

# 重构内部宏块在视频转换编码中的应用

徐岩 李桂苓 刘昱 郭映

(天津大学电子信息工程学院, 天津 300072)

**摘要** 视频空间分辨率下变换转换编码过程的运算量较大,受实时性、输出码率的限制,转换后的图像质量会有一定程度的下降。为了改善转换后码流的质量,同时又不会大幅增加计算量和输出码率,提出了综合考虑相邻宏块的运动矢量的集簇状态及其残差信息的情况,判断下变换后宏块纹理信息和运动矢量的失真情况,然后动态地选择一种转换编码算法。实验结果表明,使用该方案进行转换编码后,码流的质量有一定提高,同时避免了进行大范围运动搜索和引入大量内部编码的宏块。

**关键词** 转换编码 运动矢量 欧几里德距离 残差信息 重构内部宏块

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2007)02-0283-05

## Intra-refresh in Video Transcoding

XU Yan, LI Gui-ling, LIU Yu, GUO Ying

(School of Electronic and Information Engineering, Tianjin University, Tianjin 300072)

**Abstract** Video transcoding for reduced spatial resolution needs a lot of computation. Restricted by real-time demand, the quality of transcoded stream will decrease in variant degree. In order to improve the quality of transcoded stream, a new scheme is proposed in this paper which estimates the distortion degree of downscaled texture and motion vector according to the condition of neighbor macroblocks' texture and the cluster degree of their motion vectors, and then adaptively chooses one arithmetic to transcode. Experimental results show that this scheme can bring some improvement to the quality of transcoded stream, and at the same time avoid extensive motion search and introducing much of intra-coded macroblocks.

**Keywords** transcoding, motion vector, Euclidean distance, residual, intra-refresh

## 1 引言

随着多媒体和网络技术的发展,多媒体服务如电视会议、视频点播、远程教育等在生活中越来越普遍。在这些应用中,经常需要进行各种变换来适应终端约束、网络限制或某个用户的特殊需要。空间分辨率下变换转换编码是其中的一种。在下变换转换编码中,重量化、运动矢量映射和残差信息下采样都可以产生漂移误差。为减小漂移误差,运动矢量优化(MVR)和重构内部宏块(Intra-Refresh)技术被提出并应用于转换编码器中<sup>[1-5]</sup>。

MVR的过程与编码中运动矢量预测的过程相似,是个运算量大、耗时多的过程,令 $[-n, +n]$ 表示运动矢量优化时需要的搜索区的范围,运动矢量优化中需要的搜索点数则为 $4 \times n \times (n + 1)$ , $n$ 每增加1,搜索点数就增加 $8 \times (n + 1)$ 。为节省计算量,通常取 $n = 2$ ,搜索点数为 $24^{[3]}$ 。但是在文献[6]中,实验结果表明,在空间分辨率下变换中 $[-2, +2]$ 大小的搜索窗只包含了不到70%的最佳运动矢量。这说明在转换编码运动剧烈、复杂的视频序列时运动矢量优化技术提高质量的能力是有限的。

Intra-Refresh技术的提出是基于这样一个事实,即由重量化、运动矢量映射和残差信息下采样产生

基金项目:天津市自然科学基金项目(033600511)

收稿日期:2005-05-26;改回日期:2005-10-30

第一作者简介:徐岩(1977~),女,天津大学电子信息工程学院信号与信息处理专业博士研究生。现主要研究方向为数字视频、多媒体技术。E-mail: xu\_yan@yeah.net

的漂移误差只会通过非内部宏块传播。如果把某些非内部宏块变成内部宏块,就可以控制漂移误差的传播<sup>[5]</sup>。

MVR 和 Intra-Refresh 方法具有一定的互补性: MVR 方法由于运算量大,搜索窗不能取得太大,一般取  $[-2, +2]$ ,适用于运动较平缓、一个宏块组中宏块的运动矢量较集中的图像区域;而对于各宏块的运动矢量发散较严重的图像区域收效甚微,这种情况下, Intra-Refresh 方法是最好的选择。

考虑到 Intra-Refresh 方法会使下变换后的码流中内部(intra)宏块数目增加,码率增大,所以应尽可能减少引入的内部宏块,如果能用 MVR 就可以达到或接近用 Intra-Refresh 的效果,应该优先选择用 MVR 的方法。

基于以上分析,提出了一种综合考虑运动矢量和残差信息的状态、动态选择转换编码方法(MVR 或 Intra-Refresh)的转换编码方案。

## 2 分 析

在空间分辨率下变换转换编码中,宏块经运动补偿后得到的残差信息,在经 DCT(离散余弦变换)后,常用来表示该宏块的空间活动性,在经 DCT 的宏块残差信息中,非零 AC(交流)系数的个数越多,表示该宏块的空间活动性越强,下变换后的运动矢量越应该向该宏块倾斜<sup>[7]</sup>。所以可以用宏块残差信息的非零 AC 系数的个数做加权系数、用加权平均的方法对相邻宏块的运动矢量进行下变换,得到低分辨率的新的运动矢量<sup>[7]</sup>。

以空间分辨率 1/2 下变换为例,令  $mv_0, mv_1, mv_2, mv_3$  表示 4 个相邻的运动矢量,  $mv_d$  表示下变换后的运动矢量,  $A_i (i=0, 1, 2, 3)$  表示相邻 4 个宏块的空间活动性(非零 AC 系数的个数),则有<sup>[7]</sup>:

$$mv_d = \frac{1}{2} \frac{\sum_{i=0}^3 A_i \times mv_i}{\sum_{i=0}^3 A_i} \quad (1)$$

当源相邻宏块的运动矢量呈集簇状态或只有轻微的发散的时候,下变换后的源运动矢量  $\frac{1}{2}mv_i (i=0, 1, 2, 3)$  通常会位于以  $mv_d$  为中心的小范围搜索窗内,漂移误差可以通过在映射运动矢量周围的小范围内进行搜索(运动矢量优化)、找到最优运动矢量得以校正。

如果源相邻宏块的运动矢量的发散程度较大,下变换后的源运动矢量  $\frac{1}{2}mv_i (i=0, 1, 2, 3)$  也是发散的,但由于残差信息中非零 AC 系数的加权平均作用,下变换后的运动矢量有可能会位于源发散的

运动矢量的中心(或接近中心)区域。如在“mobile”序列的第 2 帧的第 45 个宏块组(相邻的 4 个宏块为一个宏块组)中,相邻 4 个宏块的空间活动性(非零 AC 系数的个数)  $A_i (i=0, 1, 2, 3)$  分别为

$$\{A_0 = 41 \quad A_1 = 49 \quad A_2 = 26 \quad A_3 = 30\} \quad (2)$$

它们的运动矢量  $mv_i (i=0, 1, 2, 3)$  分别为

$$\begin{aligned} mv_0 &= (1, -7) & mv_1 &= (-3, 1) \\ mv_2 &= (0, 0) & mv_3 &= (-3, 1) \end{aligned} \quad (3)$$

用加权平均的方法求下变换后的运动矢量  $mv_d$ <sup>[7]</sup>, 将式(2)和式(3)带入式(1), 可得

$$mv_d = (-1, -1) \quad (4)$$

比较  $mv_d$  与 1/2 空间分辨率下变换后的  $mv_i$ :

$$\frac{1}{2}mv_i = \{(0, -3) \quad (-2, 0) \quad (0, 0) \quad (-2, 0)\} \quad (5)$$

图 1 所示为  $\frac{1}{2}mv_i$  与  $mv_d$  的位置关系。

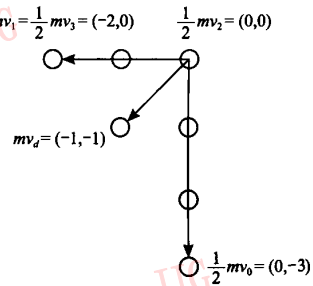


图 1 1/2 空间分辨率下变换后的  $mv_i$  与  $mv_d$  的位置关系示意图

Fig.1 Schematic of the positions of  $mv_d$  and  $1/2 mv_i$

从图 1 中可以看出,虽然源 4 个运动矢量较发散,但经加权平均下变换后的运动矢量位于其中心,4 个源运动矢量分别经 1/2 下变换后,位于以加权平均得到的运动矢量为中心的  $[-2, +2]$  的小窗内。也就是说源码流中 4 个发散得较严重的运动矢量,在经过用空间活动性进行加权平均后,新的运动矢量有可能会位于源运动矢量的中心或接近中心的位置,此时下变换后的 4 个源运动矢量分布在新运动矢量的周围,可能会包含在以新运动矢量为中心的  $[-2, +2]$  的搜索窗内,或接近该搜索窗。用运

动矢量优化即可较好地修正下变换后的运动矢量。

### 3 过程

首先,分别计算运动矢量  $mv_i$  与其他 3 个运动矢量之间的欧几里德距离(这里取内部宏块的运动矢量为 0),并把它们相加得到

$$d_i = \sum_{j=0,1,2,3}^3 \|mv_i - mv_j\| \quad (6)$$

并从中得到最大的欧几里德距离

$$d_{\max} = \max_{i=0,1,2,3} (d_i) \quad (7)$$

这时需要两个门限值:上限  $t_{\text{high}}$  和下限  $t_{\text{low}}$ 。

如果

$$d_{\max} \leq t_{\text{low}} \quad (8)$$

表示相邻的 4 个运动矢量是集簇的,只进行运动矢量和纹理信息下变换(开环结构)也可以得到较满意的图像质量。

如果

$$t_{\text{low}} < d_{\max} < t_{\text{high}} \quad (9)$$

表示相邻的 4 个运动矢量只有些轻微的发散,滤波(下变换)后,漂移误差可以通过在映射运动矢量周围的小范围内进行搜索(运动矢量优化)找到最优运动矢量得以校正。

如果

$$d_{\max} \geq t_{\text{high}} \quad (10)$$

表示相邻的 4 个运动矢量发散得很厉害<sup>[8]</sup>。这时,需要继续比较加权平均下变换后的运动矢量  $mv_d$  与  $\frac{1}{2}mv_i$

求  $mv_d$  与  $\frac{1}{2}mv_i (i=0,1,2,3)$  的平均最大绝对距离

$$d = \frac{1}{4} \sum_{i=0}^3 d_i (mv_d, \frac{1}{2}mv_i) \quad (11)$$

其中

$$d_i (mv_d, \frac{1}{2}mv_i) = \max \left( \left| mv_d^x - \frac{1}{2}mv_i^x \right|, \left| mv_d^y - \frac{1}{2}mv_i^y \right| \right) \quad (12)$$

其中,  $mv_d^x, mv_d^y$  分别表示下变换后的运动矢量  $mv_d$  的水平和垂直分量,  $mv_i^x, mv_i^y$  分别表示第  $i$  个运动矢量  $mv_i$  的水平和垂直分量。

若

$$\lfloor d \rfloor \leq n \quad (13)$$

( $n$  为运动矢量优化搜索窗的半径),那么说明对这组宏块,如果以加权平均得到的运动矢量为基准(中心),在其周围进行搜索窗为  $[-n, n]$  的运动矢量优化,可以得到最优的低分辨率运动矢量,所以对这组宏块,用运动矢量优化的方法即可得到较好的效果,而不必使用会导致码率上升的重构内部宏块的方法。

若

$$\lfloor d \rfloor \geq n + 1 \quad (14)$$

说明对该组宏块,  $1/2$  空间分辨率下变换后的 4 个源运动矢量没有落在以加权平均得到的运动矢量为中心的  $[-n, n]$  的搜索窗内,运动矢量优化也难以校正漂移误差。所以对这样的宏块组,还是需要用重构内部宏块的方法才能保证较好的图像质量,即对一个宏块组中的非内部宏块进行运动补偿,使其成为内部宏块,然后再进行滤波下变换,并设下变换后的宏块的编码模式为内部(intra)。

### 4 门限值及搜索窗的确定

因为输入码流中的运动矢量是半像素精度的,  $1/2$  下变换后,如果两个运动矢量的水平或垂直分量的距离小于 0.5,误差就可以忽略。而且因为  $d_i$  是 3 个欧几里德距离之和,而每个欧几里德距离都是两个 2 维变量之间的距离,所以可以得到<sup>[8]</sup>:

$$0.5 \times 2_{\text{dimension}} \times 3_{\text{distance}} \times 2_{\frac{1}{2}\text{downscale}} = 6 \quad (15)$$

所以,设

$$t_{\text{low}} = 6 \quad (16)$$

也就是说  $1/2$  下变换后,两个运动矢量的水平或垂直分量的平均距离的门限值是 0.5。

令  $D$  表示  $1/2$  下变换后两个运动矢量的水平或垂直分量的最大距离,于是有

$$D = d_{\max} / (2 \times 3 \times 2) \quad (17)$$

如果  $D$  小于 0.5,误差可以忽略。如果  $D$  大于 0.5,需要进行 MVR 或 Intra-Refresh 来校正漂移误差。

令  $\pm n$  表示运动矢量优化时需要的搜索区的范围。为得到最优的优化效果,  $\pm n$  应该覆盖 4 个  $1/2$  下变换的运动矢量所指向的全部点。因此可以得到:

$$n = \lceil D \rceil \quad (18)$$

考虑到式(9)中采用的规则,  $t_{\text{high}}$  越大,运动矢量优化中所需的搜索点数就越多。然而如果  $t_{\text{high}}$  太

小,输出码流就会因为含有太多内部宏块而增大码率。为平衡计算量和输出码率,取

$$n = 2 \tag{19}$$

根据式(10)、(17)、(18)

$$t_{high} = 24 \tag{20}$$

### 5 实验与结果

MVR 表示只使用运动矢量优化进行转换编码; MVBA (motion vector based adaptive transcoding technique)表示用文献[8]中提出的“求运动矢量的最大欧几里德距离来判断是否使用 Intra-Refresh”的方法进行转换编码; MVRA (motion vector and residual based adaptive transcoding technique)表示用本文提出的“综合考虑运动矢量和残差信息进行 intra-refresh”的方法进行转换编码。

上述 3 种方案均在一个 MPEG2-H.263 空间分

辨率 1/2 下变换转换编码器(基于 VC 6.0 的软件包)中被实现。这种转换编码器的输入是 MPEG-2 视频流,输出是水平、垂直空间分辨率都是输入码流的 1/2 的 H.263 视频码流。用级联转换编码器<sup>[1,2]</sup>生成的视频流作为基准分别计算 MVR、MVBA、MVRA 这 3 种结构生成的视频流的峰值信噪比 (PSNR)。

图 2 表示 6 种视频序列: foreman、mobile、football、silent、suzie 和 popple 分别用上述 3 种方法进行空间分辨率 1/2 下变换得到的码流的亮度(Y)分量的 PSNR 的比较。这 6 种序列的帧率均为 25fps,图组 (GOP) 结构为  $N = 10, M = 1$ 。其中 foreman、silent、mobile 和 football 序列是 CIF 格式的, silent、mobile 序列各有 300 帧, foreman、football 序列各有 90 帧, susie 和 popple 序列是 4CIF 格式的,各有 250 帧。输出的 H.263 码流中 I 帧和 P 帧的量化参数固定为 5。

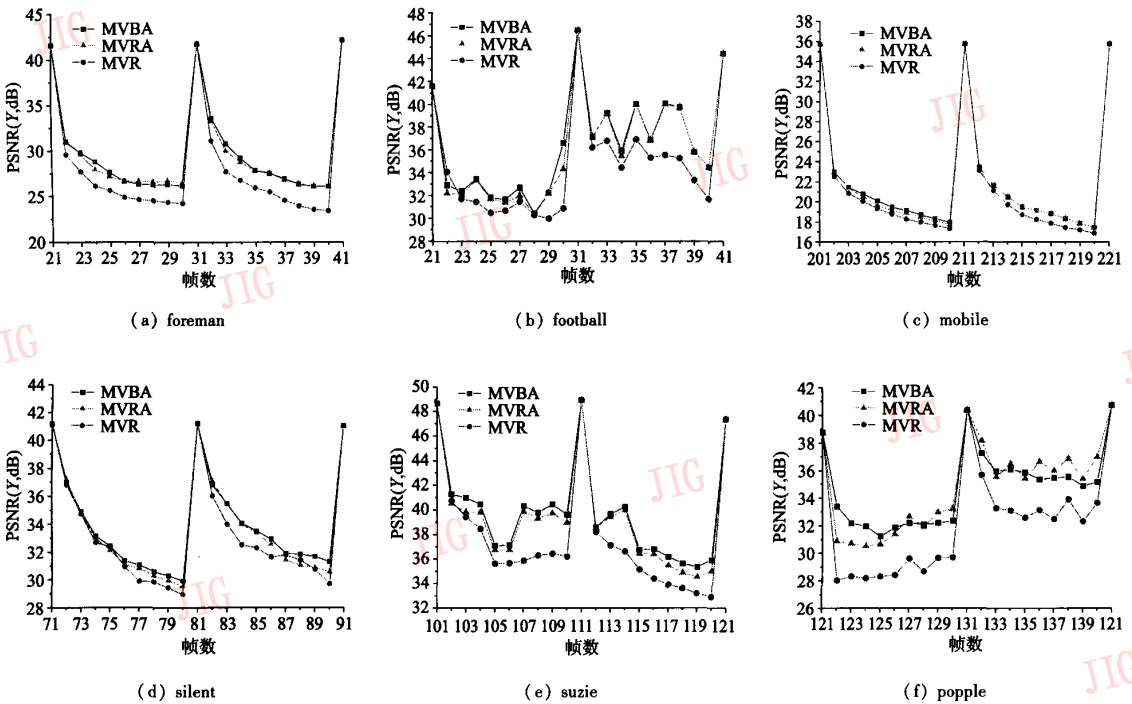


图 2 3 种方法进行空间分辨率 1/2 下变换得到的亮度分量 PSNR 的比较(以帧为单位)

Fig. 2 Frame-based comparison of PSNR of the luminance component for MVBA, MVRA and MVR architectures

表 1 表示上述 6 种序列分别用 MVR、MVBA 和 MVRA 方法进行空间分辨率 1/2 下变换后得到的码流的亮度分量(Y)的平均 PSNR 和平均码率。

从图 2 和表 1 中可以看出,用 MVBA 和 MVRA

方法得到的码流均好于用 MVR 方法得到的码流质量,用 MVRA 方法得到的码流质量与用 MVBA 方法得到的码流质量非常接近,且平均码率大大降低。

表 1 MVR、MVBA 和 MVRA 3 种方法的性能比较

Tab. 1 Comparison of the performance of MVR, MVBA and MVRA architectures

序列	格式	MVR		MVBA		MVRA	
		平均 PSNR(Y, dB)	码率(kbps)	平均 PSNR(Y, dB)	码率(kbps)	平均 PSNR(Y, dB)	码率(kbps)
Foreman	CIF→QCIF	28.196	364.4	29.393	455.55	29.347	371.11
Mobile	CIF→QCIF	30.317	1309.62	31.001	1358.67	30.992	1320
Football	CIF→QCIF	31.916	699.53	32.974	854.79	32.946	748.89
Silent	CIF→QCIF	32.843	245.3	33.885	274	33.841	246
Popple	4CIF→CIF	30.948	2248	33.075	2480	32.958	2320
suzie	4CIF→CIF	36.938	560.8	38.032	896	37.931	598.4

## 6 结 论

本文提出的方法利用输入码流中自带的运动矢量的失真情况,然后动态地选择一种合适的转换编码算法,可较方便地嵌入至转换编码模块中实现。实验结果表明,该方法有助于提高视频转换编码的编码效率,改善了转换后码流的质量。

### 参考文献 (References)

- 1 Anthony Vetro, Charilaos Christopoulos, Sun Huifang. Video transcoding architectures and techniques: an overview [J]. IEEE Signal Processing Magazine, 2003, 20(2): 18 ~ 29.
- 2 Jeongnam Youn, Sun Ming-ting, Lin Chia-wen. Motion estimation for high performance transcoding[J]. IEEE Transactions on Consumer Electronics, 1998, 44(3): 649 ~ 658.
- 3 Jeongnam Youn, Sun Ming-ting, Lin Chia-wen. Motion vector refinement for high performance transcoding[J]. IEEE Transactions on Multimedia, 1999, 1(1): 30 ~ 40.
- 4 Anthony Vetro, Yin Peng, Liu Bede, *et al.* Reduced spatio-temporal transcoding using an intra refresh technique [A]. In: Proceedings of 2002 IEEE International Symposium on Circuits and Systems [C], Scottsdale, Arizona, USA, 2002; 723 ~ 726.
- 5 Yin Peng, Anthony Vetro, Liu Bede, *et al.* Drift compensation for reduced spatial resolution transcoding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(11): 1009 ~ 1020.
- 6 Junichi Nakajima, Hiroyuki Tsuji, Yoshiyuki Yashima, *et al.* Motion vector re-estimation for fast video transcoding from MPEG-2 to MPEG-4 [A]. In: Proceedings of 2001 Workshop and Exhibition on MPEG-4 [C], San Jose, California, USA, 2001; 87 ~ 90.
- 7 Shen Bo, Ishwar K Sethi, Bhaskaran Vasudev. Adaptive motion-vector resampling for compressed video downscaling [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(6): 929 ~ 936.
- 8 Xu Yan, Meng Qing-gang, Liu Yu, *et al.* MV-based adaptive transcoding technique for reduced spatial resolution [A]. In: Proceedings of 2004 IEEE International Symposium on Consumer Electronics [C], Reading, UK, 2004; 324 ~ 327.