

分布式视频编码中一种基于混合模型的边信息生成算法

纪 涛, 张太镒

(西安交通大学 电子与信息工程学院, 陕西 西安 710049)

摘 要: 提出了一种基于块匹配模型和光流模型的混合式边信息生成方法. 该方法对块匹配效果较差的区域采用光流法重新估计其运动矢量, 得到了更为精确的运动场. 实验表明, 该算法生成的边信息质量明显高于传统的基于块匹配的运动估计内插法.

关键词: 分布式视频编码; 边信息; 块匹配; 光流

中图分类号: TN919.81

文献标识码: A

文章编号: 1000-7180(2010)09-0016-04

Hybrid Model-Based Side Information Generation Algorithm for Distributed Video Coding

JI Tao, ZHANG Tai-yi

(School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A hybrid block-based and optical flow-based method is proposed. By using the optical flow method to recompute the motion vector in the regions where block matching algorithms have poor performance, this method produces more accurate motion fields. The experimental results show that comparing to other algorithms, the proposed method can significantly improve the quality of side information.

Key words: DVC; side information; block matching; optical flow

1 引言

传统的视频编码框架已经不再适用. 分布式视频编码(Distributed video coding, DVC)开始受到关注, 它主要是在解码端完成视频帧间统计相关特性的挖掘, 具有编码复杂度低, 抗误码性能好的特点, 能够为以上应用场合提供良好的解决方案^[1-2].

在 DVC 中, 边信息的质量高低是影响编码器性能的重要因素. 因为当前的 DVC 编码器大多是基于信道编码技术的, 常用 Turbo 码, LDPC 码等作为 Slepian-Wolf 编解码器, 它们将 WZ 帧与边信息 SI 之间的相关性建模成一个虚拟信道(常用二元对称信道或 AWGN 信道)^[3-4], 于是 SI 被看作是 WZ 帧的一个加噪版本, SI 与 WZ 帧之间的相关性越高, 解码的效果就越好. 目前常用的是基于运动补偿的内插插值法和外插插值法, 以及基于 hash code 的运动估计补偿

法^[5]. 其中的运动估计部分大多是基于传统视频编码中的块匹配方法. 然而块匹配方法是基于平移运动模型的, 是用一个运动矢量来描述块内所有点的运动情况, 它对旋转、缩放, 及非刚体的变形等运动的估计效果都不理想, 产生的边信息质量并不高, 这从一定程度上影响了 DVC 编码器的性能.

为了提高边信息的估计质量, 本文引入了光流估计法, 将光流模型在估计旋转、缩放, 及非刚体的变形等运动时的天然优势与块匹配法相结合, 提出了一种基于混合模型的边信息内插生成方法.

2 基于光流模型的边信息生成方法

2.1 Horn-Schunck 光流算法

光流(optical flow)是指图像亮度模式的表观运动(apparent motion), 是实际运动场的一种近似, 通过光流计算, 可以获取对象的运动信息. 光流法基于

物体移动的光学特性提出如下假设:运动物体的灰度在很短的时间间隔内保持不变. 设 $I(x, y, t)$ 代表时刻为 t 的视频帧内点 $m = (x, y)$ 处的亮度, 设 $u(x, y), v(x, y)$ 分别是该点处光流的 x, y 方向分量, 则由该假设可得

$$I(x + u\Delta t, y + v\Delta t, t + \Delta t) = I(x, y, t) \quad (1)$$

将左式在 (x, y, t) 点用泰勒公式展开, 忽略高阶项可得

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0 \quad (2)$$

记 $I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, I_t = \frac{\partial I}{\partial t}$, 并注意到 $u = \frac{dx}{dt} = \frac{\Delta x}{\Delta t}, v = \frac{dy}{dt} = \frac{\Delta y}{\Delta t}$, 上式可化为

$$I_x u + I_y v + I_t = 0 \quad (3)$$

式(3)称为光流约束方程(Optical Flow Constraint Equation), 从该方程求解光流场是一个不适定问题, 必须引入附加的约束条件. 为此, Horn 和 Schunck^[6]提出另一假设: 给定领域内的速度向量场变化是缓慢的. 根据这一假设, 光流场在满足光流约束方程的同时还满足全局平滑性. Horn 和 Schunck 给出了速度的平滑度量:

$$|\nabla u|^2 + |\nabla v|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \quad (4)$$

将光滑性测度和光流约束方程结合起来, 有

$$E_{HS}(u, v) = \int_{\Omega} ((I_x u + I_y v + I_t)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2)) dx dy \quad (5)$$

式中, α 是控制平滑度的参数, α 越大, 则得出的光流场平滑度就越高. 使 $E_{HS}(u, v)$ 最小化就可计算得到光流场 (u, v) . 具体求解可采用 Gauss-Seidel 迭代法:

$$u^{k+1} = \bar{u}^k - I_x \frac{I_x \bar{u}^k + I_y \bar{v}^k + I_t}{\alpha + I_x^2 + I_y^2} \quad (6a)$$

$$v^{k+1} = \bar{v}^k - I_y \frac{I_x \bar{u}^k + I_y \bar{v}^k + I_t}{\alpha + I_x^2 + I_y^2} \quad (6b)$$

式中, k 为迭代次数. HS 光流法产生的运动场较为平滑, 且更加致密, 因而大多数情况下比块匹配方法产生的运动场更接近真实运动情况.

2.2 层次化的 Horn-Schunck 光流算法

直接基于原图像的光流算法在处理大的运动速度时效果不好, 为了克服这一缺点, 研究人员提出了基于图像金字塔结构的光流算法^[7], 即层次化的光

流算法. 图像的金字塔表示法是计算机视觉中常用的一种多分辨率表示法. 利用金字塔表示法, 能分析图像中不同大小的物体, 例如高分辨率的下层可用于分析细节, 低分辨率的上层可用于分析较大的物体. 同时, 通过对低分辨率、尺寸较小的上层进行分析所得到的信息还能用来指导对高分辨率、尺寸较大的下层进行分析, 从而大大简化了分析和计算.

在本文所使用的层次化的 Horn-Schunck 光流算法中, 首先构建原图像的高斯金字塔序列, 按分辨率递减表示为 $L_0^i, L_1^i, \dots, L_L^i, \dots, L_L^i (i = 1, 2; 0 \leq l \leq L)$, 然后按如下步骤完成光流场计算:

(1) 令 $l = L$, 初始化矢量场 $V_l^{(0)}$; (2) 根据 $V_l^{(0)}$ 对 L_l^1, L_l^2 进行补偿; (3) 在补偿图像上根据式(6)计算速度矢量 $V_l^{(k)}$, k 为迭代次数, 令 $V_l^{(k)} = V_l^{(k)} + V_l^{(0)}$; (4) 如果 $l = 1$, 则转至(1), 否则, 转至(5); (5) 根据映射算子 $P: V_l^{(k)} \rightarrow V_{l-1}^{(0)}$, 计算高一分辨率层的初始化矢量场 $V_{l-1}^{(0)}$; (6) $l = l - 1$, 返回(2); (7) 根据映射算子 $P: V_1^{(k)} \rightarrow V_0$, V_0 即为所求的光流场.

其中映射算子 P 可采用双线性插值算子. 本文没有在最高分辨率层上直接用 Horn-Schunck 方法计算光流场, 而是从较低分辨率层得到的光流场插值得到最后的光流场, 这样做是为了降低运算量, 同时又不至于降低光流场的精度.

2.3 边信息生成

由光流法得到前后相邻两 K 帧间的运动场后, 便可内插出边信息. 和块匹配法一样, 假设相邻两 K 帧之间的运动是沿运动轨迹恒定不变的, 则可求出中间帧的前后向运动矢量, 进而可以根据前后 K 帧内插出中间帧, 也就是解码所需要的边信息.

3 基于混合模型的边信息生成方法

相比基于块匹配的运动估计, HS 光流法可以获得更为精确的运动场, 从而产生更为精确的边信息, 但基于梯度计算的 HS 光流算法, 存在着对噪声鲁棒性差的缺点, 同时, 在运动场不连续处, 光流法的估计精度也会降低, 另外, 基于块匹配的运动估计对静止区域及刚体的平移运动有着良好的估计效果, 这些导致在图像的某些区域, 光流法的估计效果可能不如基于块匹配的运动估计, 或者改善的效果并不明显. 除此之外, 光流法的计算量大, 也使它的应用受到了很大的限制. 基于以上的考虑, 我们提出了一种新的边信息生成方法.

首先,使用传统的基于块匹配的运动补偿内插值法,生成边信息 SI_{block} . 在此,本文采用前向运动估计结合对称双向运动估计. 首先对已经得到的 K 帧 X_{2i-1} 和 X_{2i+1} 进行前向块匹配运动估计,块大小为 8×8 ,搜索方法采用全搜索,目标函数采用平均绝对误差(MAD)准则,如下式:

$$MAD_{uv} = \frac{1}{N^2} \sum_{m,n=0}^{N-1} |I_1(m-v, n-u) - I_2(m+v, n+u)| \quad (7)$$

这样的运动估计可能会产生遮挡或空白区域,因此,对每一个内插块,从前面得到的运动矢量中选取离该块中心最近的作为该块的运动矢量. 然后,对得到的运动矢量进行对称双向运动估计,以提高其精度. 对称双向运动估计以内插块的前向后向运动矢量的对称性作为前提,以前向运动估计得到的运动矢量 $MV_{initial}$ 为初值,以内插块为基准,对称地在 X_{2k-1}, X_{2k+1} 中进行小范围的块匹配搜索,以改善前面得到的运动场,如图 1 所示.

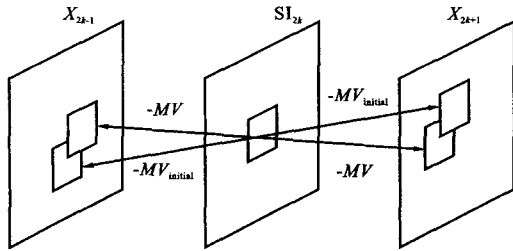


图 1 对称双向运动估计

块匹配的目标函数仍然采用平均绝对误差准则. 记录下匹配后对应两块的 $MAD \min$ 值. 由于基于块匹配的运动估计是通过搜索相邻区域中最小的 $MAD \min$ 值来进行匹配的,因而较大的 MAD 值通常意味着块匹配运动估计的效果不好,对这些块,本文采用 HS 光流法重新进行运动估计,以进一步改善得到的运动场.

具体实现中,可以选定一个阈值 T ,对 $MAD \min$ 大于 T 的块进行标记,接着采用区域生长的方法将这些标记块归并成一个个小的区域,为了便于后续处理,本文将这些区域补成矩形,如果两个矩形有重叠,便再次对它们进行合并,一直到每个矩形区域都是独立的为止. 然后,再对这些区域采用层次化 HS 光流法进行运动补偿内插,在内插的过程中,记录下每一个内插块的 MAD 值,记为 MAD_{op} ,用它来判断光流估计的准确性:

$$MAD_{op} = \frac{1}{N^2} \sum_{m,n=0}^{N-1} |I_1(m-v(m,n), n-u(m,$$

$$n)) - I_2(m+v(m,n), n+u(m,n))| \quad (8)$$

最后,如果 MAD_{op} 小于对应块的 $MAD \min$,使用光流法补偿块来替换 SI_{block} 中的对应块.

本文将该方法归纳为如下步骤:(1) 采用前向运动估计结合对称双向运动估计的方法,对关键帧 X_{2k-1} 和 X_{2k+1} 进行运动补偿内插,得到 SI_{block} ,同时记录 SI_{block} 中每一块 B_{ij} 所对应的 $MAD_{ij} \min$ 值;(2) 对 $MAD_{ij} \min$ 大于 T 的块进行标记,由标记块归并出待处理矩形区域集合 $\{region_i\}$;(3) 将 $region_i$ 在 X_{2k-1} 和 X_{2k+1} 中的对应区域向外扩展 4 个像素后,再进行光流计算,插值得到 SI_{region_i} ,并记录其中每一块 OPB_{ij} 所对应的 $MAD_{op_{ij}}$,如果 $MAD_{op_{ij}} < MAD_{ij} \min$,令 $B_{ij} = OPB_{ij}$;(4) 令 $SI_{2k} = SI_{block}$, SI_{2k} 即为所求的边信息.

本文算法采用块匹配后的平均绝对误差(MAD)作为准则,对块匹配法的匹配效果进行判定, MAD 值越大,说明匹配的效果越差. 由于块匹配效果差大部分情况下是由于该块内同时存在着多种不同方向的平移运动,同时还可能有缩放、旋转、变形等运动存在,故对该区域采用光流模型重新计算其运动场. 由于光流模型为每一个像素都分配了一个运动矢量,从而克服了块匹配模型在描述这些不同运动时所存在的固有缺陷,且 HS 光流法产生的运动场更加平滑,因而更接近真实运动,图 2 为 foreman 序列第 6 帧某一区域的运动场示意图.

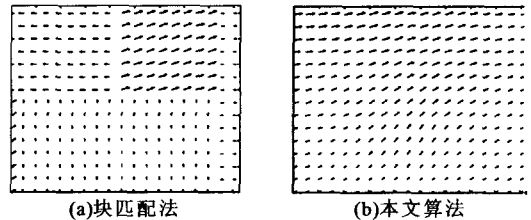


图 2 用块匹配法和本文算法分别得到的运动场

从图 2 可以看出,本文算法得到的运动场更加平滑,而块匹配法产生的运动场则有明显的块效应,与真实运动情况差距很大. 另外,只对块匹配效果较差的区域进行光流计算,并且采用层次化的 HS 光流法,在很大程度上减少了光流的运算量,同时又能发挥块匹配模型计算量小,对噪声鲁棒性强,在单一运动区域估计效果好的优势. 此外,本文算法用 MAD_{op} 来判断光流估计的准确性,从而排除了噪声引起的光流估计效果下降的块,保证了算法的稳定性. 总之,本文算法将光流模型与块匹配模型的优势相结合,得到了更加精确的运动场,从而可以达到改

进边信息生成质量,提高分布式视频编码器的整体率失真性能的目的.

4 实验结果

测试序列为 foreman 序列(30fps, QCIF),前 150 个奇数帧作为 K 帧,前 149 个偶数帧作为 WZ 帧,相邻两个奇数帧用来内插出边信息 SI . 实验使用 4 种边信息内插法,分别是前向运动估计法,双向运动估计法,前向运动估计结合对称双向运动估计法,以及本文算法.在本文算法中,令 $T=10$,光流算法中金字塔层数选为 3. 各种算法中的搜索方法都选为全搜索,其中,前向及双向运动估计的搜索范围为 ± 8 ,对称双向运动估计的搜索范围为 ± 2 . 只考虑亮度分量,计算出用不同方法得到的边信息的 PSNR 值,结果如图 3 所示.

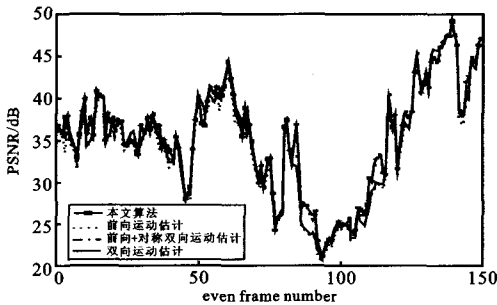


图 3 foreman 序列 149 个偶数帧(边信息)的 PSNR 曲线

从图 3 及表 1 可以看到,本文算法生成的边信息质量在整个 foreman 序列上几乎都要优于其他算法,比前向运动估计平均高将近 1dB,比双向运动估计平均高 0.4dB,比前向运动估计结合对称双向运动估计平均高 0.2dB.

表 1 foreman 序列 149 个偶数帧(边信息)的 PSNR 均值	
运动估计方法	PSNR 平均值
前向运动估计	34.18
双向运动估计	34.73
前向 + 对称双向运动估计	34.96
本文算法	35.17

图 4 是 Foreman 序列第 6 帧通过不同方法得到的边信息图像.

从图 4 中可以看出,本文算法得到的图像更加平滑,块效应明显减少,主观质量明显优于其他算法.由此可见,本文算法通过在解码端进行更为精确的运动估计,达到了提高边信息生成质量的目的.

5 结束语

本文针对块匹配模型所存在的固有缺陷,提出

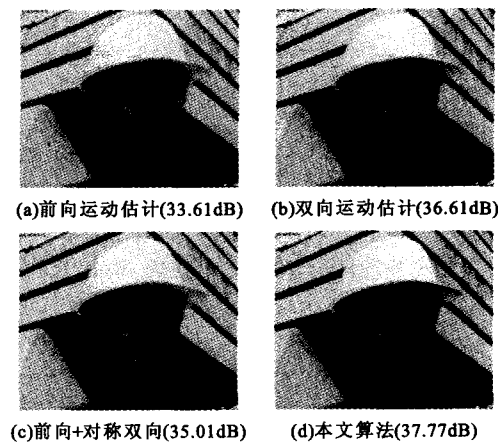


图 4 foreman 序列第 6 帧

了一种块匹配模型与光流模型相结合的混合式边信息生成方法.实验表明,使用本文算法所生成的边信息,其 PSNR 和主观质量都较传统的边信息内插方法有所提高.本文算法在没有增加编码端复杂度的前提下,在解码端充分发掘视频的帧间相关性,提高了边信息的生成质量,对提高 DVC 编码器的整体率失真性能有重要意义.

参考文献:

[1] 胡波,王宏远. MPEG 编码算法的优化[J]. 微电子学与计算机,2008,25(4):201-204.

[2] 段大高,李文正,莫倩,等. 一种改进的 H.264 码率控制算法[J]. 微电子学与计算机,2008,25(6):40-42.

[3] Slepian J D, Wolf J K. Noiseless coding of correlated information sources [J]. IEEE Transactions on Information Theory, 1973(17-19):471-480.

[4] Wyner A, Ziv J. The rate-distortion function for source coding with side information at the decoder[J]. IEEE Transactions on Information Theory,1976,IT-22(1):1-10.

[5] Girod B, Aaron A, Rane, D. Rebollo-Monedero. Distributed video coding [J]. the IEEE,2005,93(1):71-83.

[6] Berthol Horn, Brian schunck. Determining optical flow. artificial intelligence[M]. Cambridge: Massachusetts Institute of Technology, 1981:185-203.

[7] Jean-Yves Bouguet. Pyramidal implementation of the lucas kanade feature tracker - description of the algorithm [R]. Intel Corporation, Microprocessor Research Labs, OpenCV Documents, 1999.

作者简介:

纪涛 男,(1984-),硕士研究生.研究方向为视频压缩及多媒体通信.

张太镒 男,(1944-),博士生导师.研究方向为新一代移动通信技术、软件无线电、数字音视频技术.