

基于运动空间相关性和时间连续性的 运动估算快速算法

周海山 陈孝威 张 健

(贵州大学计算机科学系, 贵阳 550025)

摘要 运动估算是视频信号的帧间预测编码中的一个重要环节,其效率和精度直接影响到编码器的性能。由于全搜索算法搜索速度较低,而很少采用,故目前普遍采用三步法、交叉法等各种快速近似算法,但是这些算法匹配精度较低,而且某些情况下应用效果不好。为解决上述算法存在的问题,在对视频编码中运动物体的空间相关性和时间连续性进行分析的基础上,给出了一种利用运动物体的空间相关性和时间连续性来进行运动估算的快速算法。实验结果表明,该算法计算每个宏块运动矢量所需的平均搜索次数低于三步法,而匹配精度则非常接近于全搜索算法,并且采用该算法的编码器,其总的编码输出位数少于采用全搜索算法的编码器。

关键词 运动估算 视频压缩 视频编码 空间相关性 时间连续性

中图法分类号: TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2001)11-1125-05

A Fast Motion Estimation Algorithm Based on the Spatial Correlation and the Time Continuity of the Motion Object

ZHOU Hai-shan, CHEN Xiao-wei, ZHANG Jian

(Dept. Computer Science of Guizhou University, Guiyang 550025)

Abstract Motion estimation is a very important part in the inter prediction technical of video coding, and the performance of a video coder largely depends on the efficiency and precision of motion estimation. The full search algorithm is seldom adopted because of its enormous complexity of computation, and now the fast algorithms of motion estimation such as three-step algorithm and cross search algorithm are used widely. But the matching precision of these algorithms is low and in some cases they can not produce a good result. This paper first discusses the spatial correlation and time continuity of the motion object in video coding. Then, a new fast motion estimation algorithm based on the spatial correlation and time continuity of the motion object is presented. The average search times for computing motion vectors per macroblock of this algorithm are less than those of the three-step algorithm and the matching precision of this algorithm is very close to that of the full search algorithm by experiments. Moreover, the output bits of the coder using this algorithm are less than those of the coder using the full search algorithm.

Keywords Motion estimation, Video compression, Video coding, Spatial correlation, Time continuity

0 引 言

在现有的几个关于视频压缩编码的国际标准

中,帧间预测编码是得以减少图象冗余度的最有力的图象编码技术之一,而运动补偿技术中的运动估算又是帧间预测编码中的一个重要环节,因其效率和精度可直接影响编码器的性能,如今基于块的运

动作算法中最通用的方法是块匹配法。这种块匹配法,就是首先将当前帧均匀划分成若干个矩形块(宏块),然后在一定的匹配准则下,通过搜索上一帧,并找出同样大小的最佳匹配块来确定每一块的运动矢量。从计算量因素考虑,这种搜索通常限制在一定范围内(称为搜索窗口);最后采用全搜索算法,根据某种匹配准则,对搜索窗内的每一点进行评测,以便找出全局最佳运动矢量。虽然这种算法的匹配精度最高,但该方法相当费时,从而限制了其应用,而三步搜索、交叉搜索等快速算法又只能产生局部最优解,匹配精度较低。文献[3]中给出了一种具有全搜索匹配精度的快速算法(简记为筛选法),由于算法根据该文推出的一个不等式进行全搜索,从而避免了大量不必要的计算。

由于运动物体跨越几个相邻宏块的可能性很大,也就是说,一帧图象中绝大多数相邻宏块的运动矢量差别不会太大,因此利用好相邻宏块的运动矢量将有助于减少搜索范围和提高运动估算的速度,而要想利用相邻宏块的运动矢量,则必须首先判断出当前宏块和周围相邻宏块的相关程度(即它们的运动矢量的接近程度),只有当相关程度较好时,才能在相邻宏块运动矢量的基础上缩小搜索范围,进而提高搜索速度。文献[4]虽也考虑到了利用这种相关性,但是该文提出的块运动速度模型,实际上很难成立。另外,由于运动物体还具有时间上的连续性,这种时间上的连续性是指运动物体具有在多个连续帧之间保持相同速度和趋势的特点,因而,相邻两帧之间多数对应宏块的运动矢量也是相似的。可是三步搜索及交叉搜索等快速算法的共同的缺点恰恰是没有利用这种运动物体所具有的空间相关性和时间连续性。

本文提出了一种新的基于块匹配法的运动估算快速算法,由于它很好地利用了运动物体的空间相关性和时间连续性,因此其计算每个宏块运动矢量所需的平均搜索次数低于三步法,不仅匹配精度非常接近全搜索算法,而且采用该算法的编码器总的编码输出位数少于采用全搜索算法编码器的编码输出位数。

在块匹配法中,常用的块匹配准则有最大互相关函数、最小均方误差函数和最小绝对值误差函数(SAD)。因为SAD准则计算量小和硬件实现简单而被较多采用,所以本文采用SAD准则,其定义为

$$SAD(x, y) = \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} |f_k(m, n) - f_{k-l}(m+x, n+y)| \quad (1)$$

其中, $f_k(m, n)$ 表示第 k 帧中像素 (m, n) 的灰度值; $f_{k-l}(m+x, n+y)$ 表示第 $k-l$ 帧中像素 $(m+x, n+y)$ 的灰度值; $M_1 \times M_2$ 表示宏块的大小; x, y 分别为搜索时的水平、垂直方向的偏移量。

1 运动的空间相关性和时间连续性

1.1 空间相关性

如图1所示,当前宏块 $b(i, j)$ 的相邻宏块集 $N_{b(i,j)}$ 定义为:

$$N_{b(i,j)} = \{b(i-1, j-1), b(i-1, j), b(i-1, j+1), b(i, j-1), b(i, j+1), b(i+1, j-1), b(i+1, j), b(i+1, j+1)\} \quad (2)$$

$b(i-1, j-1)$	$b(i-1, j)$	$b(i-1, j+1)$
$b(i, j-1)$	$b(i, j)$	$b(i, j+1)$
$b(i+1, j-1)$	$b(i+1, j)$	$b(i+1, j+1)$

图1 相邻宏块的空间位置

把该集合的上、下两个半子集定义为:

$$\begin{aligned} N_{b(i,j)_{UP}} &= \{b(i-1, j-1), b(i-1, j), \\ &\quad b(i-1, j+1), b(i, j-1)\} \\ N_{b(i,j)_{LOW}} &= \{b(i, j+1), b(i+1, j-1), \\ &\quad b(i+1, j), b(i+1, j+1)\} \end{aligned} \quad (3)$$

若当前宏块 $b(i, j)$ 和相邻宏块集合 $N_{b(i,j)}$ 中某一宏块相关性较好,则它们的运动矢量差别很小。另外,宏块的相关性还分如下3种情况:

(1) $b(i, j)$ 与相邻宏块半子集 $N_{b(i,j)_{UP}}$ 中某一宏块相关性较好;

(2) $b(i, j)$ 与相邻宏块半子集 $N_{b(i,j)_{LOW}}$ 中某一宏块相关性较好,且 $b(i, j)$ 与相邻宏块半子集 $N_{b(i,j)_{UP}}$ 中所有宏块相关性较差;

(3) $b(i, j)$ 与相邻宏块集合 $N_{b(i,j)}$ 中所有宏块相关性均较差,即 $b(i, j)$ 与相邻宏块半子集 $N_{b(i,j)_{UP}}$ 中所有宏块相关性较差,且 $b(i, j)$ 与相邻子集 $N_{b(i,j)_{LOW}}$ 中所有宏块相关性也较差。

大家知道,在各种视频编码标准中,编码器是按照从上到下,从左到右的顺序逐块进行编码的,但由于在对当前宏块 $b(i, j)$ 进行编码时,只有 $N_{b(i,j)_{UP}}$ 集

合中宏块的运动矢量是已知的,而 $N_{b(i,j)_{low}}$ 集合中各宏块的运动矢量并不知道,因此,只有当宏块 $b(i,j)$ 的相关性满足情况(1)时,才可以利用相邻宏块的运动矢量来减少搜索范围。

1.2 时间连续性

时间上的连续性是指运动物体具有在多个连续帧之间保持相同速度和趋势的特点,因而,相邻两帧之间多数对应宏块的运动矢量也相似。为了利用这种连续性,设当前宏块 $b(i,j)$ 在前一帧中的对应宏块为 $b'(i,j)$, 然后就可将集合 $N_{b(i,j)_{up}}$ 修改为

$$N_{b(i,j)_{up}} = \{b(i-1,j-1), b(i-1,j), b(i-1,j+1), b(i,j-1), b'(i,j)\} \quad (4)$$

2 空间相关性和时间连续性的判断及算法描述

2.1 空间相关性和时间连续性的判断

在进行空间相关性和时间连续性判断之前,需先引入宏块复杂度的定义,这里将宏块复杂度 A 定义如下

$$A = \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} |f_k(m,n) - \bar{b}_k(m,n)| \quad (5)$$

其中, $M_1 \times M_2$ 表示宏块大小; $f_k(m,n)$ 表示第 k 帧图象中像素 (m,n) 的灰度值; $\bar{b}_k(m,n)$ 表示该宏块的平均灰度值。

$$\bar{b}_k(m,n) = \frac{1}{M_1 M_2} \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} f_k(m,n)$$

由式(5)可见,该值反映了宏块内相邻像素之间灰度的变化程度。通过对图象大量宏块 A 值的统计结果表明,一帧图象内各宏块的复杂度差别很大,其变化范围从几百至上万。

由于式(4)所表示的集合 $N_{b(i,j)_{up}}$ 已考虑了时间连续性,因此若利用当前要编码宏块 $b(i,j)$ 的空间相关性和时间连续性,只需考虑 $N_{b(i,j)_{up}}$ 集合中各宏块的空间相关性即可。

这里,首先分别将该集合中 5 个宏块的运动矢量作为当前要编码宏块的参考运动矢量,即先求出 5 个 SAD 值,记为 $SAD_1, SAD_2, SAD_3, SAD_4$ 和 SAD_5 。且令 $SAD_{min} = \min(SAD_1, SAD_2, SAD_3, SAD_4, SAD_5)$ 。很显然, SAD_{min} 反映了当前宏块与集合 $N_{b(i,j)_{up}}$ 的相关程度,即 SAD_{min} 越小,相关程度越好,再结合宏块的复杂度,就可以用 SAD_{min} 来判断当前宏块 $b(i,j)$ 的空间相关性和时间连续性;然后

再判断当前宏块与它在前一帧中的最佳匹配块的匹配情况,如果匹配得很好,那么这两块像素的灰度值分布应相近,也就是说,这两块像素灰度值的均值必然很接近,因此本文采用如下方法来判断当前宏块 $b(i,j)$ 的空间相关性和时间连续性:

若 $SAD_{min} < A + c$, 并且 $\bar{b}_k(m,n)$ 接近于 $\bar{b}'_{k-1}(m,n)$, 则 $b(i,j)$ 与相邻宏块子集 $N_{b(i,j)_{up}}$ 中某一宏块相关程度较好;否则 $b(i,j)$ 与相邻宏块子集 $N_{b(i,j)_{up}}$ 中所有宏块相关程度较差。其中, A 为当前宏块的复杂度, c 为一常量, $\bar{b}_k(m,n)$ 为当前宏块 b 的平均灰度值, $\bar{b}'_{k-1}(m,n)$ 为当前宏块 b 在上一帧(第 $k-1$ 帧)中的参考匹配宏块 b' 的平均灰度值, b' 由 SAD_{min} 所对应的宏块运动矢量来确定。

注:由于上述方法考虑了 $\bar{b}_k(m,n)$ 和 $\bar{b}'_{k-1}(m,n)$ 接近程度,故可提高判断的准确性。理由是:由于一帧图象中各宏块的复杂度 A 差别很大,小到几百,大至上万,因此对于 A 值较大的宏块,仅根据 $SAD_{min} < A + c$ 并不能保证 $b(i,j)$ 与相邻宏块子集 $N_{b(i,j)_{up}}$ 中某一宏块相关程度较好,这是因为虽然 $SAD_{min} < A + c$, 但是 SAD_{min} 与全局最佳运动矢量位置处的 SAD 值仍可能相差较大。

2.2 算法描述

本文用于表征两个宏块均值接近程度 B 的公式为

$$B = 1 - \frac{|\bar{b}_k(m,n) - \bar{b}'_{k-1}(m,n)|}{\bar{b}_k(m,n)}$$

这里,如果 B 设置两个阈值 $T_1, T_2 (T_1 < T_2)$; 并设变量 $s_{x,y}$ 表示搜索范围, δ_1, δ_2 和 δ_3 为事先确定的搜索范围, $\delta_1 < \delta_2 < \delta_3$; C 表示搜索中心; $V(SAD_{min})$ 表示 SAD_{min} 对应宏块运动矢量所指向的位置,则本文算法描述如下:

(1) 计算出当前宏块的复杂度 A, SAD_{min} 及 B 。转入第(2)步。

(2) 若 $SAD_{min} < A + c$, 且 $B > T_1$, 则当前宏块 $b(i,j)$ 与相邻宏块子集 $N_{b(i,j)_{up}}$ 中某一宏块的相关程度较好,转入第(3)步;否则,当前宏块 $b(i,j)$ 与集合 $N_{b(i,j)_{up}}$ 中宏块的相关程度较差,令搜索范围 $s_{x,y} = \delta_3$, 宏块的搜索中心 $C = (0,0)$, 转入第(4)步。

(3) 若 $B > T_2$, 则当前宏块与集合 $N_{b(i,j)_{up}}$ 中的某一宏块的相关程度非常好,令搜索范围 $s_{x,y} = \delta_1$; 否则,令 $s_{x,y} = \delta_2$ 。此时,宏块的搜索中心 $C = V(SAD_{min})$, 转入第(4)步。

(4) 以 C 为搜索中心, 以与 SAD_{\min} 对应的宏块运动矢量为初始运动矢量, 在设定的 $\pm s_x$ 搜索范围内, 采用筛选法进行搜索。

由上述算法步骤可见, 由于当 $b(i, j)$ 与相邻宏块子集 $N_{m, n, j_{\text{up}}}$ 中某一宏块相关程度较好时, 就减少搜索范围; 并且根据 B , 还可把搜索范围进行细化, 因此该算法大大提高了搜索速度, 其中, $T_1, T_2, \delta_1, \delta_2, \delta_3$ 和 c 的确定, 应根据具体应用的效率和精度要求来权衡。

3 算法性能测试

本文根据 ITU-H.324 协议, 在开发用于低码速多媒体通信的可视电话软件时, 采用了该算法, 得到了很好的效果, 实践结果证明了该算法的可靠性和有效性。

为了进一步证实该算法的可靠性和有效性, 本文还作了如下对比测试实验。

本文测试所用的图象序列是从摄像头摄入的, 图象格式为 QCIF, 宏块的大小为 16×16 ; 测试平台为符合 ITU-T H.263 建议的视频编码器; 算法中的参数设定为: $T_1 = 0.85, T_2 = 0.95, \delta_1 = 2, \delta_2 = 5, \delta_3 = 15, c = 500$ 。本文为评价算法的性能, 采用了以下的测度参数:

SNR_Y : 重建帧中亮度信号 Y 的信噪比;

SNR_{C_B} : 重建帧中色差信号 C_B 的信噪比;

SNR_{C_R} : 重建帧中色差信号 C_R 的信噪比;

$ASNP$: 计算每个宏块运动矢量平均所需的搜索次数 (计算 SAD 值的次数);

$ABSP$: 每帧图象编码输出的平均位数 (包括各种头信息和运动矢量的信息等);

AYP : 每帧亮度信号编码平均输出位数;

ACP : 每帧色差信号编码平均输出位数;

MV : 每帧运动矢量的编码平均输出位数。

其中, 信噪比 SNR (以亮度信号为例) 的计算公式如下:

$$SNR_Y = 10 \log_{10} \left\{ \frac{255 \times 255}{\sum_{m=1}^M \sum_{n=1}^N [f(m, n) - f'(m, n)]^2} \right\}$$

式中, $f(m, n)$ 表示原始图象中, 象素 (m, n) 的灰度值; $f'(m, n)$ 表示解码重建图象中, 象素 (m, n) 的灰度值; $M \times N$ 表示图象的大小。

本文通过大量图象序列的测试结果统计发现, 每帧图象中约 90% 以上宏块的搜索范围都在 $\pm \delta_1$ 之内。表 1 列出了采用本文算法的编码器对输入 3 个不同图象序列进行编码时 (量化步长 $Q=10$), 处于 $\pm \delta_1, \pm \delta_2$ 及 $\pm \delta_3$ 搜索范围中的平均宏块数目。由表 1 可见, 每帧图象中, 约 90% 以上宏块的搜索范围都在 $\pm \delta_1$ 之内。由于搜索速度的快慢很大程度上依赖于搜索范围较小的宏块数目的多少, 因而该算法极大地提高了搜索速度。

表 1

图象序列	宏块数目 (搜索范围: $\pm \delta_1$)	宏块数目 (搜索范围: $\pm \delta_2$)	宏块数目 (搜索范围: $\pm \delta_3$)
序列 1	94.9	2.6	1.4
序列 2	93.6	3.9	1.4
序列 3	89.3	9.1	0.6

表 2 中所列的数据, 是当量化步长 $Q=10$, 编码器按 10 帧/s 从每个序列中取 60 帧进行编码时, 有关测度参数的结果, 表中筛选法为一种具有全搜索匹配精度的快速算法。

表 2

图象序列	算法	SNR_Y	SNR_{C_R}	SNR_{C_B}	$ASNP$	$ABSP$	AYP	ACP	MV
1 ($Q=10$)	筛选法	33.71	36.75	37.65	215.6	2498	1527	40	467
	本文算法	33.70	36.77	37.60	21	2425	1561	35	368
2 ($Q=10$)	筛选法	33.29	36.84	37.49	202.3	3324	2346	58	470
	本文算法	33.27	36.88	37.52	20.6	3269	2396	63	373
3 ($Q=10$)	筛选法	35.36	39.95	40.30	283.4	743	143	4.5	285
	本文算法	35.29	39.90	40.47	22	684	161	5.4	208

从表 2 中可以看出:

(1) 采用本文算法后, 实际计算每个宏块 SAD 值的平均次数较筛选法有了大幅度的减少, 且低于三步法的搜索次数 (33)^[3]。

(2) 由于这种减少是利用了运动空间相关性和时间连续性的结果, 因而, 编码数据 Y, C_B, C_R 及重建帧的信噪比几乎没有受到影响。

(3) 采用本文算法的编码器, 每帧图象编码平

均输出的位数 $ABSP$ 比采用筛选法(全搜索匹配精度)的编码器有所减少,这是因为编码器的输出数据中除了压缩的图象数据外,还包括各种头数据和运动矢量的数据,而且利用全搜索算法按绝对值误差(SAD)准则找出最佳匹配块来进行编码,它只能保证压缩的图象数据最少,并不能保证总的输出数据最少^[5]。实际上,运动矢量数据所占的比例并不小,它仅次于图象的亮度数据。H. 263 是通过对比宏块的运动矢量和该宏块候选矢量的差值来进行 Huffman 编码,进而产生运动矢量编码数据的。按照 H. 263

建议,候选矢量为当前宏块左边、上边和右上角 3 个宏块运动矢量取中值的结果,而本文的算法,由于利用了空间相关性,并使相邻宏块的运动矢量趋于一致的可能性加大,因而减少了运动矢量的编码数据。

表 3 列出了对应于不同量化步长 Q 值的有关测度参数的数据,表中用于对比的算法仍为具有全搜索匹配精度的筛选法,该表中数据表明在不同的量化步长下(图象的重建质量不同),该算法都保持了较好的性能。

表 3

算法	Q=5		Q=10		Q=15		Q=20		Q=25	
	SNR-Y	AYP	SNR-Y	AYP	SNR-Y	AYP	SNR-Y	AYP	SNR-Y	AYP
筛选法	36.84	5572	33.71	1527	32.13	677	31.00	358	29.39	215
本文算法	36.83	5615	33.70	1564	32.13	691	30.97	359	29.35	218

4 结 语

在可视电话和视频会议等应用环境中,由于带宽的限制和实时性的要求,因而迫切需要快速和精准的运动估算算法。通过研究发现,在这些应用环境中,运动物体往往具有较高的空间相关性和时间连续性,由于本文算法较好地利用了这一特性,因此具有较高的搜索速度和匹配精度,且输出数据位数较低。实践证明,该算法完全能够满足可视电话或视频会议等应用中的视频编码器的实时性和精度要求,是一个性能较优、实用的块匹配算法。

参 考 文 献

- 1 Mura: A Tekalp 著. 数字视频处理. 崔之昂等译. 北京:电子工业出版社,1998.
- 2 黎洪松. 数字视频技术及其应用. 北京:清华大学出版社,1997.
- 3 Li W, Salari E. Successive elimination algorithm for motion estimation. IEEE Transactions on Image Processing, 1995, 4(1):105~107.
- 4 Fan J, Gan F. Motion estimation based on global and local uncomensability analysis. IEEE Transactions on Image Processing, 1997, 6(11):1584~1587.
- 5 Li R, Zeng B, Liou M L. A new three-step search algorithm for block motion estimation. IEEE Transactions On CS for Video Tech., 1994, 4(8):438~442.
- 6 Coban M Z, Mersereau R M. A fast exhaustive search algorithm for rate-constrained motion estimation. IEEE Transactions on Image Processing, 1998, 7(5):769~773.



周海山 1971年生,贵州大学计算机科学系计算机应用技术专业97级硕士研究生,主要研究方向为数字图象处理及计算机应用。



陈孝威 1945年生,贵州大学计算机科学系教授,主要研究方向为数字图象处理及计算机应用。



张 健 1977年生,贵州大学计算机科学系计算机应用技术专业99级硕士研究生,主要研究方向为数字图象处理及计算机应用。