

# 一种抗包丢失的自适应刷新视频编码

王进 张灿 涂国防

(中国科学院研究生院信息科学与工程学院, 北京 100039)

**摘要** 包丢失是影响 Internet 视频通信质量的一个重要问题, 为此提出了一种新颖的抗包丢失的自适应刷新方法. 该方法是通过估计接收端重建图象的失真, 在发送端自适应地确定需要刷新的宏块的位置. 模拟实验表明, 该方法与已有的顺序刷新和随机刷新<sup>[4]</sup>等方法相比, 不仅峰值信噪比(*PSNR*)有明显提高(0.3~2.5dB), 而且计算复杂度较低, 有利于实时实现.

**关键词** 图象处理(510·4050) 包丢失 自适应刷新 视频通信

**中图分类号:** TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2003)11-1352-04

## Adaptive Intra Update for Packet Loss Resilience

WANG Jin, ZHANG Can, TU Guo-fang

(School of Info. Sci. & Eng., Graduate School of Chinese Academy of Sciences, Beijing 100039)

**Abstract** Because of buffer overflow or delay timeout, packet loss is very common in current internet data transmission. The situation is even worse for real time internet video communications. Due to the prediction property of most state of the art video coding methods, data loss in a single frame can also deteriorate several successive frames. Intra update is well known as an effective method to reduce this error propagation. In this paper, we propose a novel adaptive intra update method. Taking decoder error concealment into account, the distortion at the decoder can be precisely estimated at the encoder by using the Recursive Optimal Per-pixel Estimate (ROPE). Then, image regions with maximum distortion are intra updated. Thus the damaged image regions at the decoder can be recovered in time. Simulation results demonstrate that our method achieves substantial and consistent gains in *PSNR* (0.3~2.5dB) over the non-adaptive methods such as raster-order intra update and random intra update for different video coding rates and packet loss rates. The complexity of this method is also modest, and it can be easily incorporated into current video communication system while maintaining compatibility.

**Keywords** Packet loss, Adaptive intra update, Video communications

## 0 引言

随着 Internet 的飞速发展, 网络视频通信及其服务质量(QoS)控制成为研究的热点. 由于现有的网络通常只提供有限的端到端(end-to-end)的服务质量保证, 因此在传输过程中, 数据包会因为缓存溢出或者等待超时而被丢弃, 甚至当网络拥塞时, Internet 包丢失率可达 20%<sup>[1]</sup>. 由于视频通信系统通常是采用预测编码, 因此当发生包丢失时, 解码错误会在时间和空间上扩散和积累, 以致当包丢失较

严重时, 图象质量会急剧下降, 甚至失去同步, 使通信中断. 针对这个问题, Wang 等提出了大量的增强视频通信系统稳健性的方法<sup>[2][3]</sup>, 如帧内编码刷新就是一种行之有效的方法. 由于帧内编码切断了当前宏块对前帧图象的依赖, 从而能有效阻止错误扩散. 同时, 由于帧内编码刷新(Intra Update)不需要修改比特流的语法结构, 因而与现有标准和系统完全兼容. 不过, 由于过多地采用帧内编码将会因压缩效率的下降而增加通信延时, 因此必须合理选择刷新方案. Cote 等人提出按光栅扫描顺序(raster order)来周期性刷新全帧或者周期性刷新随机分组

基金项目: 国家自然科学基金(60272056); 国家 863 计划项目(2002AA121070)

收稿日期: 2002-04-25; 改回日期: 2003-06-20

块<sup>[4]</sup>,但这类方法没有考虑图象的特点,而将图象各部分统一处理. Judy 等人提出错误敏感度准则<sup>[5]</sup>,并根据图象各部分对错误的敏感度不同来进行刷新,该方法是针对无线信道中的比特错误而提出的;另一种更直接的方法就是在率失真最佳意义下进行模式选择<sup>[4,6~8]</sup>. 由于这类方法综合考虑了解码端错误隐藏、信道特性以及传输协议的影响,因而能通过了解码端失真的估计来获得率失真最佳意义下的编码模式,但这类算法复杂度高,难以实时实现.

本文提出了一种新颖的自适应刷新方法,即在已知解码端错误隐藏方法和网络包丢失率的条件下,通过对解码端失真的准确估计来及时刷新包丢失或因错误扩散而导致的重建质量严重下降的图象区域. 与顺序刷新和随机刷新<sup>[4]</sup>等非自适应方法相比,本方法不仅其平均峰值信噪比 (PSNR) 提高了 0.3~2.5dB,而且计算复杂度较低,有利于实时实现.

## 1 算法描述

根据 H. 263 协议<sup>[9]</sup>,码流结构自上而下分为图象层 (Picture Layer)、块组 (Group of Block) 层 (GOB Layer)、宏块层 (Macroblock Layer) 和块层 (Block Layer) 4 个层次. 每个宏块 (MB) 为  $16 \times 16$  个像素,由 4 个亮度信号块和两个色度信号块组成. 在本实验中,一帧图象中每一行组成一个块组 GOB,每个 GOB 可相互独立解码. 本文采用实时传输协议/用户数据报协议/互联网协议 (RTP/UDP/IP) 协议<sup>[10]</sup>来传输视频流,由于每一个数据包 (Packet) 包含一个 GOB,因此,在这种设置下,就可根据包的丢失率来求像素的均方误差值. 假定包丢失率  $p$  已知,则其可以通过传输协议,如 Real-time Transport Control Protocol (RTCP)<sup>[10]</sup>来实时获得.

解码端是采用一种有效的错误隐藏方法——基于运动补偿的时间预测 (Motion compensated temporal prediction)<sup>[11]</sup>,即丢失的宏块的运动矢量由上一行对应位置的宏块的运动矢量代替,如果上一行宏块也丢失了,则置为零. 这样丢失的宏块数据就由运动矢量指向的前帧宏块数据代替.

### 1.1 解码端失真的估计

Rui 等人提出了 ROPE (Recursive Optimal Per-pixel Estimate) 算法<sup>[7]</sup>,其解码端的失真可通过迭代计算来逐个像素进行精确估计. 由于在没有反馈的条件下,解码端的像素值对于编码端来说是一

个随机变量,所以当采用均方误差 (MSE) 准则时,解码端每个像素的失真可表示如下:

$$d_n^{(i)} = E\{(x_n^{(i)} - \tilde{x}_n^{(i)})^2\} = (x_n^{(i)})^2 - 2x_n^{(i)}E\{\tilde{x}_n^{(i)}\} + E\{(\tilde{x}_n^{(i)})^2\} \quad (1)$$

其中,  $i$  表示像素位置,  $x_n^{(i)}$  表示第  $n$  帧第  $i$  个像素的原始值,  $\tilde{x}_n^{(i)}$  表示对应像素在解码端的重建值,为一随机变量.

$d_n^{(i)}$  的计算关键在于  $\tilde{x}_n^{(i)}$  的一阶矩和二阶矩,分两种情况讨论如下:

(1) 像素位于帧内编码宏块 如果像素  $x_n^{(i)}$  被正确接收,则解码端重建值为  $\tilde{x}_n^{(i)}$  (等于编码端的重建值); 如果当前 GOB 丢失,而前一个 GOB 被正确接收,则解码端用上一个宏块的运动矢量所指向的前帧第  $k$  个像素值  $\tilde{x}_n^{(k)}$  代替当前像素; 如果前一个 GOB 也丢失了,则解码端用前帧对应位置的第  $i$  个像素  $\tilde{x}_{n-1}^{(i)}$  代替当前像素,即

$$E\{\tilde{x}_n^{(i)}\} = (1-p)(x_n^{(i)}) + p(1-p)E\{\tilde{x}_n^{(k)}\} + p^2E\{\tilde{x}_{n-1}^{(i)}\} \quad (2)$$

$$E\{(\tilde{x}_n^{(i)})^2\} = (1-p)(x_n^{(i)})^2 + p(1-p)E\{(\tilde{x}_n^{(k)})^2\} + p^2E\{(\tilde{x}_{n-1}^{(i)})^2\} \quad (3)$$

其中,  $i, j, k$  表示像素位置 (序号),  $p$  为包丢失率.

(2) 像素位于帧间编码宏块 对于帧间编码宏块,实际传输的是量化误差  $\hat{e}_n^{(i)}$  和运动矢量. 如果当前 GOB 被正确接收,则解码端重建像素值为

$$\tilde{x}_n^{(i)} = \hat{e}_n^{(i)} + \tilde{x}_{n-1}^{(j)} \quad (4)$$

其中,  $i, j$  表示像素位置,  $\tilde{x}_{n-1}^{(j)}$  表示正确接收的运动矢量所指向的前帧像素值. 如果当前 GOB 丢失了,则解码端执行与帧内编码宏块相同的错误隐藏方法,即

$$E\{\tilde{x}_n^{(i)}\} = (1-p)(\hat{e}_n^{(i)} + E\{\tilde{x}_{n-1}^{(j)}\}) + p(1-p)E\{\tilde{x}_n^{(k)}\} + p^2E\{\tilde{x}_{n-1}^{(i)}\} \quad (5)$$

$$E\{(\tilde{x}_n^{(i)})^2\} = (1-p)E\{(\hat{e}_n^{(i)} + \tilde{x}_{n-1}^{(j)})^2\} + p(1-p)E\{(\tilde{x}_n^{(k)})^2\} + p^2E\{(\tilde{x}_{n-1}^{(i)})^2\} = (1-p)(\langle \hat{e}_n^{(i)} \rangle^2 + 2e_n^{(i)}E\{\tilde{x}_{n-1}^{(j)}\} + E\{(\tilde{x}_{n-1}^{(j)})^2\}) + p(1-p)E\{(\tilde{x}_n^{(k)})^2\} + p^2E\{(\tilde{x}_{n-1}^{(i)})^2\} \quad (6)$$

### 1.2 刷新策略

Rui 等人将 ROPE 算法<sup>[7]</sup>应用于率失真最佳模式选择 (Rate Distortion Optimized Mode Selection), 并取得了很好的效果, 由于但计算复杂度高, 难以实时实现. 因此本文提出了一种低复杂度的刷新方法, 即在编码第  $n+1$  帧前, 首先估计出第

$n$  帧所有宏块在解码端的失真,同时计算出各宏块的量化失真;然后选择两者之差最大,即由包丢失或错误扩散引起的失真最大的  $N$  个宏块来进行帧内编码刷新, $N$  为给定常数, $M_i, M_j$  为第  $i$  个和第  $j$  个宏块, $q$  表示量化(quantization).令  $\Phi$  为被刷新宏块的集合,则

$$\Phi = \{M_i, i=1, \dots, N \mid \forall M_i \in \Phi, M_j \notin \Phi, (D_{\text{decoder}} - D_q) \text{ of } M_i > (D_{\text{decoder}} - D_q) \text{ of } M_j\} \quad (7)$$

其中,  $D_{\text{decoder}}$  为解码端失真的期望值

$$D_{\text{decoder}} = \sum_{d_n^{(i)} \in M_i} d_n^{(i)} / 256 \quad (8)$$

$D_q$  为量化失真,  $d_n^{(i)}$  为第  $i$  个宏块中元素.

$$D_q = \sum_{x_n^{(i)} \in M_i} (x_n^{(i)} - \hat{x}_n^{(i)})^2 / 256 \quad (9)$$

## 2 实验结果

在 H. 263 TMN5 Codec 的基础上,模拟实现了本文的算法.为了对比本文算法的性能,同时也对顺序刷新和随机刷新算法<sup>[4]</sup>进行了模拟实验.顺序刷新即按照光栅扫描顺序每次刷新  $N$  个宏块,直至刷新全帧;随机刷新即首先将全部宏块分成  $1/p$  组( $p$  为包丢失率),然后每次刷新其中一组,因此每个宏块的刷新频率为  $1/p$ .

为了全面对比本文算法与顺序刷新和随机刷新算法性能,使用了 Miss-America, Grandma, Salesman, Mother and Daughter, Carphone, and Foreman 6 个 QCIF(Quarter Common Intermediate Format)测试图象序列进行实验:其中,miss America 为 150 帧,其他为 300 帧.另外还使用随机包丢失来模拟 Internet 信道,并计算解码端亮度信号在 30 种不同的包丢失模式下的平均峰值信噪比(PSNR).表 1 为比特率为 64bps,帧率 10 帧/s,包丢失率 10%时,3 种算法重建图象的平均 PSNR,表 2 为比特率为 144bps,帧率 10 帧/s,包丢失率 10%时,3 种算法重建图象的平均 PSNR.

表 1 比特率为 64bps 时 3 种算法的 PSNR 对比

序列	本算法	随机刷新	顺序刷新
Miss America	35.93	35.30	35.54
Grandma	34.25	33.52	33.48
Salesman	32.57	30.88	31.05
Mother & Daughter	32.00	30.90	31.00
Carphone	27.86	27.37	27.42
Foreman	25.19	24.40	24.90

表 2 比特率为 144kbps 时 3 种算法的 PSNR 对比

序列	本算法	随机刷新	顺序刷新
Miss-America	38.20	36.66	37.23
Grandma	37.17	35.74	35.71
Salesman	35.84	33.18	33.34
Mother & Daughter	34.07	32.49	32.83
Carphone	29.11	28.43	28.80
Foreman	26.42	25.55	26.15

以上实验中均使用整像素运动补偿,包丢失率为 10%时,每帧刷新 10 个宏块,即  $N=10$ .

从表 1 和表 2 的结果可以看出,本文提出的自适应刷新算法在不同的比特率下,其性能均明显优于顺序刷新和随机刷新算法,其重建图象的平均 PSNR 提高了 0.3~2.5dB.

从模拟实验还可以发现,图象的活动性对于算法性能影响较大,即对于背景静止、人物活动性较弱的图象序列,如 Miss-America、Grandma、Salesman 及 Mother & Daughter 序列,应用本文算法重建图象的峰值信噪比(PSNR)提高较大;而对于背景运动、人物活动性较强的图象序列,如 Carphone 和 Foreman 序列,重建图象信噪比则提高较小.由此可见,适当提高刷新频率将有助于提高这类图象的重建质量.

为了考察本文算法在不同的包丢失率下的性能,又模拟了包丢失率为 3%、5%和 20%的情况,比特率为 144bps,帧率为 10 帧/s,重建图象的 PSNR 如图 1 和图 2 所示.

由图 1 和图 2 可以看出,本文提出的自适应算

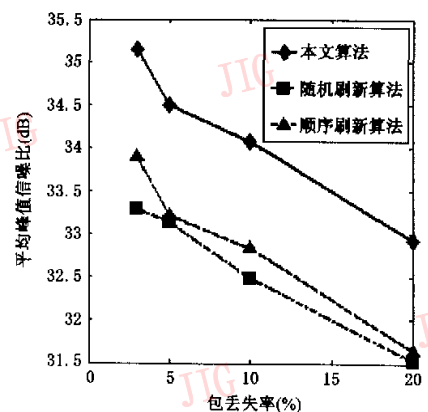


图 1 Mother & Daughter 序列 3 种算法在不同包丢失率下重建图象的 PSNR 对比

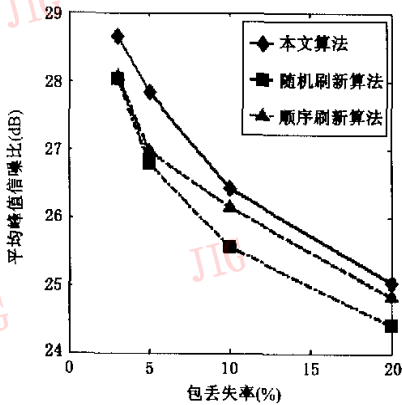


图 2 Foreman 序列 3 种算法在不同包丢失率下重建图象的 PSNR 对比

法在不同的包丢失率下,其性能均优于随机刷新和顺序刷新算法。

### 3 结 论

本文提出了一种新颖的抗包丢失的自适应刷新算法。模拟实验结果表明,在不同的比特率和包丢失率下,该算法性能均明显优于现有非自适应算法。

此外,在当前算法中,刷新率取值为  $1/p$  ( $p$  为包丢失率),从实验结果来看,这并非对所有序列都是最佳选择。根据图象的活动性和包丢失率,如何自适应地改变刷新率将是下一步的研究课题。

#### 参 考 文 献

- 1 Zhu Q F, Kerofsky L. Joint source coding, transport processing and error concealment for H. 232-based packet video[A]. In: Proceedings of SPIE Symposium on Visual Communications and Image Processing[C], San Jose, CA, USA, 1999, 3653:52~62.
- 2 Wang Y, Zhu Q F. Error control and concealment for video communication: A review[J]. Proceedings of IEEE, 1998, 86(5):974~997.
- 3 Wenger S, Knorr G, Ott J *et al.* Error resilience support in H. 263+[J]. IEEE Transactions of Circuits and Systems on Video Technology, 1998, 8(11):867~877.
- 4 Cote G, Kossentini F. Optimal intra coding of blocks for robust video communication over the internet [J]. The European Association for Signal Speech and Image Processing(EURASIP) Image Communications, Special issue on Real-Time Video over the Internet, 1999, 9(15):25~34.
- 5 Liao J Y, Villasenor J. Adaptive Intra Block Update for Robust

- Transmission of H. 263[J]. IEEE Transaction of Circuits and Systems on Video Technology, 2000, 10(2):30~35.
- 6 Ortega A, Ramchandran K. Rate-distortion methods for image and video compression[J]. IEEE Signal Processing Magazine, 1998, 15(11):23~50.
- 7 Zhang R, Regunathan S L, Rose K. Video coding with optimal Inter/Intra-mode switching for packet loss resilience[J]. IEEE Journal on Selected Area of Communications, 2000, 18(6):966~976.
- 8 Wu D P, Hou Y T, Li B *et al.* An end-to-end approach for optimal mode selection in internet video communication: theory and application [J]. IEEE Journal on Selected Area of Communications, 2000, 18(6):977~995.
- 9 ITU-T Standardization Sector. Recommendation H. 263, Video coding for low bit rate communication [S], International Telecommunications Union May 1996.
- 10 Schulzrinne H, Casner S, Frederick R *et al.* RTP: A Transport Protocol for Real-Time Applications [EB/OL]. Internet Engineering Task Force <http://www.ietf.org/rfc/rfc1889.txt>, 1996-01.
- 11 Haskell P, Messerschmitt D. Resynchronization of motion compensated video affected by ATM cell loss [A]. In: IEEE Proceedings of International Conference on Acoustics, Speeches and Signal Processing[C], San Francisco, CA, USA, 1992, 3: 545~548.



王 进 1977 年生,现为中国科学院研究生院信号与信息处理专业硕士研究生,主要研究兴趣为视频编码与传输。



张 焯 1954 年生,1978 年毕业于中国科学技术大学无线电系,现为中国科学院研究生院信息科学与工程学院研究员。主持完成科研项目 20 多项,获部级二、三等奖五项,近年来以数字通信、信息与处理为主要研究方向。



涂国防 1954 年生,1985 年获中国科学技术大学无线电系硕士学位,现为中国科学院研究生院信息科学与工程学院常务副院长,教授,博士生导师。主持完成国家自然科学基金,国家“863”计划等科研项目多项,在国内外发表论文四十多篇,被 SCI、EI 等期刊多次收录;研究方向为数字通信、图象编码与处理。