

Rate-Distortion Optimized Layered Stereoscopic Video Streaming with Raptor Codes

A. Serdar Tan*, Anil Aksay†, Cagdas Bilen†, Gozde Bozdagi Akar†, Erdal Arıkan*

*Bilkent University / Department of Electrical and Electronics Engineering, Ankara, Turkey

†Middle East Technical University / Department of Electrical and Electronics Engineering, Ankara, Turkey

Abstract—A near optimal streaming system for stereoscopic video is proposed. Initially, the stereoscopic video is separated into three layers and the approximate analytical model of the Rate-Distortion (RD) curve of each layer is calculated from sufficient number of rate and distortion samples. The analytical modeling includes the interdependency of the defined layers. Then, the analytical models are used to derive the optimal source encoding rates for a given channel bandwidth. The distortion in the quality of the stereoscopic video that is caused by losing a NAL unit from the defined layers is estimated to minimize the average distortion of a single NAL unit loss. The minimization is performed over protection rates allocated to each layer. Raptor codes are utilized as the error protection scheme due to their novelty and suitability in video transmission. The layers are protected unequally using Raptor codes according to the parity ratios allocated to the layers. Comparison of the defined scheme with two other protection allocation schemes is provided via simulations to observe the quality of stereoscopic video.

I. INTRODUCTION

The recent advances in stereoscopic coding techniques and standardization efforts caused stereoscopic video transmission to gain considerable interest. Stereoscopic video is formed by the simultaneous capture of two video sequences corresponding to left and right views of human visual system. The increase in the size of the source data due to coding more than one views can be reduced by exploiting the dependency among left and right views. However, the transmission bandwidth requirement is obviously more than the monoscopic case and specific reliability methods has to be used for efficient transmission on error prone channels.

Significant portion of data transmission is carried on lossy packet networks. Common error protection schemes on packet networks utilize retransmissions or Forward Error Correction (FEC). Retransmission method can be used in video transmission applications as in [1]. However, retransmission methods may bring large latency. FEC schemes do not utilize retransmissions to provide reliability where protection against losses is inserted before lossy transmission. In literature, FEC methods are studied for video transmission as in [2], [3] and [4].

A novel technique that is suitable for transmission in lossy packet networks is fountain codes, also called as rateless codes. Fountain coding idea is proposed in [5] and followed by practical realizations such as LT codes [6], online codes [7] and raptor codes [8]. The main idea behind fountain coding is to produce parity packets on-

the-fly as many as needed. This approach is different than the general idea of FEC codes where channel encoding is performed for a fixed channel rate and all encoded packets are generated prior to transmission. In [6], the idea is proven to be efficient for large source data sizes, as in the case of video data, and does not utilize retransmissions. Fountain codes have gained attention in video streaming area in recent years [9], [10], [11].

Stereoscopic video is more prone to losses due to increase in the predictive structure. Views are coded dependent on each other, thus the loss of a packet from one of the views during transmission may cause distortion in both of the views. Specific loss protection and distortion minimization methods has to be used to obtain error resilient and robust stereoscopic video streaming. The main contributions in our work is twofold. First part includes the definition of a layered structure, analytical modeling of the RD curve of these layers and derivation of the optimal encoder bit rates for the layers. The second part includes estimation of the distortion in video quality per lost NAL unit of a layer and minimizing the average distortion of losing a single NAL unit.

The organization of this paper is as follows. In Section II, we describe the stereoscopic codec and define the layers of the stereoscopic video. In Section III, we describe the fountain codes and describe Raptor codes and their systematization. In Section IV, first the analytical model of the RD curve of the encoder distortion is modeled for each of the layers. Then, the analytical model is used to obtain the optimal bit rates of the layers to achieve minimum distortion. In Section V, we estimate the distortion of the loss a single NAL unit from the layers. Then, the estimated distortions for layers are used to minimize the expected distortion of losing a single NAL unit to obtain optimal Unequal Error Protection (UEP) ratios for each layer. In Section VI, we present the results of the simulations of stereoscopic video streaming and compare the performances of three different protection allocation schemes among the layers. Finally, in Section VII, we conclude and state our future work.

II. STEREOSCOPIC CODEC

In our experiments, multiview video codec based on H.264 [12] is used due to its low complexity and simplified decoding procedure. In this codec, B frames are not supported. However, the results can easily be extended for JMVM codec [13]. Any other layered stereoscopic

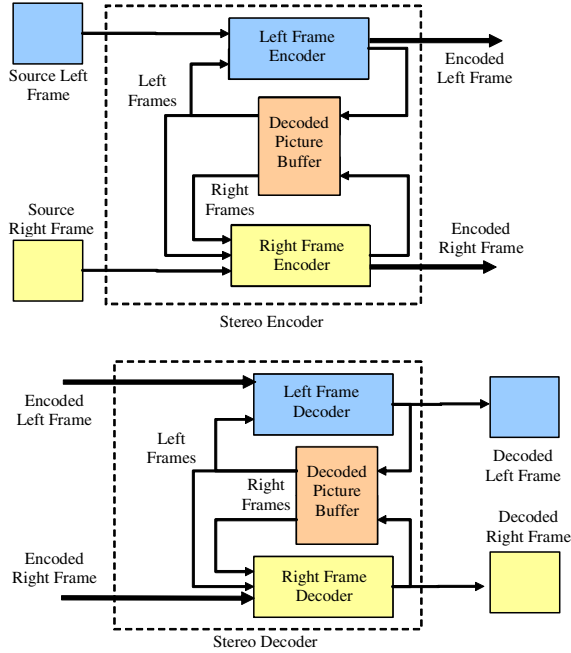


Fig. 1. Stereoscopic Encoder and Decoder Structure

codec can also be used with the proposed methods in our work. The codec in [12] uses a modified Decoded Picture Buffer (DPB) to perform both motion and disparity compensation with reduced complexity. For stereoscopic videos, a special mode allows for monoscopic compatible streams, where standard H.264 decoders can decode only left frames and stereoscopic decoder can decode both left and right frames. In monoscopic compatible mode, left frames are predicted from left frames only, whereas right frames can be predicted from both left and right frames. Right frames are always predicted from previous frames, whereas some of the left frames are encoded without prediction (i.e. I-frames). Stereoscopic encoder and decoder structure is given in Figure 1.

Denote \mathbf{I}_L , \mathbf{P}_L and \mathbf{P}_R as the set of I-frames of left view, P-frames of left views and P-frames of right views respectively. The set of frames can be written in open form as $\mathbf{I}_L = \{I_{L1}, I_{L5}, \dots\}$, $\mathbf{P}_L = \{P_{L2}, P_{L3}, \dots\}$, $\mathbf{P}_R = \{P_{R1}, P_{R2}, \dots\}$, where i denotes the frame number and L and R indicate the frames of left and right video. An illustration is given in Figure 2 where GOP size is 4.

Although this coding scheme is not layered, frames are not equal in importance. We can classify the frames according to their contribution to the overall quality and use them as layers of the video. Since losing an I-frame causes large distortions due to motion / disparity compensation and error propagation, I-frames should be protected the most. Among P-frames, left frames are more important since they are referred by both left and right frames. According to this prioritization of the frames, 3 layers are formed as shown in Figure 2. UEP protection on the defined layers will be explained in Section V.

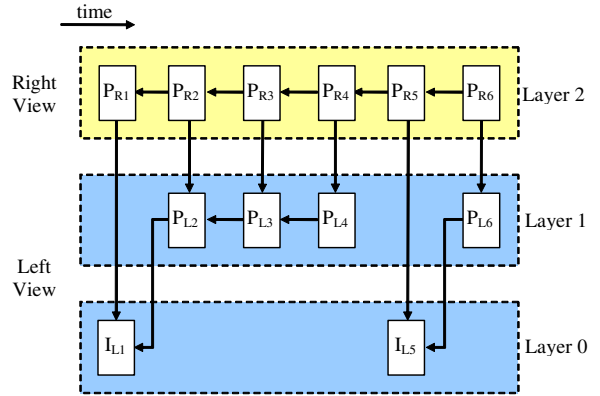


Fig. 2. Layers of stereoscopic video and referencing structure

III. FOUNTAIN CODES

A novel approach that provides retransmission free reliability, low latency and loss rate adaptability is fountain coding which is first mentioned in [5]. Fountain codes are well-suited for lossy packet networks. An ideal fountain encoder can generate potentially infinitely many encoding symbols from the original data consisting of k symbols in linear time and decoder can reconstruct the original data from *any* k -element subset of received encoding packets in linear time.

A. Raptor Codes

Luby Transform (LT) [6] codes are scalable codes proposed for reliable delivery of bulk data. Raptor codes [8] are an extension of LT codes with increased efficiency. Raptor codes are formed by two consecutive channel encodings. The input symbols are first encoded with a fixed high channel rate pre-code. The output of the pre-code is fed into an LT encoder to generate potentially limitless output symbols. The main reason behind the insertion of a pre-code is to reduce the coding overhead of the LT codes which is the main problem of LT codes for low number of input symbols as observed from the results in [6].

B. Systematic Raptor Codes

Raptor encoding can be performed as systematic or non-systematic. In the non-systematic coding schemes the encoded data is completely transformed into new symbols for protection. In the systematic case the original source data is included in the encoded data with protection symbols. The access to original data is beneficial in some cases such as video transmission where 100% reliability is not obliged. In systematic case, even if the channel decoder can not recover any lost source symbols, the video decoder still has some received parts of source data and error concealment techniques can be applied for the lost symbols. The systematization of Raptor codes is described in [11] in detail where encoding and decoding are similarly performed using Code Constraints Processing

(CCP) and LT encoder for both cases. The systematization process of raptor codes is briefly described in the following based on a similar notation in [11].

Raptor codes can be represented as linear block codes when the number of output symbols is fixed. Let the k source symbols be denoted as $\mathbf{E} = [E_1^T, E_2^T, \dots, E_k^T]^T$. The pre-coder generator matrix is denoted as \mathbf{G}_p and it has size k by s . The generated parity symbols in the pre-coder is denoted as \mathbf{D}_p and obtained as $\mathbf{D}_p = \mathbf{G}_p \mathbf{D}$. The intermediate symbols after pre-coder are formed as $\mathbf{F} = [\mathbf{D}^T, \mathbf{D}_p^T] = [F_1^T, F_2^T, \dots, F_{k+s}^T]^T$. The defined intermediate symbols are LT coded to form the final output symbols. The generator matrix of the LT encoder is given as $\mathbf{G}_{LT} = [\Gamma_1^T, \Gamma_2^T, \dots, \Gamma_n^T]^T$ where Γ_i are the row vectors with ones at positions corresponding to the index of XOR-summed input symbols. Code constraints processing is applied as $\mathbf{A}(1:k) \cdot \mathbf{F} = [0^T, E_1^T, \dots, E_k^T]^T$ to solve for the intermediate symbols where

$$\mathbf{A}(1:k) = \begin{bmatrix} \mathbf{G}_p & \mathbf{I} \\ \mathbf{G}_{LT(1:k)} \end{bmatrix}.$$

After the code constraints processing, LT encoding is applied to the intermediate symbols to obtain the output symbols as $\mathbf{G}_{LT} \mathbf{F} = [E_1^T, \dots, E_n^T]^T$ where first k rows are the source symbols and the last $n - k$ rows are the parity symbols.

The decoding of systematic Raptor codes uses the same scheme as the encoder. Let i_1, i_2, \dots, i_r denote the received output symbols. Then the code constraints processing is performed to obtain the intermediate symbols as $\mathbf{A}(i_1, i_2, \dots, i_r) \cdot \mathbf{F} = [0^T, E_{i_1}^T, \dots, E_{i_r}^T]^T$. Solving can be done by Gaussian elimination (ML decoder) but instead for lower complexity belief propagation can be used. After the intermediate symbols are obtained LT encoding is applied to obtain the source symbols as $\mathbf{C} = \mathbf{G}_{LT}(1:k) \cdot \mathbf{F}$. Raptor coding decreases the overhead of LT codes but still requires a small overhead to operate as stated in [8] which can also be shown by simulations.

IV. RD OPTIMIZATION

A. Encoder RD curve calculation

In this section, the approximate analytical model of the encoder rate distortion curve is obtained. In [14] a method that approximates the RD curve of monoscopic video is presented. The distortion measure in the RD models is the total mean squared error over all frames of a GOP of the corresponding layer. For monoscopic video the approximate analytical model of the RD curve is given with following equation,

$$D_e = \frac{\theta}{R_e - R_0} + D_0 \quad (1)$$

where D_e is the encoder distortion in MSE, R_e is the encoder rate. θ , R_0 and D_0 are the variables to be solved. Three samples of RD curve are enough to determine the stated variables. The RD curve model for the monoscopic case can also be used with layered video. The layers of stereoscopic video are defined in section II. I-frame (layer-0) is self decodable i.e, it does not depend on the

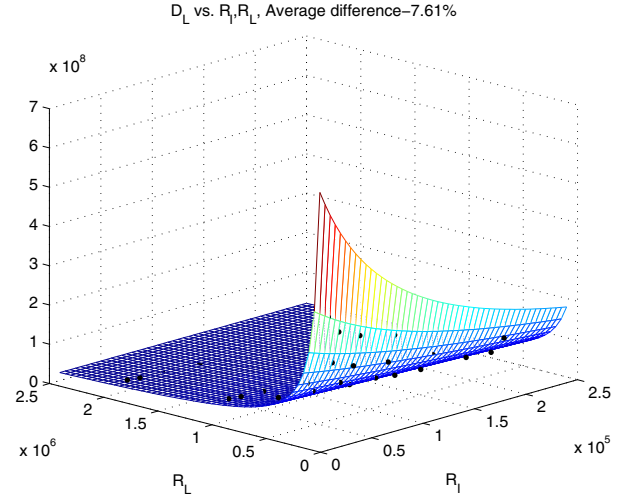


Fig. 3. RD curve fit for layer-1 (predicted L-frames)

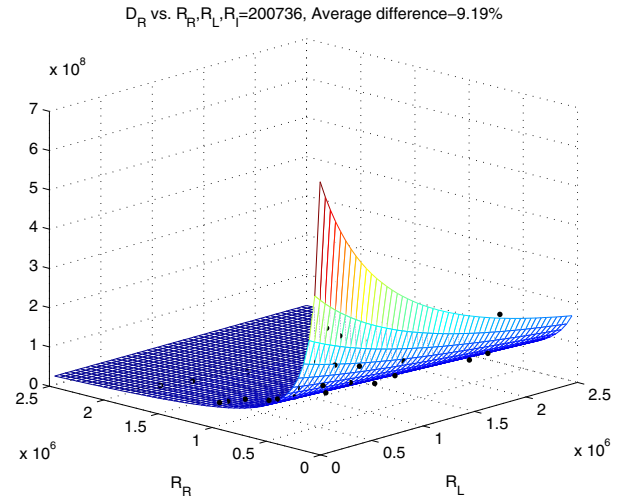


Fig. 4. RD curve fit for layer-2 (R-frames)

other layers. Thus, the RD curve model of layer-0 is given below,

$$D_I = \frac{\theta_I}{R_I - R_{0I}} + D_{0I} \quad (2)$$

where D_I is the encoder distortion in MSE, R_I is the encoder rate of I-frames. θ_I , R_{0I} and D_{0I} are the variables to be solved. 3 samples of RD curve is enough to solve for the stated variables.

Layer-1 (predicted L-frames) consists of predicted frames of left view. Layer-1 is coded dependent on the layer-0. Thus, the RD curve model is modified to handle the dependency as given below,

$$D_L = \frac{\theta_L}{R_L + c_1 R_I - R_{0L}} + D_{0L} \quad (3)$$

where D_L is the encoder distortion in MSE, R_L is the encoder rate of predicted L-frames. θ_L , c_1 , R_{0L} and D_{0L} are the variables to be solved. At least 4 samples of the RD curve is required to obtain the stated variables by curve fitting methods.

Layer-2 (R-frames) consists of the frames of right view. Layer-2 is coded dependent on layer-0 and layer-1. Thus, the RD curve model is modified to handle the dependency as given below,

$$D_R = \frac{\theta_R}{R_R + c_2 R_I + c_3 R_L - R_{0R}} + D_{0R} \quad (4)$$

where D_R is the encoder distortion in MSE, R_R is the encoder rate of predicted L-frames. θ_R , c_2 , c_3 , R_{0R} and D_{0R} are the variables to be solved. Similar to the previous case, at least 5 samples of the RD curve is required to obtain the stated variables by curve fitting methods.

The variables in the analytical models in Equations 2 to 4 are approximated by curve fitting tools separately. The match results for the 'Rena' video described in Section VI is given in Figure 3 for distortion of layer-1 and in Figure 4 for distortion of layer-2. The dots represent the experimental rate-distortion results of the encoded stereoscopic video for given bit rates for layers. The surface structure is the plot of analytical modeling of the RD curve of the video. As observed from the figures satisfactory fit is achieved where the analytical model approximates the experimental results with less than 10% error on the average.

B. Optimization on Encoder RD curve

After obtaining the approximate models of the RD curves of layers, all models are summed to obtain the overall distortion. The total distortion model is given below.

$$D_{ILR} = \frac{\theta_I}{R_I - R_{0I}} + \frac{\theta_L}{R_L + c_1 R_I - R_{0L}} + \frac{\theta_R}{R_R + c_2 R_I + c_3 R_L - R_{0R}} + D_0 \quad (5)$$

where $D_0 = D_{0I} + D_{0L} + D_{0R}$.

Using the analytical model of the RD curve, the optimal encoding rates for each of the layers can be calculated for a constant transmission bandwidth. The optimization is defined as given below.

$$\begin{aligned} \min_{(R_I, R_L, R_R)} D_{ILR} \\ \text{s.t. } R_I + R_L + R_R = (1-p)R_C \end{aligned} \quad (6)$$

In the above optimization process p and R_c denote ratio of total inserted parity packets and channel bandwidth respectively. The solution of this optimization can be calculated with the Lagrange multiplier method as following.

$$\begin{aligned} L(\lambda) &= D_{ILR} + \lambda (R_I + R_L + R_R - (1-p)R_C) \\ \frac{\partial L(\lambda)}{\partial R_I} = 0 &\Rightarrow R_I = R_{I0} + \sqrt{\frac{\theta_I}{\lambda(1-c_1-c_2+c_1c_3)}} \\ \frac{\partial L(\lambda)}{\partial R_L} = 0 &\Rightarrow R_L = R_{L0} - c_1 R_I + \sqrt{\frac{\theta_L}{\lambda(1-c_3)}} \\ \frac{\partial L(\lambda)}{\partial R_R} = 0 &\Rightarrow R_R = R_{R0} - c_2 R_I - c_3 R_L + \sqrt{\frac{\theta_R}{\lambda}} \\ \frac{\partial L(\lambda)}{\partial \lambda} = 0 &\Rightarrow R_I + R_L + R_R - R_C = 0 \\ \Rightarrow \lambda &= \left(\frac{\sqrt{\theta_I(1-c_1-c_2+c_1c_3)} + \sqrt{\theta_L(1-c_3)} + \sqrt{\theta_R}}{R_C - (1-c_1-c_2+c_1c_3)R_{I0} - (1-c_3)R_{L0} - R_{R0}} \right)^2 \end{aligned}$$

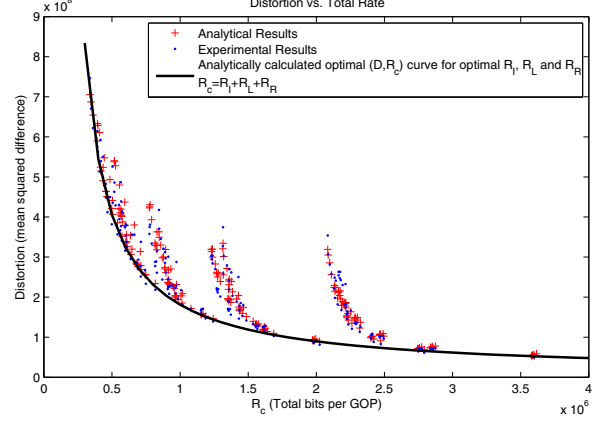


Fig. 5. Optimized RD curve for 3 layers

The optimization in Equation 6 is repeated for several values of R_c to obtain optimal values of R_I , R_L and R_R . The optimized D vs. R_c curve is obtained and plotted for lossless case in Figure 5. As clearly observed, the optimal curve accurately approximates the convex hull of the RD samples of the stereo encoder as expected.

V. UEP ALLOCATION ON LAYERS

The aim of the UEP allocation is to find the UEP rates that minimize the average distortion of losing a single NAL unit. The average distortion of losing a NAL unit from the defined layers can be calculated off-line. The losses for layers are assumed to be independent. Denote D_{Iloss} , D_{Lloss} and D_{Rloss} as the average distortion of losing a single NAL unit from layer-0, layer-1 and layer-2 and are calculated as below in Formulas 7, 8 and 9 respectively.

$$\frac{\sum_{k \in S_{MB}} \left[\sum_{x,y \in MB_k} \left(I_I(x,y,0) - \sum_{x',y' \in MB'_k} I_I(x',y',0) / N'_k \right)^2 \right]}{N_{NALU,I}} \quad (7)$$

$$\frac{\frac{1}{T-1} \sum_{i=1}^{T-1} \sum_{x,y} [I_L(x,y,i) - I_L(x,y,i-1)]^2}{N_{NALU,L}} \quad (8)$$

$$\begin{aligned} \frac{\sum_{x,y} [I_L(x,y,0) - I_R(x,y,0)]^2}{(T-1)N_{NALU,R}} \\ + \frac{\sum_{i=1}^{T-1} \sum_{x,y} \left[\left(\frac{I_R(x,y,i-1) + I_L(x,y,i)}{2} \right) - I_R(x,y,i) \right]^2}{(T-1)N_{NALU,R}} \end{aligned} \quad (9)$$

In the above equations S_{MB} denotes the set of macroblocks. $I(x,y,i)$ denotes the pixel in position (x,y) of i^{th} frame of corresponding layer of the original video. $N_{NALU,I}$, $N_{NALU,L}$, $N_{NALU,R}$ represent the number of

TABLE I
CALCULATED ENCODER BIT RATES AND UEP RATIOS FOR $R_c=500000$ BITS

Total Inserted Protection (p)	Encoder Bit Rates (Kbps)			Average Loss (p_e)								
	R_I	R_L	R_R	%3			%5			%10		
				p_I	p_L	p_R	p_I	p_L	p_R	p_I	p_L	p_R
5%	39.6	213.3	222.0	0.091	0.064	0.034	0.126	0.093	0.000	0.000	0.117	0.000
10%	37.3	202.4	210.1	0.197	0.108	0.099	0.180	0.114	0.096	0.291	0.193	0.000
20%	32.9	180.4	186.5	0.855	0.209	0.182	0.855	0.209	0.182	0.405	0.243	0.228
30%	28.4	158.5	162.9	0.891	0.396	0.379	0.891	0.396	0.379	0.891	0.396	0.379

NAL units of layer-0,1 and 2 respectively. T is the intra period. MB_i , MB'_i and N'_i represent the i^{th} macroblock, the set of i^{th} MB's neighbors and the number of neighbors of i^{th} MB respectively. The distortion in I-frames is approximated by spatial error concealment and the distortion in other frames are approximated by temporal error concealment.

The mean, μ , and variance, σ^2 , statistics of the losses in the channel is assumed to be known and assumed to be Gaussian. Denote p_I , p_L and p_R as the added protection ratio of layer-0, layer-1 and layer-2 respectively. The optimization is defined as below where D_{loss} is the average distortion of losing a single NAL unit.

$$\begin{aligned} \min_{(p_I, p_L, p_R)} D_{loss} &= D_{Iloss} P(p_e > p_I) \frac{R_I}{(1-p)R_C} \quad (10) \\ &+ D_{Lloss} P(p_e > p_L) \frac{R_L}{(1-p)R_C} \\ &+ D_{Rloss} P(p_e > p_R) \frac{R_R}{(1-p)R_C} \\ \text{s.t. } &p_I R_I + p_L R_L + p_R R_R = p R_C \end{aligned}$$

In the above optimization $\frac{R_I}{(1-p)R_C}$ denotes the probability that the NAL unit is of layer-0, p_e denotes loss rate per block and $P(p_e > p_I)$ denotes probability that there are more losses than the FEC scheme can recover. p_e is assumed to be Gaussian with mean μ and variance σ^2 as stated before. The optimization is performed to obtain a total parity bit rate that is equal to pR_C .

The optimal UEP values obtained by the minimization defined in Equation 10 for different p and p_e values are given in Table I for $R_C = 500Kb$. Generally, layer-1 is protected with highest channel rate. When the inserted parity ratio is less than the loss rate only layer-1 is protected. In other cases, layer-1 is protected better than the other layers. As the inserted parity ratio increases, the scheme tends to allocate more protection for layer-0, and layer-2 is protected close to the protection of layer-1. The optimal encoder bit rates of layers are also given in Table I whose values are calculated by the encoder distortion minimization given in Equation 6.

VI. EXPERIMENTAL RESULTS

The proposed streaming systems are evaluated via simulations. Three protection schemes are compared. First one is the proposed scheme described in Section V. The second scheme uses the results of [15] where best protection scheme for stereoscopic video is observed via simulations. The results in [15] state that when channel

protection is not enough to recover losses all protection should be allocated to I-frames. When the protection is enough, equal protection should be allocated to I-frames and L-frames, and no protection should be allocated to R-frames. The third scheme is EEP where all layers receive same parity ratio equal to p . The results of no-loss and no-protection cases are also presented. The no-loss case represents the quality of the video when the stereoscopic video is encoded with all available channel bandwidth and no transmission occurs. The no-protection case represents the transmission of the video of no-loss case in the lossy channel without any channel protection.

The results are provided for stereoscopic video pair *Rena* (Camera 38, 39) (640×480 , first 30 frames). I-frames are inserted every 30 frames. The proposed schemes for transmission of stereo H.264 /AVC streams are evaluated based on the ITU-VCEG loss patterns [16] and loss simulator [17]. As mentioned previously systematic Raptor codes are used based on their suitability for our case as explained in Sec. III-B. The encoded packets are generated according to the UEP method given in Sec. V. Since Raptor codes are probabilistic codes, loss simulation is repeated 100 times by changing the initial point of the loss pattern each time. NAL unit size is fixed to 250 bytes.

We provide a comparative analysis of the three different protection schemes for layered stereoscopic video streaming. In Figure 6, the results for the cases when $p \sim p_e$ are presented where protection is not enough to recover the losses due to the required small overhead in Raptor codes. In this case proposed UEP and the results stated in [15] yields similar results where the proposed UEP is slightly better. Both cases allocate the whole protection only to left view. In the EEP case, the insufficient protection is distributed equally to the layers which degrades the quality significantly.

In Figure 7, the results for the cases when $p > p_e$ are presented where the protection is enough to recover the losses. The proposed UEP protection performs similar but better than the EEP scheme. The scheme proposed in [15] does not perform well due to the inefficient distribution of the protection over layers where no protection is allocated to layer-2.

In Figure 8, the results for the cases when $p \gg p_e$ are presented where the protection is quite adequate to recover the losses. In this case EEP and the proposed UEP scheme performs similar due to allocation of high protection to all of the layers. The scheme proposed in [15] yields lower PSNR values due to the lack of

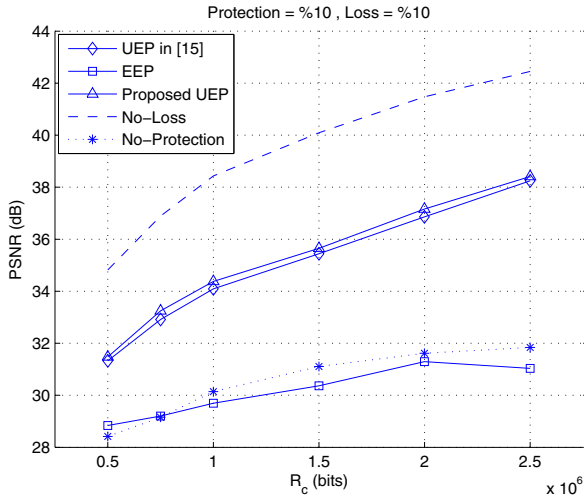


Fig. 6. PSNR results for 10% protection and 10% loss

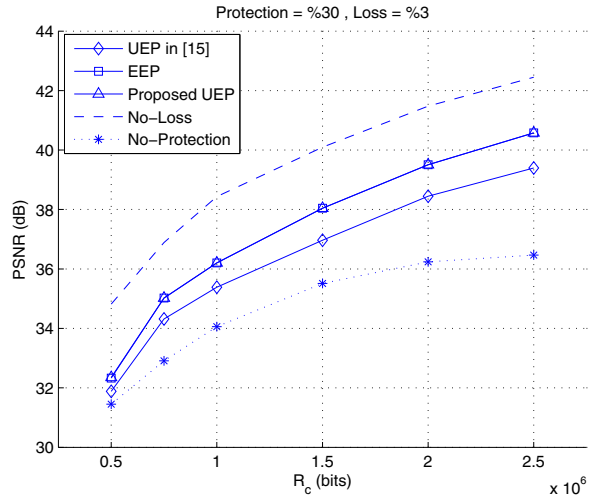


Fig. 8. PSNR results for 30% protection and 3% loss

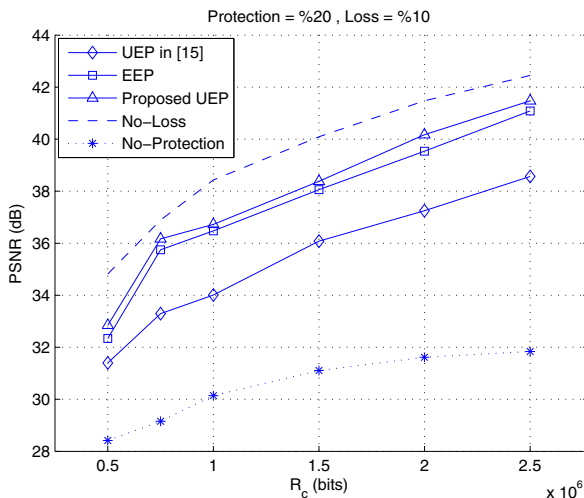


Fig. 7. PSNR results for 20% protection and 10% loss

protection for layer-2 where all protection is allocated equally to layer-0 and layer-1.

In all of the results, the quality of the no-protection scheme is also provided where only error concealment is used for error recovery. The stereoscopic video quality degrades quite significantly when adequate protection is not applied. These results clearly demonstrate the need of utilizing FEC codes, such as Raptor codes, for stereoscopic video streaming under lossy transmission channels.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we presented error resilient layered stereoscopic video streaming. We defined a layering structure specific to stereoscopic video and provided the analytical modeling of the RD curve of the layers. Using the analytical models we obtained optimal encoder bit rates to achieve minimum distortion in video quality. We also estimated the distortion in video quality per lost NAL unit and used them to find optimal UEP ratios for the layers.

Simulations are performed to observe the performance of the defined UEP schemes. The protection schemes yield different performances in different channel and protection conditions. The derived UEP scheme yields the best result in most of the cases, however significant quality gain is not observed. As a future work, the optimal values of encoder rates and UEP rates will be determined in a joint minimization to achieve more efficient stereoscopic streaming system.

ACKNOWLEDGMENT

This work was supported by the EC under contract FP6-511568 3DTV and in part by TUBITAK (Scientific and Technical Research Council of Turkey) under contract BTT-Turkiye 105E065. A.S.T, A.A and C.B are supported in part by TUBITAK.

REFERENCES

- [1] G.J. Conklin, G.S. Greenbaum, K.O. Lillevold, A.F. Lippman, Y.A. Reznik, R.N. Inc, and W.A. Seattle, "Video coding for streaming media delivery on the Internet," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, no. 3, pp. 269–281, 2001.
- [2] B. Girod, K. Stuhlmuller, M. Link, and U. Horn, "Packet loss resilient internet video streaming," in *Proc. SPIE Visual Commun. Image Processing*, 1999.
- [3] Hua Cai, Bing Zeng, Guobin Shen, Zixiang Xiong, and Shipeng Li, "Error-resilient unequal error protection of fine granularity scalable video bitstreams," *Applied Signal Processing, EURASIP Journal on*, 2006.
- [4] Y. Pei and J.W. Modestino, "H.263+ packet video over wireless IP networks using rate-compatible punctured turbo (rcpt) codes with joint source-channel coding," in *Proc. of the IEEE ICIP*, 2002.
- [5] J. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," in *Proceedings of ACM Sigcomm*, 1998.
- [6] M. Luby, "LT codes," in *Proc. of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2002, pp. 271–282.
- [7] P. Maymounkov, "Online codes," *Research Report TR2002-833, New York University*, Nov, 2002.
- [8] Amin Shokrollahi, "Raptor Codes," *Information Theory, IEEE Tans. on*, vol. 52, no. 6, pp. 2551–2567, June 2006.
- [9] J.P. Wagner, J. Chakareski, and P. Frossard, "Streaming of scalable video from multiple servers using rateless codes," in *Proc. IEEE Conf. on Multimedia and Expo (ICME)*, Toronto, Canada, July 2006.

- [10] M. Luby, T. Gasiba, T. Stockhammer, and M. Watson, "Reliable Multimedia Download Delivery in Cellular Broadcast Networks," *Broadcasting, IEEE Transactions on*, vol. 53, no. 1 Part 2, pp. 235–246, 2007.
- [11] M. Luby, M. Watson, T. Gasiba, T. Stockhammer, and W. Xu, "Raptor codes for reliable download delivery in wireless broadcast systems," in *Proc. of the IEEE CCNC*, 2006.
- [12] C. Bilen, A. Aksay, and G. Bozdagi Akar, "A multi-view video codec based on H.264," in *Proc. IEEE Conf. Image Proc. (ICIP)*, Oct. 8-11, Atlanta, USA, 2006.
- [13] A. Vetro, A. Pandit, H. Kimata, and A. Smolic, "Joint Draft 4.0 on Multiview Video Coding," *Joint Video Team*, vol. Doc. JVT-X209, 2007.
- [14] K. Stuhlmüller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 6, pp. 1012–1032, 2000.
- [15] A. Serdar Tan, Anil Aksay, Gozde Bozdagi Akar, and Erdal Arikan, "Error resilient layered stereoscopic video streaming," in *3DTV-Conference*, Kos, Greece, 2007.
- [16] S. Wenger, "Error patterns for internet experiments," in *VCEG Q15-I-16r1*, 2002.
- [17] Yi Guo, Houqiang Li, and Y.K. Wang, "SVC/AVC loss simulator donation," in *JVT-Q069*, 2005.