

一种快捷的高精度运动估计方法

陆寄远^{1,2} 朝红阳³

(中山大学信息与计算机科学学院 广州 510275)¹ (广东金融学院计算机科学与技术系 广州 510521)²
(中山大学软件学院 广州 510275)³

摘 要 因为高清数字电视对图像质量有严格的要求,所以高清数字电视编码采用了计算十分复杂的高精度运动估计方法,而非快速的简单运动估计方法,这就导致了运动估计在视频编码中占用了很多的计算资源。提出了一种结合不同宏块运动特性的方法,以便自适应地选择高精度运动估计方法或简单运动估计方法。该方法通过改变宏块编码顺序与运动向量预测顺序的关系,在减少 70% 运算时间的情况下仍能保持与高精度方法相同的搜索精度。

关键词 视频编码,运动估计,预测运动向量

中图分类号 TN919.81 **文献标识码** A

Simple and High Precision Fast Motion Estimation Algorithm

LU Ji-yuan^{1,2} CHAO Hong-yang³

(School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China)¹

(Department of Computer Science and Technology, Guangdong University of Finance, Guangzhou 510521, China)²

(School of Software, Sun Yat-Sen University, Guangzhou 510275, China)³

Abstract Because of intense requirement of image quality for digital television, high precision motion estimation algorithms were recommended instead of simple motion estimation algorithms. To benefit from both the speed of simple search and the accuracy of high precision search, a new method making use of inter macroblock characteristic was proposed. According to experiments, a reduction of 70% execution time is attained while still maintaining the same accuracy of high precision search.

Keywords Video coding, Motion estimation, Predictor

1 引言

MPEG-2 凭借其清晰的图像质量,成为了第一代数字电视视频部分的压缩标准。但随着压缩技术的不断发展以及硬件设备运算速度的持续提高, MPEG-2 在压缩率上越来越显示出不足的地方。有 3 个相互竞争标准最有机会成为下一代高清数字电视的视频压缩标准。第一,由 ISO 的 MPEG 工作组和 ITU 的 VCEG 工作组联合制定的 MPEG-4 AVC/H. 264 标准。第二,由中国于 2006 年 4 月颁布的视频压缩标准 AVS,目前主要针对数字电视、视频存储等应用领域。第三, SMPTE 颁布的 VC-1 标准,实质是微软视窗系统的 Windows Media Video-9(WMV-9)。在这 3 个压缩标准中, H. 264 一直是最权威的压缩标准,并且它的前身 MPEG-2 已经被实际应用到了数字电视领域。而 AVS 和 VC-1 都在计算速度和专利授权两方面标榜它们比 H. 264 更有优势。

视频压缩技术关键的 3 个指标就是图像质量、压缩率和计算复杂度。新一代数字电视也称作高清数字电视,它把图像质量作为首要考虑因素。此外,现在电视频道和电视节目的数量都在大幅度增长,自然对广播网络的负载大大加重。

如何提高压缩率、减少带宽,成为了第二重点考虑的问题。而且为了节省电视节目制作的成本,视频的压缩设备都会尽量采用低廉的硬件,所以对于压缩算法的实现也要尽量简单。

视频编码采用运动补偿技术来提取视频信号时域上的冗余度。运动补偿技术虽然对压缩码率的贡献尤为重要,但同时也令运动估计的计算时间大大增加。目前出现了许多快速运动估计算法^[1-3]。它们大致可以分为两类:一类是简单运动估计方法(以下简称简单方法)^[4-6]。这类方法比较注重算法的运算速度,它们利用了相邻宏块间运动向量的相似性,准确地估计预测运动向量(以下简称预测点),然后使用一些相对较小的搜索模板在该预测点的邻域进行运动估计。因为它们所使用的搜索模板较小,所以其搜索速度很快。但在运动剧烈的场景下,相邻宏块间运动向量的相似性就会变差。使用这些简单方法经常会使得运动搜索被中止在局部最优的位置上,从而影响最终的图像质量。这个原因,使得简单方法多数应用在视频会议、可视电话这类对图像质量和硬件配置都要求不高的领域。另一类是高精度的运动估计方法(以下简称高精度方法)^[7,8]。为了保证最终的图像质量,高精度方法在简单运动方法的基础上增加了许多复杂的搜索模板,以减少

到稿日期:2009-03-30 返修日期:2009-06-10 本文受国家 863 计划项目——基于动漫数据库的计算机辅助动漫创作关键技术与系统(2007AA01Z340)和广东省自然科学基金项目——新一代视频编码标准中运动估计关键技术的研究(07003728)资助。

陆寄远(1976—),男,博士生,讲师,主要研究方向为视频编码、图像传输等, E-mail: dtc005001@163.com; 朝红阳(1957—),女,教授,博士生导师,主要研究方向为视频编码、图像处理等。

由于视频中运动不规则而导致的预测运动向量不准确的问题。这类方法主要适用于高清视频播放领域,如高清电视、DVD等。这类方法虽然能提供好的图像质量,但要消耗庞大的计算量。

本文提出的一种快捷的高精度运动估计方法(以下简称快捷高精度方法),综合了高精度方法和简单方法的优点,在计算复杂度与简单方法相当的同时,可以达到高精度方法的精度。快捷高精度方法通过分析视频场景中不同区域的特性以及这些特性与运动向量的关系,在不同的区域采用不同的运动估计方法。第2节是简单方法与高精度方法的介绍;第3节是本文所提出的快捷高精度方法;第4节是实验的结果;最后是结论。

2 简单运动估计方法和高精度运动估计方法介绍

简单方法采用基于预测点的搜索技术。这类技术是在搜索区域中有条件地选定几个位置作为初始的搜索位置,即上面所提到的预测点。这些被选中的初始搜索位置也称作预测点。然后在这些预测点上应用一些简单的搜索模板进行搜索,如钻石型、方型和六角型。图1是其中一个简单方法(EPZS^[9])的示例。图中标有数字1和2的位置都是预测点,其中位置1同时也是中值运动向量。运动估计首先在这些预测点中选择率失真代价最小的位置,然后在该位置上使用钻石型模板搜索。钻石型模板一直移动到位置5停止。从图中可看出,简单方法需要检查的搜索位置很少。

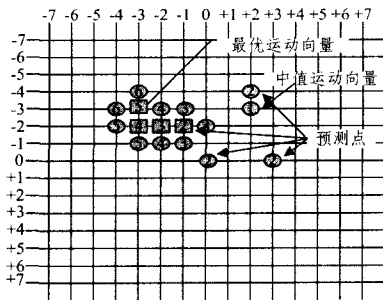


图1 简单预测运动搜索方法

简单方法找到最优运动向量的关键就是预测点的选择。目前经常使用的预测点分两类:一类是空间相邻运动向量,例如左、上和右上宏块的运动向量以及它们的中值运动向量;另一类是时域相邻的运动向量,就是相邻帧对应宏块的运动向量。这两类预测点只能在视频内容相关性高的情况下作为准确的起始搜索位置。

有3种情况会大大地影响这种简单方法的精度。首先,如果当前的运动向量与相邻的运动向量都没有关联,而预测点与实际的运动向量相差甚远时,简单的搜索方法很容易就中止在某一个局部最优的位置上。这种情况经常发生在视频场景切换的时候。其次,采用基于预测点的搜索技术,位于图像边缘的运动向量在极大程度上会影响所有运动向量的搜索。这是因为视频编码通常是采用光栅扫描的顺序,由上至下、从左到右逐行扫描。顶行的运动向量和最左列的运动向量都直接或间接地作为其余运动向量的预测点。这些基本预测点的精度同时也影响了其余运动向量的精度。第三,对特征不明显的宏块进行搜索,例如一堵墙或一块布,不同搜索位置所得到的残差值是十分接近的。在这些位置上搜索出来的运动向量会有一些的混乱性。使用这些运动向量作为其他运动向量的预测点会导致搜索不准确。

高精度方法的基本原理与简单方法十分相近,关键的区别之处就在于预测点的选择。高精度方法除了利用空间和时域相邻的运动向量外,还采用了一些全局的固定模板搜索。这些固定模板可以防止后续的搜索被中止在局部最优的位置上。图2是其中一个高精度方法(UMHexagonS^[7])的示例。首先,UMHexagonS使用数个固定的模板进行搜索。这些固定的模板分别是step2中的大十字型模板、step3-1中的大方型模板、step3-2中的大六边型模板。其次,UMHexagonS以上述固定模板搜索中的最优位置为中心,使用小六边型模板(step4-1)和钻石型模板(step4-2)细化搜索。虽然采用高精度方法能提供更好的搜索精度,但从图2中知道这也增加了大量的搜索点。在多数情况下,使用简单方法也能准确地搜索出正确的运动向量。在下一节中,本文根据不同宏块的特性混合地采用简单和高精度的方法,可达到既节省计算资源又保持高搜索精度的目的。

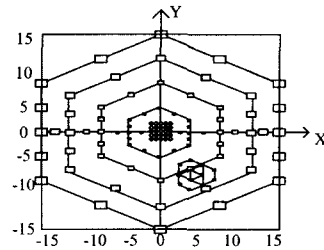


图2 高精度运动搜索方法

3 本文方法

上述的运动搜索方法都没有考虑到不同宏块的特性,对于所有的宏块都采用同样的运动估计方法。本文提出的快捷高精度方法,克服了单一使用简单方法或高精度方法的缺点,能根据不同宏块的特性自适应地选用不同的运动估计方法。

3.1 根据空间相邻运动向量的一致性选择运动估计方法

在一段内容连续的视频中,同一帧内的运动向量会表现出很强的相关性。当前宏块的运动向量可以用相邻宏块的运动向量比较准确地预测出来。简单方法就是利用这一特性快速地搜索出当前宏块的运动向量。而简单方法成功的关键是预测点必须落在与最优位置足够靠近的地方。如果某个区域运动向量的一致性比较好的话,在这个区域上使用简单方法所能达到的率失真性能会与使用高精度方法相同。本文提出使用运动向量差异(Motion Vector Discrepancy, MVD)来表示当前宏块所在邻域运动向量的一致性,其计算公式如下:

$$MVD = \frac{1}{n} \sum_{i=1}^n \frac{|MV_i - MV_{16 \times 16}|}{|MV_{16 \times 16}|} \quad (1)$$

其中, MV_i 是空间相邻运动向量,而 $MV_{16 \times 16}$ 是空间相邻运动向量的均值。MVD 越小,说明空间相邻运动向量越一致。反之亦然。图3是空间相邻运动向量和当前宏块运动向量的一个示意图。如果空间相邻运动向量的一致性比较好,即MVD比较小,那么当前宏块的运动向量也会接近这些空间相邻运动向量。MVD越小的宏块应该优先使用简单方法,而MVD越大的宏块则使用高精度方法。

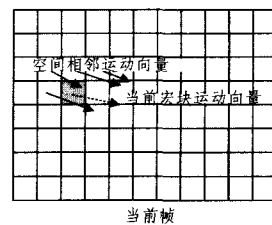


图3 当前宏块和其空间相邻运动向量

3.2 宏块的运动估计顺序

估计每幅图像的第一个宏块运动时,它没有空间相邻运动向量,即无法计算 MVD。如果使用简单方法,那么简单方法就会退化成为以中心点为起始的细化搜索。而后续的运动向量都是以它作为预测点,整体的运动搜索精确度都会因此受到影响。特别是在有全局运动的视频场景中,产生的影响就会特别明显。所以,对每一幅图像的第二个运动向量,应该使用高精度的方法进行搜索。

此外,位于每幅图像边缘的宏块,其运动向量很有可能是指向图像之外的。这时就算使用了高精度的搜索方法也不能正确得到实际运动向量的位置,所以从左到右、自上而下的光栅扫描的编码顺序就不适合作为运动向量的估计顺序。

本文提出了一种新的宏块扫描的顺序,如图 4 所示。选择第 2 行第 2 列的宏块 1 作为起始,搜索第一个运动向量,并采用高精度的运动估计方法保证其准确度。然后对宏块 2 进行简单的运动搜索。当这两个宏块完成运动搜索后,才对剩余的宏块按照通常的光栅顺序进行简单的运动搜索。这里只对第一个运动向量的搜索采用了高精度的运动估计方法,其余的宏块由其 MVD 的大小来决定要使用的运动估计方法。

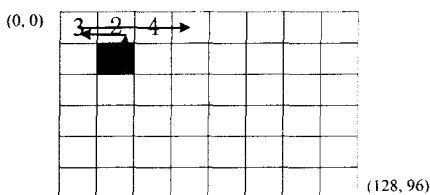


图 4 宏块搜索策略

3.3 预测点选取

如图 4 所示,宏块 1 采用高精度方法保证运动向量的准确度。高精度方法在没有任何预测点的情况下,对宏块 1 运动估计所产生的运动向量都能足够准确地反映视频内部的运动情况。而且,该运动向量可以作为一个可靠的预测点。其余的宏块可以采用空间相邻运动向量作为简单方法的预测点。采用传统的光栅扫描顺序时,空间相邻预测点只能是左、左上、上和右上这几个宏块的运动向量。但本文的运动估计的顺序已经不同于压缩编码的顺序,所以空间相邻的 8 个宏块上都可能有已经搜索出来的运动向量。这些相邻的运动向量都可以作为简单运动估计方法的预测点。例如图 4 中的宏块 2 采用的预测点就是宏块 1 的运动向量,而宏块 3 采用的预测点就是宏块 2 的运动向量。

3.4 搜索代价函数

在运动估计的过程中,搜索代价函数是评定最优运动向量的标准。目前大多数压缩器为了提高压缩性能,都采用了率失真优化技术。这种技术对于运动搜索最优点的判断并不仅仅根据图像块之间的残差值,同时考虑了运动向量本身所占码率的权重。式(2)就是利用拉格朗日乘法决定率失真优化代价函数:

$$J(m, \lambda_{\text{MOTION}}) = \text{SAD}(s, c(m)) + \lambda_{\text{MOTION}} \cdot R(m - p) \quad (2)$$

运动搜索的过程就是为了找到使这个代价函数的值最小的运动向量 m 。其中 λ_{MOTION} 是与量化参数相关的乘子, s 和 c 分别表示当前帧和参考帧, m 和 p 分别表示当前的运动向量和预测运动向量, $R(m - p)$ 表示对运动向量进行编码所消耗码流的位数, SAD 表示当前块与预测块的绝对残差和 (Sum

of Absolute Differences)。

估计图 4 中宏块 1 运动时采用的预测运动向量 p 为 $(0, 0)$, 而估计图 4 中宏块 2 运动时采用的预测运动向量 p 就是宏块 1 的运动向量。因为宏块 1 使用的是高精度搜索方法, 所以其运动向量表示的是视频真实的运动轨迹。宏块 3 采用的预测运动向量 p 就是宏块 2 的运动向量。其余宏块的预测运动向量仍根据不同压缩标准里的定义计算得出。选择宏块 1 的运动向量作为顶行宏块的初始预测运动向量, 可以使顶行的运动向量都有偏向于视频真实运动的趋势。

4 实验结果

本文使用的测试平台是 H. 264 的参考模型 (JM) 15. 2。其中除了第一帧为 I 帧外, 余下的都作为 P 帧。运动估计使用了 JM 中提供的快速算法。运动估计范围是 ± 8 像素, 不使用哈达玛变换, 并且应用了快速的率失真优化的快速选项。参考帧的数目为 JM 中可以选择的最大数目 (5 帧), 帧间的运动估计只使用 16×16 的模式。编码方式采用上下文自适应算术编码方式, 运动向量的精度为 $1/4$ 像素。

图 5 是分别采用高精度方法、简单点方法和本文的快捷高精度方法对 Stafan 序列进行逐帧质量比较的结果。其中高精度算法采用了 JM 中提供的 UMHexagonS 算法^[7], 简单预测点方法采用的是 EPZS 的改进算法^[4]。从图 5 可以看出, 本文的方法在质量上十分接近高精度方法, 大大高于简单方法。

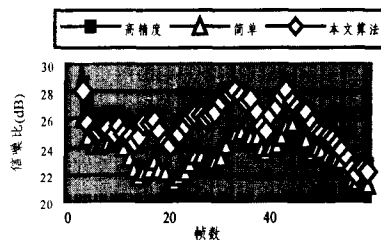


图 5 逐帧质量比较

表 1 提供了本文算法与 H. 264 JM 中 3 种运动估计算法 Full Search, UMHexagonS, Simplified UMHexagonS (简称 S-UMH) 的比较结果。实验所采用的测试序列是 CIF 格式的图像。运动情况从剧烈到平缓, 分别采用了 Mobile, Stanfan, Foreman, News 视频序列。

表 1 本文算法与 H. 264 JM 快速运动估计算法的比较结果

视频序列	运动估计方法	平均搜索点数	峰值信噪比(dB)	码率(bps)
Mobile	Full Search	289	39.01	2549.14
	UMHexagonS	116	39.32	2555.71
	S-UMH	101	38.99	2561.31
	本文算法	88	39.62	2563.12
Stafan	Full Search	289	42.01	2046.96
	UMHexagonS	216	41.32	2049.01
	S-UMH	201	40.76	2048.01
	本文算法	78	41.62	2048.02
Foreman	Full Search	289	39.07	1023.95
	UMHexagonS	216	38.81	1024.02
	S-UMH	187	38.86	1023.94
	本文算法	65	38.85	1022.99
News	Full Search	289	46.01	767.01
	UMHexagonS	216	45.27	768.81
	S-UMH	112	44.99	768.03
	本文算法	62	45.40	768.90

从表 1 的数据可以看出,本文的算法在速度上比 JM 提供的标准算法有很大程度的提高,同时在信噪比和码率上最高只有 0.03dB 和 0.12% 的增加。另外,从表中的数据还可以知道,本文的算法在运动比较平缓的视频序列 News 上有更好的速度上的优势,这是因为在静止的视频序列上应用简单算法是有效的。

结束语 本文通过结合高精度运动估计算法和简单预测点运动估计算法的优势,提出了一种能在速度上与简单算法相当、在准确度上与高精度算法媲美的快速运动估计算法。本算法通过计算空间相邻运动向量差异(MVD),自适应地判断采用何种运动估计方法。同时,提出了一种新的宏块运动估计顺序,其代替传统的宏块编码顺序作为新的运动估计的顺序,克服了位于图像边缘宏块运动估计的不准确性。根据实验表明,本文提出的快捷高精度运动估计方法,在相同码率和相同质量的情况下,能比原来的算法节省 70% 左右的搜索时间。

参 考 文 献

[1] Xu X, He Y. Improvements on Fast Motion Estimation Strategy for H. 264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(3): 285-293
 [2] Luo J, et al. Motion Estimation for Content Adaptive Video Compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(7): 900-909

[3] Chen T-C, et al. Fast Algorithm and Architecture Design of Low-power Integer Motion Estimation for H. 264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(5): 568-577
 [4] Ahmad I, et al. A fast adaptive motion estimation algorithm[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3): 420
 [5] Lee Y G, Ra J B. Fast Motion Estimation Robust to Random Motions Based on a Distance Prediction[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(7): 869-875
 [6] Qaralleh E A A, Chang T S. Fast Variable Block Size Motion Estimation by Adaptive Early Termination[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(8): 1021-1026
 [7] Chen Z, et al. Fast integer-pel and fractional-pel motion estimation for H. 264/AVC[J]. Journal of Visual Communication and Image Representation, 2006, 17(2): 264
 [8] Chao H, Lu J, Fisher P. Hybrid method for fast block based motion estimation in video coding[J]. Journal of Computational Information Systems, 2005, 1(2): 309
 [9] Tourapis A M, Cheong H-Y, Topiwala P. Fast ME in the JM reference software, JVT-P026. doc. in Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 16th Meeting[S]. Poznan, Poland, 2005

(上接第 231 页)

表 3 不同方法的实验结果

	Group one			Group two			Group three		
	Features	SVs	Accuracy	Features	SVs	Accuracy	Features	SVs	Accuracy
Ionosphere	34	171	0.942	9	127	0.953	9	91	0.943
Breast Cancer	31	279	0.978	8	114	0.976	8	88	0.970
Heart Disease	14	122	0.921	5	83	0.927	5	69	0.922

结束语 本文针对支持向量机方法对于高维大规模数据无法直接处理和对异常样本敏感的问题,提出了一种基于邻域粗糙集模型的改进支持向量机,从属性约简和样本选取两个方面对样本数据进行预处理。最终试验结果证实本方法在提高支持向量机效率,改善支持向量机结构方面是可行的。由于在本方法中涉及了多个参数的选取,因此需要进一步研究各参数的选取和分析它们之间的关系。

参 考 文 献

[1] 李国正,王猛. 支持向量机导论[M]. 北京:电子工业出版社, 2005
 [2] Shin H, Cho S. Fast pattern selection for support vector classifiers[J]. Lecture Notes in Artificial Intelligence, 2003, 2637: 376-387
 [3] Abe S, Inoue T. Fast training of support vector machine by extracting boundary data[C]// Proceeding of the International Conference on Artificial Neural Networks (ICANN). 2001: 308-313
 [4] 安金龙,王正欧. 预抽取支持向量的支持向量机[J]. 计算机工

程, 2004, 30(10): 9-11
 [5] 李红莲,王春花,袁保宗. 一种改进的支持向量机 NN-SVM[J]. 计算机学报, 2003, 26(8): 1015-1019
 [6] Hu Q H, Yu D R, Xie Z X. Neighborhood Classifiers[J]. Expert System with Applications, 2008, 34(2): 866-876
 [7] Hsu Chih-wei, Chang Chih-chung, Lin Chih-jen. A Practical Guide to Support Vector Classification[EB/OL]. <http://home.eng.iastate.edu/~julied/classes/ee547/Handouts/SVM-seguide.pdf>
 [8] Lin T Y. Granular Computing on binary relations I: data mining and neighborhood system[C]// Proc. of Rough Sets in Knowledge Discovery. Heidelberg, Germany: Physica-Verlag, 1998: 107-121
 [9] Li Ye, Cai Yun-ze, Li Yuan-gui, et al. Rough Set Method for SVM Data Processing[C]// Proceeding of the 2004 IEEE Conference on Cybernetics and Intelligent System Singapore. 2004: 1039-1042
 [10] 邓九英,杜启亮,毛宗源,等. 基于粗糙集与支持向量机的分类算法[J]. 华南理工大学学报, 2008, 36(5): 123-127
 [11] 张建明,曾建武,谢磊,等. 基于粗糙集的支持向量机故障诊断[J]. 清华大学学报, 2007, 47(S2): 1774-1777
 [12] 胡清华,赵辉,于达仁. 基于粗糙集的符号与数值属性的快速约简算法[C]//第七届中国 Rough 集与软计算学术会议. 山西,太原, 2007
 [13] 王凯,张永祥,姚晓山,等. 支持向量机惩罚参数的自适应调整方法[J]. 计算机工程与应用, 2008, 44(26): 45-47
 [14] Shin Hyunjung, Cho Sungzoon. Invariance of Neighborhood Relation Under Input Space to Feature Space Mapping[J]. Pattern Recognition Letters, 2004, 26(6): 707-718