GENERALIZED SOURCE-CHANNEL PREDICTION FOR ERROR RESILIENT VIDEO CODING

*Hua Yang**

Corporate Research
Thomson Inc.
Princeton, NJ 08540

Kenneth Rose†

ECE Department
University of California
Santa Barbara, CA 93106-9560

ABSTRACT

This paper proposes an error-resilient modification of conventional (encoder-reconstruction based) prediction framework in video coding. The technique is called generalized source-channel prediction (GSCP) and generates a prediction reference for the next frame as a *weighted sum* of the current frame reconstruction and the prediction reference of the last frame. Compared to existing leaky prediction, GSCP achieves better coding efficiency in single layer video coding, and specifically better exploits the robustness benefits offered by Intra coded macroblocks in past frames so as to reduce *error propagation in the future frames*. Significant performance gains were observed in simulations and support the effectiveness of GSCP.

1. INTRODUCTION

Motion compensated prediction (MCP) is a commonly used technique to effectively remove temporal redundancy from original source video signals. Video coding standards generally adopt the classical predictive quantization framework, which uses past *encoder-reconstructed* frames for prediction. As is well known, this conventional framework was primarily designed for improving source coding efficiency, and generally ignores possible loss in the channel. Past reconstruction is used for prediction (closed loop), rather than the original (open loop), so that there is no prediction mismatch between encoder and decoder [1]. However, in the case of lossy communications, encoder and decoder mismatch is inevitable, and a revised paradigm is needed. This paper considers the fundamental problem of the optimal prediction scheme for lossy video transmission.

It is noteworthy that most error resilient video coding techniques, such as slicing, reference picture selection (RPS), and multi-frame motion compensation (MFMC) in H.263 and/or H.264, leave the conventional prediction framework intact. In [2], a source-channel prediction (SCP) scheme was proposed, which uses the *expected decoder reconstruction* of past frames for prediction. Hence, the problem addressed by SCP is that of optimizing the encoder prediction given the decoder and the information on the channel or network conditions.

Herein, we extend this line of research, and attack the more general problem: given the freedom to re-design the prediction mechanism at *both the encoder and the decoder*, how can we maximize the overall performance. For this problem, earlier contributions revolve around the concept of leaky prediction. The basic idea of leaky prediction is to combine predictive coding (in video – Inter coding) and non-predictive coding (Intra coding) and thereby facilitate the decay of propagating errors. In the context of lossy networks this may establish a better trade-off between coding efficiency and error resilience. Earlier efforts date back to work on error resilient DPCM coding systems [3] [4]. Leaky prediction was introduced into video coding in [5]. As it generally compromises video coding efficiency, existing practical schemes primarily focus on applying leaky prediction in layered video coding [6] [7] [8]. In this case, it is assumed that the base layer is subject to no (or minimal) loss, hence, a crude reconstruction is always available and may be readily integrated into leaky prediction. The resulting method offers error resilience at lower cost in terms of source coding efficiency. Recent analysis of leaky prediction in single or multiple layer video coding can be found in [9] and [10], respectively.

In this work, we propose a new solution to the problem of error resilient prediction, where the prediction reference of the next frame is composed as an appropriately weighted sum of the current frame reconstruction and the *prediction reference of the previous frame*. The proposed framework is in fact a generalization of our original SCP scheme [2], hence the name GSCP. GSCP involves no leaking, and generally yields better coding efficiency than leaky prediction in single layer coding scenarios. More importantly, GSCP offers an efficient means to exploit the potential of *Intra coded MB's in the past frames* to reduce the *error propagation effect in the following frames*. As shown via simulation, GSCP consistently yields

*Work performed while the first author was at the University of California, Santa Barbara.

†This work is supported in part by the NSF under grant EIA-0080134, the University of California MICRO Program, Applied Signal Technology, Inc., Dolby Laboratories, Inc., and Qualcomm, Inc.

significantly better overall system performance than that of leaky prediction under various testing scenarios.

2. THE LEAKY PREDICTION SCHEME

Rather than employ pure prediction, leaky prediction averages it with a constant term, resulting in decrease of the error propagation effect, and a corresponding improvement of error resilience. Typically, leaky prediction is defined as

$$\check{f}_n = \alpha \cdot \hat{f}_n + (1-\alpha) \cdot C. \quad (1)$$

Here, $\check{f}_n$ denotes the prediction reference frame, which uses the reconstruction of frame n to predict frame $n+1$. Coefficient α is called the leaky factor, and C is a constant. In many applications C is set to zero, but in video coding, where the pixel value ranges from 0 to 255, it is common to select the mid range level of 128 [4].

The error resilience offered by leaky prediction comes at some cost in video coding efficiency, due to the degradation in the prediction. This cost is reduced in layered video coding, where indeed most applications of leaky prediction can be found [6] [7] [8]. In this case, a high quality base layer reconstruction (usually involving no loss) is always available, and can be used to replace the constant component and implement leaky prediction with reduced compromise of coding efficiency. The leaky prediction reference for enhancement layers is defined as

$$\check{f}_{E,n} = \alpha \cdot \hat{f}_{E,n} + (1-\alpha) \cdot \hat{f}_{B,n}. \quad (2)$$

Here, B and E denote the base and enhancement layer, respectively. As adapting α has been found not to be cost effective in general, in most existing schemes, α takes a constant heuristic value, typically in the range $0.8 - 0.9$ [5] [7] [8]. Theoretical analysis on optimal leaky factor can be found in [10] and [9].

3. GENERALIZED SOURCE-CHANNEL PREDICTION

The GSCP framework for error resilient video coding, can be stated as

$$\check{f}^i_n = \alpha \cdot \hat{f}^i_n + (1-\alpha) \cdot \check{f}^i_{n-1}. \quad (3)$$

For some initial intuition, we note that in the case of $\alpha = 1$, GSCP is exactly the conventional pure prediction. At the other extreme, with $\alpha = 0$, GSCP is equivalent to always using the first frame (usually an I-frame) for prediction. Clearly, α controls the tradeoff between robustness and prediction quality.

Moreover, we identify that (3) can be viewed as a generalization of the SCP scheme proposed in [2]. In SCP, one uses $\alpha = (1-p)$ (where p is packet loss rate) so that the prediction becomes the expected reconstructed frame at the decoder. Note further that SCP employs this modified prediction only at the encoder. GSCP, on the other hand, offers a more flexible weighting of the two terms and modifies both encoder and decoder. (Due to space limitations, the reader is referred to [2] for details on SCP.) In practice, experiments have suggested a good choice of α that achieves optimal or near optimal performance is

$$\alpha = 1 - p - H, \quad (4)$$

where, H is a constant in the range of $0.1 - 0.2$. The rule is hence close but not exactly that of SCP encoding, and is applied in both encoder and decoder. We assume here that p is available at the decoder, in which case there is no need to transmit α to the decoder, and the proposed scheme is standard compatible.

Note that the only difference between GSCP in (3) and leaky prediction of (1) is in the second term, where GSCP uses the previous prediction frame rather than a constant. Thanks to this modification (3) is no longer a leaky filter and possesses significantly different properties.

Specifically, as will be demonstrated later, the use of $\check{f}_{n-1}$ instead of the constant C of (3), results in better prediction performance than leaky prediction in the important case of single layer video coding, and thus, yields higher coding efficiency. In terms of error robustness, an important advantage of GSCP is that it more effectively exploits *existent Intra coded MB's in the past frames* to reduce the *error propagation effect in the future frames*. Clearly, an intra-coded MB will yield less encoder-decoder mismatch that an inter-coded one, due to the effects of error propagation. By combining with $\check{f}_{n-1}$ GSCP effectively propagates a past Intra coded MB into the prediction. Note that in GSCP, contribution from past Intra coded MB's in the previous frame is guaranteed by the weighting factor $(1-\alpha)$. Simulation results show that GSCP consistently achieves better overall system performance than leaky prediction.

4. SIMULATION RESULTS

Our simulation setting builds on the latest JM9.0 H.264 codec. Herein, we used constrained Intra prediction and CAVLC for entropy coding. We adopted the rate control scheme from the JM codec and set one common quantization scale to all the MB's of one row. For each sequence, only the 1st frame was coded as I-frame, and all the rest were coded as P-frames. At the decoder, for each packet loss rate, 200 randomly generated packet loss patterns were applied, and average luminance PSNR was computed to measure the system performance.

The competing methods were tested under two extreme Intra updating scenarios: random Intra updating and optimal Intra updating. In random Intra updating, given packet loss rate p, a fraction p of MBs in each frame are selected for Intra coding. (The Intra MB's are selected according to the implementation in the JM9.0 encoder.) In optimal Intra updating,

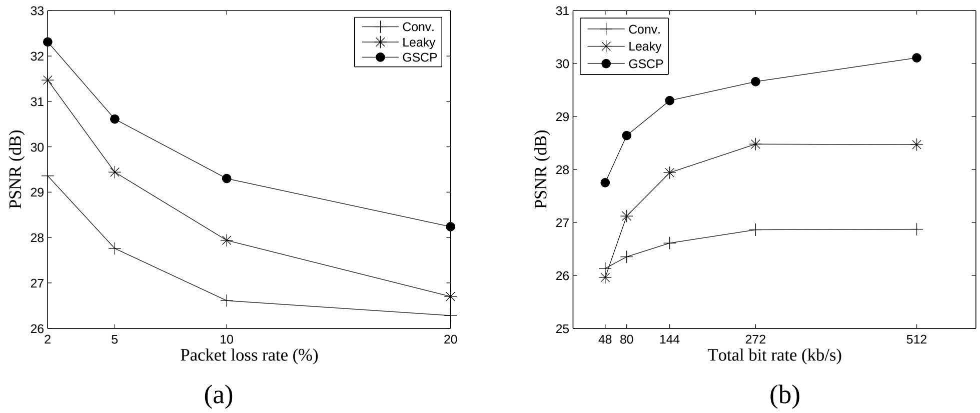


Fig. 1. Performance with random Intra updating. Carphone, QCIF, $10f/s$. (a) PSNR vs. Packet loss rate. $144kb/s$. (b) PSNR vs. Total bit rate. $p = 10\%$.

Table 1. Performance with various sequences in the scenario of random Intra updating. ($10f/s$, $144kb/s$, $p = 10\%$)

$PSNR\,(dB)$	Foreman	Carphone	Coastguard	Salesman	Mthr_dotr	Miss_am
Conv.	23.01	26.61	24.00	33.37	31.82	33.49
Leaky	24.38	27.94	25.29	31.61	32.30	33.08
GSCP	25.86	29.30	26.71	34.93	33.64	35.90

the coding mode of a MB is optimally selected from all the available coding mode options via an RD optimization framework. The involved Lagrangian multiplier is as suggested in the JM codec, and end-to-end distortion is accurately estimated using the recursive per-pixel method proposed in [11]. Herein, to eliminate from the results any possible irrelevant accuracy issues in the distortion estimation, only full-pixel motion estimation is conducted.

We compared the performance of our proposed GSCP scheme ("GSCP"), the existing leaky prediction scheme ("Leaky"), and the conventional prediction scheme ("Conv."). Herein, α of GSCP is as defined in (4) with $H = 0.13$, and α of leaky prediction in (1) is set to 0.95.

Fig. 1 gives the performance results in the scenario of random Intra updating. It is easy to see that while both GSCP and leaky prediction may achieve significantly better performance than that of the conventional scheme, GSCP consistently outperforms both the leaky prediction and the conventional scheme at all packet loss rates and total bit rates. For example, in Fig. 1 (a), the average gain from GSCP over leaky prediction is $1.96dB$. One can also see from Fig. 1 (b) that even at the low bit rate of $48kb/s$, where the performance of leaky prediction is a little worse than that of the conventional prediction, significant performance gain can still be achieved by GSCP.

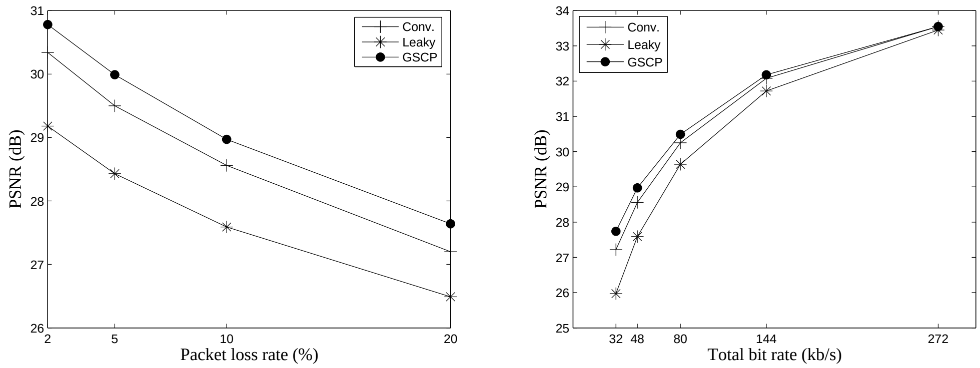
Further performance results with various testing sequences are also provided in Tab. 1. It is easy to verify that the aforementioned observations also hold here. Particularly, we can see that significant performance gain can be achieved by GSCP over conventional prediction for a variety of high and low motion sequences. The performance gains range from $1.56dB$ to $2.85dB$.

Fig. 2 presents the results in the scenario of optimal Intra updating. One obvious change of the results is that, due to the effectiveness of optimized Intra refreshing, the error propagation effect is already considerably mitigated, which implies a largely reduced scope for further performance enhancement. In spite of this, we can see from the figure that GSCP is still able to achieve significant performance gains, especially at low bit rates, e.g., below $48kb/s$. In contrast, the performance of leaky prediction is considerably worse than that of the conventional prediction scheme. Similar results can also be observed in Tab. 2. Therein, the average performance gain of GSCP over conventional prediction ranges from $0.17dB$ to $0.36dB$, while the PSNR performance *drop* due to leaky prediction reaches $5.37dB$ as for the sequence of "salesman".

In summary, we conclude that GSCP consistently outperforms both leaky prediction and conventional prediction in all the testing scenarios, which supports the analysis of Section 3, and provides evidence for the effectiveness of the proposed GSCP scheme.

5. REFERENCES

[1] A. Gersho and R. M. Gray, "Vector quantization and signal compression," *Kluwer Academic Publishers*,



(a) (b)

Fig. 2. Performance with optimal Intra updating. Carphone, QCIF, $10f/s$. (a) PSNR vs. Packet loss rate. $48kb/s$. (b) PSNR vs. Total bit rate. $p = 10\%$.

Table 2. Performance with various sequences in the scenario of optimal Intra updating. ($10f/s$, $48kb/s$, $p = 10\%$)

$PSNR\,(dB)$	Foreman	Carphone	Coastguard	Salesman	Mthr_dotr	Miss_am
Conv.	25.72	28.56	25.36	33.24	31.99	36.19
Leaky	25.31	27.59	25.17	27.87	30.37	34.34
GSCP	25.89	28.97	25.55	33.42	32.35	36.54

pp. 206-14, 2001.

[2] H. Yang and K. Rose, "Source-channel prediction in error resilient video coding," *Proc. of IEEE ICME 2003*, Baltimore, MA, USA, Aug. 2003.

[3] M. C. W. Van Buul, "Hybrid D-PCM, a combination of PCM and DPCM," *IEEE Trans. Commun.*, vol. COM-26, no. 3, pp. 362-8, Mar. 1978.

[4] D. J. Connor, "Techniques for reducing the visibility of transmission errors in digitally encoded video signals," *IEEE Trans. Commun.*, vol. COM-21, no. 6, pp. 695-706, 1973.

[5] M. Ghanbari and V. Seferidis, "Efficient H.261-based two-layer video codecs for ATM networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 2, pp. 171-5, Apr. 1995.

[6] H. C. Huang, C. N. Wang, and T. Chiang, "A robust fine granularity scalability using trellis-based predictive leak," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 372-85, June 2002.

[7] S. Han and B. Girod, "Robust and efficient scalable video coding with leaky prediction," *Proc. of IEEE ICIP 2002*, vol. 2, pp. 41-4, Rochester, NY, Sept. 2002.

[8] W.-H. Peng and Y.-K. Chen, "Error drifting reduction in enhanced fine granularity scalability," *Proc. of IEEE ICIP 2002*, vol. 2, pp. 61-4, Rochester, NY, Sept. 2002.

[9] Z. Li and E. J. Delp, "Channel-aware rate-distortion optimized leaky motion prediction," *Proc. of IEEE ICIP 2004*, pp. 2079-82, Singapore, Oct. 2004.

[10] Y. Liu, J. Prades-Nebot, P. Salama, and E. J. Delp, "Rate distortion analysis of leaky prediction layered video coding using quantization noise modelling," *Proc. IEEE ICIP 2004*, pp. 801-4, Singapore, Oct. 2004.

[11] 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, pp. 966-76, June 2000.