

基于 DCT 域的 MPEG-4 编码器改进方法研究

万源, 陈建业, 吴传生

(武汉理工大学理学院, 武汉 430070)

摘要:在对 MPEG-4 标准的分析和研究的基础上, 对 MPEG-4 编码器提出了一种基于离散余弦变换(DCT)域运动估计的改进方案。这种编码结构直接在 DCT 变换域中进行运动估计, 不需要将 DCT 系数变换回空间域。因此它能有效地减少编码的运算次数和编码结构的复杂度, 满足实时编码的要求。

关键词: MPEG-4; 离散余弦变换; 运动估计; 视频压缩

Research on Improved Approach of MPEG-4 Encoder Based on DCT

WAN Yuan, CHEN Jianye, WU Chuansheng

(Science Institute, Wuhan University of Technology, Wuhan 430070)

【Abstract】According to the analysis of MPEG-4, this paper presents an improved approach of MPEG-4 encoder based on DCT. Motion estimation and transform coding can be completely combined in the transform domain without the need to convert the DCT coefficients back to the spatial domain for motion estimation and compensation. So the approach can reduce the complexity of the encoder computation, meet the real-time need.

【Key words】MPEG-4; DCT; Motion estimation; Video compression

MPEG-4 视频编码标准^[2]提供了一套强大的视频编码工具。使用这些编码工具, 可以得到不同比特率的高质量的视频, 满足各种视频应用的需要。但是在实际应用中, MPEG-4 编码器存在一些问题: (1)编码过程中的数据运算量比较大, 不适合实时编码的需要; (2)编码过程中(进行运动估计和运动补偿时)需要缓存大量的图像数据, 不利于硬件实现。为了满足一些实际的应用, 本文引入了一种新的编码结构来解决这些问题, 编码后得到的比特流可以不经任何修改, 直接被符合 MPEG-4 标准的解码器进行解码。

1 MPEG4 视频编码

MPEG-4 采用了视频压缩编码标准中流行的运动补偿与变换编码相结合的混合编码方法。该编码器由 DCT 变换、量化、反量化、反 DCT 变换、运动估计与补偿和变长编码(VLC)几个部分组成。其中 DCT 变换、量化、反量化、反 DCT 变换、运动估计与补偿构成了反馈环。这种编码结构中采用的运动估计是在空间域中进行的, 称为空间域运动估计(Spatial Domain Motion Estimation, SDME), 通常采用全搜索块匹配(Block-Matching, BKM)的方法完成运动估计。但是, 这个反馈环不仅增加了整个编码器的复杂度, 而且限制了编码器的编码效率, 是实现实时编码的瓶颈。

因此, 本文引入了一种与传统的混合编码结构相似的编码结构, 反馈环中只有量化、反量化和运动估计 3 个部分, 去掉 DCT 变换和反 DCT 变换过程。与传统的混合编码结构不同, 这种新的编码方法中的运动估计和运动补偿是在 DCT 变换域中进行的, 称为变换域运动估计(Transform Domain Motion Estimation, TDME)。传统的混合编码结构中对于预测帧(P-VOP 和 B-VOP)进行编码的时候, 反馈环中的 DCT 变换是用来计算运动补偿后的预测误差的 DCT 系数。但是对于基于 DCT 变换的运动补偿, 运动补偿是直接变换域中进行的, 因此可以不需要计算这些系数。从这些分析可以看出,

新的编码器不仅降低了编码器的复杂程度, 且解决了实时编码的瓶颈问题, 不需要大量的帧缓存空间, 有利于硬件实现。

2 伪相位技术

Ut.Koc 和 Ray Liu^[3]提出了一种利用傅立叶变换的相位信息在傅立叶变换域中进行运动估计的方法——DCT伪相位技术。

设信号 $\{x_1(n), n=0, 1, \dots, N-1\}$ 为一维离散信号。对它右移 m 位($m>0$)得到信号 $\{x_2(n), n=0, 1, \dots, N-1\}$ 。假设信号 $x_1(n)$ 在定义域 D 以外的值都为 0, 因此有

$$x_2(n) = \begin{cases} x_1(n-m) & (n-m) \in D \\ 0 & (n-m) \notin D \end{cases}$$

将它的 DCT 变换和 DST(离散正弦变换)展开成下式:

$$X_2^c(k) = \frac{2}{N} C(k) \sum_0^{N-1} x_2(n) \cdot \cos\left[\frac{k\pi}{N}(n+0.5)\right] \quad k \in \{0, \dots, N-1\}$$

$$X_2^s(k) = \frac{2}{N} C(k) \sum_0^{N-1} x_2(n) \cdot \sin\left[\frac{k\pi}{N}(n+0.5)\right] \quad k \in \{0, \dots, N\}$$

$$Z_1^c(k) = \frac{2}{N} C(k) \sum_0^{N-1} x_1(n) \cdot \cos\left[\frac{k\pi}{N} \cdot n\right] \quad k \in \{0, \dots, N\}$$

$$Z_1^s(k) = \frac{2}{N} C(k) \sum_0^{N-1} x_1(n) \cdot \sin\left[\frac{k\pi}{N} \cdot n\right] \quad k \in \{0, \dots, N-1\}$$

其中

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k=0 \text{ 或 } 1 \\ 1, & k \text{ 为其它值} \end{cases}$$

Z_1^s 和 Z_1^c 分别表示信号 $x_1(n)$ 的 DST 变换(DST-1)和 DCT 变换(DCT-1), X_2^s 和 X_2^c 分别表示信号 $x_2(n)$ 的 DST 变换(DST-2)和 DCT 变换(DCT-2)。定义

基金项目:国家自然科学基金资助项目(60473081)

作者简介:万源(1976-), 女, 讲师、硕士, 主研方向: 图像处理; 陈建业, 讲师; 吴传生, 教授

收稿日期: 2006-04-24 **E-mail:** wanyuan@mail.whut.edu.cn

$$g_m^s(k) \stackrel{\Delta}{=} \sin\left[\frac{k\pi}{N}\left(m + \frac{1}{2}\right)\right] \text{ 和 } g_m^c(k) \stackrel{\Delta}{=} \cos\left[\frac{k\pi}{N}\left(m + \frac{1}{2}\right)\right]$$

为伪相位。

3 基于 DCT 变换的运动估计算法

我们可以将文献[3]介绍的通过伪相位技术求一维信号的位移信息的方法推广到二维空间,由此得到在 DCT 域中求物体的二维平移运动的方法。图 1 是这种在 DCT 变换域中进行运动估计的算法的流程图,图 2 是算法实现的结构图。

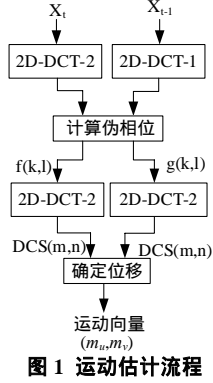


图 1 运动估计流程

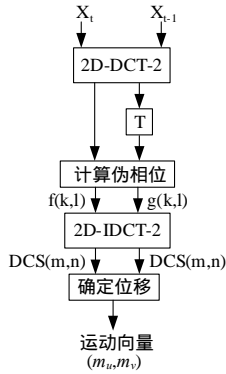


图 2 运动估计实现结构

3.1 二维 DCT 变换和伪相位的计算

首先将前一帧图像 X_{t-1} 和当前帧 X_t 分别进行二维 DCT-2 (2D-DCT-2) 和二维 DCT-1 (2D-DCT-1) 变换。二维 DCT-2 变换计算 4 个系数,每个系数都是由 DCT-2/DST-2 变换核定义的一个二维可分离的函数,定义如下:

$$\begin{aligned} X_t^{cc}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_t(m,n) \cos\left[\frac{k\pi}{N}\left(m + 0.5\right)\right] \cos\left[\frac{l\pi}{N}\left(n + 0.5\right)\right] \\ X_t^{cs}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_t(m,n) \cos\left[\frac{k\pi}{N}\left(m + 0.5\right)\right] \sin\left[\frac{l\pi}{N}\left(n + 0.5\right)\right] \\ X_t^{sc}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_t(m,n) \sin\left[\frac{k\pi}{N}\left(m + 0.5\right)\right] \cos\left[\frac{l\pi}{N}\left(n + 0.5\right)\right] \\ X_t^{ss}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_t(m,n) \sin\left[\frac{k\pi}{N}\left(m + 0.5\right)\right] \sin\left[\frac{l\pi}{N}\left(n + 0.5\right)\right] \end{aligned}$$

同样,二维 DCT-1 是由 DCT-1/DST-1 的变换核定义的两个函数:

$$\begin{aligned} Z_{t-1}^{cc}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_{t-1}(m,n) \cos\left(\frac{k\pi}{N} \cdot m\right) \cos\left(\frac{l\pi}{N} \cdot n\right) \\ Z_{t-1}^{cs}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_{t-1}(m,n) \cos\left(\frac{k\pi}{N} \cdot m\right) \sin\left(\frac{l\pi}{N} \cdot n\right) \\ Z_{t-1}^{sc}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_{t-1}(m,n) \sin\left(\frac{k\pi}{N} \cdot m\right) \cos\left(\frac{l\pi}{N} \cdot n\right) \\ Z_{t-1}^{ss}(k,l) &= \frac{4}{N^2} C(k)C(l) \sum_{m,n=0}^{N-1} x_{t-1}(m,n) \sin\left(\frac{k\pi}{N} \cdot m\right) \sin\left(\frac{l\pi}{N} \cdot n\right) \end{aligned}$$

与一维情况相似,假定运动只有平移运动,设 m_u 为沿 X 轴正方向的平移值, m_v 为沿 Y 轴正方向的平移值。在二维空

间中,一个物体可能向 4 个方向运动,东北 ($m_u > 0, m_v > 0$), 西北 ($m_u < 0, m_v > 0$), 东南 ($m_u > 0, m_v < 0$) 和西南 ($m_u < 0, m_v < 0$)。运用 Ut.Koc 和 Ray Liu 提出的利用傅立叶变换的相位信息在傅立叶变换域中进行运动估计的方法——DCT 伪相位技术^[3]求得两个伪相位函数:

$$\begin{aligned} f_{m_u m_v}(k,l) &= \begin{cases} g_{n_u n_v}^{cc}(k,l), k,l \in \{1, \dots, N-1\} \\ \frac{1}{\sqrt{2}} \cdot \frac{Z_{t-1}^{cc}(k,l)X_t^{cs}(k,l) - Z_{t-1}^{cs}(k,l)X_t^{cc}(k,l)}{[Z_{t-1}^{cs}(k,l)]^2 + [Z_t^{cc}(k,l)]^2}, k=0, l \in \{1, \dots, N-1\} \\ \frac{1}{\sqrt{2}} \cdot \frac{Z_{t-1}^{cc}(k,l)X_t^{cs}(k,l) + Z_{t-1}^{sc}(k,l)X_t^{ss}(k,l)}{[Z_{t-1}^{cc}(k,l)]^2 + [Z_t^{sc}(k,l)]^2}, l=N, k \in \{1, \dots, N-1\} \\ \frac{1}{2} \cdot \frac{X_t^{cc}(k,l)}{Z_{t-1}^{cc}(k,l)}, k=0, l=N \end{cases} \\ g_{m_u m_v}(k,l) &= \begin{cases} g_{n_u n_v}^{sc}(k,l), k,l \in \{1, \dots, N-1\} \\ \frac{1}{\sqrt{2}} \cdot \frac{Z_{t-1}^{sc}(k,l)X_t^{cs}(k,l) - Z_{t-1}^{cs}(k,l)X_t^{sc}(k,l)}{[Z_{t-1}^{cs}(k,l)]^2 + [Z_t^{sc}(k,l)]^2}, k=0, l \in \{1, \dots, N-1\} \\ \frac{1}{\sqrt{2}} \cdot \frac{Z_{t-1}^{sc}(k,l)X_t^{cs}(k,l) + Z_{t-1}^{cs}(k,l)X_t^{ss}(k,l)}{[Z_{t-1}^{sc}(k,l)]^2 + [Z_t^{cs}(k,l)]^2}, l=N, k \in \{1, \dots, N-1\} \\ \frac{1}{2} \cdot \frac{X_t^{sc}(k,l)}{Z_{t-1}^{sc}(k,l)}, k=0, l=N \end{cases} \end{aligned}$$

3.2 位移的计算

在计算 $f_{m_u m_v}(k,l)$ 和 $g_{m_u m_v}(k,l)$ 时,如果绝对值大于 1,则该值属于一个病态情况。对于这种值应舍弃,一般将病态值赋为零。对此计算得到的这两个伪相位进行二维离散余弦变换运算得到两个函数 DCS 和 DSC。然后利用 DCT-2 和 DST-2 的正交原理,有

$$\begin{aligned} DCS(m,n) &= \frac{4}{N^2} \sum_{k=0}^{N-1} \sum_{l=1}^N C(k)C(l) f_{m_u m_v}(k,l) \cos\left(\frac{k\pi}{N}\left(m + \frac{1}{2}\right)\right) \sin\left(\frac{l\pi}{N}\left(n + \frac{1}{2}\right)\right) \\ &= [\delta(m - m_u) + \delta(m + m_u + 1)] \cdot [\delta(n - m_v) - \delta(n + m_v + 1)] \\ DSC(m,n) &= \frac{4}{N^2} \sum_{k=0}^{N-1} \sum_{l=1}^N C(k)C(l) g_{m_u m_v}(k,l) \cos\left(\frac{k\pi}{N}\left(m + \frac{1}{2}\right)\right) \sin\left(\frac{l\pi}{N}\left(n + \frac{1}{2}\right)\right) \\ &= [\delta(m - m_u) - \delta(m + m_u + 1)] \cdot [\delta(n - m_v) + \delta(n + m_v + 1)] \end{aligned}$$

如果 m_u 为正,则

$$DSC(m,n) = \delta(m - m_u) \cdot [\delta(n - m_v) + \delta(n + m_v + 1)]$$

无论 m_v 为正还是为负, $DSC(m,n)$ 都为正。

如果 m_u 为负,则

$$DSC(m,n) = -\delta(m + m_u + 1) \cdot [\delta(n - m_v) + \delta(n + m_v + 1)]$$

无论 m_v 为正还是为负, $DSC(m,n)$ 都为负,因此 DSC 的最大值的正负决定 m_u 的正负。同理, DCS 的正负决定 m_v 的正负。表 1 列出了如何根据 DCS 和 DSC 的最大值的正负来判断运动的方向。

表 1 位移运动方向

DSC 峰值的符号	DCS 峰值的符号	峰值的位置	运动方向
+	+	(m_u, m_v)	东北
+	-	($m_u, -(m_v+1)$)	东南
-	+	($-(m_u+1), m_v$)	西北
-	-	($-(m_u+1), -(m_v+1)$)	西南

最后用下面的等式估计运动向量 $\hat{d} = (\hat{m}_u, \hat{m}_v)$:

$$\begin{aligned} \hat{m}_u &= \begin{cases} i_{DSC} = i_{DCS}, & \text{if } DSC(i_{DSC}, j_{DSC}) > 0 \\ -(i_{DSC} + 1) = -(i_{DCS} + 1), & \text{if } DSC(i_{DSC}, j_{DSC}) < 0 \end{cases} \\ \hat{m}_v &= \begin{cases} j_{DSC} = j_{DCS}, & \text{if } DSC(i_{DCS}, j_{DCS}) > 0 \\ -(j_{DCS} + 1) = -(j_{DSC} + 1), & \text{if } DSC(i_{DCS}, j_{DCS}) < 0 \end{cases} \end{aligned}$$

其中, (i_{DCS}, j_{DCS}) 是 $DCS(m,n)$ 最大值的位置, (i_{DSC}, j_{DSC}) 是 $DSC(m,n)$ 最大值的位置。一般来说这两个最大值的位置是相

同的,但是在有噪声的情况下,它们的值可能不同。在这种情况下选择非峰值与峰值比(Non Peak to Peak Ratio, NPR)小的那个为运动向量。判决函数如下:

$$(i_D, j_D) = \begin{cases} (i_{DSC}, j_{DSC}) & \text{if } NPR(DSC) < NPR(DCS) \\ (i_{DCS}, j_{DCS}) & \text{if } NPR(DSC) > NPR(DCS) \end{cases}$$

4 运算复杂度分析与实验结果

从图1可以看出,基于DCT域的运动估计算法需要在标准的DCT变换(2D-DCT-2)的基础上加入单独的2D-DCT-1变换。这样会增加程序实现的复杂度。通过分析2D-DCT-1变换和2D-DCT-2变换的系数,我们发现它们之间有如下关系:

$$Z_{t-1}^{cc}(k,l) = \left[+\cos\frac{k\pi}{2N}\cos\frac{l\pi}{2N} + \cos\frac{k\pi}{2N}\cos\frac{l\pi}{2N} + \cos\frac{k\pi}{2N}\cos\frac{l\pi}{2N} + \cos\frac{k\pi}{2N}\cos\frac{l\pi}{2N} \right] \times X_{t-1}^{cc}(k,l)$$

对 $Z_{t-1}^{cs}(k,l)$, $Z_{t-1}^{sc}(k,l)$, $Z_{t-1}^{ss}(k,l)$ 同理可得。

其中, $Z_{t-1}^{cc}(k,l)$, $Z_{t-1}^{cs}(k,l)$, $Z_{t-1}^{sc}(k,l)$, $Z_{t-1}^{ss}(k,l)$ 是前一帧的2D-DCT-1系数, $X_{t-1}^{cc}(k,l)$, $X_{t-1}^{cs}(k,l)$, $X_{t-1}^{sc}(k,l)$, $X_{t-1}^{ss}(k,l)$ 是前一帧的2D-DCT-2系数。因此,如图2所示,可以增加一个系数变换矩阵T,将2D-DCT-2系数转换为2D-DCT-1系数。这样,系统实现中就不需要实现2D-DCT-1变换。

伪相位计算和峰值搜索运算的复杂度为 $O(N^2)$,如果DCT计算的复杂度是 O_{dct} ,整个基于DCT变换的运动估计算法的复杂度 $O(N^2)+O_{dct}^{[4]}$ 。在计算伪相位函数f和g时,可以在计算二维系数的时候得到这些系数。对于二维DCT系数的计算,可以采用lattice结构的实现,这样实现的DCT变换的运算复杂度为 $O(N)^{[5]}$ 。总的来说,基于DCT域的运动估计算法的复杂度是 $O(N^2)$,而全搜索的块匹配算法的运算的复杂度是 $O(N^4)$ 。



图4 DCT 运动估计与补偿后的图像

下面对基于DCT变化的运动估计算法和传统空间域运动

估计算法的性能进行比较。实验中,采用码流为128K的标准MPEG4码流,25fps。从第2帧起,每一帧都以前一帧作为参考图像进行运动估计与运动补偿。现行的视频编码标准中,都采用对图像的Y分量进行运动估计,而U和V分量的运动向量直接由Y分量得到的运动向量进行相应的处理得到。因此视频图像的Y分量分别采用传统的块匹配算法和基于DCT变换的算法进行运动估计与补偿。在块匹配算法中,采用了运算速度较快的三步法。在两种算法中运动补偿采用的都是空间域运动补偿算法,如果采用了基于DCT域的运动补偿^[4],则可以进一步提高编码器的速度。图3是对第2帧采用三步搜索算法进行运动估计与补偿以后重建的图像,图4是对第2帧采用基于DCT变换运动估计算法进行运动估计与补偿以后重建的图像。从图3和图4的主观质量比较中,可以看出这两种算法进行运动估计与补偿后的效果没有大的区别。对于人的视觉上来说,都是可以接受的。

5 结束语

本文对MPEG-4编码器提出了一种基于DCT域运动估计的改进方案。这种编码结构直接在DCT变换域中进行运动估计,不需要将DCT系数换回空间域,因此它能有效地减少编码的运算次数和编码结构的复杂度,解决了实时编码的瓶颈问题,而且不需要大量的帧缓存空间,有利于硬件实现。虽然编码器采用了基于变换域的运动估计算法,但是它适合于所有基于DCT变换和空间域运动补偿相结合的混合视频编码标准,如H.26X系列和MPEG系列标准。编码后的比特流可以被符合这些标准的解码器直接进行解码。

与传统的空间域运动估计和补偿的方法相比较,这种编码结构也有它的缺点。这种算法直接用原始图像而不是重建后的图像作为参考图像进行运动估计,因此编码后的图像的质量会有所下降。

参考文献

- 1 钟玉琢,王琪. 基于对象的多媒体数据压缩编码国际标准MPEG-4及其校验模型[M]. 北京: 科学出版社, 2000.
- 2 ISO/IEC JTC1/SC29/WG11 Document N3908. MPEG-4 Video Verification Model Version 18.0[S]. 2003-01.
- 3 Koc U, Liu R. DCT Based Motion Estimation[J]. IEEE Trans. on Image Processing, 1998, 7(7): 948-965.
- 4 Ling Guan. Multimedia Image and Video Processing[M]. Boca Raton: CRC Press LLC, 2003.
- 5 Chiu C T, Liu K L. Real Time Parallel and Fully Pipelined Two Dimensional DCT Lattice Structures with Applications to HDTV Systems[J]. IEEE Trans. on Circuits and Systems for Video Technology, 1999, 2(2): 25-67.

(上接第194页)

- 2 Harris B, Hunt R. TCP/IP Security Threats and Attack Methods[J]. Computer Communications, 1999, 22(10): 885-879.
- 3 Vickers B, Albuquerque C, Suda T. Source Adaptive Multi-layered Multicast Algorithms for Real-time Video Distribution[J]. IEEE/ACM

Trans. on Networking, 2000, 6(5): 720-733.

- 4 Kim T, Ammar M H. A Comparison of Layering and Stream Replication Video Multicast Schemes[C]//Proc. of Network and Operating System Support for Digital Audio and Video. 2001: 63-72.