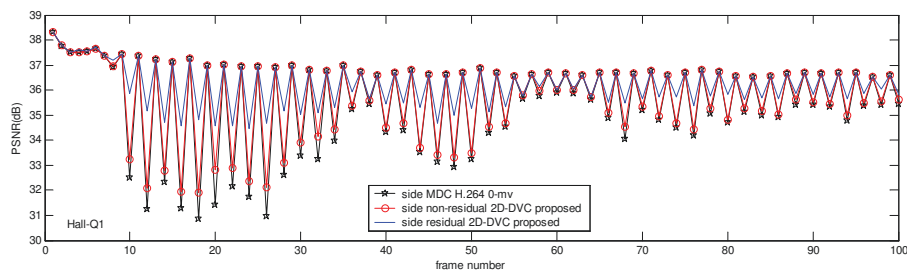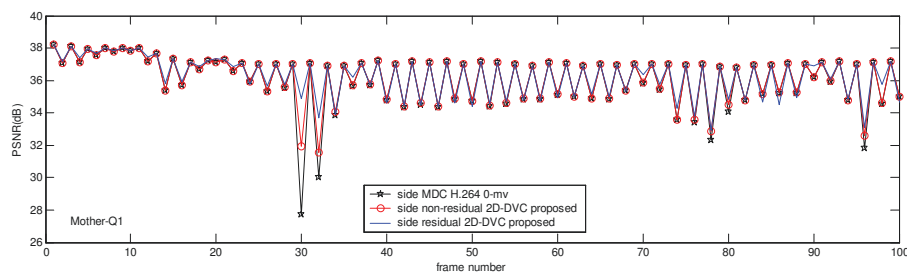


(a) Hall: MDC H.264-0mv with 56 kbps, non-residual 2D-DVC with extra 29kbps and residual 2D-DVC with extra 7 kbps
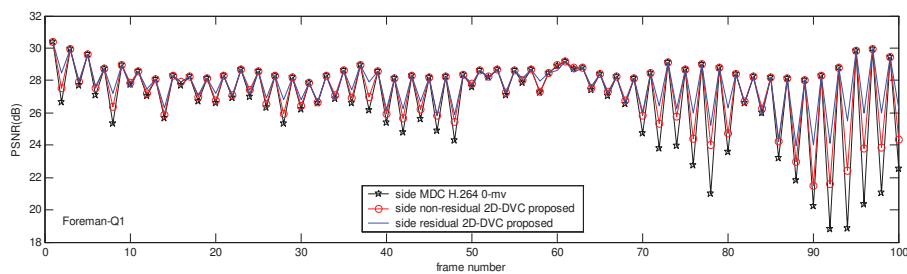
(b) Foreman: MDC H.264-0mv with 297 kbps, 2D-DVC with extra 35kbps and residual 2D-DVC with extra 29 kbps
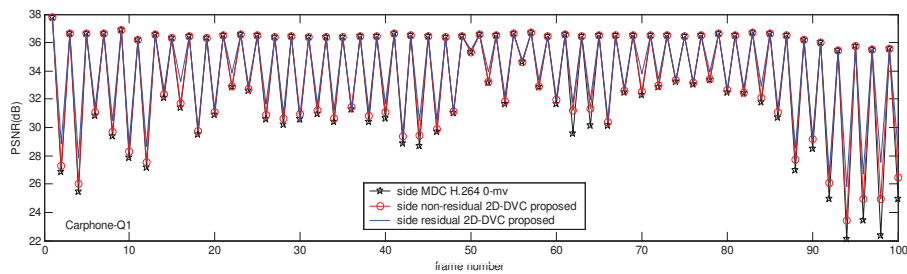
(c) Carphone:MDC H.264-0mv with 148 kbps, non-residual 2D-DVC with extra 28kbps and residual 2D-DVC with extra 17 kbps

(d) Mother: MDC H.264-0mv with 63 kbps, 2D-DVC with extra 15kbps and residual 2D-DVC with extra 5 kbps

**Figure 7 Frame PSNR comparison for Q1 point (a)** *Hall* **sequence; (b)** *Foreman* **sequence; (c)** *Carphone* **sequence; (d)** *Mother-daughter* **sequence**.

2D-DVC and non-residual 2D-DVC get better quality consistency than MDC zero-motion-based H.264 due to the WZ bitstream added. Moreover, the residual 2D-DVC performs the best.

For further comparison, we compute the variance values of frame side-PSNR for each rate-distortion point according to

$$\text{Var} = \sum_{i=1}^{n} \left( P_{\text{SNR}}(i) - E(P_{\text{SNR}}) \right)^2 \qquad (7)$$

where $P_{\text{SNR}}(i)$ is the PSNR value of $i$th frame, and $n$ is the total frame number. $E(P_{\text{SNR}})$ is the average value of PSNR on all frames. Table 2 shows the variance value for all rate-distortion points. It can be seen that the residual and non-residual 2D-DVC achieve smaller variance value, which means they have better consistency in frame side-PSNR.

### 4.2. Encoding complexity

The proposed 2D-DVC schemes have lower complexity encoding compared with the scheme in [5]. In our schemes, each description consists of two parts, zero-motion-based H.264 encoding and SW-SPIHT encoding. Since each part has computation less than or similar to intra-frame coding, 2D-DVC's encoding complexity is similar to the conventional intra-frame model. However, in [5], each description also has two parts, where both of them apply motion compensated temporal filter so the encoding complexity of the system is similar to the conventional inter-frame model.

Finally, the encoding time is measured in millisecond (ms). The hardware used is an HP notebook nx6330, Intel 2 processor at 1.66 GHz with 1.0 GB of RAM. The software condition is Windows XP operative system with VC6.0 released version. The average encoding time of each frame for *Hall* sequence are 162, 175, 183, and 199 ms for Q1, Q2, Q3, and Q4, respectively.

**Table 2 Variance value of frame side-PSNR**

|  |  | *Hall* | *Foreman* | *Mother* | *Carphone* |
|---|---|---|---|---|---|
| Q1 | MDC H.264 0-mv | 2.7664 | 5.6716 | 2.3155 | 17.1559 |
|  | non 2D-DVC | 2.0889 | 2.9850 | 1.9093 | 14.214 |
|  | resi. 2D-DVC | 0.6919 | 1.9208 | 1.6863 | 11.4474 |
| Q2 | MDC H.264 0-mv | 2.7664 | 5.6716 | 2.3155 | 17.1559 |
|  | non 2D-DVC | 4.9373 | 30.6142 | 7.1029 | 38.6589 |
|  | resi. 2D-DVC | 1.4101 | 19.2249 | 4.8161 | 29.7952 |
| Q3 | MDC H.264 0-mv | 11.3975 | 59.0705 | 17.9744 | 69.3289 |
|  | non 2D-DVC | 7.3947 | 39.1868 | 14.7978 | 56.5313 |
|  | resi. 2D-DVC | 4.4998 | 29.3181 | 14.2172 | 40.9057 |
| Q4 | MDC H.264 0-mv | 15.1514 | 101.3118 | 35.0827 | 109.4687 |
|  | non 2D-DVC | 10.4177 | 75.6918 | 30.7555 | 93.3305 |
|  | resi. 2D-DVC | 2.9302 | 61.2109 | 20.1534 | 46.0834 |

### 5. Conclusion

This article has proposed two 2D-DVC schemes for robust transmission with low-complexity encoding. The video is first separated into two subsequences by odd/even frame splitting. Then in the first 2D-DVC, each description is composed of two parts, part 1 being a zero-motion-based H.264 stream and part 2 a WZ stream, which maintain some redundancy to produce an acceptable quality reconstruction when one description is lost. In the second scheme, a residual 2D-DVC is proposed to reduce the redundancy, where the difference of the two sub-sequences is WZ encoded and used as part 2 in each description. The amount of redundancy can be controlled using an optimization scheme. The experimental results have shown that the proposed schemes can achieve better or comparable rate-distortion performance compared with the referenced one especially when the motion is low. Moreover, our schemes maintain low-complexity encoding so they are suitable for applications of portable video communication devices with very limited power and storage, such as mobile cameras, wireless low-power surveillance devices.

### Author details
[1]School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China [2]Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China [3]Research Institute of China Ordnance Industry, No. 5 Room No. 207, Taiyuan 030001, China

### Competing interests
The authors declare that they have no competing interests.

### References
1. B Girod, A Aaron, S Rane, D Rebollo-Monedero, Distributed video coding. Proc IEEE. **93**(1), 71–83 (2005)
2. R Puri, K Ramchandran, PRISM: an uplink-friendly multimedia coding paradigm, in *Proc International Conference on Acoustics, Speech, and Signal Processing,* Hongkong, China, 856–859 (April 2003)
3. X Artigas, J Ascenso, M Dalai, S Klomp, D Kubasoy, M Ouaret, The DISCOVER codec: architecture, techniques and evaluation, in *Proc Picture Coding Symposium,* Lisbon, Portugal, 1950–1953 (November 2007)
4. S Rane, A Aaron, B Girod, Systematic lossy forward error protection for error-resilient digital video broadcasting–a WZ coding approach, in *Proc International Conference on Image Processing,* Singapore, 609–612 (October 2004)
5. O Crave, C Guillemot, B Pesquet-Popescu, C Tillier, Distributed temporal multiple description coding for robust video transmission. EURASIP J Wirel Commun Netw (2008). Article ID 183536

6.   S Rane, A Aaron, B Girod, Error-resilient video transmission using multiple embedded Wyner-Ziv descriptions, in *IEEE International Conference on Image Processing (ICIP)*. **2**, 666–669 (2005)

7.   A Aaron, D Varodayan, B Girod, Wyner-Ziv residual coding of video, in *Proc Picture Coding Symposium,* Beijing, China, 28–32 (2006)

8.   A Wyner, J Ziv, The rate-distortion function for source coding with side information at the decoder. IEEE Trans Inf Theory. **22**(1), 1–10 (1976). doi:10.1109/TIT.1976.1055508

9.   D Slepian, JK Wolf, Noiseless coding of correlated sources. IEEE Trans Inf Theory. **19**(4), 471–480 (1973). doi:10.1109/TIT.1973.1055037

10.  C Tang, N Cheung, A Ortega, C Raghavendra, Efficient inter-band prediction and wavelet based compression for hyperspectral imagery: a distributed source coding approach, in *Proc Data Compression Conference*, 437–446 (March 2005)

11.  X Guo, Y Lu, F Wu, W Gao, S Li, Wyner-Ziv video coding based on set partitioning in hierarchical tree[C], in *Proceedings of International conference on Image Processing (ICIP)*, 601–604 (2006)

12.  A Wang, Y Zhao, JS Pan, Residual distributed video coding based on LQR-hash. Chin J Electron. **18**(1), 109–112 (2009)

13.  D Varodayan, A Aaron, B Girod, Rate-adaptive distributed source coding using low-density parity-check codes. EURASIP Signal Process. **86**, 3123–3130 (2006)