Journal of Southeast University (English Edition) Vol. 20 No. 2 June 2004 ISSN 1003—7985

A novel intra refresh scheme for wireless video transmission

Feng Xiubo Xie Jianying

(Department of Automation, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: The burst error of wireless channels and the two-state Markov wireless model are analyzed. Based on this model and the coding modes of the video encoder, the channel distortion of inter-coding and intra-coding due to burst error is deduced. Then we propose a novel intra refresh scheme in rate-distortion (R-D) framework. This scheme optimizes the error resilience and coding efficiency of wireless video transmission system. It can also stop error propagation and reduce channel distortion effectively. Simulations under different channel conditions verify the improvements of the proposed scheme with respect to error resilience for wireless video communication.

Key words: intra refresh; rate-distortion; error resilience; Markov channel model; feedback

Prevailing video coding standards such as MPEG-4[1] and H. 263 +[2] employ motion estimation (ME) and motion compensation (MC) to reduce temporal redundancy, discrete cosine transform (DCT) to reduce spatial redundancy, and variable length code (VLC) to reduce statistic redundancy. These techniques lead to error propagation and loss of synchronization in wireless channels with burst errors. A variety of error resilience tools have been proposed to combat wireless transmission errors[3]. Intra refresh is an effective scheme to stop error propagation by switching the encoder from inter mode to intra mode and stopping dependence on the reference frame. However, due to its low compression efficiency, intra refresh should not be employed frequently. The tradeoff between compression efficiency and error resilient performance of intra refresh techniques should be considered.

Basically there are three types of intra refresh techniques. The first is the heuristic intra refresh scheme[4,5]. The heuristic techniques predefined the number of intra coded macroblocks (MBs) and did not take varying channel condition into account. The second is the intra refresh approach based on error tracking[6,7]. The encoder reconstructed the error distribution in the current frame via feedback information from the decoder. Severely affected MBs were selected to be intra coded. The error tracking methods require complicated computation and a great deal of memory. The last intra refresh approach incorporates an intra refresh scheme into the R-D framework and optimizes the overall performance. Cote introduced channel error into an R-D framework in Ref. [8]. Different error concealment schemes and corresponding methods to evaluate distortion due to error-prone networks have been discussed in Refs. [9, 10]. However, Refs. [8 − 10] addressed on video over wire channel with random errors and could not be applied to fading wireless channel directly. Again, the mismatch of packet loss rate of channel model and actual channel loss rate was not considered.

In this paper, we propose a novel intra refresh scheme in a wireless scenario. Though the feedback channel usually introduces additional delay and complexity, the combination of source and channel coding as well as the exploitation of feedback in the video transmission system, might lead to significant improvements in the overall system performance. In our proposed intra refresh scheme, with the aids of two-state Markov wireless channel model, feedback channel introduces no additional delay. The proposed scheme is standard compatible and appropriate to real-time end-to-end wireless video communications. Simulations show that the intra refresh scheme yields significant video quality improvement for fading wireless channel with burst errors.

1 Wireless Video Transmission Protocol and Wireless Channel Model

Fig.1 is the block diagram of an end-to-end wireless video communication system with feedback. In real time video service over wireless channels, two protocol stacks are of major interest. They are RTP/UDP/IP[11] for IP-switched mobile service and ITU Recommendation H. 324M[12] for circuit-switched channels. We adopt RTP/UDP/IP protocol stack for wireless video transmission. A typical RTP/UDP/IP packet consists of header information for IP, UDP, RTP, and RTP payload, as well as the payload of the compressed video bit stream itself. The head information

Received 2003-07-25.

Biographies: Feng Xiubo (1975—), male, graduate; Xie Jianying (corresponding author), male, professor, xiejianying@sjtu.edu.cn.

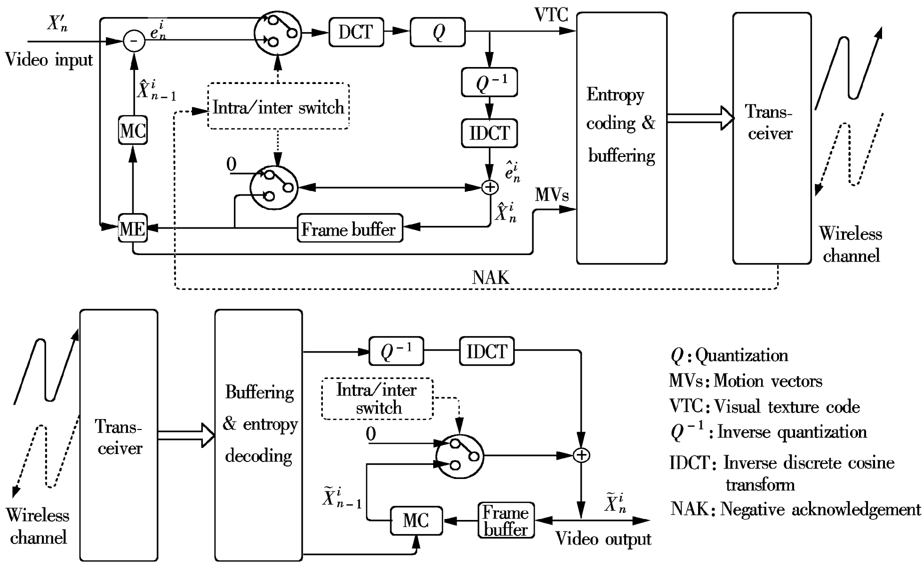


Fig.1 Wireless video transmission block diagram with intra/inter mode switch

of each packet is 40 bytes, in which 20 bytes for IP, 8 bytes for UDP and 12 bytes for RTP. RTCP is a companion protocol with RTP, which designed to provide quality of service (QoS) feedback to the participants of an RTP session. RTCP keeps the total control packets to 5% of the total session bandwidth. Among the control packets, 25% are allocated to the sender reports and 75% to the receiver reports[13].

In order to focus on the fading and bursting nature, we adopt a two-state Markov model[14] to describe the RTP/UDP/IP packet loss property over the wireless channel. There are two states in the Markov model: a good state (S_0) and a bad state (S_1). Video packets are transmitted undamaged in S_0 and errors occur in state S_1. S_0 and S_1 can switch to each other, as shown in Fig.2. The state-transition probability matrix is

$$\boldsymbol{P}=\begin{bmatrix} 1-P_{01} & P_{01} \\ P_{10} & 1-P_{10} \end{bmatrix} \tag{1}$$

where P_{10} is the transition probability from S_1 to S_0, and P_{01} vice versa. P_{01} and P_{10} can be calculated by average burst length and packet error rate as in Ref. [15]. Current channel states can be estimated through matrix $\boldsymbol{P}$ and initial states, i.e.,

$$\lfloor \pi_0^{\mathrm{c}} \quad \pi_1^{\mathrm{c}} \rfloor = \lfloor \pi_0^{\mathrm{i}} \quad \pi_1^{\mathrm{i}} \rfloor \boldsymbol{P}^d \tag{2}$$

where $\pi_0^{\mathrm{c}}, \pi_1^{\mathrm{c}}, \pi_0^{\mathrm{i}}, \pi_1^{\mathrm{i}}$ are the channel state probabilities of S_0 and S_1 at current time n and initial time $n-d$, respectively. Here, the time period is normalized with the time to transmit a video packet.

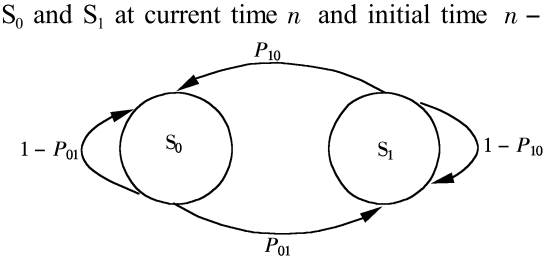


Fig.2 Two-state Marko model for fading wireless channel

2 Proposed Intra Refresh Scheme

2.1 Analysis of channel distortion

Various error concealment techniques can be applied at the decoder. If the encoder is aware of the error concealment scheme at the decoder, it can evaluate the distortion at the decoder. For the sake of less complexity and real time communication, we apply a simple but effective scheme error concealment in the proposed scheme. Only intra and inter modes are implemented at the encoder. If the decoder detects that uncorrectable error(s) occurred in a packet, the decoder simply copies the MBs at the same location of the erroneous packet from the previous frame.

As shown in Fig. 1, let X_n^i denote the original pixel value of pixel i in frame n; $\hat{X}_n^i$ and $\tilde{X}_n^i$ denote the encoder reconstruction pixel value and the decoded value of pixel i in frame n at the decoder, respectively. For the pixel of intra mode, if no error occurs during the video transmission, $\tilde{X}_n^i=\hat{X}_n^i$. Otherwise, $\tilde{X}_n^i=\tilde{X}_{n-1}^i$. So the expectation of distortions $d_n^i(I)$ in intra mode is

$$\begin{aligned} d_n^i(I) &= \| \hat{X}_n^i-\tilde{X}_n^i \|^2 = \pi_0 \| \hat{X}_n^i-\tilde{X}_n^i \|^2 + \\ &\pi_1 \| \hat{X}_n^i-\tilde{X}_{n-1}^i \|^2 = \pi_1 [\| \hat{X}_n^i-\hat{X}_{n-1}^i \|^2 + \\ &\| \hat{X}_{n-1}^i-\tilde{X}_{n-1}^i \|^2] = \pi_1 \| \hat{X}_n^i-\hat{X}_{n-1}^i \|^2 + \\ &\pi_1 d_{n-1}^i(I) \end{aligned} \tag{3}$$

where π_1 is the probability of channel states S_1 in the two-state Markov model. Note that the result of (3) is based on the assumption that the frame difference and channel distortion are not correlated to each other.

For the pixel of inter mode, if no error occurs during transmission, $\tilde{X}_n^i = \tilde{X}_{n-1}^j + \hat{e}_n^i$, where $\tilde{X}_{n-1}^j$ is the reference pixel of $\tilde{X}_n^i$ and $\hat{e}_n^i$ is the residual error after MC and ME. If uncorrectable errors are detected at the decoder, $\tilde{X}_n^i = \tilde{X}_{n-1}^i$, which is copied from previous decoded frame. Thus the expectation of distortions $d_n^i(P)$ in inter mode is

$$d_n^i(P) = \| \hat{X}_n^i - \tilde{X}_n^i \|^2 = \pi_0 \| (\hat{X}_{n-1}^j + \hat{e}_n^i) - (\tilde{X}_{n-1}^j + \hat{e}_n^i) \|^2 + \pi_1 \| \hat{X}_n^i - \tilde{X}_{n-1}^i \|^2 = \pi_0 d_{n-1}^j(I) + d_n^i(I) \quad (4)$$

where π_0 and π_1 are the probabilities of channel states S_0 and S_1 in the two-state Markov model, respectively. It should be noted that the second component in the result of (4) is the same as the distortion of intra refresh in (3). Therefore, $d_n^i(P)$ is larger than $d_n^i(I)$. This is the reason why the encoder should implement intra refresh scheme.

2.2 Feedback-based intra refresh scheme

In an error-free environment, given the distortion of video is D, the transmission rate is R and the rate constraint of the channel is R_c. The goal of video transmission system can be described as an optimization problem:

$$\min\{D\} \qquad \text{subject to } R \leqslant R_c \quad (5)$$

The optimization in (5) can be converted to an unconstrained problem via Lagrange multiplier λ as follows:

$$\min\{J\} \qquad J = D + \lambda R \quad (6)$$

In (6), the distortion is only from the quantization part in Fig.1.

In the error prone wireless channel, the channel distortion should be taken into account. We get

$$J = \sum_{i \in \mathrm{MB}} \| X_n^i - \hat{X}_n^i \|^2 + \sum_{i \in \mathrm{MB}} d_n^i(M) + \lambda_{\mathrm{mode}} R(M) \quad (7)$$

where $M \in (I, P)$. The first term in (7) represents the quantization error, the second is the channel error and the third is the bit rate weighted by the Lagrange multiplier. Using the Lagrange multiplier $\lambda_{\mathrm{mode}} = 0.85\ \mathrm{QP}^2$ has been proven to provide good R-D tradeoffs in an error free channel as in Ref. [16], where QP is the quantization parameter. We utilize the same value in our scheme. Therefore, the optimal process of (7) becomes straightforward computation. The task of intra refresh is switching the encode mode that minimizes J and optimizes the transmission performance.

We further simplify the mode selection criterion in (7). π_0 and π_1 can be assumed to be constant during the transmission of one packet due to short average bits in a packet. Let J_{intra} be the Lagrange value of intra mode and J_{intra} be the inter mode, we get

$$\Delta J = J_{\mathrm{intra}} - J_{\mathrm{intra}} = \pi_0 \sum_{i \in \mathrm{MB}} d_{n-1}^j(I) + \lambda_{\mathrm{mode}}(R(P) - R(I)) = \pi_0 \mathrm{SSD}(d_{n-1}^j(I)) + \lambda_{\mathrm{mode}} \Delta R \quad (8)$$

where π_0 and π_1 are channel states probabilities of S_0 and S_1, respectively; $\mathrm{SSD}(d_{n-1}^j(I))$ is the sum of the square distortion of the reference pixels in frame $n-1$; ΔR stands for the bit rate difference of inter and intra mode and can be precisely calculated at the encoder. Before the current MB is encoded, we calculate ΔJ via (8). If $\Delta J > 0$, the current MB should be intra refreshed. Note that if $\mathrm{SSD}(d_{n-1}^j(I))$ is significant, $\Delta J > 0$ will hold and intra refresh will be implemented. Therefore the error propagation can be eliminated effectively.

As mentioned before, the feedback mechanism is widely used in wireless video communication. In the proposed scheme, if the compressed packets transmitted from the video encoder in one frame cannot be decoded successfully, a negative acknowledgement (NAK) packet is formed and transmitted to the encoder, as the dashed line in Fig.1. The NAK packet contains the temporal and spatial information of the corrupted MBs. The encoder evaluates the packet loss rate from the NAK packet via QoS mechanism. We assume that the feedback channel is error free, i.e., the feedback packets can arrive at the encoder properly. This assumption is reasonable because the feedback packet usually contains a small amount of bits and can be applied enough protection.

When the encoder processes the MBs of the n-th packet, the feedback information of packet $n-d$, NAK_{n-d}, arrives at the encoder due to the feedback delay, where d is the so-called round trip delay (RTD). From NAK_{n-d}, the encoder can obtain $\mathrm{SSD}(d_{n-d}^j(I))$ as well as the channel state at time $n-d$. In order to decide whether an MB at time n requires to be intra refreshed, π_0, π_1 at time n and $\mathrm{SSD}(d_{n-1}^j(I))$ at time $n-1$ must be available. However, NAK_{n-1} and NAK_n cannot arrive at the encoder due to round trip delay. The actual values of π_0, π_1 at time n and $\mathrm{SSD}(d_{n-1}^j(I))$ at time $n-1$ are not available. Let the channel state at time $n-d$ as the initial channel state, we can predict the channel states at time $n-d+1$ to n according to (2). Also, from (3) and (4), $\mathrm{SSD}(d_{n-1}^j(I))$ is dependent on the distortion in frame $n-2$; $\mathrm{SSD}(d_{n-2}^j(I))$ is dependent on the distortion frame $n-3$, and so on. Therefore, $\mathrm{SSD}(d_{n-1}^j(I))$ can be calculated from the exact value $\mathrm{SSD}(d_{n-d}^j(I))$ according to (3) and (4). Due to the continuous update of initial states of the wireless model and the initial distortion information of video packet by NAK packets, errors accumulated by the two-state Markov wireless model mismatch can be reduced.

3 Simulation Results

To verify the improvement of our feedback-based intra refresh scheme, simulations were performed for different video sequences under two different channel conditions as in (9). The code was modified from TMN10 developed by UBC Image Processing Laboratory. No option mode was used in simulations. TMN8 rate control was used to obtain 64 kbit/s target bit rate. The test sequences in our experiments were 4∶2∶0 YUV format at QCIF(176 × 144) resolution. The test sequences were "Miss America" and "Foreman". The former was a sequence with average motion and the later with heavy motion as well as background change. The original frame rate was 30 frame/s. Two out of three consecutive frames were skipped and the resulting frame rate was 10 frame/s.

$$\boldsymbol{P}_1 = \begin{bmatrix} 0.987 & 0.013 \\ 0.250 & 0.750 \end{bmatrix},\ \boldsymbol{P}_2 = \begin{bmatrix} 0.988 & 0.012 \\ 0.111 & 0.889 \end{bmatrix} \quad (9)$$

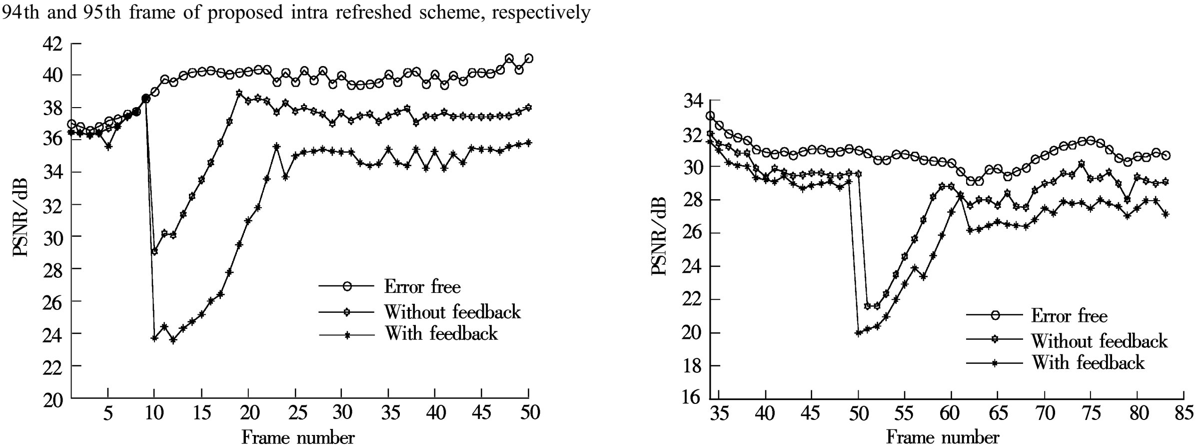
The MBs in the same row composed a group of block (GOB). Each GOB is transmitted in one RTP/UDP/IP packet. We generated error pattern files according to the channel parameters in (9) to simulate the error patterns of wireless channel. The RTP packets were corrupted via the generated error pattern files and then decoded. The feedback information can also be achieved from the error pattern files according to different RTDs. We also encoded the test sequence without the feedback information from the error pattern files as well as in error free conditions for comparison.

Fig.3 shows the subjective improvements of the proposed intra refreshed scheme. Figs.3(a) and (b) are the decoded 94th and 95th frames of sequence "Foreman" without feedback from decoder to encoder. Figs. 3(c) and (d) are the corresponding decoded frames of the proposed scheme. It is clear that the reconstructed picture quality of the proposed scheme is more acceptable. Experiments with sequence "Miss America" obtained the same results.

The frame-by-frame peak signal noise ratio (PSNR) of the first 150 original frames of "Miss America" and 150 frames (starting from frame 100 of original sequence) of "Foreman" under condition of $\boldsymbol{P}_1$ with feedback delay 300 ms is shown in Fig.4. The results of $\boldsymbol{P}_2$ and other delay conditions yielded similar results. Because the proposed scheme can get the error information in past frames after round trip time and predicts the distortion that might occur in the current frame via the feedback information and wireless channel model, it can recover the PSNR more rapidly in both sequences. Higher average PSNR is also achieved via the feedback information.

(a) (b) (c) (d)

Fig.3 Decoded frames of "Foreman" at 64 kbit/s. (a) and (b) are the 94th and 95th frames without feedback; (c) and (d) are the 94th and 95th frame of proposed intra refreshed scheme, respectively



(a) (b)

Fig.4 Comparison of PSNR under condition of $\boldsymbol{P}_1$. (a) Miss America; (b) Foreman

4 Conclusion

To combat fading wireless channel errors, we have proposed a more reliable intra refresh scheme, which can be applied in real time end-to-end wireless video communication. The two-state Markov model is adopted to simulate the fading wireless channel. The channel distortion is deduced based on the channel model and error concealment scheme. Then the optimized intra refreshed scheme can be decided in the R-D framework. Feedback information is utilized to reduce the channel model mismatch. Simulations demonstrate the improvement of the proposed scheme.

References

[1] IEO/JTC1/SC29/WG11. ISO-IEC 14496-2 Information technology-generic coding of audio-visual objects —part 2: visual [S]. Vancouver, MPEG, 1999.

[2] Cote G,Erol B, Gallant M. H. 263 +: video coding at low bit rates [J]. *IEEE Trans Circuits Syst Video Technol*, **1998, 8**(7): 849 - 866.

[3] Wang Yao, Wenger S, Wen Jiantao, et al. Error resilient video coding techniques [J]. *IEEE Signal Processing Magazine*, **2000, 17**(4): 61 - 82.

[4] ITU-T/SG15. Q15-D-65d1. Video codec test model, TMN10 [S]. Tampere, Finland, 1998.

[5] Worrall S T. Motion adaptive intra refresh for MPEG-4 [J]. *Electronics Letters*, **2000, 36**(23): 1924 - 1925.

[6] Chang Pao-Chi, Lee Tien-Hsu. Precise and fast error tracking for error-resilient transmission of H. 263 video [J]. *IEEE Trans Circuits Syst Video Technol*, **2000, 10**(4): 600 - 607.

[7] Steinbach E, Farber N. Standard compatible extension of H. 263 for robust video transmission in mobile environments [J]. *IEEE Trans Circuits Syst Video Technol*, **1997, 7**(6): 872 - 881.

[8] Cote G, Kossentini Fao. Optimal intra coding of blocks for robust video communication over the Internet [J]. *Signal Processing: Image Communication*, **2000, 15**(1): 25 - 34.

[9] Cote G, Shahram S. Optimal mode selection and synchronization for robust video communication over error prone networks [J]. *IEEE Journal on Selected Area in Communications*, **2000, 18**(6): 952 - 965.

[10] Zhang R, Regunathan S L. Video coding with optimal inter/intra-mode switching for packet loss resilience [J]. *IEEE Journal on Selected Area in Communications*, **2000, 18**(6): 966 - 976.

[11] Schulzrinne H, Casner S, Frederick R, et al. RTP: a transport protocol for real-time applications [EB/OL]. http://www.faqs.org/rfcs/rfc1889.html.1996/2003-07-01.

[12] Farber N, Girod B. Extensions of ITU-T recommendation H. 324 for error-resilient video transmission [J]. *IEEE Communications Magazine*, **2000, 36**(6): 120 - 128.

[13] Wu Dapeng, Hou Yiwei Thomas. On end-to-end architecture for transporting MPEG-4 video over the Internet [J]. *IEEE Trans Circuits Syst Video Technol*, **2000, 10**(6): 923 - 941.

[14] Aramvith S, Pao I M, Sun M T. A rate-control scheme for video transport over wireless channels [J]. *IEEE Trans Circuits Syst Video Technol*, **2001, 11**(5): 569 - 580.

[15] Elliott E O. Estimates of error rates for codes on burst-noise channels [J]. *Bell Syst Tech J*, **1963, 42**(9): 1977 - 1997.

[16] Sullivan G J, Wiegand T. Rate-distortion optimization for video compression [J]. *IEEE Signal Processing Magazine*, **1998, 15**(6): 74 - 90.

一种新的无线视频传输帧内刷新策略

冯秀波　　谢剑英

(上海交通大学自动化系, 上海 200030)

摘要: 首先分析了无线信道突发性出错和二阶 Markov 无线信道模型. 根据该模型和视频编码器的编码方式推导了突发性出错引起的帧内编码和帧间编码的信道失真. 然后提出了一种新的速率-失真率(R-D)框架内的帧内刷新策略. 该策略一方面对无线视频传输系统的容错性能和编码效率进行了优化, 另一方面可有效阻止帧间错误传播, 减小信道失真. 在不同传输信道中的仿真结果验证了该策略在无线视频传输容错方面的性能改进.

关键词: 帧内刷新; 速率-失真率; 容错; Markov 信道模型; 反馈

中图分类号: TN919.81