

一种遗传搜索块匹配运动估计算法

许悦雷 毕笃彦 毛柏鑫 马林华

(空军工程大学工程学院航空电子工程系, 西安 710038)

摘要 运动估计是帧间视频编码中的关键技术,但现有的快速搜索算法中大都为次优算法,且易陷于局部极小点.针对此问题,提出了一种将遗传算法应用于块运动估计中的遗传搜索块匹配运动估计算法(GSAME).该方法把块运动向量作为遗传染色体,经过杂交、变异等操作,以便得到全局意义上的最优解,并与经典的全局搜索法和三步搜索法进行了比较.实验结果显示,该算法不仅有效地解决了局部极小问题,而且计算量也较少.

关键词 遗传算法 块运动匹配 运动估计

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2001)02-0164-04

A Genetic Search Algorithm for Motion Estimation

XU Yue-lei, BI Du-yan, MAO Bai-xin, MA Lin-hua

(Electronic Department of Air Force Institute of Engineering, Xi'an 710038)

Abstract Motion estimation is essential for many interframe video coding techniques, block-matching algorithms, such as FSA and TSS, have been widely used for motion estimation. The easiest implementation is the FSA, which evaluates all the blocks in the search window and has the highest computational cost. Therefore, many fast search algorithm including TSS, have been proposed to reduce the computational complexity, but most of them are based on the assumption that there should be only one optimal solution in the search window, however, in normal cases, there always exist multitudinous local optima, so they will miss the global optima, but get a suboptimal solution. In this paper, we propose a genetic search algorithm for motion estimation (GSAME) which applies genetic operation to motion estimation. We also introduce a scheme called competition evolution, which can bring the better solutions into the next evolution, and can accelerate the iteration process converging. In this method, the motion vector of block is defined as chromosome, after crossover, mutation and competition evolution, the global optimal solutions will be got. Last we compare the GASME to TSS, FSA, and the result shows that the method not only solve the problem of being trapped to local optima, but also have speed close to that of TSS.

Keywords Genetic algorithms, Block matching, Motion estimation.

0 引言

在图象压缩编码系统中,采用块运动匹配算法(BMA)可以有效地去除相邻帧之间的时间相关性.其一个典型的BMA算法就是,每帧图象被分成许多互不交叠的 $M \times M$ 大小的方块,然后在搜索窗范围内,对参考图象进行搜索,再以给定的匹配标准找到最佳匹配块,从而得到对应方块的运动向量.这里,若设纵向横向最大搜索位移为 W ,则搜索窗的大小为 $(2W+1) \times (2W+1)$.其中,最直接使用BMA的

是全局搜索法(FSA),由于该算法是对搜索窗内的所有点进行计算,所以可找到全局最优解,但却需高达 $(2W+1)^2$ 次的匹配计算量,因此很难应用于实时视频压缩中.为减少计算量,现在又出现了很多新的改进算法,如三步搜索算法(TSS)^[1]、交叉搜索算法(CS)^[2]以及二维对数搜索法^[3]等等,尽管这些算法在减少计算量方面有了很大的进步,并在实际工作中已被广泛采用,但它们大都有着各种限制条件和假设,其主要体现在:①匹配函数单调变化,如当搜索位置远离最优点时,块扭曲度(BDM)将单调增长;②一般情况下,每次搜索只有一个点被选中,这

就意味着当前点总比前面搜索到的点要优。但在实际应用中,匹配函数不一定是单调的,并且正确的最优优点往往隐含于搜索过程中的次优位置,这样会使算法仅找到一个局部极小点,而丢掉全局最优优点。

为解决局部极小问题,已出现了几种基于遗传算法(GA)的块运动匹配算法。其中,文献[4]提出的一种 GSA 算法,由于具有很高的迭代次数,故其计算耗时非常大,接近 FSA,所以很难应用于图象视频编码。文献[5]中提出的 LGSA 算法,虽然计算时间大大减少,但由于它未采用交叉操作,仅仅利用生存竞争策略控制下的高变异率去寻找全局最优,因此这样也会使算法质量变得不稳定。

本文提出的基于遗传算法的块匹配算法,不仅能有效地解决局部极小问题,而且具有很高的运行效率和较好的鲁棒性。

1 遗传算法的基本原理

遗传算法是对生物界自然选择和自然遗传机制进化过程的模拟,最初是由 Holland 在 70 年代提出来的^[6]。由于遗传算法能利用简单的编码技术和繁殖机制来表现复杂的现象,从而可以解决非常困难的问题,而且由于它不受搜索空间限制性假设的约束,不要求诸如单调、连续等假设,因此能以很大的概率快速找到全局最优解,如今它以其固有的计算并行性,已广泛应用于问题优化、模型识别、并行处理等领域。

在遗传算法中,问题的求解过程实质上就是一个迭代搜索过程,其重点在于适应规划和适应度度量方面。这里,适应规划用于指导算法怎样在空间进行搜索,常采用遗传操作,如交配(Crossover)和变异(Mutation)以及自然选择机制,而适应度度量则常采用计算适应值的方法,来评估一个候选解的优劣。用遗传算法求解问题有如下 5 个基本要素:①候选解的表示方式;②适应度函数的定义;③算法的控制参数与变量;④选择方法的确定;⑤算法终止准则。

2 基于 GA 的块匹配运动估值方法

在块运动估计中,寻找最佳匹配点实质上就是求问题最优解的过程,但经典的 BMA 大都存在局部最优和速度问题,而应用遗传算法则可以很好地达到全局寻优和快速高效的目的。

2.1 染色体的编码表示

这里,染色体采用二进制编码,由于每个运动向量有两个分量,故每个染色体均有两个元素,且每个元素的编码长度 l 由搜索窗的大小决定。若搜索范围为 $(-w, w-1)$,则 $l = \log_2(2w)$ 。

设第 i 个染色体为 C_i ,则

$$C_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} m_{i,l-1} & m_{i,l-2} & \cdots & m_{i,1} & m_{i,0} \\ n_{i,l-1} & n_{i,l-2} & \cdots & n_{i,1} & n_{i,0} \end{bmatrix} \quad (1)$$

$$m_{i,j} = \left\lfloor \frac{x_i + 2w}{2^j} \right\rfloor \bmod 2 \quad n_{i,j} = \left\lfloor \frac{y_i + 2w}{2^j} \right\rfloor \bmod 2 \quad (2)$$

其中, $i = 0, 1, \dots, N-1$; $j = 0, 1, \dots, l-1$

这里, $\lfloor \cdot \rfloor$ 表示取整时,只舍不入。mod 是取模操作, N 为群体个数。

2.2 适应度函数的确定

在经典运动估值算法中,常用的匹配标准有均方误差(MSE)和平均绝对误差(MAD),其中 MAD 不需作乘法运算就可得到与 MSE 相近的效果,故得到广泛应用,本文采用 MAD 值作为寻优度量,但 MAD 不符合遗传算法中适应度函数的条件,且其最优解不对应目标函数的最大值,但可以由 MAD 得到如下适应度函数

$$f(i) = d_{\max} - d_i + \delta_{d_{\max} - d_i} \quad (3)$$

其中, d 即为 MAD 值,且有

$$d_i = \frac{1}{N^2} \sum_{n_x, n_y \in B} |S_t(n_x, n_y) - S_{t-1}(n_x + x_i, n_y + y_i)| \quad (4)$$

式中, B 为当前要匹配的块, $d_{\max}$ 是 N 个 MAD 值的最大值, δ 是单位冲击函数。由式(3)可见, MAD 值越小,对应的适应值越大,相应的个体越优良,被选中去繁殖后代的机会也越大。适应函数中引入 $\delta_{d_{\max} - d_i}$ 是为了防止适应值出现零值。

2.3 控制参数的选择

GA 的控制参数主要有群体规模 N , 杂交率 p_c 和变异率 p_m 。其中,群体规模的选择会在很大程度上影响算法的质量,且对于运动估计问题, N 的选择与搜索窗口的大小有关,因此实现算法时,既要考虑到减少算法的计算量,又要防止过早收敛于局部最优。杂交过程是遗传算法中很重要的部分,杂交率的选择将直接关系到算法的性能,即 p_c 过高,群体中个体的更新越快,则高性能个体的破坏也越快,而 p_c 过低,相对来说,搜索范围就会变小,易造成算法停滞不前。变异过程一方面可以使在当前解附近找

到更好的解成为可能,另一方面又可保持群体的多样性,以防止陷入局部最优,一般选取较小的变异率以保证算法的质量.

2.4 选择方法的确定

选择策略要体现出如下原则,即高适应值的个体将有更多的机会去繁殖后代. 本文采用一种称为“转轮法”的方法来实现这种按比例选择,其基本步骤为:首先求出群体中所有个体的适应值及总和,然后产生一个在 0 与总和之间的随机数 r ;最后从群体中第一个个体开始,将其适应值与后代个体适应值相加,直到累加和大于等于 r ,这时将最后一个加入的个体选入交配池.

2.5 竞争进化策略的引入

在经典的 GA 算法中,是将得到的 N 个新个体,用来完全代替 N 个父代个体,这样容易造成振荡或收敛太慢. 本文的做法是:将完成遗传操作后得到的 N 个子代个体与原来的 N 个父代个体,按适应值从大到小排队,取前 N 个个体作为新一代群体. 这样就能把父代中适应值大的个体保留下来,从而可以加速算法的收敛.

2.6 终止准则的确定

迭代终止的条件有如下两个,且满足其中任何一个即可:其一是当迭代次数等于 $I(I = \log_2 w)$ 时终止;其二是当前群体平均适应值与上一代群体平均适应值的比值在区间 $[1, \alpha]$ 时,即终止迭代. 其中,在图象视觉应用中, α 的最优值为 1.005.

2.7 算法的具体步骤

(1)在搜索窗内选择 N 个初始点,这些点的位置是固定的(如图 1 所示).

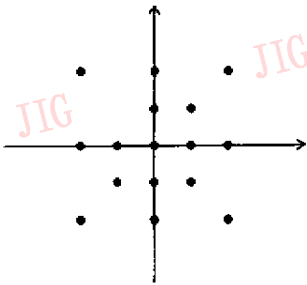


图 1 初始点位置

其中心点到其它点的向量就是初始群体的各个体,即 $\{C_1, C_2, \dots, C_{N-1}\}$.

(2)计算各个体的适应值 $f(i)$,以及适应值总和,其中, $i = 0, 1, \dots, N - 1, j$ 为迭代次数, $0 \leq j \leq I, j = 0$ 为初始状态.

(3)用转轮法选择 N 个用来繁殖后代的个体,并放入交配池.

(4)交配 从交配池中任选两个个体作为父、母代个体,根据预设的杂交率 p_c 进行交配或复制,重复该过程,最后得到 N 个子代个体.

(5)变异 根据变异概率 p_m ,对任一个体的任一位进行取反操作,最后得到后代群体.

(6)竞争进化 将得到的 N 个后代个体与 N 个父代个体共 $2N$ 个个体,按其适应值从大到小依次排列,取前 N 个个体形成下一代群体.

(7)判断是否满足终止条件,不满足条件时,则转到步骤(3),否则到下一步.

(8)得到最优匹配点.

3 实验结果及分析

本文在试验中采用的参数为:图象块大小为 16×16 pixel;杂交率 $p_c = 0.8$;变异率 $p_m = 0.04$,搜索范围为 $(-16, 15)$;最大迭代次数 $I = 4$;个体编码长度 $l = 5$. 为了测试本文提出算法的性能,分别使用了 3 种标准图象序列(Tennis、Flower 和 Miss-A)的前 80 帧进行试验,并将结果与经典的 BMA 算法 FSA 和 TSS 作了比较.

表 1 中列出了对 3 个图象序列分别应用 3 种算法所得到的平均 PSNR 值.由表 1 可见,本算法要优于 TSS 算法,接近或达到 FSA 的效果.这是由于本算法采用遗传操作,又引入了竞争进化策略,这样既可以保持在搜索空间内的不同区域进行多个点的搜索,使得群体具有多样性,同时又保证了群体的进化收敛,从而可避免陷入局部最优,以得到全局最优解.而 TSS 只是对于固定点的搜索,易陷于局部极小,所以效果要差于本算法.

表 1 三种搜索算法的 PSNR 值比较

搜索算法	图象序列			
	Tennis	Flower	Miss-A	Average
FSA	29.10	23.52	36.28	29.63
TSS	28.24	22.18	35.60	28.67
GSAME	28.64	22.50	35.75	28.96

本文分别将 3 种算法应用到 MPEG 编码器中,并采用软件编码方式进行编码,它们花费的运行时间关系如图 2 所示.

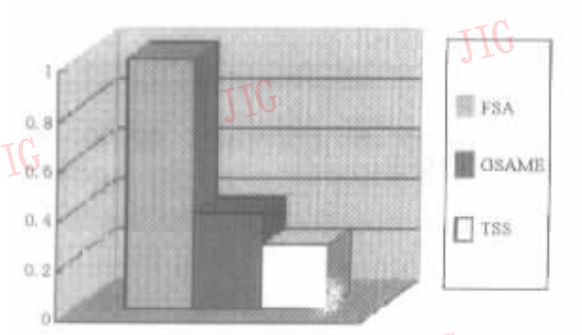


图 2 3 种算法在 MPEG 编码器中执行时间比较
由于 TSS 的搜索点固定为 25 个点,而本算法中每个图象块的根据实验得到的平均迭代次数约为 2.5,故其平均搜索点数为 40 左右,因此 TSS 的计算时间要比本算法稍快,不过本算法的计算复杂度与 FSA 相比要小得多,相应计算时间也要少得多.

4 结 论

本文将 GA 算法应用于块运动估计中,提出了一种基于 GA 的块匹配运动估计算法,并利用 GA 固有的全局寻优的特点,很好地解决了经典算法中搜索易陷于局部极小的问题. 尽管算法的速度低于 TSS,但其性能却接近 FSA,计算效率也比 FSA 要高很多. 从实验结果看,本文提出的 GSAME 算法优于两种经典算法,可以很好地应用于 MPEG 图象编码中,因而具有潜在的应用价值. 如果在初始点选择时考虑到运动矢量,并减少点数,可望进一步提高算法的运行效率.

参 考 文 献

1 Koga T, Iinuma K, Hirano A *et al.* Motion-compensated interframe coding for video conferencing. In :Proc. Nat. Telecommun Conf., New Orleans, LA, 1981 :G5.3.2G5.3.5.
2 Ghanbari M. The cross-search algorithm for motion estimation. IEEE Trans. Commun., 1990, 38 :950953.

3 Jain J R, Jain A K. Displacement measurement and its application in interframe image coding. IEEE Trans. Commun., 1981, COM-29 :17991808.
4 Chow K H.-K, Liou M L. Genetic motion search algorithm for video compression. IEEE Trans. Circuits Syst. Video Technol., 1993, 3 :440-445.
5 Lin Chun-Hung, Wu Ja-Ling. A lightweight genetic block-matching algorithm for video coding. IEEE Trans. Circuits Syst. Video Technol., 1998, 8 :386392.
6 Holland J H. Adaptation in natural and artificial systems (2nd ed)., Cambridge, MA :MIT press, 1992.

许悦雷 1998 年获空军工程学院航空电子工程系硕士学位,现为空军工程大学工程学院信号与信息处理专业硕士研究生. 研究方向为图象处理与模式识别.

毕笃彦 1986 年获国防科技大学电子技术系硕士学位,1997 年获法国图尔大学博士学位,现为空军工程大学工程学院航空电子工程系教授. 研究兴趣为图象处理、图象分析、数据压缩、模式识别.

毛柏鑫 1964 年毕业于南开大学数学系,现为空军工程大学工程学院航空电子工程系教授,从事信号与信息处理学科的教学和科研工作. 主要感兴趣的领域是图象处理、模式识别、数据压缩、数字信号处理.

马林华 1991 年获空军工程学院(现为空军工程大学工程学院)航空电子工程系硕士学位,现为空军工程大学工程学院航空电子工程系副教授,西安电子科技大学在职博士生. 研究方向为图象处理、图象分析、模式识别、多媒体技术.