International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.8 (2016), pp.253-262
http://dx.doi.org/10.14257/ijsip.2016.9.8.22

Stereo Video Transmission Distortion Derived Algorithm Based On Data Partitioning In H.264/AVC

Xiaodong Wang, Jiao Ke, Hanzhang Wang and Lianjun Zhang

Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China
wangxiaodong@nbu.edu.cn

Abstract

H.264/AVC coding standard slices the encoded Data using the Data Partitioning (DP) based on syntax priority to improve the error resistance for H.264 bit stream. According to this feature, considering the temporal and spatial correlation in stereoscopic video, network performance and terminal error concealment technology. A derived algorithm for determining the transmission distortion based on different DP parts in a Group-of-Pictures (GoP) is proposed. The algorithm takes both the distortion introduced by error diffusion and error concealment into account, which can be estimated by the stereo video quality at encode side under the condition of the specific network. Experimental results show that the algorithm used to calculate the distortion basically coincided with simulation results, and the algorithm can be applied to different sequences with various movement characteristics.

***Keywords**: Data Partitioning; stereo video; transmission distortion; derived algorithm*

1. Introduction

As people's growing demand for visual sense, the related technology of stereo video has gradually become a hot issue of extensive research. However, stereoscopic video coding scheme with high compression rate makes the stream extremely fragile. The packet loss is inevitable when network congestion occurs, which will cause a sharp dropped perceived quality for users. In order to guarantee the quality of stereoscopic video, it is necessary to study the relationship between packet loss and video perceived quality in heterogeneous networks. adopting the specific error resilient technology, and establishing a suitable network transmission distortion model.

The existing transmission distortion estimation model can be roughly classified into three different kinds. (1) Distortion modeling on pixel level. The most representative one is Zhang who proposed the transmission distortion estimation method based on pixel recursive (ROPE) [1]. It recursively computes the distortion of each decoding reconstruction pixel at the current moment and the next moment as the total distortion, under specific, packet loss rate and error concealment methods. [2-3] improve the ROPE method to make it applicable to half pixel accuracy motion compensation and simplify the calculation of correlation coefficient in ROPE. (2) Distortion modeling on frame level. [4-6] put forward a video streaming media distortion model on frame level. The value of the involved parameters is between 0 and 1, which will affect the accuracy of the model to a certain extent.Wang puts forward a kind of frame-level transmission distortion estimation model based on recursive thought [7]. It takes into account the packet loss rate, coding mode, frame refresh rate, and sub-pixel motion vector, what's more, H.264 intra-frame prediction and loop filter this two new features, and error diffusion factor are introduced. The model introduced some parameters based on assumption. All above frame-level

ISSN: 2005-4254 IJSIP

distortion models are applied to single view condition, Zhou propose a multi-view transmission distortion oriented model based IP network [8]. The model analyzed the error diffusion theory and principle of coding technology in different coding mode in detail, computed the distortion of each correctly received frame and each lost frame in different coding mode respectively. (3) Distortion modeling on macro block level. [9] analyzed the relation among some key parameters in H.264 and the video content features. Based on this, it advances a kind of weight estimation of macro block distortion model. [10] also estimated the total distortion on the macro block level, whether the former frame is received correctly or not will determine the distortion of current macro block. Zhang propose a fast distortion estimation method on macro block level for single view video [11]. The computational complexity of it is low that we can estimate the transmission distortion at the coding side quickly. Besides, [12] propose a hierarchical clustering algorithm to solve the signal missing problem in transmission.

In the above-mentioned models, the accuracy of estimation models based on pixel level is high, but the computational complexity is high as well. The computational complexity of estimation model based on frame level is low, but it cannot satisfy the requirement of real-time application. Computational complexity and real-time is a pair of contradictory community, estimation method based on macro block level can solve the contradiction between the computational complexity and real-time problem, and it is a good compromise solution. But now these macro block-level distortion estimation methods are focus on single view video, macro block level transmission distortion estimation model for stereo video has not been reported, and less consideration of bit scream error resistance. For this reason, on the basis of analyzing the characteristics of the stereo video, this article presents a macro block level transmission distortion model for stereo video in detail, and DP technology is used to reduce the bit error rate.

2. Transmission Distortion Modeling Based on DP

2.1. Data Partion (DP) Technology

Typically, the data within a macro blocks is stored together and form the slice, DP reassembles the semantically related macro block data of a slice. H.264/AVC defines the data partitioning which divide coded data into Partition A (DPA), Partition B (DPB), Partition C (DPC) according to the syntax elements. Specially, DPA contains slice header information, macro block types, quantitative parameters, prediction mode and motion vector and so on. DPB contains the prediction residual information of intra-frame coding MB. DPC contains the prediction residual information of inter-frame coding MB. The relation between them is that DPA can be decoded independently. When DPA is lost, the partition data of DPB and DPC cannot be decoded independently, we use the same location MB data in previous frame to conceal the lost information. The information in DPB and DPC can only be decoded depend on the DPA. Similarily, When DPB is lost, intra-frame coding MB use the same location MB data in the previous frame to conceal it. When DPC is lost, inter-frame coding MB use the motion vector resolved by DPA to acquire predicted data of the reference frame so as to conceal it. That is to say, only if the DPA received correctly, DPB and DPC can be decoded correspondingly. So it is obviously that DPA is the most important, and DPB is usually more important than DPC because the intra-frame coding MB suppressed the influence of error diffusion.

2.2. Modeling of Stereo Video Transmission Distortion

We define that $F(s,n,k,i)$, $\hat{F}(s,n,k,i)$, and $\tilde{F}(s,n,k,i)$ are the original pixel values, reconstruction values in encoder, and reconstruction values in decoder of the view s, frame n, pixel i, and the data partitioning part k respectively. $D_c(s,n,k)$ is the transmission distortion of view s, frame n and data partitioning part k. We adopt mean square error (MSE) as the measurement of transmission distortion. According to the definition, the channel distortion can be calculated by the following formula

$$D_c(s,n,k) = E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n,k,i)]^2\} \tag{1}$$

When coding data received correctly, that is $\hat{F}(s,n,k,i) = \tilde{F}(s,n,k,i)$. While when packet loss occur, we use error concealment technology for recovery at the decoder.

The modeling of stereo video transmission distortion based on DP includes the modeling of transmission distortion in left view and right view. the detailed derivation will be introduced in the following two parts.

2.2.1. Transmission Distortion Modeling in Left View: The left view distortion modeling based on DP includes intra-frame and inter-frame distortion modeling. The derivation of inter-frame distortion in the left view is similar to modeling in one view. Following is the computing process of distortion in each part. Each DP part in left view can be divided into intra-frame mode and inter-frame mode. We Assume that β is the intra-frame mode rate . D_L^I and D_L^P are the total transmission distortion of intra-frame mode and the total distortion of inter-frame mode respectively. So the overall transmission distortion of the left view can be expressed as:

$$D_L = \beta \times D_L^I + (1-\beta) \times D_L^P \tag{2}$$

(1) Intra-frame distortion modeling based on DP in the left view

For each part of DP in left view, transmission distortion can be expressed as followed when being coded in intra-frame mode:

$$\begin{aligned} D_L^I &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n-1,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n,k,i)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s,n-1,k,i)]^2\} + p \times E\{[\hat{F}(s,n-1,k,i) - \tilde{F}(s,n-1,k,i)]^2\} \\ &= p \times RFD(n,n-1,k) + p \times D(s,n-1,k) \end{aligned} \tag{3}$$

where $RFD(n,n-1,k)$ is the Mean Square Error (MSE) of the DP part k in frame n and frame n-1. $D(s,n-1,k)$ is the transmission distortion of the DP part k in frame n-1 in view s. We suppose that I frame will not be lost, so the initial (first frame) distortion value in DPA, DPB and DPC of the left view is $D(0,0,k)=0$ (view number and macro block number starts from 0), and the value of $D(s,n-1,k)$ can be obtained recursively.

(2) Inter-frame distortion modeling based on DP in the left view

For each part of DP in the left view, the reconstruction value of using inter-frame coding mode can be expressed as $e(s,n,k,i) + \tilde{F}(s,n-1,k,j)$ when there is no data losing. Because of the inter-frame prediction, pixel j in DP part k in frame n-1 is the prediction reference value of pixel i in DP part k in frame n. $e(s,n,k,i)$ is the prediction residual. If the data in one macro block is lost, decoder will conceal the errors using the methods mentioned above. So the transmission distortion in inter-frame mode can be represented as

$$\begin{aligned} D_L^P &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n-1,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - e(s,n,k,i) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s,n-1,k,i)]^2\} + p \times E\{[\hat{F}(s,n-1,k,i) - \tilde{F}(s,n-1,k,i)]^2\} \\ &\quad + (1-p) \times E\{[\hat{F}(s,n-1,k,j) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times RFD(n,n-1,k) + p \times D(s,n-1,k) + (1-p) \times D(s,n-1,\overline{k}) \end{aligned} \tag{4}$$

In Equation 4, $D(s, n-1, \overline{k})$ can be expressed as

$$D(s,n-1,\overline{k}) = b \times D(s,n-1,k) \tag{5}$$

where b represents the violent intensity of sequence's movement. In a smooth moving sequence , b is a constant which is between 0 and 1 [8].

DPA and DPB adopt the intra-frame coding mode, while DPC adopts the inter-frame coding mode, so the distortion of DPA in the left view can be expressed as

$$\begin{aligned} D_{L,DPA}^I &= p \times E\{[\hat{F}(s,n) - \tilde{F}(s,n-1)]^2\} \\ &= p \times E\{[\hat{F}(s,n) - \hat{F}(s,n-1)]^2\} + p \times E\{[\hat{F}(s,n-1) - \tilde{F}(s,n-1)]^2\} \\ &= p \times RFD(n,n-1) + p \times D(s,n-1) \end{aligned} \tag{6}$$

And the distortion of DPB in left view can be expressed as

$$\begin{aligned} D_{L,DPB}^I &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n-1,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n,k,i)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s,n-1,k,i)]^2\} + p \times E\{[\hat{F}(s,n-1,k,i) - \tilde{F}(s,n-1,k,i)]^2\} \\ &= p \times RFD(n,n-1,k) + p \times D(s,n-1,k) \end{aligned} \tag{7}$$

the distortion of DPC in the left view can be expressed as

$$\begin{aligned} D_{L,DPC}^P &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n-1,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - e(s,n,k,i) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s,n-1,k,i)]^2\} + p \times E\{[\hat{F}(s,n-1,k,i) - \tilde{F}(s,n-1,k,i)]^2\} \\ &\quad + (1-p) \times E\{[\hat{F}(s,n-1,k,j) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times RFD(n,n-1,k) + p \times D(s,n-1,k) + (1-p) \times D(s,n-1,\overline{k}) \end{aligned} \tag{8}$$

2.2.2. Transmission Distortion Modeling in Right View: The intra-frame prediction in right view is as same as its in left view, but the inter-frame is not. The inter-frame prediction in the right view could be divided into inter-frame prediction in single view and inter-frame prediction between two views. The right view inter-frame prediction in single view refers to predicting the current view by previous frame, inter-frame prediction between two views refers to reference predicting by the another view. We suppose that γ is the inter-frame mode rate. D_R^I, D_R^{PI}, D_R^{PV} are the total distortion of intra-frame prediction in right view, the total distortion of inter-frame prediction in a single view and the total distortion of inter-frame prediction between two views respectively. So the overall distortion of the right view is

$$D_R = \beta \times D_R^I + (1-\beta)[\gamma \times D_R^{PI} + (1-\gamma) \times D_R^{PV}] \tag{9}$$

(1) Intra-frame Prediction Distortion Modeling in Right View

The intra-frame coding distortion of DP in the right view is similar with the left view, can be represented as:

$$\begin{aligned}D_R^I &= p\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n-1,k,i)]^2\}+(1-p)\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n,k,i)]^2\}\\ &= p\times E\{[\hat{F}(s,n,k,i)-\hat{F}(s,n-1,k,i)]^2\}+p\times E\{[\hat{F}(s,n-1,k,i)-\tilde{F}(s,n-1,k,i)]^2\}\\ &= p\times RFD(n,n-1,k)+p\times D(s,n-1,k)\end{aligned}\tag{10}$$

(2) Intra-view Inter-frame Prediction Distortion Modeling in Right View

For intra-view inter-frame prediction, if the macro block received correctly, the reconstruction pixel value can be expressed as $e(s,n,k,i)+\tilde{F}(s,n-1,k,j)$. Pixel i in DP part k of frame n-1 is the prediction reference value of the pixel i in DP part k of frame n, $e(s,n,k,i)$ is the prediction residual. If the DP part is lost, decoder will conceals the errors using specific error concealment methods introduced above, so the intra-view inter-frame transmission distortion can be expressed as

$$\begin{aligned}D_R^{PI} &= p\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n-1,k,i)]^2\}+(1-p)\times E\{[\hat{F}(s,n,k,i)-e(s,n,k,i)-\tilde{F}(s,n-1,k,j)]^2\}\\ &= p\times E\{[\hat{F}(s,n,k,i)-\hat{F}(s,n-1,k,i)]^2\}+p\times E\{[\hat{F}(s,n-1,k,i)-\tilde{F}(s,n-1,k,i)]^2\}\\ &\quad+(1-p)\times E\{[\hat{F}(s,n-1,k,j)-\tilde{F}(s,n-1,k,j)]^2\}\\ &= p\times RFD(n,n-1,k)+p\times D(s,n-1,k)+(1-p)\times D(s,n-1,\bar{k})\end{aligned}\tag{11}$$

(3) Inter-view Inter-frame Prediction Distortion Modeling in Right View

The inter-view prediction in right view means the prediction refers to current frame of previous view. If the macro block received correctly, the reconstruction pixel value in decoder can be expressed as $e(s,n,k,i)+\tilde{F}(s-1,n,k,j)$. Pixel j in DP part k of frame n and view s-1 is the prediction reference of pixel in DP part k in frame n of view s. $e(s,n,k,i)$ is the prediction residual. If the DP part is lost, the decoder conceals the errors using error concealment methods introduced above, so the inter-view inter-frame transmission distortion in the right view can be expressed as

$$\begin{aligned}D_R^{PV} &= p\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s-1,n,k,i)]^2\}+(1-p)\times E\{[\hat{F}(s,n,k,i)-e(s,n,k,i)-\tilde{F}(s-1,n,k,j)]^2\}\\ &= p\times E\{[\hat{F}(s,n,k,i)-\hat{F}(s-1,n,k,i)]^2\}+p\times E\{[\hat{F}(s-1,n,k,i)-\tilde{F}(s-1,n,k,i)]^2\}\\ &\quad+(1-p)\times E\{[\hat{F}(s-1,n,k,j)-\tilde{F}(s-1,n,k,j)]^2\}\\ &= p\times RFD(s,s-1,k)+p\times D(s-1,n,k)+(1-p)\times D(s-1,n,\bar{k})\end{aligned}\tag{12}$$

where $D(s-1,n,\bar{k})$ can be expressed as

$$D(s-1,n,\bar{k})=c\times D(s-1,n,k)\tag{13}$$

where c is related to spatial correlation in stereo video. In a sequence with stable movement, c is a constant which is between 0 and 1 [8].

DPA and DPB adopt the intra-frame coding mode, while DPC adopt the inter-frame coding mode,

so the distortion of DPA in right view is

$$\begin{aligned}D_{R,DPA}^I &= p\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n-1,k,i)]^2\}+(1-p)\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n,k,i)]^2\}\\ &= p\times E\{[\hat{F}(s,n,k,i)-\hat{F}(s,n-1,k,i)]^2\}+p\times E\{[\hat{F}(s,n-1,k,i)-\tilde{F}(s,n-1,k,i)]^2\}\\ &= p\times RFD(n,n-1,k)+p\times D(s,n-1,k)\end{aligned}\tag{14}$$

the distortion of DPB in right view is

$$\begin{aligned}D_{R,DPB}^I &= p\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n-1,k,i)]^2\}+(1-p)\times E\{[\hat{F}(s,n,k,i)-\tilde{F}(s,n,k,i)]^2\}\\ &= p\times E\{[\hat{F}(s,n,k,i)-\hat{F}(s,n-1,k,i)]^2\}+p\times E\{[\hat{F}(s,n-1,k,i)-\tilde{F}(s,n-1,k,i)]^2\}\\ &= p\times RFD(n,n-1,k)+p\times D(s,n-1,k)\end{aligned}\tag{15}$$

the intra-view inter-frame prediction distortion of DPC in right view is

$$\begin{aligned} D_{R,DPC}^{PI} &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s,n-1,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - e(s,n,k,i) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s,n-1,k,i)]^2\} + p \times E\{[\hat{F}(s,n-1,k,i) - \tilde{F}(s,n-1,k,i)]^2\} \\ &\quad + (1-p) \times E\{[\hat{F}(s,n-1,k,j) - \tilde{F}(s,n-1,k,j)]^2\} \\ &= p \times RFD(n,n-1,k) + p \times D(n-1,n,k) + (1-p) \times D(s,n-1,\bar{k}) \end{aligned} \tag{16}$$

The inter-view inter-frame prediction distortion of DPC in right view is

$$\begin{aligned} D_{R,DPC}^{PV} &= p \times E\{[\hat{F}(s,n,k,i) - \tilde{F}(s-1,n,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s,n,k,i) - e(s,n,k,i) - \tilde{F}(s-1,n,k,j)]^2\} \\ &= p \times E\{[\hat{F}(s,n,k,i) - \hat{F}(s-1,n,k,i)]^2\} + (1-p) \times E\{[\hat{F}(s-1,n,k,j) - \tilde{F}(s-1,n,k,j)]^2\} \\ &= p \times RFD(s,s-1,k) + p \times D(s-1,n,k) + (1-p) \times D(s-1,n,\bar{k}) \end{aligned} \tag{17}$$

The total distortion of DPC in the right view is

$$\begin{aligned} D_{R,DPC}^{P} &= \gamma D_{R,DPC}^{PI} + (1-\gamma) D_{R,DPC}^{PV} \\ &= p \times \{\gamma \times RFD(n,n-1,k) + (1-\gamma) \times RFD(s,s-1,k) + \gamma \times D(s,n-1,k) + (1-\gamma) \times D(s-1,n,k)\} \\ &\quad + (1-p) \times \left\{\gamma \times D(s,n-1,\bar{k}) + (1-\gamma) \times D(s-1,n,\bar{k})\right\} \end{aligned} \tag{18}$$

In summary, the total transmission distortion includes the transmission distortion diffuse from the previous coding frame and the error concealment distortion by recoverying the packet loss. The transmission distortion of current block is the weighted sum of these two parts, and the weighting factors depend on packet loss rate and coding mode rate.

3. Simulation Results

The loss of DPA will induce the whole slice information permanently damaged because it contains the data header information. While DPB contains only intra-frame information which account for rather small proportion especially for the sequences with non-violent movement, so we do not modeling the transmission of DPA and DPB, while only build up stereo video transmission distortion model based on DPC. In order to verify the validity and accuracy of the distortion model, testing has gone on sequences with different movement at different packet loss rate on JM code platform. This paper adopts IPPP encoding structure, and we assume that I frame will not be lost. Encoded bit stream go through the analog simulator at different packet loss rate, and using error concealment technology to recover the distorted video, then the concealed sequences are compared with corresponding encoding reconstruction sequences to get the actual distortion values by MSE. Meanwhile, we can obtain the theoretical distortion values under same condition. This paper selects the sequence of Book which has a gentle movement and Flamenco2 which has a violent movement to do the test on the condition that the packet loss rate equals to 8% and the length of GOP is 12. Table 1, and Table 2, list distortion values of Book and Flamenco2 in 12 frames respectively. Figure 1, and Figure 2, show the estimating distortion and actual distortion of sequence Book and Flamenco2 respectively.

Table 1. The Transmission Distortion of Book Sequence at 8% Packet Loss Rate (MSE)

Frame Number	Left View(View 0)		Right View(View 1)	
	Estimation	Actual	Estimation	Actual
1	0	0	45.39	49.3
3	11.05	8.95	33.02	29.88
6	32.31	28.93	44.74	36.12
9	74.47	69.95	65.99	72.87
12	103	99.88	123.7	118.4

Table 2. The Transmission Distortion of Flamenco2 Sequence at 8% Package Loss Rate (MSE)

Frame Number	Left View(View 0)		Right View(View 1)	
	Estimation	Actual	Estimation	Actual
1	0	0	73.17	62.62
3	32.47	27.29	100.5	94.06
6	73.86	67.1	128.5	121.4
9	107	100.3	143.1	134.3
12	142.3	140.4	167	163.6

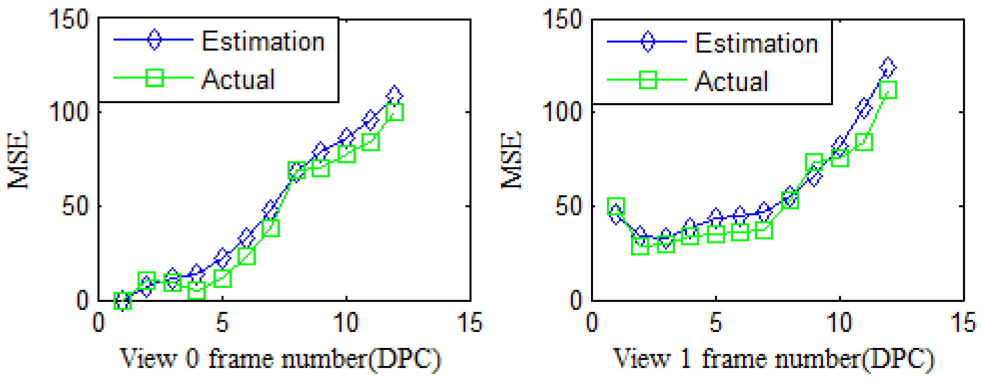

Figure 1. Average Distortion of DPC in Sequence Book at p= 8% (MSE)

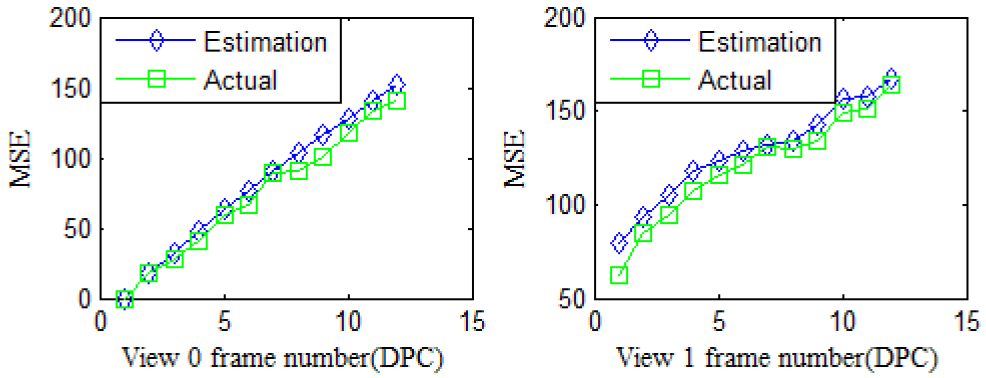

Figure 2. Average Distortion of DPC in Sequence Flamenco2 at p= 8% (MSE)

The diamond lines and square lines represent the modeling prediction distortion and the actual distortion respectively. As shown in Figure 1, and Figure 2, we can see that as the frame number increases, the left and right distortion values of the two sequences increase accordingly. Since the data loss cause not only the distortion in current frame, but also the diffusion to subsequent frames, and the accumulation of errors leads to the growing distortion. We repeat the random packet loss for 100 times and use the average of it to obtain the actual transmission distortion. The experimental results show that the proposed algorithm can be applied to different packet loss rate and the sequence with different movement intensity.

Besides, we repeat the test on sequences with different characteristics to verify the applicability of the proposed distortion model, and we define the average prediction error of viewpoint to verify the accuracy of the model as following:

$$\mathrm{E}(s) = \frac{\sum_{t=0}^{N} \left| D_E(s,t) - D_A(s,t) \right|}{\sum_{t=0}^{N} D_A(s,t)} \times 100\% \qquad (s = 0,1) \tag{19}$$

where $D_E(s,t)$ is the simulated distortion at moment t in view s， $D_A(s,t)$ is the actual distortion. Table 1 shows the average prediction error of sequences with different characteristics in different network environment.

Table 3. The Average Prediction Errors of Different Sequences

frame number	Book		Crowd		Alt		Flamenco2	
	View 0	View 1	View 0	View 1	View 0	View 1	View 0	View 1
3	2.86	0.73	1.82	2.51	2.26	0.61	7.58	7.08
5	7.99	5.58	3.44	0.43	0.34	0.64	8.98	4.57
8	1.77	6.56	1.59	1.29	2.77	0.35	2.28	2.11
10	4.71	5.60	3.56	1.15	7.50	3.94	3.97	7.57
Average	4.33	4.61	2.60	1.34	3.21	1.38	5.70	5.33

Book is a sequence with slow movement and its camera distance is 6.5 centimeter. Alt is a sequence with intense movement and its camera distance is 6.5 centimeter. Crowd is a sequence with intense movement and its camera distance is 20 centimeter. Flamenco2 is a rotating sequence with intense movement and its camera distance is 20 centimeter. Table 1 shows us the average distortion error of different sequences with various characteristic. The average distortion error of left view of these sequence is respectively 4.33, 2.60, 3.21, 5.70, and the average distortion error of the right view is respectively 4.61, 1.34, 1.38, 5.33. Thus it can be concluded that the proposed model is applicable for sequence with different movement intensity and different parallax at different packet loss rate.

4. Conclusion

The paper propose a recursive algorithm to determine the transmission distortion caused the loss of different DP parts, it not only considers the impact of DP losses on the video quality of current reconstructed frame, but also on the video quality of the subsequent frames in the GOP. With the proposed method, we can accurately and efficiently estimate the transmission distortion at the encoding end. What’s more, the algorithm can be further utilized in application layer for providing efficient error protection for stereo video transmission and achieve the improved transmission performance even though packet loss is inevitable.

Acknowledgements

The work was supported by the National Science-technology Support Plan (No. 2012BAH67F01), National Nature Science Foundation of China (No. 60832003, No.

61071120), Zhejiang Science bureau/Innovation team design project (No. 2012R10009-08), Zhejiang Nature Science Foundation (No. Y1110161), Ningbo Science-technology Innovation team research project (No. 2011B81002), Zhejiang Education Bureau Scientific research project (No. Y201327703).

References

[1] R. Zhang, S. L. Regunathan and K. Rose, "Video coding with optimal intra/ inter mode switching for packet loss resilience", IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, **(2000)**, pp. 966–976.

[2] H. Yang and K. Rose, "Advances in recursive per-pixel end-to-end distortion estimation for robust video coding in H.264/AVC", IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 7, **(2007)**, pp. 845–856.

[3] Z. F. Chen, P. V. Pahalawatta and A. M. Tourapis, "Improved estimation of transmission distortion for error-resilient video coding", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 4, **(2012)**, pp. 636–647.

[4] J. L. Zhou, Z. L. Ma and Y. H. Fan, "Analysis of Video Transmitting Distortion", mini-micro systems, vol. 27, no. 2, **(2006)**, pp. 264-266.

[5] M. H. You and B. Wang, "Analytic Channel Distortion for Video over Wireless Network", Wireless Communication Technology, vol. 3, **(2003)**, pp. 44-47.

[6] F. Zhao, L. Y. Zhong and T. Jing, "Joint Coding Scheme for Wireless Video Based on End-to-End Distortion Optimization", Journal of Beijing Jiaotong University, vol. 34, no. 6, **(2010)**, pp. 71-75.

[7] Y. Wang, Z. Y. Wu and J. Boyce, "Modeling of transmission-loss-induced distortion in decoded video", IEEE Transaction on Circuits and Systems for Video Technology, vol. 16, no. 6, **(2006)**, pp. 716-732.

[8] Y. Zhou, "Channel Distortion Analysis and Modeling for Multi-View 3D Video Transmission over IP Netwoks", Tianjing University Doctoral Thesis, **(2011)**.

[9] G. Sun, "Research on Content-aware Network Transmitted Video Quality Estimation", Zhejiang University Doctoral Thesis, **(2009)**.

[10] G. Cote, S. Shirani and F. Kossentini, "Optimal mode selection and synchronization for robust video communications over error-prone networks", IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, **(2000)**, pp. 952–965.

[11] D. Zhang, J. Liang and I. Singh, "Fast transmission distortion estimation and adaptive error protection for H.264/AVC-based embedded video conferencing systems", Signal Processing :Image Communication, **(2013)**.

[12] R. Jin, L. Hong, C. Wang, L. Wu and W. Si, "A hilerarchical clustering community algorithm which missed the signal in the process of transmission", Review of computer engineering studies, vol. 2, no. 3, **(20l5)**, pp. 27-34.