

一种提高 H.261 编码性能的实现方法¹

薛向阳 吴立德

(复旦大学计算机系 上海 200433)

摘 要 研究表明采用最小均方误差或绝对误差准则的块匹配运动估值算法对 H.261 编码器来说不是最好的。本文提出了一种改进的块匹配运动估值算法。它所采用的准则不仅考虑了预测误差能量的大小,还考虑了运动矢量信息以及帧间预测误差的编码比特数目的多少。实验结果表明新的准则能够显著地改善 H.261 的编码性能。

关键词 运动估值, 视频编码

中图分类号 TN919.8

1 引 言

在活动图象编码^[1,2]中,运动估值和运动补偿技术^[3-5]是提高活动图象压缩比的有效手段。传统的块匹配运动估值算法(BMA)的基本原理是将图象分割成规则的方块(称为宏块(macroblock)),然后对当前帧中每一块在前一帧图象中进行匹配,基于均方误差或绝对误差准则寻找最相似的宏块,从而得到最佳位移矢量。但是由于实际图象中存在大量噪声,因而传统的块匹配运动估值算法的性能受到噪声的严重影响,估计出的运动矢量不能反映物体的真实运动,并且还使得运动矢量间的相关性也因为噪声而减弱,最终造成运动矢量的编码比特数的增加,预测图象中方块效应明显。研究还表明即使当前图象块的帧间预测误差能量达到最小,对它进行编码所需要的总的编码比特数不一定最少。因此,需要使用新的运动估值算法和更好的准则。

本文提出一种新的准则来代替传统的绝对误差或均方误差准则,用改进的运动估值算法实现整个 H.261 编译码器,实验结果表明新方案提高了 H.261 的编码性能。

2 改进的运动估值算法和新的匹配准则

在 H.261 中,如果当前宏块采用帧间(Inter)+运动补偿(MC)编码模式,那么对该宏块编码的比特数目主要由两部分组成:运动矢量的编码比特数和帧间预测误差的编码比特数。传统的运动估值算法仅仅考虑宏块的预测误差能量大小,本文认为好的运动估值算法应该能更直接地反映宏块的编码比特数目,基于这种思想我们给出改进的运动估值算法和新的匹配准则:

(1) 设当前帧中当前待编码宏块为 C , 前一帧中由运动矢量 (mv_x, mv_y) 指定的相应位置处的宏块为 $P(mv_x, mv_y)$, 因为一个宏块由四个 8×8 大小的子块组成,所以下面给出的运算实际上是对各个 8×8 子块进行的,这里用 C_k 、 P_k 代表各个子块, $k=0,1,2,3$ 。首先计算每个子块的预测误差 $E_k(mv_x, mv_y) = C_k - P_k(mv_x, mv_y)$, 然后对 E_k 进行 DCT 运算,即 $DE_k = \text{DCT}(E_k) = \text{DCT}(C_k - P_k(mv_x, mv_y))$ 。

¹ 1996-05-22 收到, 1997-02-25 定稿

国家自然科学基金(69602003)、国家 863 项目和中国博士后科学基金资助课题

(2) 对 DCT 变换后的预测差值 DE_k 中每个系数进行量化得到量化子块 QE_k , $QE_k = Q(DE_k)$, $Q(\cdot)$ 表示量化函数, 量化台阶为 QUANT, 具体量化算法见 H.261 建议.

(3) 计算新的匹配函数 $CF(mv_x, mv_y)$:

$$CF(mv_x, mv_y) = \sum_{k=0,1,2,3} \sum_{0 \leq i, j \leq 7} \text{abs}(QE_k(i, j)) + (\text{abs}(mv_x - mv_{x\text{prev}}) + \text{abs}(mv_y - mv_{y\text{prev}})) \times \alpha \quad (1)$$

式中 (i, j) 表示 8×8 子块中每个象素的坐标, α 是加权系数, $\text{abs}(\cdot)$ 表示取绝对值运算, (mv_x, mv_y) 代表宏块 C 的运动矢量, $(mv_{x\text{prev}}, mv_{y\text{prev}})$ 代表前一个相邻宏块的运动矢量.

(4) 以一定方式 (例如三步搜索算法、对数搜索算法、共轭梯度搜索算法等^[3,4]) 搜索最佳运动矢量 (mv_x, mv_y) , 如果 $CF(mv_x, mv_y)$ 达到最小值, 则认为其对应的运动矢量就是最佳的.

从 (1) 式可以看出, 新的匹配函数主要有两项组成, 第一项是宏块的预测误差经 DCT 变换、量化后的能量大小, 即 QE_k 的能量大小, 在 H.261 中熵编码器的输入就是 QE_k , 如果 QE_k 中非 0 系数越多或系数的绝对值越大, 那么熵编码器输出的编码比特数就越多, 因此第一项的大小直接反映编码比特数的多少; 第二项表示运动矢量的相关程度, 由于 H.261 对运动矢量采用一维差分预测和 Huffman 编码, 因此第二项越小就意味着运动矢量编码所需要的编码比特数越少. 所以说, 准则 CF 意味着以下几个方面含义: 帧间预测误差能量较小、预测误差的编码比特数目较少、运动矢量的编码比特数目较少.

因此, 改进的运动估值算法和新的匹配准则能够更直接地反映宏块的编码比特数目的多少, 但是新算法引入了巨大的运算复杂度, 为减少复杂度我们需要对上述算法进行简化.

3 简化的运动估值算法

为了减少 DCT 和量化 $Q(\cdot)$ 引入的巨大运算复杂度, 我们首先计算运动矢量为 $(0,0)$ 时代价函数 $CF(0,0)$. 如果 $CF(0,0) = 0$, 则认为 $(0,0)$ 就是最佳运动矢量, 否则再搜索其它矢量. $CF(0,0) = 0$ 意味着当前宏块不需要编码. 实验表明, 这种简化方法对会议电视类图象序列 (例如 Miss America 和 Clarie) 非常有效, 我们发现, 当量化步长 QUANT = 6 时, 每幅图象中约 50% 以上的宏块不必进行搜索运算; 当 QUANT=12 时, 70% 以上的宏块不必进行搜索运算, 因而大大减轻了运算负担.

当 $CF(0,0) \neq 0$ 时, 我们将 (1) 式改为下式:

$$CF(mv_x, mv_y) = \sum_{k=0,1,2,3} \sum_{0 \leq i, j \leq 7} \text{abs}(C_k(i, j) - P_k(i, j)) + (\text{abs}(mv_x - mv_{x\text{prev}}) + \text{abs}(mv_y - mv_{y\text{prev}})) \times \alpha \quad (2)$$

由于 (2) 式避免了 DCT 和量化运算, 所以运算复杂度大大减小, 它实际上是在传统的块匹配算法上增加了运动相关性约束, 不仅具有抗噪声能力, 还能保证运动矢量编码比特数较少.

4 模拟结果

为验证算法性能,我们测试了两种 CIF 格式、30 帧/s 的序列图象,即 Miss America 和 Clarie. 量化步长 QUANT 分别取 6 和 12, 图象中所有宏块的编码模式都定为 Inter+MC, 不采用环路滤波器, 每隔 15 帧刷新一次. 图 1 至图 4 分别给出采用两种运动估值算法 (传统的全搜索块匹配运动估值算法和本文算法) 时编码结果.

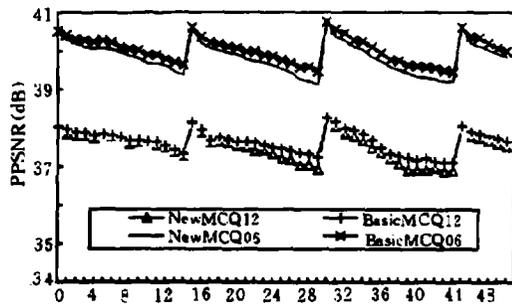


图 1 Miss America 的峰值信噪比曲线

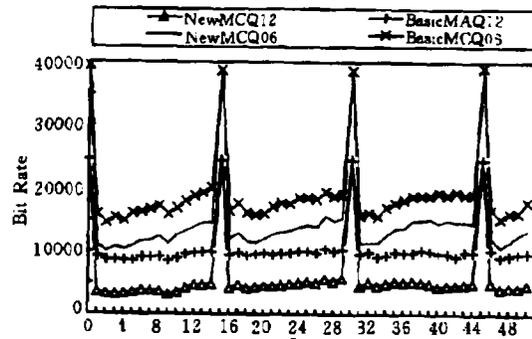


图 2 Miss America 的编码比特数曲线

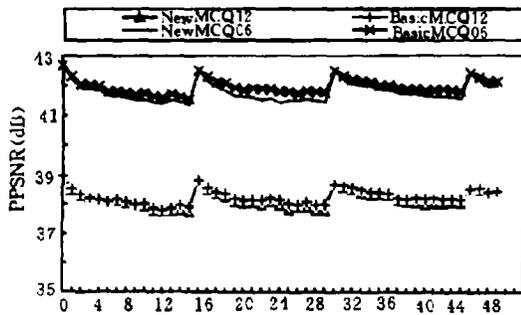


图 3 Clarie 的峰值信噪比曲线

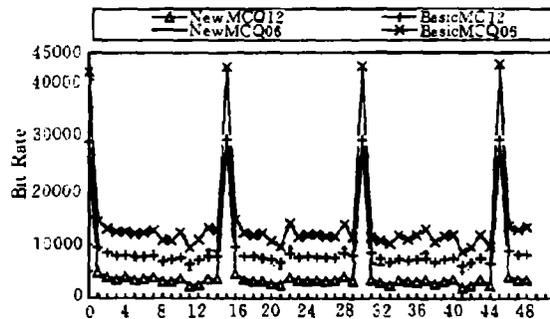


图 4 Clarie 的编码比特数曲线

注意, 在图 1 至图 4 中, 各个标注的含义是:

NewMCQ06 为新的运动估值算法, 量化步长 QUANT=6;

NewMCQ12 为新的运动估值算法, 量化步长 QUANT=12;

BasicMCQ06 为传统的运动估值算法, 量化步长 QUANT=6;

BasicMCQ12 为传统的运动估值算法, 量化步长 QUANT=12.

图 1 表明两种运动补偿算法的峰峰值信噪比 (PPSNR) 曲线非常接近, 新算法的 PPSNR 仅比传统算法差 0.133dB (QUANT=12) 和 0.197dB (QUANT=6). 图 2 表明新算法的每帧平均编码比特数比传统算法节省 4734bit (QUANT=12) 和 4085bit (QUANT=6); 如果换算成压缩比, 则新算法的平均压缩比要比基本算法提高很多, 即从 80.32:1 提高到 175.12:1 (QUANT=12), 从 44.81:1 提高到 60.05:1 (QUANT=6).

图 3 表明新算法的 PPSNR 仅比基本算法差 0.092dB (QUANT=12) 和 0.184dB (QUANT=6).

图 4 表明新算法的平均每帧编码比特数比传统算法节省 3928bit (QUANT=12) 和 3782bit (QUANT=6); 如果换算成压缩比, 则新算法的平均压缩比比传统算法提高很多, 即从 101.28:1 提高到 234.43:1 (QUANT=12)、从 65.90:1 提高到 102.44:1 (QUANT=6)。

上述结果表明新算法较大幅度地提高了 H.261 的压缩比, 图象质量损失很小, 同时运算复杂度增加很少。

5 结论

本文提出了一种改进的运动估值算法, 它很好地改善了 H.261 的编码性能, 同时计算复杂度增加不多。

参 考 文 献

- [1] CCITT Recommendation H.261, Video Codec for Audiovisual Services at $P \times 64$ kbits/s, Geneva: 1990.
- [2] MPEG, ISO CD11172.2: Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s, Nov. 1991.
- [3] Jain J R, Jain A K. Displacement measure and its application in interframe image coding. IEEE Trans. on COM, 1981, COM-29(12):1799-1806.
- [4] Musmann H G, Pirsch P, Grallert H J. Advances in picture coding, Proc. IEEE, 1985, 73(4):523-548.
- [5] Jong H M, Chen L G, Chiueh T D. Accuracy improvement and cost reduction of 3-step search block matching algorithm for video coding, IEEE Trans. on CASVT, 1994, 4(1): 88-90.

AN IMPLEMENTATION METHOD FOR IMPROVING H.261 VIDEO CODER

Xue Xiangyang Wu Lide

(Department of Computer Science, Fudan University, Shanghai 200433)

Abstract This paper proposes a method to improve H.261 video coder by modifying the block matching motion estimation algorithm. It is thought that the traditional block matching algorithm based on the least mean square error or absolute error criteria is not suitable for H.261 video coder. The good criteria must consider not only the error energy, but also the bit number for encoding motion vectors as well as the interframe prediction error. Experimental results show that the performance of H.261 coder is improved greatly by using the proposed implementation method.

Key words Motion estimation, Video coding

薛向阳: 男, 1968 年生, 博士, 现从事图象编码, 计算机视觉和图象检索等研究工作。

吴立德: 男, 1937 年生, 教授, 现从事计算机视觉, 图象检索和自然语言理解等工作。