

简析块匹配的估计算法

廖颖民

(南宁市科技馆 广西 南宁 530022)

摘要：块匹配的运动估计算法，是在一定的匹配准则下，对两个图像帧之间的像素域通过使用搜索算法来寻找最优的运动矢量估算值。在块匹配搜索过程中，区别不同算法的是搜索路径上的差异。当前在全搜索算法的基础上，陆续发展出三步搜索算法、四步搜索算法等新方法。

关键词：全搜索算法；三步搜索算法；四步搜索算法

引言

块匹配的运动估计算法，是在一定的匹配准则下，对两个图像帧之间的像素域通过使用搜索算法来寻找最优的运动矢量估算值。块匹配的算法在实现时，不同算法在最佳匹配准则、匹配块搜索过程及选择块的大小等方面会有不同，其中分块大小的选择主要体现在 H.264/AVC 的编码过程中，最佳匹配准则在不同运动估计算法中对精度的影响不大。在块匹配的搜索过程中，区别不同算法的是搜索路径上的差异。当前在全搜索算法基础上，发展出了三步搜索算法、四步搜索算法等新方法，本文对这三种算法进行详细分析。

1 全搜索算法

全搜索算法又称穷尽搜索法，该算法要求对搜索窗内所有的像素点都计算 MAD 值，从而计算出最小平均绝对值误差 (MBD)，其对应的位移量就是所求的运动矢量。该法具有简单可靠、精度高，全局最优点肯定能找到等优点，但这种算法也存在计算量大，难以实时处理等不足。实际应用中全搜索法通常只作为一种基准算法，或者将其搜索结果作为参照标准去衡量其它算法的优劣，它的方法如下：

第一步：在给定的搜索窗口内，按扫描式的搜索方式对块内所有的像素点进行搜索，分别计算各个像素点的平均绝对误差值并记录下来，直到将搜窗口内的所有像素点搜索完成。

第二步：在第一步所有记录下来的平均绝对误差值中寻找 MAD 值最小的点所在位置，这个位置的像素点就是最小块匹配误差点，该点所对应位置的偏移量就是运动估计算法所求的运动矢量。

全搜索法是所有块匹配算法中最简单、最有效的，由于它要对搜索窗口内的所有像素点都进行匹配计算，若单从匹

配的角度看，其必定能找到全局最优解，因而可作为其它快速块匹配算法性能比较的参考标准。但它要求对所有的像素点都进行搜索计算，这在视频压缩编码的实现过程中会占用到较大的运算量，难以在视频实时压缩中实际应用。为尽可能降低视频实时压缩时的运算量，在全搜索算法基础上，优化演变出了多种快速运动估计算法，如较有代表性的三步搜索算法及四步搜索算法。

2 三步搜索算法

三步搜索法是一种典型的快速运动估计块匹配算法。因早期的搜索范围为 ± 7 ，该法只需经三步计算即可结束，故称其为三步搜索法。在应用中若对其扩大搜索范围，该法的实际搜索过程可能会多于三步。该算法先从范围较大的区域进行搜索，再进行小范围搜索，即从原点 (0, 0) 作为搜索起始点，按一定的步长，取原点周围 8 个点构成一个每次需要搜索的点群，然后进匹配计算，跟踪 MBD 点，它的方法如下：

第一步：从搜索窗口原点 (0, 0) 位置开始搜索，以最大搜索长度的一半作为搜索步长，若所得的数值是小数，则取整。搜索方式按从左向右，从上到下的顺序对原点四周的 8 个像素点进行搜索，计算中心点及周围八个相邻点的 MAD 值，并记录下 MBD 点。

第二步：以上一步搜索到的 MBD 点作为下一步搜索的搜索中心，将步长减半，在缩小的方形上的 9 个像素点中寻找 MBD 点，以此类推，直到搜索步长减为 1。若步长为 1，则该点对应的位置就是所求的运动矢量。

三步搜索法能在搜索的速度与精度之间达到适当的平衡，是较早提出的快速算法。由于该算法的搜索步骤简单固定，且易于硬件实现，因而在很多视频压缩中得到了广泛应

用。如在一个 16×16 的像素子块, 使用三步搜索法只需搜索 25 个点, 而全局搜索法则需搜索 $15 \times 15=225$ 个点, 可以看出三步搜索法在搜索的点数上明显少于全搜索法, 这减少了视频压缩编码中的运算量。但三步搜索法有个较明显的不足, 就是在第一步搜索时, 由于搜索过程过大, 当搜索范围较大时, 其初始搜索步长也将会变得很长, 需要搜索的面积很大, 易造成因搜索范围过大而跳出运动矢量存在可能性较大的区域。这样将导致算法在搜索过程中陷入局部最优的问题, 这对搜索精度会造成不利影响。为解决该问题, 人们经过对三步搜索法的改进, 出现了四步搜索算法。

3 四步搜索算法

四步搜索算法对三步搜索算法中第一步搜索步长过大的不足进行了改进, 它综合考虑了块的中心偏移及大范围运动特性。四步搜索法在每一步的搜索范围上, 都是由上一步的最佳匹配位置决定的。该法在搜索前三步的步长是固定的, 到了最后一步才对步长进行修改, 以此得到最优匹配位置, 它的方法如下:

第一步: 从搜索窗口的原点 $(0, 0)$ 位置开始, 以 2 为搜索步长, 计算中心位置及周围等步长的 5×5 方形上 9 个点的 MAD 值。如得到的最小 MAD 值的点在窗口的中心位置, 或搜索点到达预定搜索窗口的边界上, 则跳转到第三步, 否则执行第二步。

第二步: 搜索窗口的位置依照上一步的结果而定, 如上一步得出的 MBD 点在方形的四个顶点上, 则需计算额外 5 个点的 MAD 值; 如第一步得出的 MBD 点在方形的四边上, 则需计算额外 3 个点的 MAD 值。若这一步得出的 MBD 点在窗口的中心位置, 则直接转到第三步, 否则重复执行第二步。

第三步: 将搜索步长改为 1, 计算 3×3 方形上的 9 个点的 MAD 值, 得出的 MBD 点即为最终的匹配位置, 算法

终止, 此位置即为所求的运动矢量。

四步搜索法在三步搜索法的基础上进行了优化, 其搜索过程中最多所搜索的点数也仅比三步搜索法多 5 个点, 但四步搜索算法考虑了块的中心偏置特性, 只有极少数情况下需要计算 30 个像素点, 且前两步是进行定步长搜索, 解决了三步搜索法存在的陷入局部最优问题, 且该法在搜索的同时又兼顾了块的大范围运动特性, 因而该算法处理的图像质量会优于三步搜索法。

4 结束语

块运动估计算法的评价依据主要是看算法的匹配效果及搜索时间的长短。匹配的效果可通过人眼对重构图像的质量进行评价, 但这种评价往往具有较大的主观性, 不同人的评价差异较大, 很难进行定量的分析比较。为此需要对图像质量进行更为客观的评价, 在实际应用中通常是采用图像的平均峰值信噪比 (PSNR) 来衡量, 其计算公式为:

$$PSNR = 10 \log_{10} \left(\frac{\phi_{\max}^2}{MSE} \right)$$

其中 ϕ_{\max} 是视频信号的峰值强度, 普通的彩色 8 比特视频, 其 $\phi_{\max}=255$ 。搜索时间的复杂程度主要体现在搜索速度上, 可通过搜索点数及搜索时间来进行对比。但往往同一种算法在不同硬件平台上的搜索时间也是不一致的, 因而在实际应用中一般通过搜索点数来进行对比, 即对比不同算法在搜索过程中进行匹配的次数来评价该算法的搜索速度。

参考文献:

- [1] 贾永红. 数字图像处理. [M]. 武汉: 武汉大学出版社, 2015
- [2] 陈天华. 数字图像处理及应用 [M]. 北京: 清华大学出版社, 2018
- [3] 杨帆, 王志陶, 张华. 精通图像处理经典算法 [M]. 北京: 北京航空航天大学出版社, 2018