J. Vis. Commun. Image R. 31 (2015) 294–304

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Adaptive intra-refresh for low-delay error-resilient video coding[☆,☆☆]

Haoming Chen [a,*], Chen Zhao [a,1], Ming-Ting Sun [a], Aaron Drake [b,2]

[a] Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA
[b] T-Mobile USA, Bellevue, WA 12920, USA

ARTICLE INFO

Article history:
Received 23 February 2015
Accepted 30 June 2015
Available online 6 July 2015

Keywords:
Error-resilience
Intra-refresh coding
Low delay
Video coding
IPPP coding
Refresh cycle-size
Refresh order
H.264

ABSTRACT

Low-delay and error-resilient video coding is critical for real-time video communication over wireless networks. Intra-refresh coding, which embeds intra coded regions into inter frames can achieve a relatively smooth bit-rate and terminate the error propagation caused by the transmission loss. In this paper, we proposed a novel linear model for the intra-refresh cycle-size selection adapting to the network packet loss rates and the motions in the video content. We also analyze issues in designing the intra-refresh coding pattern and the refresh order, and propose a strategy which can adapt to different cycle-size and obtain better R–D performance compared with traditional random intra-refresh and vertical-partition intra-refresh. Experimental results show that the linear cycle-size selection model works effectively, where a 3 dB improvement can be achieved compared with a fixed cycle-size. Also, with the proposed intra-refresh order, a 2.0% bitrate reduction is obtained in average compared with the vertical-partition intra-refresh.

1. Introduction

LOW-delay and error-resilient video coding is critical for real-time video chat applications over wireless networks. On the delay aspect, given a fixed network throughput, the delay induced by the video codec mostly depends on the video encoder buffer size. On the error-resilience aspect, the error induced by the packet loss will propagate to the subsequent frames due to the nature of the predictive coding framework. To terminate an error-propagation, conventional video coding schemes encode an intra-coded frame (I-frame) in each Group-Of-Pictures (GOP) [1]. An I-frame is encoded without any reference to other frames, and so, it is not affected by errors in the previous frames. For a GOP with a size of K frames, in the worst case when the beginning I-frame is lost, error propagation is limited to within the following $K-1$ Predictive frames (P-frames), assuming for low-delay applications the B-frames (Bi-directional predicted frames) are not used. However, for the low-delay aspect, the GOP coding structure may not be the best choice. Since an I-frame only exploits the spatial redundancy within itself, it generates much more bits than a P-frame. The resulting bit-stream usually needs to be smoothed by a relatively large encoder buffer to ensure it does not exceed the network transport capability. This rate-smoothing encoder buffer could cause relatively long delay. Since the I-frame appears periodically, this delay is induced not only at the initial stage, but also in the following video transmission.

Note that besides the insertion of I-frame, some other error-resilient schemes including robust entropy coding [2,3], forward error correction (FEC) [4] and unequal error protection [5] are proposed. In [6], some error-resilient tools used in H.263 and MPEG-4 are reviewed. A more recent review of the error-resilient coding tools is in [7]. This review paper focuses on the error-resilient intra-prediction, which is more relevant to our work. Some techniques mentioned in [7] are also discussed in this section.

Intra-refresh (or intra-slice) coding schemes can provide low-delay and good error resilience features. In intra-refresh coding, instead of encoding the whole frame as an I-frame, a subset of Macro Blocks (MBs) in each frame can be forced into intra-coded MBs, so that after a cycle of frames, the whole frame is completely refreshed. This spreading of intra MBs into the P-frames can produce a quite uniform bit-rate. With the relatively uniform bit-rate, the encoder buffer could be avoided or kept to a minimal size to result in a low-delay video codec. Besides offering

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.
☆☆ Part of the work appears at the Asia–Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2014.

* Corresponding author.
E-mail addresses: eehmchen@ee.washington.edu (H. Chen), zhaochen@pku.edu.cn (C. Zhao), sun@ee.washington.edu (M.-T. Sun), Aaron.Drake@T-Mobile.com (A. Drake).
1 Present address: Peking University, Beijing, China.
2 The views presented are as individuals and do not necessarily reflect any position of T-Mobile USA.

http://dx.doi.org/10.1016/j.jvcir.2015.06.018
1047-3203/

low-delay, intra-refresh coding schemes also provide good error-resilience performance. Since each frame is completely refreshed after a cycle of frames, the parts of the picture affected by the transmission errors will be constantly refreshed. A vertical partition intra-refresh scheme as shown in Fig. 1 is applied in x264 [8], a popular open-source software, encoding videos into the H.264/MPEG-4 AVC format. In the vertical partition intra-refresh scheme, given the intra-refresh cycle-size of N frames, the whole frame is split into N regions vertically. In the mode decision process, the MBs within each intra-refresh region (shaded area in Fig. 1) are forced to be intra-coded.

To achieve a better performance under different network conditions, a fixed cycle-size intra-refresh scheme is not optimal. In [9], a rate-distortion model is proposed considering the channel (network packet loss rates) and the source (the intra-coded MB percentage) jointly. Based on this model, given the network condition, the optimal number of intra-coded MBs can be derived. However, some empirical and sequence-dependent parameters make the derivation of the optimal intra-coded MB percentage (and thus the number of frames in an intra-refresh cycle) difficult.

Besides the selection of the number of frames in an intra-refresh cycle, intra-refresh coding can be applied to MBs in different grouping patterns, resulting in different rate-distortion performance. One method is to choose the intra-refresh MBs randomly, which is the method used in [9,10] and the default intra-refresh method in the latest JM reference software JM18.6 [11]. In the random intra-refresh, the percentage of MBs refreshed in each frame is set to $1/N$. To avoid duplicated intra-refreshing, refreshed MBs are tracked so that one MB is not refreshed twice within one intra-refresh cycle. In this way, it guarantees that all MBs are refreshed once in a cycle. A method based on rate-distortion (R–D) cost (considering packet loss rate) of each Macroblock was proposed in [12] to improve the coding efficiency. However, the random or R–D-based MB selection may cause dislocation artifacts [13] when errors occur next to the intra-coded parts. Moreover, the scattered distribution of intra-coded MBs decreases the compression efficiency due to the constrained intra prediction: to make the intra-coded region not be affected by the error propagation from the inter-coded region, the intra-prediction pixels cannot come from the neighboring inter-coded region [14]. An attention-based adaptive intra-refresh coding scheme is proposed in [13] which employs an attention area (or region of interest) extraction algorithm and applies the intra-refresh on these grouped attention MBs. This scheme shows good subjective quality of the transmitted video over an error-prone network. However, the scheme does not cover the whole picture and errors out of the attention area may propagate for a long time. An isolated-region-based method is proposed in [15]. In the high packet loss rate scenario, this method works well. However, it has lower coding performance when the packet loss rate is low, since the prediction efficiency is not considered.

Some cyclic intra-refresh algorithms that can cover the whole frame are proposed in [16–20]. To guarantee a complete and efficient error recovery, the refreshed region in a cycle cannot predict from an unrefreshed part, otherwise the error may be propagated into the refreshed part and the intra-refresh scheme becomes ineffective. This protection principle for the refreshed region restricts the potential motion compensation candidates and cause a compression efficiency drop, especially when the refresh direction (e.g., left to right in Fig. 1) is opposite to the motion direction in the frames. This restriction affects the bitrate significantly. A bad intra-refresh order could increase the total bitrate as much as 10% compared with a good intra-refresh order [16]. Considering this problem, a motion adaptive intra-refresh scheme is proposed in [16]. In this scheme, every frame is split into 3×4 partitions and the refresh order of these 12 regions is obtained by training

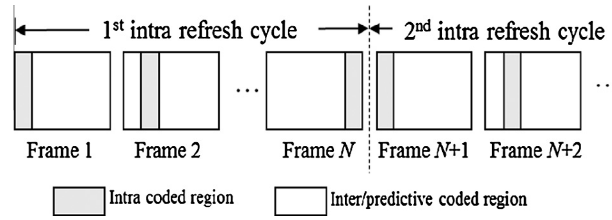

Fig. 1. Vertical partition of intra-refresh regions.

based on the motion in each partition region. This scheme can reduce the motion vector restriction effectively, but to be adaptive to the network packet loss rates, more refresh patterns and different cycle-sizes are needed.

In this paper, we propose a new intra-refresh coding scheme. In the first part, a content-based linear cycle-size selection model is proposed. Given the packet loss rate and the motion information, the best intra-refresh cycle-size is estimated. Compared with [9], our linear model is much simpler and has no empirical and sequence-dependent parameter. Also, since our model is a simple linear model, it gives better insights about how the video content affects the cycle size. It also relates the model parameters with the content and the motion of the video. All parameters in our model are trained from a set of real sequences and verified on a different set of testing sequences. In the second part, the design of intra-refresh region partition and refresh order is analyzed. Based on considerations in intra-refresh coding, we propose a rectangular partition for intra-refresh. Moreover, we modify and simplify the refresh order patterns in [16] so that they can be applied to different cycle-sizes. Experimental results show that the cycle-size model is effective and the proposed intra-refresh order outperforms the vertical cyclic intra-refresh scheme. Compared to the motion-adaptive intra-refresh scheme in [16], the proposed intra-refresh order has the flexibility that it can be applied to different cycle-sizes. The linear model in the first part is partially published in our previous work [21]. The major extension to [21] is the intra-refresh pattern design. In [21], we proposed the cycle number selection model. In this work, based on this model, we propose the intra-refresh pattern design, considering issues affecting the coding performance and the visual artifacts.

The rest of this paper is organized as follows. In Section 2, a linear model is proposed for the adaptive selection of the cycle-size. In Section 3, the design of the intra-refresh partition and the order are analyzed. Section 4 shows the simulation results. Section 5 concludes this paper and describes possible future works.

2. Adaptive content-based cycle-size selection

Fig. 2 compares different refresh cycle-sizes N ($N = 4$, 8, 16, and 32) under different packet loss rates (10^{-4}–10^{-1}). A smaller cycle is better for recovering from the errors more quickly when the packet loss rate is high, and a large refresh cycle is better for a low packet loss rate since it gives better coding gains with a smaller intra-coded area in each frame. Thus, for optimal performance, the refresh cycle-size N should be adaptive to the packet loss rates.

In this section, we proposed a linear model to predict the optimal intra-refresh rate based on the network packet loss rate and the motion in the video content.

2.1. A linear model for the intra-refresh rate based on the network packet loss rate

In [9], a joint end-to-end distortion model is proposed w.r.t. the intra-refresh rate and the packet loss rate: