

减少低比特率视频编码计算复杂性的一种新方法

郭红星¹, 余胜生¹, 周敬利¹, 保宗悌²

(1. 华中科技大学计算机科学与工程系, 湖北武汉 430074; 2. 武汉大学电子信息学院, 湖北武汉 430072)

摘要: 提出了一种改进的预先判断视频编码中全零系数的判别准则, 用于减少低比特率视频编码过程中的冗余计算及运动估计算法的计算复杂性. 实验结果表明, 对典型的可视电话视频序列(如 Claire), 有 50% 以上的亮度块被判别为全零系数块, 运动估计的平均搜索次数也减少了 30% 以上. 该方法可被用于减少软件编码器的运算量及其硬件电路的功耗.

关键词: 离散余弦变换; 量化; 运动估计; 视频编码

中图分类号: TN91.21 **文献标识码:** A **文章编号:** 0372-2112(2002)10-1445-05

A New Method for Reducing Computational Complexity in Low Bit Rate Video Coding

GUO Hong-xing¹, YU Sheng-sheng¹, ZHOU Jing-li¹, BAO Zong-ti²

(1. Department of Computer Science and Engineering, Huazhong University of Science & Technology, Wuhan 430074, China;

2. College of Electronic Information, Wuhan University, Wuhan 430072, China)

Abstract: An improved method for predetermining the blocks whose quantized DCT coefficients are completely zero is proposed. It is applied to eliminate redundant computations in the low bit rate video coding process and reduce the computational complexity of motion estimation algorithms. Experimental results show that 50% luminance blocks of the typical videophone sequence can be determined to be all zero coefficients, and the average search times for motion estimation are reduced by > 30%. It is very useful for reducing the computational load in software coders and the power consumption in hardware coders.

Key words: DCT; quantization; motion estimation; video coding

1 引言

近年来, 数字视频应用受到广泛关注. 视频会议、视频邮件及可视电话等数字视频通信应用的需求与日俱增. 为了缓解视频数据量与传输带宽之间的矛盾, 在传输视频信号前必须进行压缩. 视频编码是通过减少视频系列的空间冗余和时间冗余来压缩视频信号的. 运动估计与补偿是减少时间冗余的有效方法. 该技术可分为四类^[1]: (1) 梯度技术, (2) 像素递归技术, (3) 块匹配算法(BMA)及(4) 频域技术. 其中, 算法简单的块匹配算法在视频编码中得到广泛应用. 变换技术被用来减少空间冗余信息, 其中, 离散余弦变换(DCT)因为性能接近最佳变换(K-L变换), 但计算量比K-L变换少并且能够进行快速计算, 成为目前事实上的工业标准. 离散小波变换(DWT)则是被看好的技术.

采用运动估计与补偿及变换技术可以有效地压缩视频信号, 但计算量相当大. 为了减少视频编码所需的计算量, 产生了很多快速算法, 包括计算 DCT 的快速算法, 替代全搜索(FS)运动估计算法的快速搜索算法^[2], 如交叉搜索, 二维对数搜索, 三步法, 新三步法, 四步法及最近邻域搜索法(NNS)^[3]

等. 本文提出一个能够减少低比特率视频编码所需计算量的新方法. 该方法利用运动估计阶段算得的失真值来决定是否继续进行搜索过程, 并且检查可以跳过随后的编/解码阶段的块, 从而减少视频编码器所需的计算量.

2 低比特率视频编码过程

H. 263 是国际电讯联盟制定的低比特率视频编码标准^[4], 其支持基于运动估计与补偿的帧间预测. 采用帧间预测的编码模式叫做帧间模式(P 帧). 在该模式下, 只需对预测误差帧, 即原来的帧与运动补偿预测帧之差及运动矢量进行编码. 不采用帧间预测的编码方式即为帧内模式(I 帧). 在帧间编码模式下, 编码器首先进行运动补偿预测, 接着对运动补偿预测的帧差进行变换并量化所得的系数, 然后用变长码(VLC)对量化后的 DCT 系数, 运动矢量及其它信息进行熵编码.

运动补偿预测假定当前帧中的像素可以看成由参考帧中像素平移而来. 每一帧图像被分成许多宏块, 宏块尺寸为 16×16 像素, 每一宏块由 4 个 8×8 像素的亮度块($B = 0, \dots$,

收稿日期: 2001-01-11; 修回日期: 2002-07-09

基金项目: 中国博士后科学基金; 湖北省自然科学基金(No. 991010); 华中科技大学人才引进基金

3), 1 个 $8 \times 8C_b$ 块 ($B=4$) 及 1 个 $8 \times 8C_r$ 块 ($B=5$) 构成. 对每一宏块的 4 个亮度块进行块匹配, 即在参考帧搜索窗 (SW) 内, 找到与当前帧中宏块最匹配的宏块, 可以算出运动矢量 (v_x, v_y) . 绝对差之和 (SAD) 被广泛用于度量匹配程度, 即:

$$SAD(u, v) = \sum_{B=0}^3 SAD^B(u, v) \quad SAD(v_x, v_y) \leq SAD(u, v) \quad \forall (u, v) \in SW$$

$$SAD^B(u, v) = \sum_{x=0}^7 \sum_{y=0}^7 |f_{i,j}^{t,B}(x, y) - f_{i-u, j-v}^{r-1,B}(x, y)| \quad 0 \leq B \leq 3 \quad (1)$$

其中 $f_{i,j}^{t,B}(x, y)$, $f_{i-u, j-v}^{r-1,B}(x, y)$ 分别表示当前帧中 (i, j) 处的宏块的第 B 个亮度块及参考帧中 $(i-u, j-v)$ 处宏块的第 B 个亮度块, (u, v) 为位移矢量.

在基线 H. 263

中, 每个宏块对应一个用于运动补偿的运动矢量. 由于图像内容的帧间位移与采样网格不相关, 因此需进行亚像素运动估计. 这样可以有效地减少帧间误差的程度. 在 H. 263 中, 半像素运动估计被采用, 如图 1 所示. 其中, 圆

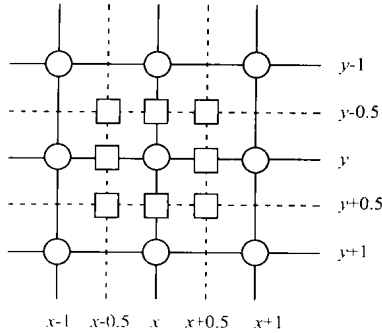


图 1 整像素及半像素位移

形代表整像素搜索点, 方形代表以整像素点为中心的 8 个半像素搜索点. 但是, 在进行半像素运动估计时, 要进行插值运算并比较其 SAD 值, 使得计算量进一步增加, 因而减少运动估计算法的计算复杂性是一个非常重要的课题.

8×8 DCT 的目的是对原来像素或运动补偿像素的 8×8 块去相关, 以尽可能地将其能量紧缩到少数系数上. 8×8 DCT 定义如下:

$$F^B(m, n) = \frac{K(m)K(n)}{4} \sum_{x=0}^7 \sum_{y=0}^7 [f_{i,j}^{t,B}(x, y) - f_{i-u, j-v}^{r-1,B}(x, y)] \cdot C(x, m) C(y, n) \quad 0 \leq B \leq 5 \quad (2)$$

其中 $F^B(m, n)$ 表示宏块中第 B 块的变换系数, 且

$$K(\Phi) = \begin{cases} \frac{1}{\sqrt{2}} & \text{当 } \Phi = 0 \text{ 时} \\ 1 & \text{当 } \Phi \neq 0 \text{ 时} \end{cases}, \quad C(z, \Psi) = \cos\left(\frac{(2z+1)\Psi\pi}{16}\right)$$

上述 8×8 DCT 形式简单, 效率较高且可以进行快速计算, 但其计算量仍然是 H. 263 编/解码中的主要部分之一.

3 消除冗余计算的方法

在低码率视频通信场合, 图像往往活动比较缓慢, 如人物的头肩部, 由于这种图像运动估计的效果较好, 其运动补偿的数据量通常很小, 量化后往往所有系数都为零. 因此, 如果能够事先判断 DCT 系数全为零, 则被判断为零的块可以省去许多冗余计算. 对运动平缓的图像序列, 这种块是很常见的, 故根据这种方法省去的计算量是很可观的. Yu 首先提出了一种在进行 DCT 和量化前预先判断 DCT 系数全为零的方法, 则被判断为零的块可以省去 DCT 和量化运算^[5]. 周对 Yu 提

出的判据进行了改进^[6]. 但这两种方法只讨论了可以节省的 DCT 及量化的计算量. 本文对周提出的判据作了进一步的改进, 以最大限度地清除 DCT 及量化中的冗余计算. 同时, 考察了其在减少运动估计算法计算复杂性中的应用.

3.1 Yu 及周的判断方法

Yu 方法的判据为: 设量化步长为 L (以下同), 则当某个块满足

$$|\sum_{x=0}^7 \sum_{y=0}^7 (f_{i,j}^{t,B}(x, y) - f_{i-u, j-v}^{r-1,B}(x, y))| < 8L \quad (3)$$

时, 被判为零系数块. 但是, 这种判据是根据 DCT 系数的统计特征来判断的, 并非充分条件, 因而往往会产生误判现象, 从而导致编码图像质量下降. 另外, 该算法还需要附加运算. 为了避免误判, 周从 DCT 公式出发, 给出了一个预先判断全零系数的充分条件, 即

$$SAD^B(v_x, v_y) < 8L \quad (4)$$

此判据为充分条件, 故不存在误判, 同时, 由于各块的绝对误差和在运动估计时已经计算过了, 只需将结果保存下来, 判断时就可以直接引用. 因此, 该方法不需任何额外运算.

3.2 改进的判断方法

经过分析发现, 周提出的判别准则并非最优, 还可以进一步改进为:

$$SAD^B(v_x, v_y) < 8L \cos^{-2}\left(\frac{\pi}{16}\right) \quad (5)$$

证明如下:

系数 $F^B(m, n)$ 量化后全为零的充要条件为*

$$\frac{|F^B(m, n)|}{2L} < 1 \Leftrightarrow |F^B(m, n)| < 2L \quad 0 \leq m, n \leq 7 \quad (6)$$

由(2)式可得

$$|F^B(m, n)| \leq \frac{K(m)K(n)}{4} \sum_{x=0}^7 \sum_{y=0}^7 |f_{i,j}^{t,B}(x, y) - f_{i-u, j-v}^{r-1,B}(x, y)| C(x, m) C(y, n) \quad (7)$$

$$\Rightarrow |F^B(m, n)| \leq \begin{cases} \frac{1}{8} SAD^B(v_x, v_y) & (m=0, n=0) \\ \frac{1}{4\sqrt{2}} \cos\left(\frac{\pi}{16}\right) SAD^B(v_x, v_y) & (m=0, n \neq 0) \cup (m \neq 0, n=0) \\ \frac{1}{4} \cos\left(\frac{\pi}{16}\right) SAD^B(v_x, v_y) & (m \neq 0, n \neq 0) \end{cases} \quad (8)$$

综合(6), (8) 两式可得, DCT 系数量化后全为零的充分条件为

$$\frac{1}{4} \cos^2\left(\frac{\pi}{16}\right) SAD^B(v_x, v_y) < 2L \quad (9)$$

式(5)得证.

注意到, 式(9)左端并非 $F^B(m, n)$ 的上确界, 该上确界由运动补偿预测帧差的空间分布决定. 但是, 如果直接确定 $F^B(m, n)$ 的上确界, 则又回到了先进行 DCT 变换然后量化的

* H. 263 标准测试模型 TMN10 中帧间系数的实际量化公式为 $(|F^B(m, n)| - L/2)/2L$, 此时充要条件可进一步改为 $|F^B(m, n)| < 5L/2$

老路. 由式(8)可知, $\frac{1}{8} SAD^B(v_x, v_y) \leq \sup(F^B(m, n)) \leq \frac{1}{4} \cos^2(\frac{\pi}{16}) SAD^B(v_x, v_y)$, 因此, 式(9)左端的取值能够保证所有情形下 DCT 系数量化后全为零, 使得(5)式为一个充分条件, 但并不必要, 因而不可避免地存在漏判. 一个极端的情形是, 当 $F^B(m, n)$ 中只存在直流量, 即当且仅当 $m=0, n=0$; $F^B(m, n) \neq 0$ 时, 有 $\sup(F^B(m, n)) = \frac{1}{8} SAD^B(v_x, v_y)$, 因而式(5)的条件可进一步放宽为 $SAD^B(v_x, v_y) < 16L$. 不过, 由于 $\cos^2(\frac{\pi}{16}) > 1$, 式(5)与式(4)相比, 前者的右端值更大, 说明改进的判据条件更宽, 从而能够减少漏判, 节省更多的计算时间.

3.3 消除冗余 DCT 计算

消除冗余 DCT 计算的方法是在原来的 H.263 编码器中增加一个两路开关控制视频编码过程. 当开关置于(a)位置时, 进行常规编码; 当开关置于(b)位置时, 就不用进行相减、DCT、量化、逆 DCT、反量化及相加运算, 而在进行运动补偿时, 只需按照运动矢量将相应块移过来即可. 对每一宏块, 首先确定当前使用的量化步长, 然后计算 $SAD(v_x, v_y)$, 并将 $SAD^B(v_x, v_y)$ 的值保存起来. 对每一亮度块 ($0 \leq B \leq 3$), 如果满足式(5), 则开关置(b)位置, 并向 VLC 单元传一个全零块; 否则, 开关置(a)位置, 进行正常编码. 这种编码工作方式不影响编码结果, 但可以节省计算量, 对硬件视频编码器也大有好处, 它可以降低功耗.

3.4 加速运动估计过程

目前, 运动估计仍然是视频编码中的一个研究热点, 但大多数算法往往只是着眼于寻找与全搜索算法得到的运动估计结果最接近的运动矢量. 实际上, 在视频编码框架中, 需将运动估计信息(运动矢量)及预测误差信息两者同时传送, 解码器才能进行运动补偿. 运动估计越精确, 预测误差就越小, 但需要更多的比特数来编码运动矢量; 反之, 运动估计越粗略, 又导致预测误差增加, 需要更多的比特数对其进行编码. 因此, 运动估计技术一方面必须提供良好的预测, 但更重要的是要保证传送的信息尽可能少, 即必须在运动信息和预测误差信息之间进行折衷. 运动估计技术的目的应该是逼近以上两项和的全局最小值. 从这个意义上讲, 全搜索运动估计算法并不是最优的^[1]. 编码运动矢量与预测误差所需比特数的这种消涨关系与具体场景有关, 目前尚无可靠的预测方法. 不过可以设想, 在进行运动估计时, 当误差值低于某一阈值中止运动估计, 此时用于编码运动矢量所节省的比特数会与编码预测误差所需增加的比特数大致相当, 使得总的比特数基本没有变化, 但可以节省计算量. 这里正是以此来加速运动估计过

程的, 具体方法(如图 2 所示)是在原来的运动估计算法的基础上增加下面的步骤:

(1) 每次运动估计前计算阈值 $T_{DCT} = 8L \cos^{-2}(\frac{\pi}{16})$, 以后当量化步长改变时, 重新计算 T_{DCT} , 否则使用原来的 T_{DCT} .

(2) 对每一搜索位置, 计算 $SAD = \sum_{B=0}^3 SAD^B$ 后, 如果 $SAD^B < T_{DCT}$ ($0 \leq B \leq 3$) 不能同时成立, 则继续原来的搜索过程 (MV_{Al}, MV_{AF}), 否则转(3).

(3) 如果当前正在进行整像素精度搜索(ME(INT)), 则在当前位置处的矢量(MV_{PI})作为运动矢量的整像素部分, 并将半像素部分置为零, 不进行半像素精度运动估计而直接结束本次运动估计过程; 如果当前正在进行半像素精度搜索(ME(HALF)), 则以当前半像素位置处的矢量(MV_{PF})作为运动矢量的半像素部分, 并结束本次运动估计过程.

由上面所增加的步骤可以看出, 这样得到的运动矢量(MV_F)并非原来算法所得到的最优矢量(MV_{AF}), 而是一个预测值. 这样在运动补偿过程中, 当前帧间差的 DCT 系数量化后是没有误差的, 但由于运动矢量为预测值, 所以与常规编码方式相比, 运动补偿过程会有误差的积累. 因此, 这里的加速运动估计的方法并非是通过清除冗余计算来完成的, 而只能算是清除准冗余计算.

4 实验结果与讨论

将上述方法分别用于 H.263 标准测试模型 TMN10^[7] 中, 对 2 个 QCIF 格式(176×144, Y: U: V= 4:2:0) 标准测试视频序列 Claire(494 帧)及 Caphone(382 帧)进行测试. 运动估计搜索范围为 -15/+15, 每 132 帧作一次帧内编码, 无任何高级选项.

4.1 节省的冗余 DCT 计算量

表 1 给出了采用本文方法(记为 NM)及周提出的方法(记为 OM)节省的不做 DCT 的块所占的百分比. 值得指出的是, 这里所清除的是冗余计算, 因而对编码器的输出码流无任何影响. 由表 1 可以看出随量化级增加, 采用本文方法和周方法, 判出的全零块都增加. 这是因为两者都具有随量化级增加自适应调整判断门限的优点. 对各种不同的情况, 采用本文方法比周方法判断不处理的块高出 0.57-2.50%, 即 2.26-9.90 个块. 而需额外增加的计算量为在编码前进行一次乘法运算计算 T_{DCT} 及以后当量化步改变时重新进行一次乘法运算计算 T_{DCT} . 因此本文方法与周方法相比, 可以进一步节省相当的计算量. 对不同的运动估计算法 FS 及 NNS, 不作处理的块所占百分比变化不大. 大多数情况下, FS 的比 NNS 的略有增加, 这是因为 FS 比 NNS 估计更准确的缘故. 当采用 NNS 方法时, 对运动量较大的视频序列 Caphone, 用本文方法判断不作处理的块所占百分比为 4.59-23.78%; 对运动量较小的视频序列 Claire, 则为 33.97-60.01. 在实际中, 经常采用的量化步长为 13, 此时对 Claire 系列, 不做处理的块达 50% 以上. 而在可视电话应用中, 说话人的运动量通常与 Claire 系列差不多. 由此可见, 采用本文方法可以大大提高低比特率视频编

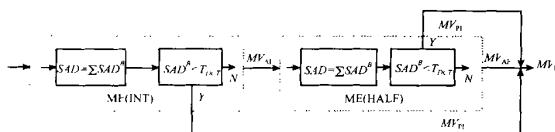


图 2 改进的运动估计模块

码速度. 本文的方法不能直接判断 2 个色度块是否不作处理. 实际应用中的一个简单处理方法是当一个宏块中的 4 个亮度块都被判为不作处理时, 则也不对 2 个色度块作处理.

再来考察该判据的效率. 表 2 给出了以 NNS 运动估计算法为例算得的实际全零 DCT 系数亮度块所占的百分比. 从中

表 1 不作处理的亮度块所占的百分比(%)

估计 算法	视频序列	量化步长											
		7		11		13		15		19		23	
		OM	MN	OM	NM	OM	MN	OM	NM	OM	NM	OM	NM
NNS	Claire	31.48	33.97	47.05	48.53	51.06	52.10	54.46	55.21	57.45	58.03	59.40	60.01
	Carphone	3.93	4.59	9.94	10.78	12.45	13.47	14.07	14.96	19.34	20.42	22.71	23.78
FS	Claire	32.58	35.08	48.32	49.73	52.17	53.07	55.16	55.85	57.79	58.36	59.89	60.48
	Carphone	4.38	5.02	10.29	11.20	12.41	13.43	15.29	16.21	20.08	21.19	23.76	24.87

表 2 实际全零 DCT 系数亮度块所占的百分比(%)

视频序列	量化步长					
	7	11	13	15	19	23
Claire	87.84	92.84	94.33	95.42	97.00	97.94
Carphone	56.03	70.12	75.13	79.11	84.87	88.60

4.2 加速运动估计的效果

在运动估计(以 NNS 算法为例)中, 采用本文所提出的方法(记为 AD)所需要的平均搜索次数(AS_T)与不采用该方法(记为 NA)的实验结果如表 3 所示. 由于本方法只是一种消除冗余计算的方法, 对输出码流有影响, 需对输出码流的变化作进一步的研究, 故表 3 也给出了输出码流的变化情况, 包括码流长度(BSL, 单位为字节)及重建视频序列的平均 PSNR(单

表 3 加速运动估计的实验结果

测试 指标	视频序列	量化步长											
		7		11		13		15		19		23	
		NA	AD	NA	AD	NA	AD	NA	AD	NA	AD	NA	AD
AS _T	Claire	23.93	21.85	24.38	17.87	24.67	16.63	24.74	15.55	24.80	15.02	24.80	14.65
	Carphone	23.23	22.99	23.42	22.45	23.42	22.15	23.49	21.94	23.52	21.42	23.66	20.97
BSL	Claire	89435	89514	51047	50947	42060	41947	36248	35873	29066	28501	25174	24501
	Carphone	328216	328124	173266	173144	135945	136145	110736	110888	80496	80568	63728	63756
NNS	Claire	38.363	38.351	35.850	35.822	34.911	34.875	34.188	34.150	32.999	32.933	32.176	32.088
	Carphone	34.930	34.927	32.294	32.285	31.357	31.354	30.618	30.612	29.437	29.422	28.527	28.523

下面考察本方法对编码器输出码流的影响程度. 首先, 由表 3 可以发现一个非常有趣的现象, 即采用本方法后, Claire 和 Carphone 序列的输出码流长度相对于原方法的变化量呈现出完全相反的变化规律. 与没有采用本方法所输出的对应码流相比, 对 Claire 序列, 当量化步长为 7 时, 码流长度有所增加, 此后一直下降, 总的趋势是先升后降, 下降的幅度随量化步长增加而加速; 对 Carphone 序列, 当量化步长为 7 和 11 时, 码流长度下降, 此后一直上升, 总的趋势是先降后升, 且上升的幅度随量化步长增加而衰减. 进一步的实验发现, 对于运动量小的序列 Claire, 运动矢量主要分布在零矢量附近, 而且预测误差很小. 当量化步长较小时, 编码预测误差所增加的比特数会超出编码运动矢量所减少的比特数; 当量化步长增加时, 预测误差经 DCT 变换后的系数大部分甚至全部为零, 所需增加的编码比特数不明显, 而用于编码运动矢量的比特数则加

可以看出, 对运动量小的序列 Claire, 大部分的全零系数块能被判别出来; 而对运动量大的序列 Carphone, 则只有很小的一部分全零系数块被正确判出. 这说明该判据更适用于处理缓变序列.

位为 dB).

由表 3 可以看出, 当不采用本文的方法时, NNS 的平均搜索次数在 23-26 间变化. 采用本文方法后, 平均搜索次数随量化步长增加而减少, 这是由于所采用的加速运动估计的终止阈值随量化步长增加而增加的缘故. 其中, 对运动量较小的序列 Claire, 这种效果尤为明显. 当量化步长大于 13 时, 需要的搜索次数降至原来的 2/3 以下. 对运动量较大的序列, 效果不如运动量小的序列明显, 其原因是因为 Carphone 序列不仅运动量大, 而且运动比较复杂, 使得运动估计的预测误差较大, 低于判断阈值的可能性小. 但是, 由于在低比特率视频通信中, 绝大多数情况下传送的是运动量较小的视频, 因此, 采用本方法可以明显降低视频编码器的计算复杂性.

速下降, 从而使得总的编码比特数下降, 而且减少量呈上升趋势. 而对于运动量大的序列 Carphone, 运动矢量分布在一个较大的范围内, 且预测误差很大. 当量化步长较小时, 编码预测误差的比特数的增加量不如编码运动矢量的减少量明显; 当量化步长增加时, 会引起预测误差大幅度增加, 导致所需的编码比特数增加量超出编码运动矢量的减少量, 但前者的增加量不如后者减少量变化快, 从而使得总的编码比特数增加, 而且增加量呈下降趋势. 其次, 从表 3 可以看出, 采用本方法前后的平均 PSNR 的差别非常细微, 最大差值仅为 0.088dB. 在实验中发现, 帧与帧的 PSNR 也总几乎是相同的. 采用本方法后的平均 PSNR 比原来的非常小的下降也主要是由于输出比特流字节数减少的缘故. 不过, 与码流长度变化规律不同的是, 采用本方法后的平均 PSNR 总是比原来的低, 这说明从 PSNR 的角度看, 运动估计精度越高, 则压缩后图像质量越好.

本方法除能够大幅度地减少编码器的运算量外,还具有两个潜在的优点。其一,抗信道错误的鲁棒性更强。由上面的分析可知,采用本方法后,用于运动矢量编码的比特数下降。数据分片模式(Data Partitioned Slice Mode DPS)将被发展作为 H. 263 的一个选项,以增强其错误恢复(Error Resilience)能力(特别是传输过程中码流内容局部被破坏的情形)。它将头及运动矢量信息与 DCT 系数数据分开,并用可逆表示(Reversible Representation)来保护运动矢量数据^[9]。既然放在运动矢量标记(Motion Vector Marker MVM)前的运动数据的长度减少了,那么运动信息被信道错误损坏的可能性也就减小了,因而使得错误清除(Error Concealment)技术会更有效。其二,使率控制的精度更高。一般而言,率控制是通过控制 DCT 系数的量化步长及跳帧数来实现的。为使控制更精确,应尽量减少使用跳帧。采用本方法后,输出码流中 DCT 系数数据所占比例更大,因而使得率控制更容易实现。

5 结论

提出了一种改进的预先判断全零 DCT 系数的方法,将该方法应用于 H. 263 中比原来的方法可以省下最多达 10 个亮度块的相减、DCT、量化、反量化、逆 DCT 及相加运算,使得总的不作上述处理的亮度块达 50% 以上。同时,将该判别准则用于运动估计中,提出一种加速运动估计的新方法,使其计算复杂性显著降低,并增加了码流抗信道错误的鲁棒性及使得率控制更精确。上述两种方法既可单独使用,也可同时使用,能显著地减少 H. 263 视频编码器的运算量。

最近,Ribas Corbera 和 Neuhoff 针对块匹配编码方案,提出最小化总的编码码率的一个寻找最优运动估计精度的理论框架^[9]。由该框可以算出最优的运动矢量精度。理论分析表明在纹理越丰富处要求运动矢量的精度越高,当由于量化步长变大或噪声、遮挡等因素使预测不准确时,所需精度较低。该方法需要计算纹理、噪声等参数来确定所需的运动矢量精度,使得计算量增加。本文提出的加速运动估计的方法可以看成是该理论框架下的一种快速实现算法,但在判断全零 DCT 系数时,还存在较严重的漏判。因此,对其中阈值的选择,特别是如何根据 DCT 系数的变化规律来提高判别效率^[10,11],还值得进一步研究。

参考文献:

[1] Dufaux F, Moscheni F. Motion estimation techniques for digital TV: a review and a new contribution [J]. Proc. IEEE, 1995, 83(6): 858-879.

- [2] Tabatabai A J, Jasinski R S, Naveen T. Motion estimation methods for video compression a review [J]. J. Franklin Inst., 1998, 335B(8): 1411-1441.
- [3] Gallant M, Cote G, Kossentini F. An efficient computation constrained block based motion estimation algorithm for low bit rate video coding [J]. IEEE Trans. Image Processing, 1999, 8(12): 1816-1823.
- [4] ITU-T SG16. Video coding for low bit rate communication [S].
- [5] Yu A, Lee R, Flynn M. Early detection of all zero coefficients in H. 263 [A]. PCS' 97 [C]. Berlin, Germany: PCS, 1997. 159-164.
- [6] 周璇, 谭径微, 余松煜. H. 263 中预先判别全零系数的新方法 [J]. 上海交通大学学报, 1998, 32(9): 107-109.
- [7] ITU-T SG16. Video Codec Test Model Near-Term Version 10(TMN10) Draft 1 [S].
- [8] ITU-T SG16. Draft For "H. 263+ + " Annexes U, V, and W To Recommendation H. 263 [S].
- [9] Ribas Corbera J, Neuhoff D L. Optimizing motion vector accuracy in block based video coding [J]. IEEE Trans. CSVT, 2001, 11(4): 497-511.
- [10] Pao I, Sun M. Modeling DCT coefficients for fast video encoding [J]. IEEE Trans. CSVT, 1999, 9(4): 608-616.
- [11] Docef A, Kossentini F, Nguuyer Phi K, et al. The quantized DCT and its application to DCF-based video coding [J]. IEEE Trans Image Processing, 2002, 11(3): 177-187.

作者简介:



wguohx@hust.edu.cn

郭红星 男, 1971 年生于湖北红安, 分别于 1994 年, 1996 年, 1999 年获武汉大学电子工程学士, 电路与系统硕士, 理学博士学位, 1999 年 6 月至 2001 年 5 月在华中科技大学计算科学与技术博士后科研流动站工作, 现为华中科技大学计算机科学与工程系副教授, 主要从事信号处理与多媒体, 可重构计算与演化硬件方面的研究。email:

余胜生 1944 年生, 华中科技大学计算机科学与工程系教授, 博士生导师, 现任外存储系统国家专业实验室主任, 国务院学位委员会计算机学科评议组成员, 主要研究方向为计算系统结构与网络通讯技术。

周敬利 1946 年生, 华中科技大学计算机科学与工程系教授, 博士生导师, 现任计算机学院副院长, 主要研究方向为网络存储与多媒体通讯技术。

保宗梯 1933 年生, 武汉大学电子信息学院教授, 博士生导师, 主要研究方向为空间物理与信息处理。