

# 一种自适应运动估值视频代码转换的新方法

李晓辉

(安徽大学计算智能与信息处理教育部重点实验室, 安徽大学电子工程与信息科学系, 合肥 230039)

**摘要** 针对窄带网络的视频信号传输问题, 分析了传统视频代码转换帧速率转换时, 由于运动矢量非最佳化所造成的图象质量下降的原因, 并提出了一种基于量化误差的自适应运动矢量模型, 从而减小了搜索域, 使最佳化输出运动矢量能进行快速运动估值; 同时根据灰度系统理论, 提出了一种有效的灰度预测搜索方法. 另外, 又根据DCT系数理论模型, 采用自适应快速视频编码方法, 进一步提高了编码速度. 实验结果表明: 该方法不仅改善了视频图象质量, 而且计算复杂度也大大减小.

**关键词** 视频 代码转换 自适应 运动估值

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2001)02-0182-04

## A New Method of Video Transcoding Using Adaptive Motion Estimation

LI Xiao-hui

(The Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education,  
Dept. of Electronic Engineering and Information Science, Anhui Univ., Hefei 230039)

**Abstract** To the video signal transmission on narrow-band networks, this paper analyzes the quality degradation that cause the non-optimized motion vectors during the traditional video transcoding when the frame-rate conversion. We proposed a model of adaptive motion vectors based on the sum of the differential quantization errors, thus reduced search area. It is possible to achieve fast motion estimation for near-optimal outgoing motion vectors. Based on the gray system theory, an efficient gray prediction search(GPS) algorithm for block motion estimation is proposed. The GPS can determine the motion vectors of image blocks quickly and correctly. We speed up the encoding in further using a method of adaptive fast video encoding based on the theoretical model for discrete cosine transform(DCT) coefficients. We also present a fast DCT algorithm to speed up the calculation of DCT further when the quantization step size is large. The experimental result show that the proposed scheme improves the video quality, with a minimal increase in computational complexity.

**Keywords** Video, Transcoding, Adaptive, Motion estimation

## 0 引言

由于视频代码转换可将一种已压缩的高速视频码流转换成低速率的视频码流, 以能保证其在不同网络中的正确传输, 因此可为不同网络上的用户提供不同服务质量的视频服务. 对于用公共交换电话网和无线网络等窄带传输的视频信号来说, 为了获得低速率视频信号, 就必须对视频信号进行高倍压缩. 这时, 如采用全帧速率对视频信号进行编码, 其

图象质量下降是很严重的, 因此, 代码转换中可采用丢帧的方法来降低帧速率, 并将多余的数据流分配给保留帧, 从而达到改善图象质量的目的.

在多数视频压缩编码标准中, 包括 MPEG1、H.261 和 H.263 等, 运动估值是在基于平均绝对误差(MAD)的亮度宏块上进行的<sup>[1]</sup>. 为了获得当前宏块的运动矢量, 必须在先前参考帧预先定义的搜索域S中, 搜索使MAD最小的匹配块. 在代码转换中, 用于输出数据流的最佳运动矢量虽然可以用全域全搜索运动估值的方法来获得, 但该方法存在着计算

复杂性大的缺点.为了减少计算复杂性,通常采用输入运动矢量重复使用模式,而该模式在代码转换期间,由于直接使用了输入码流中的运动矢量和宏块模式,从而避免了复杂的运动估值和宏块模式判决过程,并使运算量大大降低.由于这些数据,尤其是运动矢量对新的目标码率而言不是最佳的,因而会引起编码效率的下降,而且在有帧速率转换发生,位速率剧烈下降时,常常还会引起严重的视频质量下降,由此可见,按此数据重建的图象质量往往是不能接受的.因此,当位速率过分下降时,采用精确的运动矢量以防止剧烈的视频质量下降是必要的.

### 1 自适应运动矢量

视频代码转换原理框图如图 1 所示,它是由一个解码器和一个编码器级联而成,即先对编码码流(码率为  $R_1$ )进行解码,然后按新的目标码率  $R_2$  对重建图象进行重新编码,而且其通常满足  $R_2 < R_1$ ,即输入视频信号流通过视频代码转换机后,速率得以降低<sup>[2]</sup>.

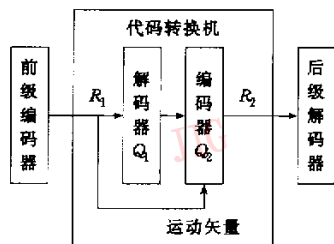


图 1 视频代码转换结构框图

设  $n - k - 2, n - k - 1, n - k, \dots, n - 1, n$  为一图象帧序列,其中,若图象序列从  $n - k$  帧到  $n - 1$  帧被丢掉,则输出运动矢量  $(O_x, O_y)$  可以从先前重构参考帧的第  $(n - k - 1)$  帧来获得.丢帧期间输入的运动矢量为  $(V_x, V_y)_{n-k}, (V_x, V_y)_{n-k+1}, \dots, (V_x, V_y)_{n-1}$ ,而从当前帧中抽取的运动矢量为  $(V_x, V_y)_n$ ,第  $n$  帧的输出运动矢量  $(O_x, O_y)$  为

$$(O_x, O_y)_n = \left[ \sum_{d=1}^{k+1} (V_x)_{n-d+1}, \sum_{d=1}^{k+1} (V_y)_{n-d+1} \right] \quad (1)$$

在代码转换期间,由于解码和编码时的量化步长不同,而引入了量化误差,因而使得输入的运动矢量不能再作为最佳化运动矢量输出和重建图象,并使最佳运动矢量在原来输入运动矢量周围的一个很小范围内发生了飘移.

为此,本文引入基于多种量化误差总和(SDQE)的自适应运动矢量模式.在这种情况下,先将从输入视频流中获得的运动矢量定义为基运动矢量  $(B_x, B_y)$ ,然后在基运动矢量周围很小的范围内进行运动估值,并将增量运动矢量  $(D_x, D_y)$  定义为基运动矢量和最佳运动矢量的差值,即如果给出了基运动矢量和增量运动矢量,则最佳运动矢量  $(O_x, O_y)$  为基运动矢量  $(B_x, B_y)$  与增量运动矢量  $(D_x, D_y)$  之和.

在代码转换期间,如果有帧速率转换发生,如运动图象系列从  $n - k$  帧到  $n - 1$  帧被丢掉,则基运动矢量可以用下式获得

$$(B_x, B_y)_n = \left[ \sum_{d=1}^{k+1} (V_x)_{n-d+1}, \sum_{d=1}^{k+1} (V_y)_{n-d+1} \right] \quad (2)$$

在一个小的搜索域  $S^R$  中,增量运动矢量  $(D_x, D_y)$  可以采用单变量一阶线性动态灰度模型 GM(1, 1) 的灰度预测算法<sup>[3]</sup>来进行计算.通常,由实验获得的原始数据由于缺少相关性,且不足以建立一个灰度模型,因此需要由原始数据产生一个更有规律的数据序列.设原始数据序列  $X = [x(1), x(2), \dots, x(n)]$  其中  $x(i)$  表示  $i$  时刻的输出,新数据序列  $X' = [x'(1), x'(2), \dots, x'(n)]$  由下式得到

$$x'(k) = \sum_{m=1}^k x(m) \quad k \in \{1, 2, \dots, n\} \quad (3)$$

灰度系统的预测函数为

$$\hat{x}'(k+1) = \left[ x(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \quad (4)$$

$$\hat{x}(k+1) = \hat{x}'(k+1) - \hat{x}'(k) \quad (5)$$

式中,  $k > 0$ ,  $\hat{x}$  为  $x$  的预测值.  $a, b$  可由下式得到

$$\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T y_n \quad (6)$$

$$B = \begin{bmatrix} -0.5(x'(1) + x'(2)) & 1 \\ -0.5(x'(2) + x'(3)) & 1 \\ \dots & \dots \\ -0.5(x'(n-1) + x'(n)) & 1 \end{bmatrix} \quad (7)$$

$$y_n = [x(2), x(3), \dots, x(n)]^T \quad (8)$$

基于上面的讨论,如果 SDQE 较前级编码器的平均绝对误差更小,则表明由输入运动矢量的重复使用而造成的质量下降并不是很严重的,因而就有可能设计出基于 SDQE 的自适应运动矢量模式.如果当前宏块的 SDQE 值低于某一门限值,则可不采用改进的运动矢量,而可直接将输入的运动矢量用

作输出运动矢量。

## 2 自适应快速视频编码

### 2.1 理论依据

DCT 变换的矩阵形式为

$$F = AfA^T \quad (9)$$

式中,  $A$  的第  $u$  列向量为  $\frac{1}{2} \alpha(u) \cos\left[\frac{(2m+1)u\pi}{16}\right]$ 。

DCT 输入子块的象素值可以用均值为零, 协方差  $r(x, y) = \sigma_f^2 \rho^{|x-y|}$  的拉普拉斯分布来近似<sup>[4]</sup>。

DCT 系数第  $(u, v)$  元素的方差  $\sigma_f^2(u, v)$  可以表示成

$$\sigma_f^2(u, v) = \sigma_f^2 [ARA^T]_{u,u} [ARA^T]_{v,v} \quad (10)$$

式中  $[\cdot]_{u,u}$  表示矩阵第  $(u, u)$  个元素。自协方差矩阵的元素有如下关系

$$[R]_{ij} = \rho^{|i-j|} \quad (11)$$

式中,  $\rho$  是相关系数  $0 \leq \rho \leq 1$ ;  $i, j = 0, 1, \dots, 7$ 。也即

$$R = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^7 \\ \rho & 1 & \rho & & \\ \rho^2 & \rho & 1 & & \\ \dots & & & \ddots & \\ \rho^7 & \dots & & & 1 \end{bmatrix} \quad (12)$$

以上等式表明, 可以通过 DCT 输入子块象素值的方差来估算 DCT 系数的方差, 同时表明了直流系数方差大于交流系数方差, 这意味着直流系数被量化成零的概率远低于交流系数, 所以很容易将直流系数同交流系数分离开来。在某些情况下, 只需计算直流系数分量, 而所有交流系数都可置为零。

因为  $MAD$  值在运动估值时很容易求得, 且  $MAD$  值处于  $(-3\sigma, 3\sigma)$  区间的概率为 99%, 所以当量化步长满足下式时,

$$(2S_Q + Z) > 3\sigma_f \quad (13)$$

则 99% 的 DCT 系数将被量化成零。式中,  $S_Q$  为量化参数,  $Z$  为量化器的死区。

从以上讨论中可知, 由给定的  $MAD$  和  $S_Q$ , 即可以确定 DCT 系数量化为零的概率, 这样即可通过仅仅计算那些不会量化为零的系数来减少编码的计算量。

### 2.2 算法实现

通过比较不同量化参数时 DCT 系数的方差, 即可估计出经量化后 DCT 系数变为零的概率, 而且从式 (10) (12) 即可得出阈值确定的判决标准

$$MAD < Cof(u, v) \times S_Q + Z \quad (14)$$

其中,  $Cof(u, v) = \frac{C}{\sqrt{[ARA^T]_{u,u} [ARA^T]_{v,v}}}$ ,  $C$  是常数。

随着  $Cof(u, v)$  的变化, 对应 DCT 系数量化为零的概率也不同。根据以上讨论, 现提出一种自适应方案以用于减少在标准视频编码器中 DCT、IDCT、量化、反量化模块的计算<sup>[5]</sup>。

这里, 设  $MAD < Cof(0, 0) \times S_Q + Z$  为第 1 个阈值;  $MAD < Cof(1, 0) \times S_Q + Z$  为第 2 个阈值;  $MAD < Cof(2, 0) \times S_Q + Z$  为第 3 个阈值, 设  $MAD < Cof(4, 0) \times S_Q + Z$  为第 4 个阈值。

为了进一步减少计算量, 当量化参数很大时, 由于变化后的系数被粗量化, 因而 DCT 不需要太精确。为此, 本文提出了一种快速算法来近似  $4 \times 4$  低频 DCT 系数的计算。

标准的一维 8 象素 DCT 为

$$F(u) = \frac{1}{2} \alpha(u) \sum_{x=0}^7 f(x) \cos\left[\frac{(2x+1)u\pi}{16}\right] \quad (15)$$

当  $u = 0$  时,  $\alpha(u) = 1/\sqrt{2}$ ;  $u \neq 0$  时,  $\alpha(u) = 1$ 。

式 (14) 中, 第 1 个余弦项是  $\cos\left(\frac{u\pi}{16}\right)$ ; 第 2 个余弦项是  $\cos\left(\frac{3u\pi}{16}\right)$ ; 第 3 个余弦项是  $\cos\left(\frac{5u\pi}{16}\right)$ ; 第 4 个余弦项是  $\cos\left(\frac{7u\pi}{16}\right)$ , 等等。当量化参数较大时, 系数无需精确表示, 可用  $\cos\left(\frac{2u\pi}{16}\right)$  来替代式 (14) 的第 1 个和第 2 个余弦项; 用  $\cos\left(\frac{6u\pi}{16}\right)$  来替代第 3 个和第 4 个余弦项, 可得

$$F(u) = \frac{1}{2} \alpha(u) \times \sum_{x=0}^7 f(x) \cos\left(\frac{2 \times (2 \times [x/2] + 1)u\pi}{16}\right) \quad (16)$$

式中,  $[\cdot]$  表示小于、等于  $x$  的最大整数。

## 3 实验结果

在 PIII 500 PC 机上, 对 foreman 测试序列进行了模拟, 测试序列以 128kbps 的速率编码, 帧速率为 30 帧/s, 通过代码转换机转换成 32kbps, 帧速率为 15 帧/s, 并进行自适应快速编码。

采用全域全搜索运动估值模式、输入运动矢量的重复使用模式和自适应运动矢量估值模式模拟获得图象的峰值信噪比 (PSNR) 性能如图 2 所示, 其中采用自适应运动矢量估值模式获得的图象质量与采

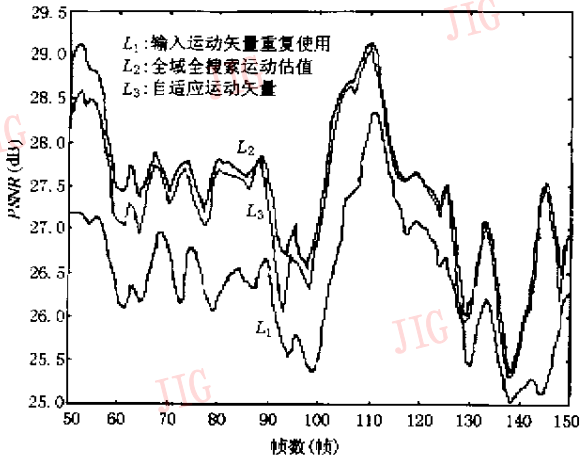


图 2 foreman 测试序列不同模式的 PSNR 比较

用全域全搜索运动估值模式所获得的图象质量基本一样,由于自适应运动矢量估值模式中,增量运动矢量的搜索域为  $\pm 4$  个整象素,因而与全域全搜索运动估值模式的计算复杂度相比是很小的,3 种模式运动估值所需时间如表 1 所示。

表 1 3 种模式运动估值所需时间比较

运动估值模式	运行时间(s)
$L_1$	1.6
$L_2$	3.9
$L_3$	2.2

同时,由于在 DCT、IDCT、量化、反量化模块中采用了自适应快速算法,其峰值信噪比虽有略微下降,但编码时间却明显减少了。采用自适应快速编码方法与 H.263 标准的 TMN8 方法编码所需时间如表 2 所示。

表 2 不同编码方法所需时间比较

编码方法	运行时间(s)
TMN8	7.4
自适应快速编码	6.3

## 4 结 论

对于有帧速率转换的视频代码转换,采用多种量化误差总和的自适应模式,可减小搜索域和计算的复杂性,同时,采用自适应快速编码方法,可加快 DCT 计算,从而提高了编码速度。模拟实验表明:采用该模式进行视频代码转换可以使视频图象质量改善到和采用全域全搜索运动估值模式获得的图象相似的程度,而计算复杂度却大大减少了。

## 参 考 文 献

- 1 Video coding for low bitrate communication. ITU-T Recommendation H.263, 1995.
- 2 Bjork, Christopoulos. Transcoder architectures for video coding. IEEE Trans. On CE, 1998 44(1) 8898.
- 3 Jou Jer Min, Chen Pei-Yin, Sun Jian-Ming. The gray prediction search algorithm for block motion estimation. IEEE Trans. On Circuits Syst. Video Technol., 1999 9(6) 843848.
- 4 Pao I M, Sun M T. Approximation of calculations for forward discrete cosine transform. IEEE Trans. On Circuits Syst. Video Technol., 1998 8(2) 264268.
- 5 Video codecs for audiovisual services at p x 64 kb/s. ITU-T Recommendation H.261, 1993.

李晓辉 1961 年生,硕士,副教授,安徽大学电子工程与信息科学系主任,主要从事图象与视频编码、多媒体通信等领域的研究工作,发表论文 40 余篇。