

一种有效的视频信号编码块效应消减技术

何润中 王仁华 郭武

(中国科学技术大学 电子工程与信息科学系, 合肥 230027)

摘要 针对视频信号编码中存在的块效应,提出了一个能有效提高解码图象质量的编码方案。首先,给出一种按系数自适应 DCT 编码算法的新的逻辑实现结构;然后,用一种简单而有效的自适应控制策略来代替极其复杂费时的严格的自适应方法。同时,为了克服低比特率下粗糙量化导致自适应性能下降的缺点,设计了一种具有良好性能的运动估计快速算法用来提高适应能力。实验结果表明,与常规的基于 DCT 的编码方法相比,这种编码方案性能有显著的提高。

关键词 块效应,自适应,运动估计

1 引言

块效应(Blocking effect)又称为块噪声(Blocking noise),是分块变换编码最主要的缺点,是对低比特率视频编码性能的最主要的限制。

现有的视频编码标准均采用 DCT 变换来去除图象数据中的空间冗余,把图象分为许多互相不重叠的大小相同的块(通常为 8×8 或者 16×16),然后对各块独立地进行 DCT 变换、量化、编码以及传输。块效应的产生就是由于各个图象块的独立处理以及量化操作对 DCT 变换的高频系数的损害。

块效应在帧内编码时较明显,而在采用运动补偿技术的帧间编码中虽然不那么严重,但也是不容忽视的,特别是在以低比特率编码时。因此消减块效应是提高图象和视频编码性能的一个重要方面。如何提高 DCT 编码的性能,尤其是采用自适应的方法,成为近年来视频信号处理研究的热点。

2 块效应消减技术的发展

消减图象编码中的块效应有很多方法,这些方法按照原理大体上可以分成三类:滤波、块重叠以及自适应的方法。

滤波方法,采用这种方法不需要改变编码端的处理过程,只是在接收端用一个低通滤波器对块边界处的象素进行平滑。这种方法不会增加编码输出的比特率,但是会使得那些跨越块边界的信号变模糊,因为低通滤波器的低通特性,使得一些高频信息受到损失。H. S. Malvar 提出一种可以避免边界变模糊的滤波方法,其做法是在发送端也增加一个滤波器^[1]。

块重叠方法,采用这种方法相邻的图象块轻轻地重叠,这样就能保证传送足够的关于块边界处图象内容的信息。在接收端则对相邻块重叠区域的象素重建值进行平均。这种方法的最大缺点是增加了需要处理的采样点的总数目,从而增大了输出的比特率。近年来出现了一种叫 LOT 变换(Lapped Orthogonal Transform)^[2]的方法,它能消减块效应,同时可避免编码比特率增加。

自适应的方法,如变块尺寸变换编码、交替 DCT/DST 变换编码、自适应量化等等。在交替 DCT/DST 变换编码^[3]中,图象被分为所谓的黑块和白块。一个黑块周围有 4 个白块,一个白块周围同样有 4 个黑块。首先对所有的黑块进行 DCT 变换、量化和编码。对于白块,从它周围的 4 个黑块的解码结果中提取一个像素宽的边界,利用某种模型如 AR 模型,求出边界响应,实际编码和传送的是白块

与边界响应的差的 DST 变换系数。由于大多数图象都可近似为一阶马尔可夫场,这种方法能在一定程度上消减块效应。

3. 块效应消减技术在视频编码中的应用

上面介绍的消减块效应的技术各有缺点,大多是针对静图象编码的,要么不适合或者很难应用到视频信号编码中去,或者会使视频编码的复杂度大大提高。

例如以 LOT 变换代替 DCT 变换,它把图象划分为相互重叠的块,而后进行变换编码。这样做在静图象编码中没有问题,但在视频序列编码中却不容易应用。因为视频编码涉及帧内和帧间两种模式,并不是所有的图象块都被编码和传送,一个传送块的某个或某几个相邻块可能并没有传送,或者相邻的块的编码模式可能不一样。所有这些都使得 LOT 变换不容易应用到视频编码中去。上节提到的交替 DCT/DST 变换编码同样面临上述 LOT 变换面临的问题,而且它的编码顺序为先黑块后白块,与现存标准采用的分层编码结构不一样。另外它的编码过程较复杂。

Hirohisa Yamaguchi 提出一种被称作按系数自适应 DCT 编码(Coefficient-by-coefficient adaptive DCT coding)的算法^[4]。这种算法能有效地提高编码的性能,消减块效应,同时不会造成解码图象块边界模糊和输出比特流增大。本文提出了一种新的逻辑结构框图,新的结构克服了原有结构的缺点,能使自适应的优越性充分发挥。同时考虑到实时实现的需要,提出了一种简单但有效的自适应控制策略—VBADCT 以代替复杂费时的严格的自适应。针对低比特率下由于粗糙量化导致的自适应性能下降,还提出了一种改进的运动估计快速算法,提高运动估计的精度,从而提高自适应的性能。

4 按系数自适应 DCT 编码算法的新实现方案

按系数自适应 DCT 编码算法基于这样一个现象:考察图象块和它的匹配块的对应的 DCT 变换系数之间的相关性,可发现低阶 DCT 系数间的相关性较像素域里的相关性要强,而那些高阶的系数间的相关性较像素域里的弱。由此可以想到,存在一

个确定的界限,划分出一个 DCT 系数的子集,在该子集内,对 DCT 系数进行帧间编码能够有效地去除相关性,即在系数域里求差值,对差值进行量化和编码;而在该区域之外,对系数进行帧内编码更有效,即直接对系数进行量化和编码。这就是按系数自适应 DCT 编码的基本原理。

令 X 表示当前编码的源图象块, Y 为 X 的预测(参考块), Q 和 Q^{-1} 分别表示量化和逆量化过程,Adaptive 表示自适应过程。则常规的基于 DCT 的视频信号编码方法如 H. 261^[5]以式(1)作为 X 的重建结果:

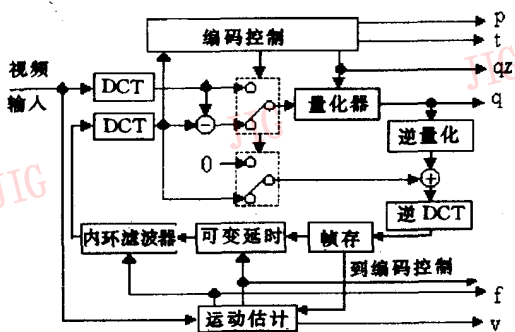
$$Rec(X) = IDCT\{Q^{-1}[Q(DCT(X - Y))]\} + Y \tag{1}$$

而文献[4]中的编码器以式(2)的结果作为 X 的重建结果:

$$Rec(X) = IDCT\{Q^{-1}[Adaptive(DCT(X), DCT(Y))]\} + Y \tag{2}$$

显然,从流程上看,它的解码过程不是编码的逆过程,它用 Adaptive(DCT(X), DCT(Y))代替了常规编码器的帧差信号的变换 $DCT(X - Y)$ 。这是一种近似,必将引入误差,影响到按系数自适应 DCT 编码算法的优越性的充分发挥。

因此,本文设计了一种新的逻辑结构,如图 1 所示。



p -帧内/帧间标识; t -传送/非传送标识; q -量化索引;
 qz -量化特性; v -运动向量; f -内环滤波器/开关信号

图 1 本文编码方案的逻辑实现框图

在新的逻辑结构中,解码过程与编码过程从流程上看是互为可逆的,如式(3)所示:

$$Rec(X) = IDCT\{Adaptive^{-1}[Q^{-1}(Q(Adaptive(DCT(X), DCT(Y))))]\} \tag{3}$$

从而克服了文献[4]中逻辑结构的缺点。

严格的自适应要根据帧内自相关系数、运动向

量、参考块 DCT 系数幅度大小以及参考块交流能量的大小等因素的不同量化值,计算出对应的自适应帧内/帧间编码的子集划分。这是一项十分复杂的工作,而且产生的自适应表十分庞大,控制过程复杂费时,无法实时实现。因此,在对大量的自适应子集划分进行观察的基础上,设计了一种简单的变界限的自适应控制策略—VBADCT。其主要思想是对不同的运动向量,选用不同的固定界限来划分自适应帧内/帧间编码子集。设计遵循下列原则:(1)对具相同运动向量的块,采用相同的自适应帧内/帧间编码子集划分;(2)存在较大运动的情况下,由于源块和参考块对应的 DCT 像素间相关性较小,相应进行帧间 DCT 编码的系数子集较小。(3)子集划分采用

简单的划分方式,对低阶系数进行帧间编码,高阶系数进行帧内编码。

表 1 给出了图 1 所示编码方案采用 VBADCT 控制策略时,与 H. 261 编码器的性能比较。其中,PSNR 为编码的平均峰值信噪比, \bar{Q} 反映平均量化步长大小(Q =量化步长/2)。可以看出,在较高比特率下(高于 300kbits/s),性能较 H. 261 有显著的提高,PSNR 提高超过 1.10dB;而在低比特率(如 64k bits/s),性能改善不明显。出现这种现象的原因在于量化,速率控制策略实际上是自适应地调整量化步长的过程,以低比特率编码时,平均量化步长加大,粗糙的量化阻碍了自适应性能的发挥。

表 1 本文方案与 H. 261 的性能比较(100 帧 QCIF 格式 Miss America 序列)运动估计采用全搜索算法, $-8 \leq V_x, V_y \leq +8$

比特率(bit/s)	1.2M		600k		300k		64k	
	H. 261	本文方案	H. 261	本文方案	H. 261	本文方案	H. 261	本文方案
编码方案	H. 261	本文方案	H. 261	本文方案	H. 261	本文方案	H. 261	本文方案
PSNR(dB)	44.760	45.888	41.851	42.989	40.776	41.064	35.458	35.458
\bar{Q}	3.256 7	2.906 7	6.406 7	5.233 3	9.196 7	7.573 3	15.936 7	15.993 3
实际输出比特流大小(bytes)	479 257	478 576	240 458	239 841	120 416	120 450	26 686	26 687

5 MCSA-改进的运动估计快速算法

按系数自适应 DCT 编码算法的性能跟运动向量的精度密切相关。运动向量的精度越高,越能代表实际的运动,依据它进行自适应帧内/帧间编码的子集划分就能更有效地去除相关。

传统的运动估计快速算法,如 TDL (Two-Dimensional Logarithmic) 算法、TSS (Three-Step Search) 算法、CDS (Conjugate Direction Search) 算法、OSA (Orthogonal Search Algorithm) 算法、CSA (Cross-Search Algorithm) 算法^[6]等,都是从当前块在参考帧中的对应位置开始,在一定的搜索窗口内,分步骤地进行一种局部最优的搜索。每步搜索都是在一些有选择性的位置上进行,搜索结果都是局部最优的,都有可能产生错误的搜索方向。我们不妨设想:如果能够从一个比较接近于实际运动的位置开始搜索,这样搜索结果成为全局最优或接近于全局最优的可能性就会大大增大,而且搜索可以只在一个较小的搜索窗口内进行。这样既能减少块匹配操作的次数,从整体上提高运动估计的速度,而且可以降低错误判决的概率,从而提高了运动估计结果的可靠性。考虑到序列中的运动相关性,即一个块的运

动总是与它相邻块的运动相近或相同,在进行搜索之前,参考已完成运动估计的相邻块(正左方、正上方)的运动向量,找出一个合适的位置作为搜索的起点,就可以达到这一目的。

我们把图象块分成 3 类:(1)较稳定的块 Cs(2)存在轻微运动的块 Cm(3)存在较大运动的块 Cr。对于 Cs 类的图象块,由于运动较小,可以只在一个较小的窗口内搜索;对于 Cr 类图象块,运动较剧烈,须采用较大的搜索窗口;而对于 Cm 类图象块,搜索窗口的大小介于上述 2 种情况之间。这样不仅能够提高预测的速度,而且可以在一定程度上提高预测的精度。

传统的运动估计块匹配函数只关心产生最小的预测残差,从而使得分配给预测残差的比特数最小。而运动补偿的真正目的是使得编码的比特流总体上最小。因此,好的编码系统应该综合考虑预测残差和作为边信息传送的运动向量二者比特率分配的联合优化。对块匹配函数修正如下:

$$\phi(i, j) = MAE(i, j) + k \cdot R_d(i, j) \quad (4)$$

其中, (i, j) 表示搜索过程中候选的运动向量, MAE (i, j) 表示在最小绝对值误差准则下的块匹配函数值。上式的第二项即修正项。 k 为一个常系数。

$R_c(i, j)$ 代表对候选的运动向量 (i, j) 进行编码所要占用的比特数。这一个准则综合考虑了对预测残差和运动向量编码占用的比特数,因而能比MAE达到更好的效果。

这种观点实际上不难理解。在对预测残差进行量化后,很多残差信号的量化索引很小,经常有不少块编码只产生很少的比特数,而其运动向量编码却可能需要多达11个比特。显然,这样编码产生的比特流不能算是最优的。

以上的讨论可以应用到各种快速搜索算法,以提高其性能。下面对改进的CSA算法—MCSA简单描述如下:

(1)比较当前宏块与参考帧中相同位置的宏块,求出块匹配函数值 $\phi(0, 0)$ 。如果 $\phi(0, 0) < T_1$,认为当前宏块没有运动,搜索结束;否则,进行下一步。

(2)初始化最小块失真 $\phi = \phi(0, 0)$, $(V_x, V_y) = (0, 0)$ 。如果 $\phi(0, 0) < T_s$,搜索步长 $p = 1$;否则,如果 $\phi(0, 0) < T_M$, $p = \frac{w}{4}$;其它, $p = \frac{w}{2}$ 。 w 为允许的最大运动偏移。

(3)选取一个合适的接近于实际运动的位置作为搜索的起点。令当前块正左边和正上方的两个图象块的运动向量分别为 (V_x^l, V_y^l) 和 (V_x^u, V_y^u) ,则:

计算 $\phi(V_x^l, V_y^l)$,如果 $\phi(V_x^l, V_y^l) < \phi$ 则置 $\phi = \phi(V_x^l, V_y^l)$, $(V_x, V_y) = (V_x^l, V_y^l)$, $p = 1$ 。如果 $\phi < T_2$,搜索结束;否则:

计算 $\phi(V_x^u, V_y^u)$,如果 $\phi(V_x^u, V_y^u) < \phi$,则置 $\phi = \phi(V_x^u, V_y^u)$, $(V_x, V_y) = (V_x^u, V_y^u)$, $p = 1$ 。如果 $\phi < T_2$,搜索结束;否则,进行下一步。

(4)以 (V_x, V_y) 为起始位置, p 为步长,按照标准CSA算法进行搜索。

门限 T_1, T_2 一方面可以限制不必要的搜索次数,另一方面也体现了预测残差和运动向量编码比特率联合优化的思想。

实验采用不同类型的序列,以检验改进后的算法性能。它们是Miss America序列(100帧, QCIF格式)和Table-Tennis序列(21帧, QCIF格式)。其中后者由7个连续帧重复组成,某些帧间存在较大的运动。实验在一台486DX/66 PC机上进行, MCSA的参数取值如下: $k = 2.5, T_1 = T_2 = 256, w = 8$ ($-8 \leq V_x, V_y \leq +8$)。

表2和图2给出了Miss America序列以64kbts/s的比特率编码时,采用不同运动估计算法的实验结果。可以看出, MCSA在性能上的改善是

很显著的。CSA的性能比全搜索算法(FSA)略差,而MCSA的性能不但使得平均峰值信噪比PSNR较

表2 本文编码方案采用不同运动估计算法时的性能
Miss America序列, 100帧, QCIF格式, 64K bits/s

搜索算法	PSNR(dB)	\bar{Q}	实际输出码流大小(Bytes)	运动估计耗时(s)
FSA	35.458	15.993 3	26 687	148
CSA	35.358	16.083 3	26 685	20
MCSA	36.232	13.883 3	26 633	17

CSA提高了0.874dB,而且较全搜索也提高了0.774dB,速度亦有所提高,而编码产生的比特流更小。同时,可以看出,采用MCSA后性能较H.261采用FSA时性能提高了0.774dB。

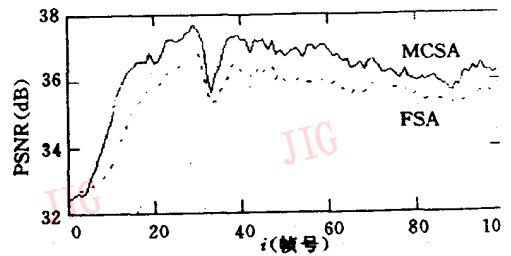


图2 MCSA和FSA的性能比较
(Miss America序列, 64k bits/s)

表3给出了目标速率为600kbts/s时Table-Tennis序列的实验结果。对于Table-Tennis序列, MCSA带来的性能提高没有那么突出,但仍然较CSA有0.413dB的提高,接近全搜索的性能。

表3 本文编码方案采用不同运动估计算法时的性能
Table-Tennis序列, 21帧, QCIF格式, 64k bits/s

搜索算法	PSNR(dB)	\bar{Q}	实际输出码流大小(Bytes)	运动估计耗时(s)
FSA	29.454	14.222 2	44 916	29
CSA	29.021	14.365 0	45 016	4
MCSA	29.434	14.222 2	44 857	4

6 结论

困扰低比特率视频信号编码的一个问题是解码

图象的质量问题。本文介绍了一种采用按系数自适应 DCT 编码算法的新的逻辑结构,克服了原有结构的缺点。还提出一种简单有效的自适应控制策略——VBADCT 代替极其复杂耗时的严格的自适应控制。实验结果表明,与常规的基于 DCT 的编码算法相比,尤其是在较高比特率下,这种方案能有效地提高解码图象质量。另外还提出了一种性能优良的运动估计快速算法——MCSA。MCSA 能有效提高编码的性能,特别是使新编码方案在低比特率下的自适应能力得以充分发挥,性能较常规编码方法提高较显著。

参考文献

- 1 Malvar H S. A pre- and post-filtering technique for the reduction of blocking effects. Picture Coding Symp., Stockholm, Sweden, 1987. 6. 2 Malvar H S, Staelin H. The LOT; Transform cod-

- ing without blocking effects. IEEE Trans. ASSP, 1989, 37(4).
- 3 Rose K, Heiman A. DCT/DST alternate-transform image coding. IEEE Trans. Commun., 1990, 38(1).
- 4 Hirohisa Yamaguchi, Adaptive DCT coding of video signals. IEEE Trans. Commun., 1993, 4:1534~1543.
- 5 ITU-T Recommendation H. 261—Video codec for audiovisual services at p×6 kbits, ITU-T draft H. 261, 03/93.
- 6 Ghanbari M. The cross-search algorithm for motion estimation. IEEE Trans. Commun., 1990, 38(7).



何润中 中国科技大学电子工程与信息科学系硕士研究生,主要研究方向为多媒体通信,图象编码。



王仁华 中国科学技术大学电子工程与信息科学系教授,博士生导师,中国通信学会理事、会士。多年从事语音通信、多媒体通信研究。



郭武 中国科学技术大学电子工程与信息科学系硕士研究生,研究方向是多媒体通信技术及视频编码。

An Efficient Approach to Reduce the Blocking Effect of Video Signal Coding

He Runzhong, Wang Renhua, Guo wu

(Dept of Elec Eng&Info su Univ of sci&Tech of China, Hefei 230027)

Abstract In this article, an efficient video encoding approach is proposed to reduce the blocking effect. It employs a new logical realization structure of coefficient-by-coefficient adaptive DCT coding algorithm. Meanwhile, a simple but efficient adaptive controlling strategy is proposed to replace the exactly adaptation which is very complex. Also, to overcome the shortcoming that the performance of the adaptation decreases at low bit rates, an excellent fast algorithm for motion estimation which can efficiently improve the adaptive ability is proposed. The experimental results indicate that the performance of the new encoding approach has a marked improvement over the conventional DCT-based codec.

Keywords Blockng effect, Adaptive, Motion estimation