DESIGNING CODING STRUCTURES WITH MERGE FRAMES FOR INTERACTIVE MULTIVIEW VIDEO STREAMING

Benedicte Motz [*], *Gene Cheung* [#], *Ngai-Man Cheung* [$]

[*] The Graduate University for Advanced Studies, [#] National Institute of Informatics,
[$] Singapore University of Technology and Design

ABSTRACT

In interactive multiview video streaming (IMVS), a client periodically requests switches to neighboring views for uninterrupted temporal video playback from a server storing a large number of pre-encoded views. The technical challenge is that the navigation path taken by a client is not known at encoding time, and thus it is difficult to employ differential coding to lower code rate without knowing exactly what frames are available at the client buffer as predictors. In a previous work, a new frame type called merge frame was designed to efficiently merge different side information (SI) frames $\mathbf{S}^n$ from different possible decoding paths into a unique construction $\mathbf{M}$, so that the following frame(s) in time can be differentially coded using $\mathbf{M}$ as predictor without coding drift. In this paper, we design new coding structures using two variants of merge frame for different view-switching probabilities and desired rate-distortion (RD) tradeoff points. Experimental results show that our proposed frame structure designs outperform view-switching mechanisms in the literature, such as SP-frames in H.264, in RD performance.

Index Terms— Multiview video coding, interactive streaming, rate-distortion optimization

1. INTRODUCTION

Advances in image sensing technologies mean that a dynamic 3D scene can now be readily captured by an array of closely spaced cameras synchronized in time, so users can individually choose from which viewpoints to observe the scene. In an *interactive multiview video streaming* (IMVS) system [1, 2], such view interaction can take place between a server and a client connected via high-speed networks: a server pre-encodes and stores multiview video contents *a priori*, and at stream time a client periodically requests switches to neighboring views as the video is played back in time. See in Fig. 1 a *picture interactive graph* (PIG) for a video with three views that illustrates possible navigation paths chosen by users as the streaming video is played back in time uninterrupted.

Because the flexibility afforded by IMVS means that a client can take any one of many possible view navigation paths, at encoding time the server does not know *a priori*
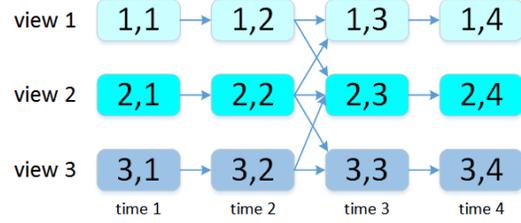

Fig. 1: Example of a *picture intereactive graph* (PIG) for IMVS, where a user can observe one of three available views at a given time, and can request a switch to a neighboring view every two frames as the video is played back in time.
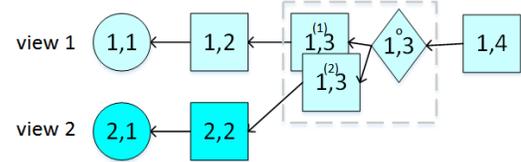

Fig. 2: Example of coding structure using M-frame proposed in [3, 4] for IMVS, where P-frames $P_{1,3}^{(1)}$ and $P_{1,3}^{(2)}$ predicted from previous frames of different views are first encoded, then an *optimized target* M-frame $M_{1,3}^o$ of the same view and instant is encoded to merge their differences to an identical reconstruction. I-, P- and M-frames are represented as circles, squares and diamonds respectively.

which frames will be available at the decoder buffer. This makes differential coding difficult to employ to reduce bitrate. The technical challenge is thus how to facilitate view-switching at stream time while still performing differential coding at encoding time for good compression efficiency.

In a recent work [3, 4], using *piecewise constant* (PWC) functions as operators, the authors proposed a new frame type called *merge frame* (M-frame) to efficiently merge slightly different *side information* (SI) frames $\mathbf{S}^n$ from different decoding paths into a unique construction $\mathbf{M}$, so that subsequent frames in time can use the identically reconstructed $\mathbf{M}$ as a common predictor for differential coding. As an example, in Fig. 2, two P-frames $P_{1,3}(1)$ and $P_{1,3}(2)$ of view 1 and time instant 3—these are the SI frames—are first predicted from $P_{1,2}$ and $P_{2,2}$ respectively. An M-frame $M_{1,3}^o$ of the same time instant is then encoded so that any one of SI frames $P_{1,3}(1)$ and $P_{1,3}(2)$ plus $M_{1,3}^o$ can result in an

identical reconstruction. Subsequent P-frame $P_{1,4}$ can then use $M_{1,3}^o$ as predictor for differential coding. M-frame thus provides a solution to facilitate view-switches (server sends combo of $(P_{1,3}(1), M_{1,3}^o)$ or combo of $(P_{1,3}(2), M_{1,3}^o)$ depending on user's chosen path) while permitting differential coding to lower bitrate. Experiments in [3, 4] show that M-frame can outperform existing switching mechanisms in the literature such as SP-frames [5] in expected and worst-case transmission rate when the probabilities of switching to any views are uniform.

To reconstruct M-frame, [4] proposed two methodologies. The first is called *optimized target merging*, where the distortion of the reconstructed M-frame $\mathbf{M}$ can be traded off with the encoding rate of $\mathbf{M}$, so long as the identical reconstruction condition from any SI frame $\mathbf{S}^n$ is met. However, the optimized target M-frame is nonetheless larger than a primary SP-frame, which is coded essentially the same as a P-frame with an extra quantization step. Notice that in the coding structure in Fig. 2, an M-frame is required for transmission for all decoding paths. Hence if the view-switching probabilities are skewed—*e.g.*, the probability of staying in the same view is very high—then the expected transmission rate of the view-switching structure in Fig. 2 can be larger than an SP-frame scheme that employs a primary SP-frame for the most likely decoding path.

In this paper, we propose alternative coding structures using M-frames that are suitable for different view-switching probability distributions. Specifically, we use also the second M-frame design methodology in [4] to reconstruct a *fixed target* M-frame, so that SI frames $\mathbf{S}^n$ are merged identically to a pre-specified target. Though fixed target M-frames are in general larger than optimized target M-frames, we show that in our structures, P-frames, optimized and fixed target M-frames can be used in novel combinations to minimize transmission costs for the more likely navigation paths, while still enabling view-switches for the less likely paths. Experimental results show that our proposed frame structures outperform view-switching mechanisms in the literature, including SP-frames in H.264 [5] and a previous M-frame structure [4], in RD performance.

The outline of the paper is as follows. We first discuss related works in Section 2. We overview the two variants of the previously proposed M-frame in Section 3. We describe our new designs of coding structures for different view switching probability distributions in Section 4. Finally, experimental results and conclusion are presented in Section 5 and 6, respectively.

2. RELATED WORK

In this section, we review some previous works on coding for IMVS. Beyond conventional I-, P- and B-frames, H.264 video coding standard introduced the concept of *SP-frames* [5] for stream-switching and interactive videos. SP-frames use loss-
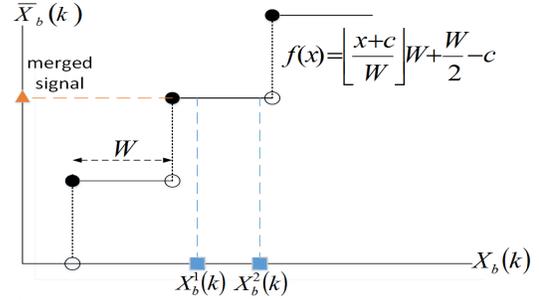

Fig. 3: Example of a piecewise constant (PWC) function $f(x)$ with two parameters (step size W and shift c) to merge two quantized coefficients $X_b^1(k)$ and $X_b^2(k)$ to an identical reconstruction $\bar{X}_b(k)$.

less coding of residue signal and are significantly larger than conventional P-frames. Thus, SP-frames suffer from inefficient coding. In particular, for each SI frame, an additional secondary SP-frame is needed. Thus SP-frames scale poorly with the number of SI frames. Alternatively, *distributed source coding* (DSC) has been a popular concept in designing interactive and stream-switching mechanisms [6–9]. However, partly due to the computation complexity required for bit-plane and channel coding, DSC is neither widely used nor adopted into any video coding standards.

Recently, the authors in [4] propose a coding construct based on the concept of "signal merging". Their work uses only quantization (PWC function) and arithmetic coding of horizontal shifts, both of which are simple, intuitive and well understood in the coding community. [4] has demonstrated coding gain over SP-frames and an implementation of DSC [9] using their approach. In addition, redundant P-frames and representation have been proposed for IMVS [1, 10]. The focus of these works is on frame structure optimization rather design of new coding tools.

3. MERGE FRAME OVERVIEW

In order to merge N slightly different SI frames $\mathbf{S}^n$ to a unique reconstruction $\mathbf{M}$, the key idea in [3, 4] is to employ a PWC function as a merge operator, whose parameters are explicitly coded in the M-frame, to merge quantized transform coefficients from different SI frames to the same values. Specifically, an SI frame $\mathbf{S}^n$ is first divided into fixed-size blocks of K pixels. Each pixel block b is transformed into the DCT domain and quantized into coefficients $X_b^n(k) \in \mathbb{I}$. Correct decoding of an M-frame means that the decoder, given only one set of $X_b^n(k)$ from an SI frame $\mathbf{S}^n$, can merge $X_b^n(k)$ to identical reconstruction $\bar{X}_b(k)$ via the use of specified PWC functions.

Suppose the `floor` function $f(x)$ with *shift parameter* c and *step size* W is used to merge coefficient $X_b^n(k)$ of block b from any SI frame $\mathbf{S}^n$ to a unique value $\bar{X}_b(k)$:

$$f(x) = \left\lfloor \frac{x + c}{W} \right\rfloor W + \frac{W}{2} - c \qquad (1)$$

That means *any* $X_b^n(k)$ of SI frame $\mathbf{S}^n$ must be floored to the same value:

$$\left\lfloor \frac{X_b^1(k) + c}{W} \right\rfloor = \left\lfloor \frac{X_b^n(k) + c}{W} \right\rfloor, \ \forall n \in \{2, \dots, N\} \quad (2)$$

(2) is known as the *identical merging* condition. Graphically, this also means that $X_b^n(k)$'s fall on the same step of the `floor` function, as illustrated in Fig. 3. The optimization is thus to select shift c and step size W for each coefficient k of block b so that (2) is satisfied.

In [4], the authors proposed two methodologies to select c and W to satisfy (2). In *fixed target merging*, a desired target value $X_b^0(k)$ is first selected *a priori*, and `floor` function parameters c and W are selected to ensure that $X_b^0(k) = f(X_b^n(k)), \forall n \in \{1, \dots, N\}$; *i.e.*, coefficients $X_b^n(k)$ of SI frames $\mathbf{S}^n$ merge identically to pre-selected $\bar{X}_b(k)$. It is proven [4] that step W must be chosen so that $W \geq 2 \max_n \left| X_b^0(k) - X_b^n(k) \right|$, and $c = X_b^0(k) \bmod W$. Because step W is typically chosen per-frequency while shift c is chosen per-block per-frequency, the overhead in coding c (via arithmetic coding) dominates the coding cost of an M-frame. Since c is the remainder of target $X_b^0(k)$, its probability distribution is roughly uniform in $[0, W)$, and thus the coding cost of a fixed target M-frame is relatively high.

In *optimized target merging*, there is no pre-selected target value for coefficient $X_b^n(k)$ to converge to: the converged value is selected based on an RD criteria. In this case, it is shown [4] that step W must now satisfy $W \geq \max_{n,m \in \{1,\dots,N\}} \left| X_b^n(k) - X_b^m(k) \right|$, while c is chosen to optimize an RD objective among all values that ensure identical merging condition in (2). This flexibility means that the probability distribution $Pr(c)$ can be designed to be skewed (not uniform), resulting in a low coding rate using arithmetic coding. Thus, in general, an optimized target M-frame is smaller than a fixed target M-frame.

4. DESIGNING FRAME STRUCTURES

We first overview different frame structures proposed in the IMVS literature to facilitate view-switching during video playback. We then describe two newly proposed structures using M-frames, which are useful when the view-switching probabilities are skewed. As done in previous works [1, 9–11], we will assume a user receives one single video view at a given time among many stored at a server, and can request a switch to a neighboring view every T frames.

4.1. Previous Frame Structures for IMVS

The simplest method to facilitate view-switches is to employ an intra-coded I-frame at each designated view-switching point [10]. As an example, in Fig. 4 we use an I-frame $I_{1,3}$ of view 1 and instant 3 to facilitate switches from P-frames $P_{1,2}$ and $P_{2,2}$ of view 1 and 2 respectively. However, because
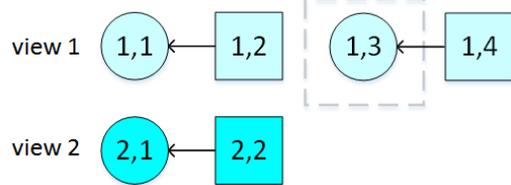

Fig. 4: Example of using I-frames for view-switching, where I-frame $I_{1,3}$ is encoded to facilitate switches from $P_{1,2}$ and $P_{2,2}$. I- and P-frames are represented as circles and squares respectively.
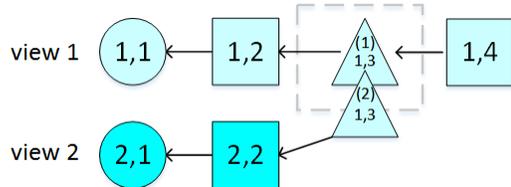

Fig. 5: Example of using SP-frames for view-switching, where primary and secondary SP-frames $SP_{1,3}^{(1)}$ and $SP_{1,3}^{(2)}$ are encoded to facilitate switches from $P_{1,2}$ and $P_{2,2}$ respectively. I-, P- and SP-frames are represented as circles, squares and triangles respectively.

I-frame requires a large coding overhead, it is not an efficient view-switching mechanism.

An alternative is SP-frames in H.264 video coding standard [5]. There are two kinds of SP-frames: *primary* and *secondary* SP-frames. A primary SP-frame is coded like a P-frame, with an extra quantization step after motion compensation so that transform coefficients of each fixed-size code block are quantized to integers. A secondary SP-frame is *losslessly* coded after motion compensation to reconstruct *exactly* the quantized coefficients of the primary SP-frame. Fig. 5 illustrates an example where a primary SP-frame $SP_{1,3}^{(1)}$ is encoded to enable switch from view 1 to 1, and a secondary SP-frame $SP_{1,3}^{(2)}$ is encoded to enable switch from view 2 to 1 at instant 3. The problem with SP-frames is that the lossless coding employed in secondary SP-frames means that the sizes of secondary SP-frames can be very large—often larger than I-frames. In the case where the probabilities of decoding paths using secondary SP-frames are non-negligible, the expected transmission cost can be significant.

4.2. Merge Frame Structures for IMVS

In [3, 4], the authors proposed to first predictively encode P-frames $P_i(j)$ for target picture of view i using decoded frames of different views j from different possible decoding paths as predictors, then encode an optimized target M-frame M_i^o to merge their differences to an identical reconstruction. We call this the *uniform probability merge* (UPM) structure. During an IMVS session, when a user requests a switch from view j to i, corresponding P-frame $P_i(j)$ and M-frame M_i^o are transmitted. The expected transmission cost C_{UPM} given view-
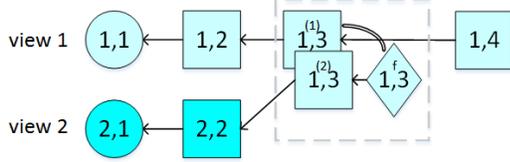
Fig. 6: Example of our proposed structure for high-probability view-switch, where P-frames $P_{1,3}^{(1)}$ and $P_{1,3}^{(2)}$ predicted from previous frames of different views are first encoded, then a *fixed target* M-frame $M_{1,3}^{f}$ with $P_{1,3}^{(1)}$ as target is encoded to merge SI frames ($P_{1,3}^{(2)}$ in this case) identically to $P_{1,3}^{(1)}$. I-, P- and M-frames are represented as circles, squares and diamonds respectively.

switching probabilities $p_{j,i}$ from view j to i is:

$$C_{\text{UPM}} = \sum_{j} p_{j,i} \left(P_i(j) + M_i^o \right) \tag{3}$$

An example is shown in Fig. 2, where P-frames $P_{1,3}(1)$ and $P_{1,3}(2)$ are first encoded, then an optimized target M-frame $M_{1,3}^o$ is encoded to merge their differences. UPM structure works well when the view-switching probabilities $p_{j,i}$ are roughly uniform, but because transmission of an M-frame is required for all decoding paths, when the distribution is skewed, C_{UPM} can be large.

In the case where one particular view-switch is highly probable, we propose a frame structure called *high probability merge* (HPM), shown in Fig. 6. The key idea is to lower the transmission cost of the most likely decoding path of view h by using a P-frame $P_i(h)$ only, then use a fixed target M-frame M_i^f to merge P-frames from other paths identically to $P_i(h)$. The expected transmission cost C_{HPM} is:

$$C_{\text{HPM}} = p_{h,i} P_i(h) + \sum_{j \neq h} p_{j,i} \left(P_i(j) + M_i^f \right) \tag{4}$$

In the example, assuming that the probability of staying in view 1 is very high, we encode a P-frame $P_{1,3}(1)$ for this decoding path, then a fixed target M-frame $M_{1,3}^f$ to merge SI frame $P_{1,3}(2)$ to $P_{1,3}(1)$. Note that the transmission cost of this decoding path from view 2 to 1 is *more* expensive than one in Fig. 2 because a fixed target M-frame is typically larger than an optimized target M-frame. However, because the highly probable path from view 1 to 1 is now less expensive, the expected transmission cost is lowered.

We now propose another structure called *low probability merge* (LPM) for the opposite case, where one or more switches from views $j \in \mathcal{L}$ are highly unlikely; an example is shown in Fig. 7. We first encode a P-frame (SI frame) for each possible decoding path—$P_{2,3}(1)$, $P_{2,3}(2)$ and $P_{2,3}(3)$ in this example. For the SI frames corresponding to likely decoding paths—$P_{2,3}(2)$ and $P_{2,3}(3)$—we first encode an optimized target M-frame $M_{2,3}^o$ to merge their differences to an RD-optimized identical reconstruction at low coding cost. Then for the SI frames corresponding to unlikely decoding
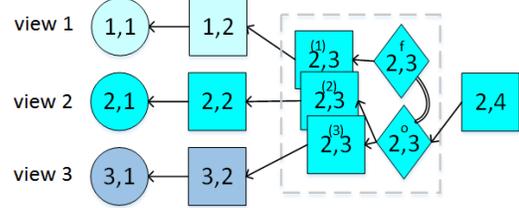

Fig. 7: Example of our proposed structure using M-frame for low-probability view-switch, where P-frames $P_{2,3}^{(1)}$, $P_{2,3}^{(2)}$ and $P_{2,3}^{(3)}$ predicted from previous frames of different views are first encoded. Then an *optimized target* M-frame $M_{2,3}^o$ is encoded to merge SI frames $P_{2,3}^{(2)}$ and $P_{2,3}^{(3)}$ to an identical reconstruction. Then a *fixed target* M-frame $M_{2,3}^f$ with $M_{2,3}^o$ as target is encoded to merge SI frame $P_{2,3}^{(1)}$ identically to $M_{2,3}^o$. I-, P- and M-frames are represented as circles, squares and diamonds respectively.

paths—$P_{2,3}^{(1)}$—we construct a fixed target M-frame $M_{2,3}^f$ to merge their differences to the target $M_{2,3}^o$. This means that for a likely decoding path, the server will transmit a P-frame and an optimized target M-frame, while for an unlikely decoding path, the server will transmit a P-frame and a fixed target M-frame. The expected transmission cost C_{LPM} is:

$$C_{\text{LPM}} = \sum_{j \in \mathcal{L}} p_{l,i} \left(P_i(j) + M_i^f \right) + \sum_{j \notin \mathcal{L}} p_{j,i} \left(P_i(j) + M_i^o \right) \tag{5}$$

The advantage of this structure over the one in Fig. 2 is as follows. The optimized target M-frame is now computed using *fewer* SI frames, simplifying the signal merging problem and resulting in a smaller size. This means that the likely decoding paths that require transmission of the optimized taget M-frame now enjoy a lower transmission cost. In contrast, the unlikely decoding paths now require the transmission of a fixed target M-frame, which is larger. However, because these paths are highly unlikely by assumption, they contribute little to the expected transmission cost.

5. EXPERIMENTATION

5.1. Experimental Setup

To demonstrate the coding performance of our proposed structures, we tested different structures in a scenario where the client can switch among video streams of different viewpoints. In particular, we used two different multiview video sequences with resolution 1024×768, Balloons and Newspaper, where the chosen views for encoding were $\{1, 3, 5\}$ and $\{3, 4, 5\}$ respectively. Using HEVC-15.0 [12] as the underlying codec, we encoded three streams for the chosen three viewpoints, with the second view designated as the target view; *i.e.*, the two side views can switch to the middle view at each view-switching point. The parameters for HEVC

were set as follows: QP ranged from 26 to 46, maximum CU size was 64×64, motion search range was 64, and entropy coding used was CABAC.

We encoded fixed and optimized target M-frames using the M-frame codec in [4] as follows. Code block size was fixed at 16×16. For fixed target M-frames, we set M-frames' QP to be the same as the SI frames' QP, as done in [4]. For optimized target M-frames, since the number of spikes in the probability distribution $Pr(c)$ of shift c of the `floor` function (merge operator) used for arithmetic coding ultimately determines the coding rate, we set QP for optimized target M-frame to be 1 and varied the number of spikes to induce different RD tradeoffs, as done in [4].

Note that in the experiments, for the sequence `Balloons` the optimized target M-frame and the fixed target M-frame are around 4 and 8 times bigger than a typical P-frame respectively. Similarly for the sequence `NewsPaper`, the optimized target M-frame and the fixed target M-frame are around 3 and 6 times bigger than a typical P-frame.

We compared the coding performance of our proposed HPM and LPM structures against SP-frames [5] in H.264, and the UPM structure using optimized target M-frames as previously proposed in [3,4]. To evaluate the performance of our proposal, we present plots of PSNR of the identically reconstructed M-frames / SP-frames versus expected transmission rate for given view-switching probabilities.

5.2. Experimental Results

At each switching instant, we considered three different view-switching probability distributions: i) unequal probability distribution where the client has a high probability to remain in the same stream ($p_{1,2} = p_{3,2} = 0.1$ and $p_{2,2} = 0.8$), and i) unequal probability distribution where the client has a small probability to switch from view 1 ($p_{1,2} = 0.04$ and $p_{2,2} = p_{3,2} = 0.48$), and iii) equal probability distribution where the client has the same probability to switch or stay in any stream ($p_{1,2} = p_{2,2} = p_{3,2} = 1/3$).

Plots of PSNR versus expected rate for the `Balloon` sequence for the three probability distributions are shown in Fig. 8. We observe that for all three probability distributions, UPM / HPM / LPM frame structures using M-frames have better RD performance than SP-frames (up to 65% at PSNR around 36dB), demonstrating the merits of M-frames. Further, depending on the probability distribution different structures among UPM / HPM / LPM are optimal at PSNR around 36dB. For distribution i, we see in Fig. 8a that HPM reduces bitrate by 60% compared to UPM, and by 50% compared to LPM. For distribution ii, we see in Fig. 8b that LPM reduces bitrate by 28% compared to UPM, and by 35% compared to HPM. Thus we have demonstrate the advantages of our newly proposed HPM and LPM structures. Finally, for distribution iii, we see in Fig. 8c that UPM reduces bitrate by 15% compared to HPM, and by 10% compared to LPM. The good
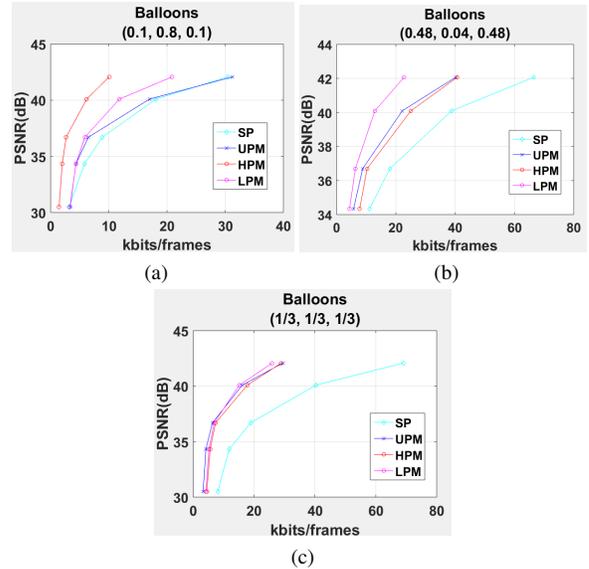

Fig. 8: PSNR vs. expected transmission rate for SP-frames, UPM, HPM and LPM frame structures for sequences `Balloons`

performance of UPM for uniform view-switching probability distribution is consistent with the results presented in [3,4].

We conducted the same experiment using sequence `Newspaper`; the results are presented in Fig. 9b. Similar to `Balloons`, structures using M-frames result in vastly better RD performance than SP-frames. The bitrate reduction numbers for different structures are summarized in Table 1. We note that unlike `Balloons`, for equal probability distribution, UPM reduces bitrate more significantly: 38% compared to HPM and 27% compared to LPM.

Table 1: Bitrate reduction using different coding structures

Prob.	(.1, .8, .1)		(.48, .04, .48)		(1/3, 1/3, 1/3)	
Struct.	H vs U	H vs L	L vs H	L vs U	U vs H	U vs L
Ball.	-60%	-50%	-35%	-28%	-15%	-10%
News.	-49%	-40%	-36%	-22%	-38%	-27%

6. CONCLUSION

Designing efficient coding schemes for interactive multiview video streaming (IMVS) systems—where a client can periodically request view-switches from server to navigate to neighboring views as the video is played back in time—is difficult, because at encoder time the server does not know with certainty what frames at the decoder buffer can serve as predictor for differential coding. While previous works [3,4] provided a partial solution by designing a new merge frame (M-frame) to merge different side information (SI) frames to an identical reconstruction, previous structures using M-frames are not efficient when the view-switching probabilities are skewed. In
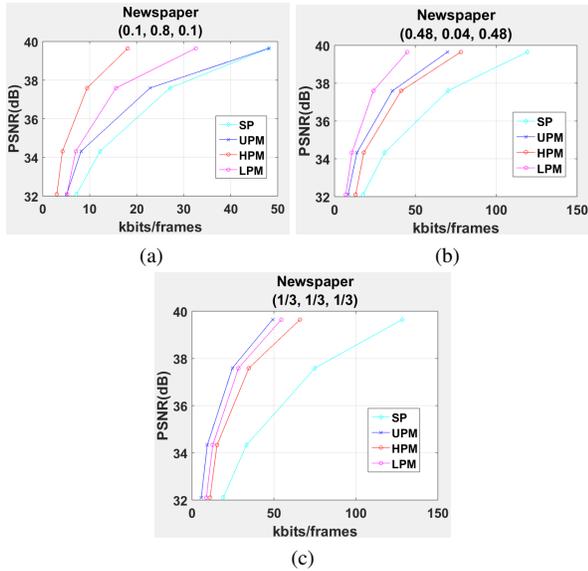
Fig. 9: PSNR vs. expected transmission rate for SP-frames, UPM, HPM and LPM frame structures for sequences `Newspaper`

this paper, we focus on the design of coding structures using two variants of M-frame (optimized target and fixed target) for applications with different view-switching probabilities. Our experiment results suggest that, for applications with skewed view-switching probabilities, considerable bitrate reduction can be achieved using our new frame structures.

7. REFERENCES

[1] G. Cheung, A. Ortega, and N.-M. Cheung, "Interactive streaming of stored multiview video using redundant frame structures," in *IEEE Transactions on Image Processing*, March 2011, vol. 20, no.3, pp. 744–761.

[2] X. Xiu, G. Cheung, and J. Liang, "Delay-cognizant interactive multiview video with free viewpoint synthesis," in *IEEE Transactions on Multimedia*, August 2012, vol. 14, no.4, pp. 1109–1126.

[3] W. Dai, G. Cheung, N.-M. Cheung, A. Ortega, and O. Au, "Rate-distortion optimized merge frame using piecewise constant functions," in *IEEE International Conference on Image Processing*, Melbourne, Australia, September 2013.

[4] W. Dai, G. Cheung, N.-M. Cheung, A. Ortega, and O. Au, "Merge frame design for video stream switching using piecewise constant functions," September 2015, http://arxiv.org/abs/1509.02995.

[5] M. Karczewicz and R. Kurceren, "The SP- and SI-frames design for H.264/AVC," in *IEEE Transactions on Circuits and Systems for Video Technology*, July 2003, vol. 13, no.7, pp. 637–644.

[6] M. Guo, Y. Lu, F. Wu, D. Zhao, and W. Gao, "Wyner-Ziv switching scheme for multiple bit-rate video streaming," in *IEEE Transactions on Circuits and Systems for Video Technology*, May 2008, vol. 18, no.5, pp. 569–581.

[7] A. Aaron, P. Ramanathan, and B. Girod, "Wyner-Ziv coding of light fields for random access," in *IEEE International Workshop on Multimedia Signal Processing*, Siena, Italy, September 2004.

[8] N.-M. Cheung and A. Ortega, "Compression algorithms for flexible video decoding," in *IS&T/SPIE Visual Communications and Image Processing (VCIP'08)*, San Jose, CA, January 2008.

[9] N.-M. Cheung, A. Ortega, and G. Cheung, "Distributed source coding techniques for interactive multiview video streaming," in *27th Picture Coding Symposium*, Chicago, IL, May 2009.

[10] G. Cheung, A. Ortega, and N.-M. Cheung, "Generation of redundant coding structure for interactive multiview streaming," in *Seventeenth International Packet Video Workshop*, Seattle, WA, May 2009.

[11] G. Cheung, N.-M. Cheung, and A. Ortega, "Optimized frame structure using distributed source coding for interactive multiview streaming," in *IEEE International Conference on Image Processing*, Cairo, Egypt, November 2009.

[12] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," in *IEEE Transactions on Circuits and Systems for Video Technology*, December 2012, vol. 22, no.12.