

视频转码中运动矢量自适应多模式精炼算法

沈晓东, 张太镒, 李小和

(西安交通大学 电子与信息工程学院, 陕西 西安 710049)

摘 要: 针对视频转码中运动矢量重估计精度不高的问题, 提出了一种运动矢量自适应多模式精炼(AMPR)算法. 该算法首先利用输入码流中运动矢量的相位和幅值信息构建了视频局部区域活动性模型, 通过此模型自适应确定精炼窗口的大小, 然后在菱形搜索算法(DS)和水平垂直搜索算法(HAVS)的基础上, 给出了多模式搜索策略. 仿真实验表明, 该算法取得接近于全搜索算法的视频质量, 并有效地降低了计算复杂度.

关键词: 视频转码; 运动判定; 空间下采样; 自适应精炼

中图分类号: TN919.81

文献标识码: A

文章编号: 1000-7180(2010)09-0024-05

An Adaptive Multi-Pattern Motion Vector Refinement Algorithm for Video Transcoder

SHEN Xiao-dong, ZHANG Tai-yi, LI Xiao-he

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

Abstract: An adaptive multi-pattern refinement algorithm for motion vector in video transcoding is presented. Using the consistency of angle and amplitude of intensity of local motion vectors, a motion activity model is constructed, so the search range of the macroblock is confirmed adaptively using the activity model. Then, a multi-pattern refinement strategy based on the diamond search and horizontal and vertical search is proposed. Experimental results show that the proposed method reduces the computation complexity significantly without degrading the quality of video transcoder.

Key words: video transcoding; motion estimation; space downscaling; adaptive refinement

1 引言

随着网络和多媒体技术的发展, 视频会议、视频通信和点播, 以及远程教育等多媒体业务已得到广泛的应用. 视频转码技术可以将给定的压缩码流经过动态调整, 输出较低码率的码流来适应不同应用环境的需要^[1-3]. 最简单的视频转码器是将输入码流完全解码后, 再重新进行编码. 由于它需要对运动矢量重估计, 计算复杂度大, 不适合实时处理. 最直接的方法是复用输入码流中的运动矢量^[3,4], 降低转码器的计算复杂度. 这种运动矢量的简单重组算法在许多应用中会引入较大的质量衰减^[2,4], 为了获得较高的视频质量通常需要对重组后的运动矢量

进行后续精炼.

运动矢量的精炼算法近年来取得一定的进展, 文献[4]提出了基于小搜索窗单调误差曲面特性假设的水平垂直搜索(HAVS)精炼算法, 有效地利用了误差的分布特性, 但是, 在单调误差曲面不完备的情况下, 易陷入局部最优点. 文献[5]提出的 DS 算法在传统的运动估计中取得了良好的效果, 但是在重组运动矢量所需要精练的窗口相对较小, DS 算法存在冗余搜索点的问题. 上述算法没有考虑到不同的局部区域具有不同的运动特性, 因此需要不同的精炼窗口大小.

针对以上不足, 本文对视频局部运动特性进行分析, 根据运动矢量相位和幅值特性, 构建了局部区

收稿日期: 2009-07-27; 修回日期: 2009-09-27

基金项目: 国家自然科学基金项目(60602034)

域活动性模型,自适应确定运动矢量的精炼范围,在后续的精炼过程中,根据精炼窗口大小的不同,提出了多模式快速搜索算法,从而在保持视频质量的基础上,有效地降低了计算复杂度。

2 基于区域活动模型的精炼区域自适应确定

利用压缩码流中的运动矢量重组新的运动矢量显著减小了视频转码的运算复杂度,运动矢量的重组算法得到广泛的应用^[4]。在视频场景中,背景区域和前景物体刚体运动部分各宏块运动矢量的强度和方向基本一致,通过重组得到的运动矢量精度比较高,只需要在较小的范围内精炼就可以找到最优值,而在前景和背景相邻的边界宏块以及前景物体的变形运动部分,由于运动矢量在强度和方向上都存在较大的差异,重组的运动矢量和宏块的最优运动矢量差异比较大,在较小的搜索范围内就很难找到最优点,因此需要在一个较大范围内进行搜索才能得到全局最优点。图1给出Foreman测试序列150帧在[8,8]搜索窗内使用全搜索得到的运动矢量和重组运动矢量^[3]之间最大偏差的累积直方图,从图中可以看出,大约57%的宏块不需要进行精炼,85%的宏块只需要在 ± 1 像素的搜索范围进行精炼,只有5%的宏块需要在大于 ± 2 像素的搜索范围进行精炼。

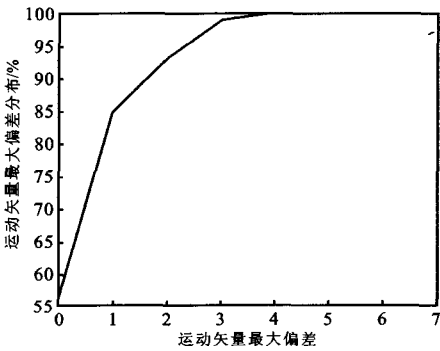


图1 全搜索和AMVR重组运动矢量最大搜索半径的分布特性

图1表明不同特性的宏块重组运动矢量和最优运动矢量之间的偏差是不一致的,因此对应的精炼范围大小也应该有所差异。

针对这些特性,借鉴视频注意力模型^[6-7],本文利用方向一致性和幅值强度构建一个局部区域活动性模型来描述视频局部的宏块信息,并依此自适应确定精炼窗口的大小。

2.1 方向一致性

用空间相位分布的熵 $\text{AngEn}(i, j)$ 来定义局部区域运动矢量的方向一致性。令 $\text{MB}_{i,j}$ 为当前帧下采样后的第 (i, j) 个宏块, i 和 j 为宏块在垂直和水平方向上的坐标,宏块 $\text{MB}_{i,j}$ 对应的局部区域定义为 $R_{i,j}$,包括需要下采样的四个宏块和周围12个相邻宏块 $\text{MB}_{k,l}$ 的运动矢量,具体表示为

$$R_{i,j} = \{\text{MB}_{k,l}\} \quad (1)$$

式中, $(2i-2) \leq k \leq (2i+1)$ $(2j-2) \leq l \leq (2j+1)$

局部区域运动矢量的一致性定义如下:

$$\text{AngEn}(i, j) = - \sum_{k=1}^n p(k) \log(p(k)) \quad (2)$$

式中, $p(k) = \frac{\text{SH}_{i,j}^w(k)}{\sum_{l=1}^n \text{SH}_{i,j}^w(l)}$ 表示第 k 个区域内运动

矢量相位分布概率, n 为角度量化的等级, $\text{SH}_{i,j}^w$ 是 $\text{MB}_{i,j}$ 支持区域内运动矢量的空间相位直方图, $\text{SH}_{i,j}^w(k)$ 是分布第 k 个区域的运动矢量的个数。 $\text{AngEn}(i, j)$ 的取值范围为 $(0, \log n]$, 当 $P(t) = 1/n, t \in [1, n]$ 时, $\text{AngEn}(i, j)$ 取得最大值 $\log n$ 。其归一化的结果表示如下:

$$\text{AngEn}'(i, j) = \frac{\text{AngEn}(i, j)}{\log n} \quad (3)$$

AngEn' 由0到1变化的过程中,意味着宏块对应的局部区域运动一致性越来越差,对应的精炼范围应该越来越大。

2.2 幅值强度

用运动矢量的活动强度 L_v 来确定局部区域的幅值强度 $\text{AmpEn}(i, j)$ 。令支持区域 $R_{i,j}$ 中宏块 $\text{MB}_{k,l}$ 的运动矢量表示为 $\mathbf{V}_{k,l} = (x_{k,l}, y_{k,l})$, 局部区域的活动强度定义为

$$L_{v,k,l} = \max(|x_{k,l}|, |y_{k,l}|) \quad (4)$$

局部区域的幅值强度表示为

$$\text{AmpEn}(i, j) = \max_{\{k,l|\text{MB}_{k,l} \in R_{i,j}\}} 2\text{th}(L_{v,k,l}) \quad (5)$$

Max2th表示局部区域的第二个最大值,有效地避免奇异点的存在。

2.3 区域活动性模型

若局部区域内运动矢量一致性变差,对应局部区域运动矢量的方向一致性会增大,需要更大的精炼范围,同时,宏块活动性增大,幅值强度也会增大,需要的精炼范围也随之增大^[7],因此可以得出局部

区域的活动性与局部区域内运动矢量的一致性和幅值强度成一定的正向比例关系,根据式(4)、式(6),局部区域的活动性模型定义如下:

$$\text{ActEn}(i, j) = \text{AngEn}'(i, j) \text{AmpEn}(i, j) \quad (6)$$

2.4 自适应精炼范围确定

重组运动矢量的精炼范围的大小主要受局部区域的运动一致性和幅值强度的影响,活动性较大的区域需要较大的精炼范围,活动性较小的区域需要较小的精炼范围.根据局部区域活动性模型,精炼窗口大小定义如下:

$$S(i, j) = \begin{cases} 1 & \text{if } \text{ActEn}(i, j) \leq 1 \\ 2 & \text{if } 1 < \text{ActEn}(i, j) \leq 4 \\ 4 & \text{if } \text{ActEn}(i, j) > 4 \end{cases}$$

式中, $S(i, j)$ 为第 (i, j) 个宏块对应精炼窗口的大小.

3 运动矢量的多模式快速精炼

运动矢量的精炼过程可以在新的搜索范围内使用全搜索的方法进行,但可以通过快速搜索算法进一步降低运算复杂度.在视频编码算法中,快速搜索算法得到了广泛的研究^[7-8],大多数搜索算法是基于单峰误差曲面假设,但是在比较大的搜索范围内,由于有许多局部最小点的存在,单峰误差曲面的假设并不总是满足.针对这些问题,本文结合局部区域活动性模型,并根据文献^[7-8]中的快速搜索算法,提出了一种多模式搜索算法.该算法降低 DS 算法^[8]中的冗余搜索点,同时克服了 HAVS 算法^[7]易于陷入局部最小点的问题.

多模式搜索算法的具体搜索策略如下:

第一部分:在局部区域活动性小于等于 2 的情况下, HAVS 算法可以得到较优的性能,但是单峰误差曲面假设不成立时,该算法易陷入局部最小点,为避免这种情况发生,本文改进 HAVS 算法.的中止条件,在水平垂直方向上搜索迭代地进行,直到绝对差之和最小的点在水平和垂直搜索方向的中心点上,才中止搜索,具体搜索策略如下:

(1) 以自适应运动矢量重组算法(AMVR)^[4]获取的运动矢量为中心,在 $[-S, +S]$ 的范围内进行搜索.

(2) 在水平方向上,先搜索左侧相邻像素点,如果搜索的像素点对应的宏块绝对值方差和(SAD)小于中心点的 SAD,则以这个点为中心,继续向同一侧进行搜索,直到中心点为最小点的情况下,转到第

三步.否则,搜索中心点右侧的像素点,如果搜索的像素点对应的宏块绝对值方差和 SAD 小于中心点的 SAD,则以这个点为中心,继续向右搜索,直到中心点为最小点的情况下,如果在垂直方向和水平方向上的一次迭代搜索循环中,中心点均没有发生偏移,停止搜索,输出最终运动矢量,否则转到第三步.

(3) 在垂直方向上,先搜索上面相邻像素点,如果搜索的像素点对应的宏块绝对值方差和 SAD 小于中心点的 SAD,则以这个点为中心,继续向同一侧进行搜索,直到中心点为最小点的情况下,转到第二步.否则向下面搜索,如果搜索的像素点对应的宏块绝对值方差和 SAD 小于中心点的 SAD,继续向下搜索,直到中心点为最小点的情况下,如果在水平方向和垂直方向上的一次迭代搜索循环中,中心点均没有发生偏移,停止搜索,输出最终运动矢量,否则转到第二步.

图 2 是上面改进的 HAVS 算法的一个搜索实例,若采用 HAVS 算法搜索到次优点 4,就中止搜索,本文算法可以顺利找到最优点 5,而只增加两个搜索点.

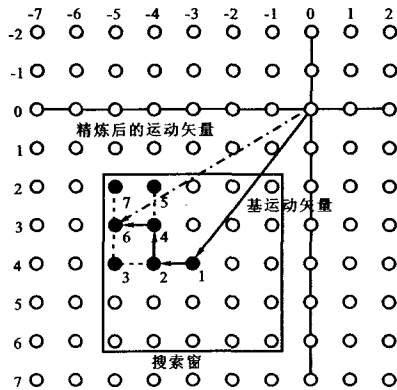


图 2 改进的 HAVS 快速搜索策略

第二部分:在局部活动性大于 4 的情况,运动矢量的重组误差比较大,运动矢量也常常偏离最优点,文献^[8]的 DS 搜索算法常得到较优的结果,在 DS 算法中,还是存在一些冗余的搜索点,可以结合 HAVS 算法对 DS 算法进一步的改进.具体策略为:首先执行大菱形搜索,如果中间的最优点距离搜索窗边界小于等于 2 时,可以认为最优点在以中间最小点为中心,在 $S=2$ 的小搜索窗范围之内,因此可以采用改进的 HAVS 搜索策略搜索最优点,具体搜索策略如下:

(1) 以 AMVR 重组的运动矢量为中心,在半径为 S 的范围内进行搜索.

(2) 首先使用大菱形进行搜索,如果绝对值方差和最小点在中心点上,在半径为1的范围内执行HAVS搜索算法,否则,检测最小点距离搜索窗边界的距离,如果距离小于等于2,则转到第三步,否则以这个最小点为中心,继续执行第二步。

(3) 以上一个最小点为中心,执行改进的HAVS快速搜索策略。

图3是上面搜索过程的一个实例。首先以点1为中心点,采用大菱形搜索,在绝对差和最小点2处,因为点2与边界的距离等于3,所以采用改进的HAVS搜索策略,顺利地找到最优点12,整个过程需要15个搜索点,如果采用文献[6]的算法,需要19个搜索点。

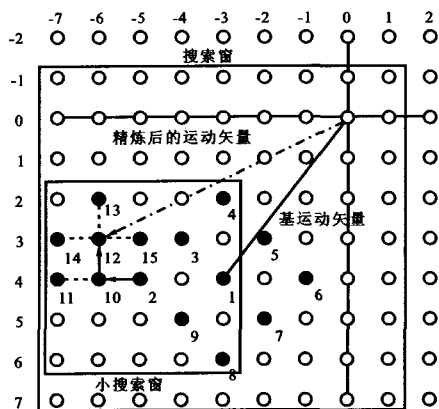


图3 改进的DS快速搜索策略

在实际应用中,对于不同的搜索窗大小,采用相应的搜索策略。在搜索过程中,超出搜索窗口的搜索点不进行搜索,这保证了精炼后的运动矢量可以用给定的码长进行编码,从而提高运动矢量编码的效率。

4 实验结果

仿真实验采用Foreman测试序列,通过视频转码分辨率由CIF格式转换为QCIF格式,输入的运动矢量是在大小为 $[-15, 15]$ 的搜索窗内通过在全搜索运动估计算法得到的。使用AMVR算法重组基运动矢量,采用本文的自适应多模式搜索算法(AMPR)对基运动矢量进行精练。为了和本文的算法进行对比,在大小 $[-2, 2]$ 的搜索窗内使用全搜索算法、HAVS算法以及DS算法得到的视频补偿结果和本文算法的结果进行性能比较,如图4所示。

图4给出不同精炼算法得到的视频质量效果。可以看出本文的算法和DS算法都接近于全搜索算

法得到的结果,在计算复杂度方面,全搜索算法即使在 ± 2 的搜索窗内每一帧都需要22 275个搜索点,AMVR算法不需要额外的搜索点,运算量很小,但效果较差。本文的算法和HAVS、DS算法的搜索点数如图5所示,可以看出本文算法每帧需要的平均搜索点数远低于全搜索算法需要的搜索点数,也明显低于HAVS和DS算法所需要的搜索点数。

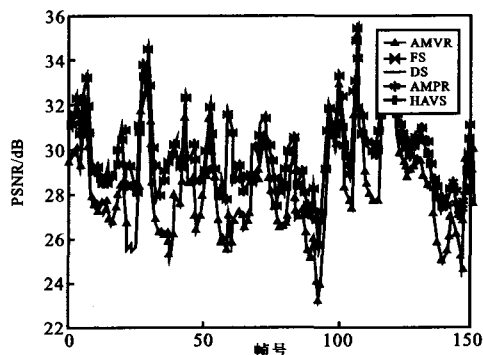


图4 AMPR算法和HAVS、DS、FS快速搜索算法以及AMVR算法的性能比较

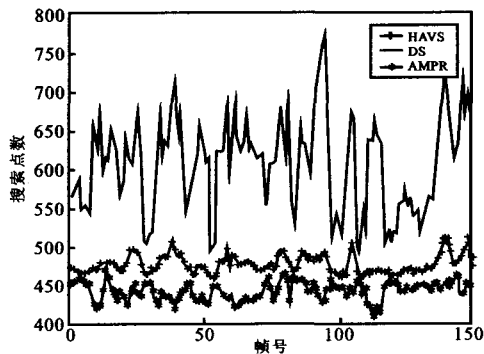


图5 AMPR算法和HAVS、DS算法的计算复杂度比较

5 结束语

本文提出了一种视频转码中运动矢量自适应多模式精炼算法,该算法根据视频局部的运动活动一致性和幅值强度构建了区域活动性模型,以此自适应的确定精炼窗口的大小;针对不同的精炼窗口,提出基于多模式的运动矢量快速搜索策略,有效地解决了HAVS易于陷入局部最小点和DS算法中冗余搜索点的问题。该算法取到与全搜索算法相近的视频质量,同时有效地降低了算法的计算复杂度。

参考文献:

- [1] Keesman G, Hellinghuizen R, Hoeksema R, et al. Transcoding of MPEG bitstreams[J]. Signal Process: Image Communi-

- cation, 1996, 8(6):481-500.
- [2] Anthony Vetro, Charilaos Christopoulos, Huifang Sun. Video transcoding architectures and technique: an overview [J]. IEEE Signal Processing Magazine, 2003(3):18-29.
- [3] Shen B, Sethi I K, Vasudev B. Adaptive motion vector re-sampling for compressed video downscaling [J]. IEEE Trans. Circuits Syst. Video Technol, 1999, 9(6):929-936.
- [4] Youn J, M T Sun, Lin C W. Motion vector refinement for high-performance transcoding[J]. IEEE Trans. On Multimedia, 1999, 1(1):30-40.
- [5] Zhu S, Ma K K. A new diamond search algorithm for fast block matching motion estimation[J]. IEEE Trans. Image Processing, 2000, 9(2):287-290.
- [6] 汪烈军, 覃锡忠. 一种改进的快速块匹配运动估计算法[J]. 微电子学与计算机, 2007, 24(6):204-206.
- [7] Ngo C W, Ma Y F, Zhang H J. Video summarization and scene detection by graph modeling[J]. IEEE Trans. on Circuits Syst for Video Technology, 2005, 15(2):296-305.
- [8] Ma Y F, Zhang H J, Li M J. A user attention model for video summarization[J]. ACM, Multimedia, 2002, 1(6):533-541.

作者简介:

沈晓东 男, (1979-), 博士研究生. 研究方向为图像处理、视频编解码等.

张太镒 男, (1944-), 教授. 研究方向为信号处理、图像视频处理与系统等.

李小和 男, (1974-), 博士研究生. 研究方向为图像处理、视频编解码等.

(上接第 23 页)

- [7] 李刚, 贺昱曜, 赵妍. 基于大津法和互信息量的路面破损图像自动识别算法[J]. 微电子学与计算机, 2009, 26(7):242-243.

作者简介:

张书真 女, (1977-), 硕士, 讲师. 研究方向为图像处理、模

式识别.

向晓燕 女, (1976-), 讲师. 研究方向为图像处理、控制理论.

赵英男 女, (1973-), 博士, 副教授. 研究方向为图像处理、模式识别.

宋海龙 男, (1977-), 硕士. 研究方向为图像处理、编码密码理论.