

一种快速全局运动补偿编码方法

贺玉文, 钟玉琢, 杨士强
(清华大学计算机系, 北京 100084)

摘 要: 全局运动补偿编码是一种基于模型的编码方法, 它可以提高编码效率, 但是目前由于全局运动补偿编码中的一些关键技术(全局运动估计)的计算量比较大, 所以使得全局运动补偿编码还不能在实时应用中得到很好应用. 本文就针对全局运动补偿编码中核心技术进行研究, 提出快速全局运动补偿编码算法, 解决全局运动补偿实时编码问题. 通过和 MPEG4 中已有的全局运动补偿编码方法实验比较可以看到本文提出的快速全局运动补偿编码方法在保证编码效率的同时编码速度得到了很大提高.

关键词: 编码; 全局运动补偿; 局部运动补偿; 运动估计

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2001) 02-0175-03

A Fast Global Motion Compensation Coding Method

HE Yu-wen, ZHONG Yu-zhuo, YANG Shi-qiang

(Computer Science and Technology Department, Tsinghua University, Beijing 100084, China)

Abstract: Global motion compensation coding is a novel coding method based on models. Its coding efficiency is higher than that of conventional coding method with motion compensation, so it is fit for very low bit-rate applications. But the key technology of global motion compensation coding is not solved well at present, such as global motion estimation, so global motion compensation coding has not been applied widely in real-time applications. This paper focuses some research on key problems of global motion compensation coding to solve real-time global motion compensation coding. The new fast method of global motion estimation is proposed and is used in global motion compensation coding, thus the speed of global motion compensation coding is accelerated. Comparative experimental results are shown in the paper. The speed of global motion compensation coding is improved compared with former method of global motion compensation coding.

Key words: coding; global motion compensation; motion estimation

1 引言

全局运动补偿编码^[1]是针对背景对象编码的特点提出的. 通常情况下, 背景对象没有自身运动, 由于摄像机运动才造成了图像序列中的背景变化, 在对这样背景的视频进行编码时适合采用全局运动补偿编码方法. 它的基本原理是在编码时, 首先通过全局运动估计技术计算出相机运动模型的模式参数, 然后通过运动模型参数进行运动预测, 最后对运动参数和补偿后的残差进行编码. 它属于 sprite 编码^[1]的一种. 从信息论的观点看, 由于这些在同一场景中拍摄的视频段中每帧之间的背景图像是相关的, 而全局运动补偿编码就是利用这些相关性进行编码. 然而在 MPEG2 运动补偿方法中只是利用了相邻帧之间的局部相关性, 而没有考虑背景整体相关性, 从这一点看全局运动补偿编码要比传统运动补偿方法要优越.

由于全局运动补偿编码要通过全局运动估计得到摄像机的运动模型参数, 而全局运动估计计算量非常大, 所以很难做

到软件实时编码, 这就限制了全局运动补偿编码这种高效的编码方法在实时应用中的应用. 而视频会议和视频会议这类实时应用, 它们场景都比较固定, 所以特别适合全局运动补偿编码方法. 如果能够解决全局运动补偿的实时编码问题, 那么将大大促进全局运动补偿编码的广泛应用, 这些都极大提高带宽的有效利用^[5,6]. 本文就对全局运动补偿编码中的关键技术全局运动估计进行了深入研究, 并提出了快速全局运动估计方法, 从而大大提高了全局运动补偿编码的编码速度.

本文第二部分将详细讲述快速全局运动补偿编码方法, 第三部分利用 MPEG4 的参考软件平台校验模型 (VM) 和 MPEG4 中已有的全局运动补偿编码算法进行了比较实验, 来验证本文算法的改进性能, 第四部分对全局运动补偿实时编码进行总结, 提出今后待研究的问题.

2 快速全局运动补偿编码方法

2.1 全局运动补偿

全局运动补偿编码方法是 sprite 编码方法的一种, 其基本

原理就是基于模型的编码。在全局运动补偿编码中,图像模型就是摄像机的运动模型,而其中分析过程就是全局运动估计,分析数据就是摄像机的运动模型参数。全局运动补偿编码是首先利用全局运动估计得到和前一帧之间的全局运动参数,然后进行运动补偿编码。它的具体编解码过程如下,编码端:(1)利用全局运动估计计算得到生成待编码帧和前一帧之间的全局运动参数;(2)根据全局运动参数和前一帧已编码图像对当前待编码帧图像进行运动补偿;(3)对补偿后的残差图像进行编码。解码端:(1)对运动参数解码,得到全局运动参数;(2)解码得到运动补偿后的残差图像;(3)根据全局运动参数得到当前解码帧预测值,和解码的残差图像相加得到该帧的重构图像。

2.2 快速全局运动补偿编码方法

全局运动补偿编码是使用运动模型参数来编码视频对象的全局运动,它的核心技术就是全局运动估计和预测,预测误差的残差编码,整个编码过程见图 1。

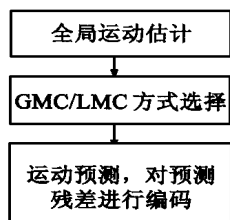


图 1 全局运动补偿编码过程

2.2.1 全局运动估计

用 $[x_i, y_i]^T$ 表示某像素点在当前图像中的位置,用 $[x_i, y_i]^T$ 表示该像素点在前一帧图像中的对应位置。它们的对应关系用六参数仿射参数模型表示:

$$\begin{cases} x_i = ax_i + by_i + c \\ y_i = dx_i + ey_i + f \end{cases} \quad (1)$$

用 $I(x, y)$ 和 $I(x', y')$ 分别表示这当前图像和根据参考帧得到的预测图像,用 $\mathbf{a} = (a, b, c, d, e, f)^T$ 表示参数矢量, $I(x, y)$ 是参数矢量的函数,则目标能量函数为: $R(\mathbf{a}) = \sum_i w_i [I(x_i, y_i) - I(x'_i, y'_i)]^2$, 全局运动估计就是求解待定参数 \mathbf{a} , 使得 $R(\mathbf{a})$ 取得最小值,从而得到参数的最佳估计。用 Gauss-Newton 方法或者 Levenberg-Marquadet 方法^[3] 可以迭代求解。假设在已有第 k 步的参数 \mathbf{a}_k , 通过目标函数 $R(\mathbf{a})$ 在该点作泰勒展开来近似 $R(\mathbf{a})$:

$$R(\mathbf{a}) \approx R(\mathbf{a}_k) + \mathbf{g}_k^T (\mathbf{a} - \mathbf{a}_k) + \frac{1}{2} (\mathbf{a} - \mathbf{a}_k)^T \mathbf{H}_k (\mathbf{a} - \mathbf{a}_k) \quad (2)$$

其中 \mathbf{g}_k 和 \mathbf{H}_k 分别是 $R(\mathbf{a})$ 在 \mathbf{a}_k 处的梯度矩阵和 Hessian 矩阵:

$$\mathbf{g}_k = \mathbf{J}_k^T \mathbf{W}_k, \quad \mathbf{H}_k = \mathbf{J}_k^T \mathbf{W}_k \mathbf{J}_k + \sum_i w_i \mathbf{H}_{ik}$$

$\mathbf{J}_k = [r_1 \ r_2 \ \dots \ r_N]^T$ 表示 \mathbf{a}_k 处的残差, $\mathbf{J}_k = \partial / \partial \mathbf{a}$, \mathbf{W} 是对角方阵且 $W_{ii} = w_i$, \mathbf{H}_{ik} 是 i 的 Hessian 矩阵,如果残差 r_i 是比较小,那么可以做如下近似: $\mathbf{H}_k \approx \mathbf{J}_k^T \mathbf{W}_k \mathbf{J}_k$, 根据方程 $\partial R(\mathbf{a}) / \partial \mathbf{a} = 0$ 由方程(2)可以得到:

$$\mathbf{J}_k^T \mathbf{W}_k \mathbf{J}_k (\mathbf{a} - \mathbf{a}_k) = - \mathbf{J}_k^T \mathbf{W}_k \mathbf{r}_k \quad (3)$$

根据方程(3)就可以得到 \mathbf{a}_k 处的增量,这样就可以求出下一步参数: $\mathbf{a}_{k+1} = \mathbf{a}_k + \Delta \mathbf{a}_k$, 这样迭代就可以逐步减小目标函数。

全局运动估计是基于密度估计(最小二乘原理)的,对外点比较敏感,而图像中既有背景全局运动也有前景物体的局部运动,这些局部运动点对于全局运动估计来说就是外点(噪

声),MPEG4 中全局运动估计算法^[1,2]中只是在迭代之前用残差直方图去除 10%残差比较大的点,在迭代过程中方程(3)中的权值矩阵为单位矩阵,不能很好地抑制外点影响,因此导致迭代收敛慢计算结果不准确。由于迭代计算中都涉及微分运算,这些算子易受噪声影响,所以在迭代计算时做了如下改进:

(1)在迭代计算时采用 $w_i = (|r_i| + 1) / (\sigma^2 + r_i^2)^2$ 加权,其中 σ 为残差的标准方差,当残差 r_i 比 σ 大许多时,权值就会很快减小,这样可以增加算法的鲁棒性,权值函数图可以见图 2。

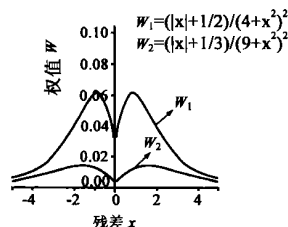


图 2 权值函数 $w_i = (|r_i| + 1) / (\sigma^2 + r_i^2)^2$, $\sigma_1 = 2.0$, $\sigma_2 = 3.0$

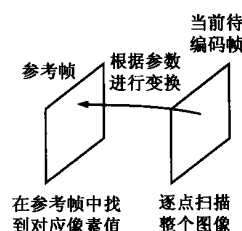


图 3 全局运动补偿编码的预测

(2)由于微分运算易受噪声影响,梯度较小时受噪声影响较大,所以在图像中梯度小的点的梯度值是不可靠的,在迭代计算中可以选取梯度较大的点进行计算,本文算法是取平均梯度 1.25 倍和 1.65 倍之间点,这是从实验统计中得出的。

(3)用三层金字塔进行快速估计,在金字塔前两层计算时采用基于直方图的方法去除外点:根据当前运动参数进行补偿后得到残差图像,然后根据残差图像直方图中 90%处得到阈值,最后就可以去除残差大于阈值的点,而由于前景区域有集中性,所以在三层金字塔计算的最后一层计算时,采用残差块的方法去除外点:首先用初始参数计算残差,然后将图像分成 16×16 的块,再计算每个块的残差值(将该块内所有点残差相加),最后就将块残差较大(1/3 左右)并且有聚集在一起(8 连通)特性的块作为外点整体去除。

2.2.2 全局运动补偿和局部运动补偿方式选择

在全局运动补偿编码中由于既有全局运动,这是由摄像机运动引起的,又有局部运动,前景物体自身运动,因此编码就要有选择性的使用全局运动补偿(GMC)和局部运动补偿(LMC),使得运动估计的残差最小,以节省码率。全局运动补偿就是使用运动参数进行预测(图 3),而局部运动补偿就是采用一般 MPEG1/MPEG2 中使用的运动预测并进行补偿,最后对残差进行编码,MPEG4^[1]中 LMC 有两种方式,一种对 16×16 宏块采用一个运动矢量进行预测,该方式用 INTER 表示,另一种对宏块采用 4 个 8×8 块的运动矢量分别对每个小块进行预测,该方式用 INTER4V 表示。下面是在对某个宏块进行编码时选取编码方式的条件^[1]:

```

If (SAD_GMC - P < SAD_LMC) Then 选择 GMC 编码模式;
Else 选择 LMC 编码模式;
End If
    
```

其中 SAD_{GMC} 表示使用 GMC 方式时这个亮度宏块和它的预测

值之间差的绝对值之和,而 SAD_{LMC} 和参数 P 的定义为:如果 LMC 采用 INTER4V 方式预测,那么:

$$SAD_{LMC} = SAD_8, \quad P = N_B \cdot Qp/16$$

SAD_8 表示 4 个 8×8 块的 SAD 之和, N_B 表示编码的像素个数, Qp 是量化参数. 如果 LMC 采用 INTER 方式预测,那么:

$$SAD_{LMC} = SAD_{16}, \quad P = N_B \cdot Qp/64$$

如果 LMC 采用 INTER 方式预测,而且运动矢量是 $(0,0)$ 那么:

$$SAD_{LMC} = SAD_{16} + (N_B/2 + 1), \quad P = N_B/2 + 1$$

3 实验结果

为了测试本文所提出的全局运动补偿算法的性能,用 MPEG4 标准测试序列进行了实验. 测试序列和测试条件见表 1, 这些实验都是在码率控制中进行的, 其中低码率控制采用 VM5 方法^[1], 高码率控制采用 TM5 码率控制方法^[1]. 实验软件平台是 MPEG4 软件测试平台 (MoMuSys PDAMI Version 0.9). 表 2 是本文所提出的改进的全局运动补偿编码算法和 MPEG-4 中已有的全局运动补偿编码的对比实验结果, 采用的参数模型都是 6 参数仿射模型.

表 1 MPEG-4 测试数据和测试条件

测试序列	格式/帧数	Bit-Rate (kbps)	编码帧率 (fps)	Rate Control
1. Foreman (Rect)	CIF (30fps) / 300	112	10	VM5
2. Coast Guard (Rect)	CIF (30fps) / 300	112	10	VM5
3. Stefan (Rect)	CIF (30fps) / 300	900	30	TM5 (15..1)
4. Stefan (VOO)	CIF (30fps) / 240	800	30	TM5 (15..1)

表 2 本文 GMC 编码和 MPEG4 中 GMC 编码算法性能比较

序列	MPEG4 GMC Bit-rate [kb/s] / PSNR- Y [dB]	本文 GMC Bit-rate [kb/s] / PSNR- Y [dB]	MPEG4 GME 算法 和本文 GME 算法 计算时间比值
Foreman (Rect)	121.62/30.10	123.42/30.02	13.04 1
Coast G. (Rect)	120.03/27.28	123.94/27.13	7.04 1
Stefan (Rect)	900.07/27.79	900.04/28.02	11.08 1
Stefan (VOO)	799.40/39.32	798.05/42.27	4.91 1

(注:表中 GMC 指全局运动补偿, GME 指全局运动估计)

表 2 是本文提出全局运动估计算法和 MPEG4 中全局运动估计算法的性能比较, 其中 GME 执行时间是指全局补偿编码算法中全局运动估计时间, 不包含编码时间, 因为这是整个 GMC 编码中的主要运算时间. 从全局运动的执行时间来看, 本文算法全局运动估计速度是 MPEG4 中算法的 5 倍以上, 对于 Foreman 和 CoastG 两个低码率序列, 本文 GMC 编码的 PSNR 略低 0.1 dB 左右, 而对于 Stefan (Rect) 和 Stefan (VOO) 两个高码率序列, 本文 GMC 算法的 PSNR 要高, 尤其是 Stefan 序列的背景 Stefan (VOO), PSNR 要高 2dB, 编码性能提高许多. 从实验结果整体来看, 本文算法在提高速度的同时编码性能没有损失太多, 和以前编码性能是可比的, 而且在某些序列还有所提高.

4 总结

本文提出了全局运动补偿编码中全局运动估计的改进算

法, 从而提高全局运动补偿编码性能. 从实验结果 (表 2) 可以看出本文所提出算法有效地提高了全局运动补偿编码的速度, 而编码效果损失不多, 另外对某些序列编码效果还有所提高. 实验结果说明本文算法对 MPEG4 中原有算法的改进是有效的. 编码速度的提高使得全局运动补偿编码在实时软件编码中应用成为可能. 本文中的全局运动补偿只是对背景进行, 如果前景也用模型编码 (如网格编码), 可以进一步提高压缩比, 改善相同码率下的图像质量. 从总体比较来看基于模型编码和传统编码方法相比有较大的优越性, 这也是下一代编码发展的趋势.

参考文献:

- [1] ISO/IEC WG11 MPEG Video Group. MPEG-4 Video Verification Model Version 15.0 [S]. the Maui meeting of ISO/IEC/JTC1/SC29/WG11, No. MPEG98/N3093, Hawaii, December 1999.
- [2] Janusz Konrad, Frederic Dufaux. Digital Equipment Corporation, Improved global motion estimation for N3 [A]. MPEG Meeting of ISO/IEC/SC29/WG11, No. MPEG97/M3096, San Jose, February 1998.
- [3] 袁亚湘, 孙文瑜. 最优化理论与方法 [M]. 科学出版社, 1997.
- [4] A. Murat Tekalp. Digital Video Processing [M]. reprinted by Tsinghua University of China, 1998.
- [5] Eung Tae Kim, Hyung-Myung Kim. Fast and robust parameter estimation method for global motion compensation in the video coder [J]. IEEE Transactions on Consumer Electronics, February 1999, 45 (1).
- [6] Aljoscha Smolic, Thomas Sikora, Jens-Rainer Ohm. Long-term global motion estimation and its application for sprite coding, content description, and segmentation [J]. IEEE Transactions on Circuits and Systems for Video Technology, December 1999, 9 (8).

作者简介:



贺玉文 1974 年出生, 清华大学计算机系人机交互与媒体集成研究所博士研究生, 主要研究领域为运动图像压缩和视频内容分析.



钟玉琢 1938 年出生, 清华大学计算机系教授, 博士生导师, 主要研究领域为多媒体数据压缩技术和分布式视频服务器.