

# 基于率失真优化的快速帧内更新算法

彭强, 邓云, 诸昌铃

(西南交通大学计算机与通信工程学院 成都 610031)

**【摘要】**提出了一种近似估计整帧图像最小量化失真度的模型和新的率失真优化的实时帧内更新准则,可使编码器直接计算出编码一帧图像的最小量化失真度,通过评估宏块帧内编码所增加的总量化失真度和减小的扩散失真度,以决定是否需要对此宏块进行帧内更新。经与传统的基于率失真的帧内更新相比和实验结果表明,更新算法在具有更小计算复杂度的同时,在大多数情况下可以取得更好的峰值信噪比值。

**关键词** 视频编码; 差错恢复; 帧内更新; 率失真优化

**中图分类号** TP391 **文献标识码** A

## A Real Time Intra-Update Method Based on Rate-Distortion Optimization

Peng Qiang, Deng Yun, Zhu Changqian

(School of Computer and Communications Engineering, Southwest Jiaotong University Chengdu 610031)

**Abstract** A rate-distortion computation model is established for approximately estimating the total quantization distortion of one frame, and a new real time Intra-update criterion is proposed. Through comparing the increased quantization distortion and reduced propagation distortion by intra-updating, the encoder can quickly decide which macroblock need to be intra-updated. Simulation results show that the performance of proposed algorithm, which reduces the computational complexity significantly, achieves the same or even better gains over traditional rate-distortion based Intra-update methods.

**Key words** video coding; error resilience; intra-update; rate-distortion optimization

差错恢复视频编码是视频传输的关键技术之一<sup>[1,2]</sup>,其中,对受错误影响区域进行帧内编码更新是一种对付视频传输差错的有效手段<sup>[3-5]</sup>。然而,这种将某些宏块(Macroblock, MB)编码模式强制改为帧内编码的方法会导致数据量的大量增加。因此,如何选择帧内更新MB的数量和位置,在数据量与失真度之间进行均衡显得十分重要。基于率失真(Rate-Distortion, R-D)优化框架的帧内更新算法是一种较好的解决办法,其典型算法通常将问题简化为在单个MB上求解,并描述为非限制的拉格朗日最小化函数<sup>[6]</sup>。文献[7,8]提出了几种对应的失真度估计算法,而对拉格朗日算子 $\lambda$ 的计算主要分为两种:一是在信源端通过对数据量和失真度进行数学推导,得出经验公式(称为信源算法)<sup>[6]</sup>;二是通过对缓冲区状态的递归计算,得到一个通过缓冲区状态来设置 $\lambda$ 值的算法(称为缓冲区算法)<sup>[9]</sup>。但这些算法都需要分别进行实际的帧内和帧间模式的编码和解码操作,得出各自的数据量与失真度,再通过拉格朗日方程决定使用哪种编码模式,故其计算复杂度极大,不能满足实时视频传输的要求。本文提出了一种新的R-D优化的帧内更新算法,用H.263+测试模型TMN8(Test Model Near-term 8.0)中的数据量与失真度模型来近似计算出编码一帧总的量化失真度,并估算出差错从前一帧扩散到当前帧的失真度,再通过计算对MB更新所增加的量化失真度与减小的扩散失真度,编

收稿日期: 2003-11-10

基金项目: 铁道部科技研究开发项目(2003X040-A)

作者简介: 彭强(1962-),男,在职博士生,副教授,主要从事视频压缩、传输和多媒体应用等方面的研究。

码器可以很快确定是否需对此MB进行帧内更新。

## 1 失真度模型

### 1.1 量化失真

文献[10]提出了低速率视频编码算法中图像的数据量与失真度关系的数学模型,并推导出在总数据量 $B$ 受限的约束条件下,使失真度最小的优化量化参数为

$$\begin{cases} D = \frac{1}{N} \sum_{i=1}^N a_i^2 \frac{Q_i^2}{12} \\ Q_i^* = \sqrt{\frac{AK}{(B-ANC)} \frac{s_i}{a_i} \sum_{k=1}^N a_k s_k} \end{cases} \quad (1)$$

式中  $D$ 为编码器上像素原始值与重建值的均方差; $N$ 为一帧中的MB个数; $A$ 为一个MB所包含的像素个数; $K$ 、 $C$ 为常数; $Q_i$ 为MB使用的量化步长; $s_i$ 为MB亮度和色度系数的标准差; $s_i$ 为MB的失真度权重。

式(1)假设了传输的信道无错,失真仅来源于量化过程,称为量化失真,记为 $D_Q$ 。基于式(1),可以推导出在总数据量 $B$ 受限的情况下,一帧总的量化失真度为

$$D_Q = \sum_{i=1}^N A a_i^2 \frac{Q_i^2}{12} = \frac{AAK}{12(B-ANC)} \left( \sum_{i=1}^N a_i s_i \right)^2 \quad (2)$$

### 1.2 扩散失真

当信道有错误时,对使用帧间预测编码的MB,如果在运动补偿过程中其预测区域中存在差错,即使解码器正确收到MB,其预测帧的错误仍然会扩散到当前帧。

用 $f_n^i$ 表示第 $n$ 帧中的第 $i$ 个像素点的原始值, $\hat{f}_n^i$ 表示其在编码器上的重建值, $\tilde{f}_n^i$ 表示其在解码器上的重建值,此重建值可能是经过差错掩盖之后的值,对于编码器,此值是一个随机变量。以 $\hat{f}_{n-1}^j$ 表示 $f_n^i$ 在进行帧间预测时所使用的第 $n-1$ 帧中的第 $j$ 个像素点的编码器重建值, $\tilde{f}_{n-1}^j$ 表示其对应的解码器重建值。

如果 $f_n^i$ 被解码器正确收到,则此像素点的失真度为

$$\begin{aligned} d_n^i &= |f_n^i - \tilde{f}_n^i|^2 = [(f_n^i - \hat{f}_n^i) + (\hat{f}_n^i - \tilde{f}_n^i)]^2 \approx (f_n^i - \hat{f}_n^i)^2 + (\hat{f}_n^i - \tilde{f}_n^i)^2 = \\ &= (f_n^i - \hat{f}_n^i)^2 + (\hat{f}_{n-1}^j - \tilde{f}_{n-1}^j)^2 = d_Q + d_P \end{aligned} \quad (3)$$

式中  $d_Q$ 表示一个像素点的量化失真度, $d_P$ 表示由其预测点扩散到当前像素点的失真,称为扩散失真。如果采用帧间预测编码的第 $k$ 个MB被正确收到,则其总的失真度为

$$D^k = \sum_{i \in MB(k)} (d_Q^i + d_P^i) = D_Q^k + D_P^k \quad (4)$$

式中  $D_P^k = A a_k^2 \frac{Q_k^2}{12}$ 为第 $k$ 个MB的扩散失真度。在具体的实现过程中,使用ROPE算法来计算扩散失真度 $D_P^k$ [7]。

### 1.3 差错掩盖失真

如上所述,如果第 $k$ 个MB在传输过程中产生了错误,解码器将采用某种特定的算法进行差错掩盖,以减小错误造成的影响。在此情况下,对掩盖过程产生的失真称为掩盖失真,记为 $D_C^k$ 。

### 1.4 总体失真

用 $p$ 表示网络丢包率,则第 $k$ 个MB的总失真度模型为

$$D^k = (1-p)(D_Q^k + D_P^k) + pD_C^k \quad (5)$$

## 2 帧内更新准则

### 2.1 模式决定准则

假设其他MB的编码模式已经确定,下面考虑对第 $k$ 个MB进行帧内更新。用 $D_i$ 表示第 $k$ 个MB使用帧内模式时一帧的总失真度,  $D_p$ 表示第 $k$ 个MB使用帧间模式时一帧的总失真度, 故有

$$D_p = (1-p)D_Q + (1-p)\sum_{i=1}^N D_p^i + p\sum_{i=1}^N D_C^i \quad (6)$$

$$D_i = (1-p)D'_Q + (1-p)\sum_{\substack{i=1 \\ i \neq k}}^N D_p^i + p\sum_{i=1}^N D_C^i \quad (7)$$

式中  $D_Q$  表示当第 $k$ 个MB使用帧间模式时一帧的量化失真;  $D'_Q$  表示第 $k$ 个MB进行帧内更新时一帧的量化失真。由于更新一个MB后一帧的总失真度应小于不更新时的失真度, 即

$$\Delta D = D_p - D_i > 0 \Rightarrow D'_Q - D_Q < D_p^k \quad (8)$$

将式(2)代入式(8), 再对扩散失真度乘以对应的权重并化简后可得到对第 $k$ 个MB进行帧内更新的条件为

$$\frac{AAK}{12(B-ANC)}(2S + \Delta \mathbf{a}_k \mathbf{s}_k) \Delta \mathbf{a}_k \mathbf{s}_k < \mathbf{a}_k D_p^k \quad (9)$$

式中  $S = \sum_{i=1}^N \mathbf{a}_i \mathbf{s}_i$ ,  $\Delta \mathbf{a}_i \mathbf{s}_i = \mathbf{a}'_i \mathbf{s}'_i - \mathbf{a}_i \mathbf{s}_i$ ,  $\mathbf{s}'_i$  和  $\mathbf{a}'_i$  为MB使用帧内模式时系数的标准差和对应的权重。式(9)已经约去了量化失真度 $D_Q$ 以及掩盖失真度 $D_C$ , 实际只需要计算扩散失真度 $D_p$ 。

### 2.2 计算复杂度

传统R-D优化的帧内更新算法需要分别进行帧内和帧间模式的编码解码过程, 计算出各自的数据量与失真度, 再算出不同模式对应的R-D值, 找出其中具有最小R-D值的编码模式, 而本文提出的帧内更新算法并不进行实际的编码和解码过程, 只需要通过少量的加法/乘法运算计算出更新一个MB时, 增加的量化失真度与减小的扩散失真度, 通过比较其大小, 快速决定是否需要对MB进行帧内模式更新, 而且TMN8中的流控算法过程本身已经预先对式(9)中的 $S$ 、 $\mathbf{a}_i$ 和 $\mathbf{s}_i$ 进行了计算, 因此相比传统算法, 本算法大大地降低了计算复杂度。

## 3 模拟实验

通过适当修改H.263的编码解码器TMN8实现了本文提出的帧内更新算法。使用实时传输协议(RTP)对H.263视频流进行打包, 每个RTP包只包含一个GOB(Group of Block)。差错掩盖算法使用简单而有效的时域掩盖算法, 对错误MB的运动向量按如下规则计算: 如果当前GOB丢失, 而当前GOB的上一GOB被正确收到, 则以当前MB的上行的三个相邻MB的运动向量的中值为当前错误MB的运动向量; 如果上一GOB也丢失, 则置当前MB的运动向量为0。对传统的信源算法, 令  $I_{\text{mode}} = 0.85 \times (QP)^2$ , 对缓冲区算法, 令  $I_{n+1} = I_n [1 + \mathbf{a}(R_n - R_{\text{target}})]$ , 其中  $\mathbf{a} = 1/5 R_{\text{target}}$ 。

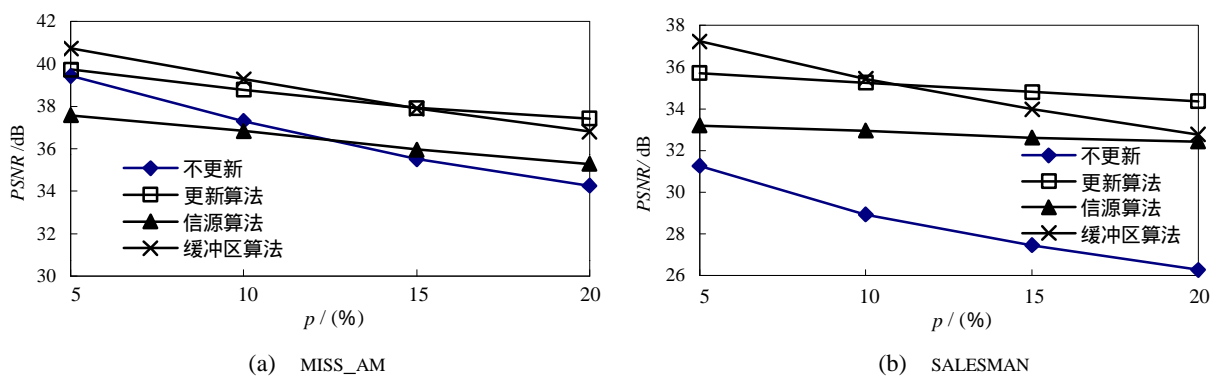


图1 4种更新算法在不同丢包率下的平均信噪比

图1为序列MISS\_AM和SALESMAN在带宽=300 Kb/s、帧速率=30 f/s时,更新算法在4种不同丢包率下的效果。从模拟结果可以看出,更新算法得出的信噪比 $PSNR$ 在各种丢包率 $p$ 下始终比信源算法高出1.5~2.0 dB,在低丢包率下略差于缓冲区算法,但在高丢包率下则比缓冲区算法略高。

## 4 结 论

本文提出了一种新的实时R-D优化帧内更新算法,增强了视频编码对抗网络错误的鲁棒性。通过比较MB帧内更新所增加的量化失真度与减少的扩散失真度,编码器可以很快决定是否需要对一个MB进行帧内更新。本算法只需要进行少量的加法/乘法操作,具有很低的计算复杂度,可以满足实时视频传输的要求。模拟结果表明,本算法比传统算法在大多情况下具有更好的更新效果。

感谢Rui Zhang博士对失真度估计算法所作的详细解释。

## 参 考 文 献

- [1] Wang Y, Zhu Q F. Error control and concealment for video communication: a review [J]. Proc. of the IEEE, 1998, 86(5): 974-997
- [2] Girod B, Farber N. Feedback-based error control for mobile video transmission [J]. Proc. of the IEEE, 1999, 87(10): 1 707-1 723
- [3] Zhu Q F, Kerofsky L. Joint source coding, transport processing and error concealment for H.323-based packet video[A]. Proc. of SPIE' 99[C], CA. USA, 1999, 3 653: 52-62
- [4] Fukunaga S, Nakai T, Inoue H. Error resilient video coding by dynamic replacing of reference pictures [A]. Proc. of IEEE Global Telecommunications Conference[C]. New York: USA, 1996, 3(11): 1 503-1 508
- [5] Liao J Y, Villasenor J D. Adaptive intra update for video coding over noisy channels [A]. International Conference on Image Processing[C]. Switzerland: 1996, 3: 763-766
- [6] Sullivan G J, Wiegand T. Rate-distortion optimization for video compression [J]. IEEE Signal Processing Mag, 1998, 15(11): 74-90
- [7] Zhang R, Regunathan S L, Rose K. Video coding with optimal intra/inter mode switching for packet loss resilience[J]. IEEE Journal of Selected Areas in Communications, 2000, 18(6): 966-976
- [8] Cote G, Shirani S, Kossentini F. Optimal mode selection and synchronization for robust video communications over error prone networks[J]. IEEE Journal on Selected Areas in Commu, 2000, 18(6): 952-965
- [9] Choi J, Park D. A stable feedback control of the buffer state using the controlled Lagrange multiplier method[J]. IEEE Trans. Image Processing, 1994, 3(9): 546-588
- [10] Jordi R C, Shawmin L. Rate control in DCT video coding for low-delay communications[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(2): 172-185

编 辑 徐培红