

一种低复杂度的多参考帧选择率失真优化方法

陈云鹏 张培仁 孙轶 高修峰

(中国科学技术大学自动化系 嵌入式系统与控制网络实验室 合肥 230027)

摘要: 视频在无线易错信道上传输, 报文丢失会导致图像质量显著下降。多参考帧选择方法可以降低误码扩散, 从而提高误码信道上传输视频的图像质量。但是参考帧选择方法也使得编码器计算复杂度大大增加。基于多参考帧选择技术提出一种低复杂度的率失真优化方法。该方法首先在编码器上重建误码图像, 使得编码器和解码器上的参考图像一致, 然后利用端到端率失真优化模型, 在重建的误码图像与编码图像的上一帧图像中选择最优的参考帧。在进行率失真计算时, 采用线性预测方法计算运动补偿后图像差值的方差, 减少了运动估计运算, 从而大大降低计算复杂度。实验结果表明相对于帧内刷新方法该方法显著提高图像质量, 相对于一般多参考帧选择方法, 在损失很小性能的情况下, 大大降低了计算复杂度。

关键词: 视频传输; 多参考帧选择; 视频容错; 率失真分析

中图分类号: TN919.81

文献标识码: A

文章编号: 1009-5896(2007)04-0967-04

One Low Complex Reference Picture Selection Optimal Method

Chen Yun-peng Zhang Pei-ren Sun Yi Gao Xiu-feng

(Embedded System & Control Networks Lab, Dept. of Automation, USTC, Hefei 230027, China)

Abstract: Packet loss led to great degradation of picture quality, when video stream transmitted in error-prone channel. RPS (Reference picture selection) could improve picture quality by selecting suitable reference picture, but RPS greatly increased computation complexity. In this paper one low complex RPS optimal method was proposed. First error concealment steps in decoder was tracked in encoder to reconstruct error picture; second only error reconstructed picture and the preceding picture were taken as optional reference pictures; third picture difference variance was computed by linear prediction when processing R-D optimization. Through above means time of motion estimation was reduced, and then computation complexity was greatly decreased. Experiment result shows the method greatly lower encoder's complexity, while only loss small picture quality.

Key words: Video transmission; Multiple reference picture selection; Error resilience; Rate distortion

1 引言

随着无线网络的发展, 无线视频业务受到广泛的关注。如何在高误码率的无线信道上传输视频数据, 并保证视频图像质量成为一项重要课题。降低信道误码对于图像质量的影响, 重要的一点是降低误码图像的失真扩散。文献[1-3]提出自适应帧内刷新方法, 通过在编码过程中针对信道误码状况有选择性的插入帧内编码, 从而降低误码失真扩散。文献[4,5]提出多参考帧选择方法, 通过选择合适的参考图像, 以中止图像误码扩散。文献[6]进一步提出一种率失真优化方法用于参考帧选择的判别。但是, 由于多参考帧选择方法需要对多个参考帧进行运动估计, 导致其在提高图像质量的同时, 也极大增加了编码器的计算复杂度。针对这一点, 本文提出一种低复杂度的多参考帧选择率失真优化方法。该方法通过减少运动估计次数以降低计算复杂度, 具体包括这样几个方面: (1)建立一种基于多参考帧选择的信道失真模型; (2)在

编码器上跟踪解码器的错误隐藏步骤, 并在编码器端重建误码图像, 仅选择误码重建图像与当前编码图像的上一帧图像作为可选的参考帧; (3)在进行率失真判别时, 采用线性预测的方法估计运动补偿后图像差值的方差。实验结果表明, 相对于文献[1]方法, 本文提出的方法提高了视频图像质量, 相对于文献[6]方法, 本文方法在损失很小性能的情况下, 极大降低了编码器的计算复杂度。

2 多参考帧选择率失真优化方法

视频图像端到端失真由两部分组成, 包括信源编码失真和信道误码失真。信源失真主要是编码量化失真, 由视频编码量化引入, 信道失真由信道误码或报文丢失引入。建立联合信源信道率失真优化模型, 需要分别计算信源和信道失真, 并在一定约束条件下, 最小化端到端图像失真。本文首先提出多参考帧模式的信道失真模型, 及信源编码失真的估计方法, 然后在码率约束条件下, 推导出多参考帧模式的率失真优化表达式。

2.1 多参考帧的信道失真分析

本文的讨论针对带反馈信道的视频通信系统, 编码器可通过反馈信道的反馈信息得到解码器端报文丢失情况。文献[7]中详细分析了如何利用反馈信息分析跟踪视频图像误码扩散, 并根据误码扩散的状况插入帧内编码, 以此来中止图像的误码扩散。本文参照文献[7]的方法, 在编码器上跟踪解码器上的误码隐藏过程。与文献[7]不同的是, 本文没有采用插入帧内编码方法, 而是根据解码器的误码隐藏过程在编码器端重建误码图像。本文实验中采用简单的错误隐藏模式, 仅从参考图像中复制丢失报文涉及的宏块。此外, 由于反馈信息的时延迟以及发送缓冲区的存在, 重建的误码图像通常不是当前编码帧的前一帧图像。如图 1, 设当前编码帧为 n , 编码器最近的误码重建图像为 m , 一般情况下 m 小于 $n-1$ 。

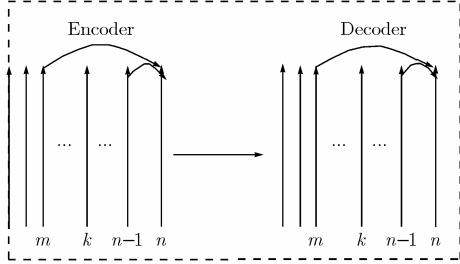


图 1 误码重建及多参考帧选择示意图

为建立率失真模型, 需分析信道误码对于参考帧选择的影响。文献[1,3]针对单参考帧系统(即仅采用 $n-1$ 作为参考帧)提出信道失真预测模型, 可根据信道状况精确预测信道失真, 本文把文献[1]模型进一步扩展到多参考帧的情形。文献[1,3]的研究表明, 信道误码失真与视频编码模式相关, 需分为帧内模式(INTRA Mode)和帧间模式(INTER Mode)讨论。当宏块编码采用帧内模式时, 由报文丢失引起的信道失真可表示为:

$$\begin{aligned} D_c^I(n) &= E\{[F(n, i) - \bar{F}(n, i)]^2\} \\ &= pE\{[F(n, i) - F(n-1, i)]^2\} + pE\{[F(n-1, i) \\ &\quad - \bar{F}(n-1, i)]^2\} \\ &= apF_d(n, n-1) + pD_c(n-1) \end{aligned} \quad (1)$$

$$F_d(n, n-1) = E\{[F(n, i) - F(n-1, i)]^2\} \quad (2)$$

式(1)中 $F(n, i)$ 表示编码器重建图像像素 i , $\bar{F}(n, i)$ 表示解码器重建图像像素 i , 若 i 所在的宏块丢失则 $F(n, i) \neq \bar{F}(n, i)$ 。 p 为信道报文丢失率, a 同图像序列相关, 本文 a 由上一帧编码数据计算得到。若宏块编码模式为帧间模式时, 信道失真 $D_c^P(n)$ 可以表示为:

$$\begin{aligned} D_c^P(n) &= E\{[F(n, i) - \bar{F}(n, i)]^2\} \\ &= (1-p)bD_c(n-1) + paF_d(n, n-1) + pD_c(n-1) \end{aligned} \quad (3)$$

式(3)中 b 为常数, 与图像序列运动剧烈程度相关。联合式(1), 式(3), 设帧内编码比率 β , 得到总信道失真:

$$\begin{aligned} D_c(n) &= \beta D_c^I(n) + (1-\beta)D_c^P(n) \\ &= [(1-\beta)(1-p)b + p]D_c(n-1) + paF_d(n, n-1) \end{aligned} \quad (4)$$

需要说明的是, 上述失真模型是在不考虑帧内预测编码的情况下得出的, 适合于 H263, MPEG4 等编码标准, 对于 H264 编码标准, 由于帧内模式也进行了预测编码, 需要对式(1), 式(4)作部分修正, 详细处理请参考文献[2]。

式(1)–式(4)表明, 当前编码帧的信道失真与参考帧失真相关, 文献[1]实际是使用参考帧的信道失真结合编码模式选择的情况估计当前编码帧可能的信道失真。这里考虑多参考帧情况, 当选定的参考帧为误码重建帧时, 由于编码器的参考图像与解码器一致, 则此时参考帧的信道失真为 0, 当选定参考帧为非误码重建帧, 则信道失真为估计值 $D_c(j)$ 。设选定 j 为参考帧, 则修改式(1)–式(4)表达式:

$$D_c^I(n) = E\{[F(n, i) - \bar{F}(j, i)]^2\} = apF_d(n, j) + pD_c(j) \quad (5)$$

$$F_d(n, j) = E\{[F(n, i) - F(j, i)]^2\} \quad (6)$$

$$D_c^P(n) = (1-p)bD_c(j) + paF_d(n, j) + pD_c(j) \quad (7)$$

$$\begin{aligned} D_c(n) &= \beta D_c^I(n) + (1-\beta)D_c^P(n) \\ &= [(1-\beta)(1-p)b + p]D_c(j) + paF_d(n, j) \end{aligned} \quad (8)$$

$$D_c(j) = \begin{cases} 0, & j \leq m \\ D_c(j), & m+1 \leq j \leq n-1 \end{cases} \quad (9)$$

2.2 多参考帧的信源失真估计

建立信源编码失真模型主要是建立信源码率、失真及量化系数之间的关系。文献[8,9]是被普遍采用的两种视频编码率失真模型, 文献[8]被 H263 参考设计所采纳, 而文献[9]出现在 MPEG4 和 H264 中, 本文讨论针对文献[8]失真模型。对第 n 帧图像, 给定码率 R_n , 图像编码失真 $D_{n,s}$, 及 i 宏块分配比特数 $B_{n,i}$:

$$D_{n,s} = \frac{1}{12N} \sum_{i=1}^N \alpha_{n,i}^2 Q_{n,i}^2 \quad (10)$$

$$B_{n,i} = A \left[K \frac{\sigma_{n,i}^2}{Q_{n,i}^2} + C \right], \quad \sum_{i=1}^N B_{n,i} \leq R_n \quad (11)$$

其中 $\alpha_{n,i}$ 为 i 宏块的失真权值(distortion weight), A 等于宏块像素数目, K 是一个与图像像素分布相关的编码参数, C 为除去 DCT 系数外其余数据的编码开销, $\sigma_{n,i}^2$ 为宏块补偿后差值方差, $Q_{n,i}$ 为宏块量化参数。

文献[8]模型以宏块的形式给出, 为便于计算分析, 简化上述模型, 令各个宏块失真权值与量化参数相等 $\alpha_{n,i} = 1$, $Q_{n,i} = Q_n$, 且 $\sigma_n^2 = \sum_{i=1}^N \sigma_{n,i}^2$, 则式(10), 式(11)化为式(12)。本文的实验表明, 利用式(12)式得到 Q_n 约等于式(10), 式(11)得到 $Q_{n,i}$ 的平均值即 $Q_n \approx \left(\sum_{i=1}^N Q_{n,i} \right) / N$:

$$D_{n,s} = \frac{Q_n^2}{12}, \quad B_n = A \left[K \frac{\sigma_n^2}{Q_n^2} + C \right] \leq R_n \quad (12)$$

从式(12)可以看出, 为得到最优量化参数 Q_n , 需要先计算 σ_n^2 , 在 H263 参考设计中, σ_n^2 需运动估计完成后才能得到。考虑多参考帧的情形, 由文献[6]方法, 需针对多个不同参考帧进行多次运动估计, 以得到针对不同参考帧 j 的 $\sigma_n^2(j)$,

从而计算出 $D_{n,s}(j)$ 。多次运动估计将大大增加编码器的计算复杂度。由此,本文仅选择误码重建帧与当前编码帧的上-帧作为可供选择的参考帧,并采用线性预测的方法估计 $\sigma_n^2(j)$,从而消除由多次运动估计引入的计算量,保持多参考帧编码器的计算量与单参考帧基本一致。如图1,可供选择的参考帧 $m, n-1$,则 $\sigma_n^2(j)$ 的预测表达式:

$$\sigma_n^2(j) = \begin{cases} a_1\sigma_m^2 + a_2, & j = m \\ b_1\sigma_{n-1}^2 + b_2, & j = n-1 \end{cases} \quad (13)$$

式(13)中 a_1, a_2, b_1, b_2 为待定系数,根据已有的 σ_i^2 统计确定。由于最后只能选择一帧参考帧, a_1, a_2, b_1, b_2 系数只能更新其中两个,其余两个系数需通过已更新的系数估计。本文令 $a_1 = b_1 + \alpha(n-1-m)$, $a_2 = b_2 + \alpha(n-1-m)$, α 为常数与图像系列运动剧烈程度相关,本文实验取 0.1。

2.3 联合信道信源多参考帧选择方法

视频编码传输端到端率失真优化需联合信道信源失真,文献[1]研究表明,端到端失真可以表示为信道信源失真之和 $D = D_c + D_s$ 。本文建立多参考帧率失真优化模型,在满足码率约束条件下,通过参考帧选择最小化端到端失真 D ,模型可表示为:

$$\text{Min: } D = D_c + D_s, \text{ s.t. } B_n \leq R_n \quad (14)$$

代入式(8),式(9),式(12),式(13),式(14)可以化为:

$$\text{Min: } D = \begin{cases} paF_d(n, m) + Q_n^2/12, & j = m \\ [(1-\beta)(1-p)b + p]D_c(n-1) \\ \quad + paF_d(n, n-1) + Q_n^2/12, & j = n-1 \end{cases} \quad (15)$$

$$\text{s.t. } A \left[K \frac{\sigma_n^2(j)}{Q_n^2} + C \right] \leq R_n, \quad j = m, n-1 \quad (16)$$

式(15),式(16)代入 Lagrange 乘子,并求偏导,令其等于 0,可以分别求得 $j = m, n-1$ 的最小端到端失真 $D_{\min}(j)$:

$$D_{\min}(j) = \begin{cases} \frac{AK\sigma_n^2(m)}{12(R-ANC)} + paF_d(n, m), & j = m \\ \frac{AK\sigma_n^2(n-1)}{12(R-ANC)} + [(1-\beta)(1-p)b + p] \\ \quad \cdot D_c(n-1) + paF_d(n, n-1), & j = n-1 \end{cases} \quad (17)$$

根据 $D_{\min}(j)$ 的大小选择 m 或 $n-1$ 作为参考帧。此外,为进一步降低计算复杂度,本文考虑发生场景变换(表明图像有比较剧烈的运动)的情况,如图1,若 k 发生场景变换,则跳过式(17)计算,优先选择 $n-1$ 作为参考帧。若选择的参考帧记为 F :

$$F = \begin{cases} n-1, & k \text{ 发生场景变换且 } m < k \leq n-1 \\ n-1, & D_{\min}(n-1) \leq D_{\min}(m) \\ m, & \text{其他} \end{cases} \quad (18)$$

2.4 实验结果及分析

由于本文的信源失真模型采用文献[8],所以实验基于

H263+视频编码器,根据2.3节的讨论修改 TMN10 参考实现。采用 QCIF 格式的 foreman 图像测试序列,编码 300 帧,编码模式 IPPP,设定帧率 25fps,解码器启用错误隐藏选项。实验采用峰值信噪比 PSNR(Peak Signal to Noise Ratio)衡量图像质量:

$$\text{PSNR} = 10 \lg(255 * 255 / \text{RMSE}) \quad (19)$$

给定 $N \times M$ 像素的第 n 帧原始图像 $f_n(x, y)$ 及解码后的图像 $f'_n(x, y)$,跳帧图像使用上一解码图像代替,即令 $f'_n(x, y) = f'_{n-1}(x, y)$ 。RMSE 为原始图像与误码传输后的解码图像的均方根差使用式(20)计算。实验中,需要多次编解码,并取 300 帧图像的 PSNR 均值。本文采用报文丢失率作为信道误码参数,报文按照 H263 的 GOB 封装,一个 GOB 封装为一个报文。解码器反馈 ACK 和 NACK 表示报文丢失或正确接收,不考虑反馈报文丢失情况。

$$\text{RMSE} = \sqrt{\frac{1}{N \times M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} [f_n(x, y) - f'_n(x, y)]^2} \quad (20)$$

实验结果与文献[1]比较,文献[1]采用帧内刷新方法抗信道误码,但由于文献[1]的信源失真模型与本文不同,本文根据文献[1]的思想,在 TMN10 上实现文献[1]算法,并比较两种方法对图像质量提高的效果。本文提出的低复杂度多参考帧选择方法是文献[6]的改进算法,实验中实现文献[6]方法,并与文献[6]做两方面比较:图像质量 PSNR;计算复杂度。本文使用在一定计算环境中,编码 300 帧图像消耗的时间衡量编码器的计算复杂度。此外,实验还给出了反馈信息延迟对于多参考帧方法有效性的影响,以反馈信息延迟帧数 $d_f = n - m$ 来表征反馈信息延迟。

实验结果如图2至图5,本文用 AIR 表示文献[1]方法, CORPS 表示文献[6]方法,本文提出的方法用 SORPS 表示。图2和图3给出反馈信息延迟 3 帧,3 种方法的图像质量曲线。其中图2给出了报文丢失率 15%,图像质量随着码率变化的曲线。从图2可以看出,在误码率 15% 的环境中, SORPS 对 AIR 有大约 0.4dB 的提高,而对于 CORPS 仅有大约 0.1dB 的性能损失。从图3中可以得到类似的结论。图4和图5给出反馈信息延迟 2 帧时的情况。当反馈信息延迟 2 帧时, SORPS 比 AIR 提高大约 1dB,比 CORPS 性能损失小于 0.1dB。此外,从图中还可以得出,多参考帧选择方法较易受反馈信息延迟影响,降低反馈信息延迟,能够有效提高图像质量,从图2-图5可以看出,反馈信息延迟 2 帧比延迟 3 帧,图像质量提高大约 0.5dB。

对于计算复杂度,实验统计所有 P 帧编码花费的时间,在 TMN10 中等效于计算函数 CodeOneOrTwo() 执行所花费时间总和,实验环境为处理器 AMD Athlon(tm) XP 2500+,操作系统 Linux。当 $d_f = 3$ 时, CORPS 花费时间为 61160 ms,本文 SORPS 花费为 22720ms,由此得到, SORPS 花费时间比 CORPS 大大减少,计算复杂度显著降低,而性能仅损失约 0.1dB。

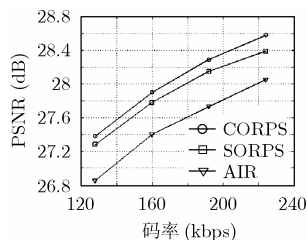


图2 $d_j=3, p=15\%$ 不同码率三种方法图像质量比较

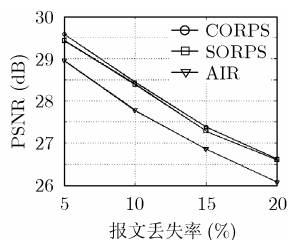


图3 $d_j=3, 128\text{kbps}$ 不同报文丢失率三种方法图像比较

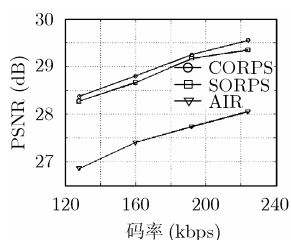


图4 $d_j=2, p=15\%$ 不同码率三种方法图像质量比较

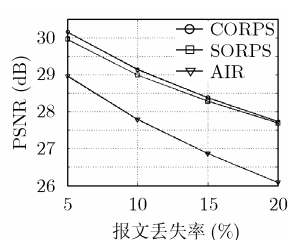


图5 $d_j=2, 128\text{kbps}$ 不同报文丢失率三种方法图像比较

3 结束语

本文针对多参考帧选择优化方法复杂度大的缺点,提出一种低复杂度的多参考帧选择率失真优化方法,目的是以很小的性能损失,换取计算复杂度极大降低。本文通过减少参考帧数目,及在参考帧选择过程中估计图像差值方差的方法,减少多参考帧选择中由于多次运动估计所引入的计算量,从而大大降低了计算复杂度。实验结果表明,相对于一般多参考帧选择率失真优化方法,本文提出的方法以很少的性能损失显著降低了编码器的计算复杂度。

参考文献

- [1] He Z. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2002, 12(6): 511-523.
- [2] 陈川, 余松煜. 联合编码模式选择的码率控制算法. 电子学

报, 2004, 32(5): 763-768.

Chen Chuan and Yu Song-yu. Joint coding mode selection and rate control algorithm. *Acta Electronica Sinica*, 2004, 32(5): 763-768.

- [3] Zhang R. Video coding with optimal inter/intra-mode switching for packet loss resilience. *IEEE Journal on Selected Areas in Communications*, 2000, 18(6): 966-976.
- [4] Girod B and Arber N. Feedback-based error control for mobile video transmission. *Proc. IEEE*, 1999, 87(10): 1707-1723.
- [5] Lin S, Mao S, and Wang Y. A reference picture selection scheme for video transmission over Ad-hoc networks using multiple paths. In Proc. of the IEEE International Conference on Multimedia and Expo (ICME), Tokyo, Japan, 2001: 96-99.
- [6] Liang Y, Flierl M, and Girod B. Low-latency video transmission over lossy packet networks using rate-distortion optimized reference picture selection. Proc. IEEE International Conference on Image Processing, Rochester, NY, 2002, 2: 18-84.
- [7] Chang P and Lee T. Precise and fast error tracking for error-resilient transmission of H.263 video. *IEEE Trans. on Circuits and Systems for Video Technology*, 2000, 10(4): 600-607.
- [8] Jordi R. Rate control in DCT video coding for low-delay communications. *IEEE Trans. on Circuits and Systems for Video Technology*, 1999, 9(1): 172-185.
- [9] Chiang T and Zhang Y. A new rate control scheme using quadratic rate distortion model. *IEEE Trans. on Circuits Syst. Video Technology*, 1997, 7(1): 246-250.

陈云鹏: 男, 1978年生, 博士生, 研究方向为视频编码、无线视频传输。

张培仁: 男, 1944年生, 教授, 主要研究方向为嵌入式系统、嵌入式网络。