

## 基于子块均值的快速块匹配算法

郭 丽,孙兴华,杨静宇

(南京理工大学 计算机科学与技术学院,江苏 南京 210094)

(liguo1975@gmail.com)

**摘 要:**根据积分图像对匹配误差函数进行变形,提出了一种新的局部匹配误差函数——基于子块均值的匹配误差函数。实验表明,基于子块均值的匹配误差函数与完全匹配误差函数相比具有相当的运动估算质量和更快的运动估算速度,与基于子抽样的匹配误差函数相比同时具有更优的运动估算质量和速度。子块划分的恒常性,也保证了基于子块均值的匹配误差函数对于不同的匹配图像所消耗的时间大致相同,非常适合于实时视频压缩等应用。

**关键词:**块匹配;积分图像;子块均值;匹配误差函数;峰值信噪比

**中图分类号:** TP391.41 **文献标识码:** A

## Fast block matching algorithm based on subblock mean

GUO Li, SUN Xing-hua, YANG Jing-yu

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing Jiangsu 210094, China)

**Abstract:** Integral image as an intermediate image representation can be used to calculate the sum of gray level in rectangle quickly. A novel partial matching error function was presented based on sub block mean. Experiments show that the matching error function based on subblock mean is superior to both the full matching error function and the matching error function based on sub-sampling in terms of the motion estimation quality and speed. The matching error function based on subblock mean with the constant subblock division costs almost the same time to different matching images, which is very suitable for the real-time applications such as video compression.

**Key words:** block matching; integral image; subblock mean; matching error function; peak signal noise ratio

块匹配技术已被广泛地用于立体视觉、视觉跟踪和视频压缩等应用。迄今为止学者们已提出多种有效的块匹配算法,主要在匹配误差函数、搜索过程和块尺寸的选择这三个方面进行了改进。为了加速每个搜索点的匹配误差计算,学者们提出了用局部匹配误差函数来代替完全匹配误差函数。在局部匹配误差函数中,一个简单的技术就是在匹配块中进行子抽样<sup>[1,2]</sup>。局部匹配误差函数的另一种技术称为局部距离的方法,通过定义局部绝对差值和函数作为局部匹配误差函数来加速匹配误差的计算<sup>[3]</sup>。最简单的搜索过程就是将“搜索窗口”中所有的像素都搜索一遍,这种完全搜索过程能够找到全局最小匹配误差,但是非常耗时。基于单峰误差平面假设的梯度下降技术,可以找到具有局部最小匹配误差的运动向量<sup>[4]</sup>。梯度下降技术可以在多精度的运动向量空间中进行由粗到细的搜索过程,这种技术包括三步搜索算法、二维对数搜索算法和交叉搜索算法。选择一个合适的块尺寸对于任何基于块的运动估算方法都是最基本的步骤,多级块匹配法提供了很好的解决方案<sup>[5]</sup>。本文根据积分图像的快速计算矩形像素灰度和的表达机制,提出了一种新的局部匹配误差函数——基于子块均值的匹配误差函数,用来进一步加速块匹配过程。

### 1 基于子块均值的快速块匹配

#### 1.1 积分图像

积分图像是对原始图像的一种中间表示<sup>[6]</sup>,积分图像上

$(x, y)$  位置的值就是原始图像的左上角到坐标 $(x, y)$ 的矩形内所有像素的灰度值之和,可以表示为  $I_f(x, y) = \sum_{i=0}^x \sum_{j=0}^y f(i, j)$ , 其中  $I_f(x, y)$  为积分图像,  $f(x, y)$  为原始图像。积分图像可以进行递归定义:

$$\begin{cases} S_f(x, y) = S_f(x, y-1) + f(x, y) \\ I_f(x, y) = I_f(x-1, y) + S_f(x, y) \end{cases}$$

其中  $S_f(x, y)$  是每一行的累积和,  $S_f(x, -1) = 0$  且  $I_f(-1, y) = 0$ 。根据递归定义,积分图像只需要经过一遍扫描即可计算完成。

在积分图像  $I_f(x, y)$  的内部表示下,可以很快计算一个矩形内所有像素的灰度值之和。已知矩形  $D$  四个顶点的坐标来求取矩形  $D$  内所有像素的灰度值之和,其中顶点 1 的坐标为  $(x_{\min}, y_{\min})$ , 顶点 4 的坐标为  $(x_{\max}, y_{\max})$ 。用  $S$  来表示一个矩形内所有像素的灰度值之和,那么在原始图像  $f(x, y)$  的表示下,有  $S(D) = \sum_{i=x_{\min}}^{x_{\max}} \sum_{j=y_{\min}}^{y_{\max}} f(i, j)$ ; 在积分图像  $I_f(x, y)$  的表示下,有  $S(D) = I_f(x_{\max}, y_{\max}) - I_f(x_{\min}-1, y_{\max}) - I_f(x_{\max}, y_{\min}-1) + I_f(x_{\min}-1, y_{\min}-1)$ 。可以很清楚地看出,  $S(D)$  的计算在原始图像的表示下需要  $(x_{\max} - x_{\min} + 1) \times (y_{\max} - y_{\min} + 1) - 1$  次加法,在积分图像的表示下只需要 3 次加(减)法。也就是说,在积分图像的表示下矩形像素灰度和的计算比在原始图像的表示下要快(除非出现矩形内不多于 4 个像素的情况,这在块匹配中出现的概率很小)。

收稿日期:2006-01-19;修订日期:2006-03-06

作者简介:郭丽(1975-),女,江苏扬州人,讲师,博士,主要研究方向:基于内容的图像检索、数字图像处理; 孙兴华(1975-),男,江苏阜宁人,副教授,博士,主要研究方向:多媒体处理、计算机视觉; 杨静宇(1941-),男,教授,博士生导师,主要研究方向:模式识别、人工智能。

## 1.2 匹配误差函数

从当前帧  $k$  到搜索帧  $k+l$ , 对于像素  $(x, y)$  和中心定位于  $(x, y)$  的  $N_1 \times N_2$  块,  $(x, y)$  的候选位移矢量为  $(u, v)$ , 匹配误差函数 (Matching Error Function, MEF) 记为  $MEF_{(x, y)}(u, v)$ , 那么运动矢量的估算被转换为求  $(u, v)$  对应匹配误差的最小值, 即  $[\hat{u} \hat{v}]_{(x, y)}^T = \arg \min_{(u, v)} MEF_{(x, y)}(u, v)$ .

### 1.2.1 完全匹配误差函数

常用的完全匹配误差函数包括平方误差和函数 (Sum of Square Error, SSE)、绝对差值和函数 (Sum of Absolute Difference, SAD) 和非匹配像素统计函数 (Non-matching Pixel Calculation, NPC), 其匹配模式就是计算匹配块中所有对应像素之间的距离, 然后逐个像素地对距离进行求和操作, 可以公式化为  $FULL\_MEF_{(x, y)}(u, v) = \sum_{i=-N_1/2}^{N_1/2} \sum_{j=-N_2/2}^{N_2/2} \zeta(f_k(x+i, y+j), f_{k+l}(x+i+u, y+j+v))$ , 其中  $\zeta(x_1, x_2)$  是距离函数, 不同的匹配误差函数对应不同的距离函数, 例如平方误差和函数对应的距离函数为  $\zeta^{SSE}(x_1, x_2) = (x_1 - x_2)^2$ .

### 1.2.2 基于全局均值的匹配误差函数

根据积分图像对匹配误差函数进行变形, 一个最简单的思想就是首先计算整个匹配块的灰度均值, 然后根据距离函数计算灰度均值距离作为匹配误差。将这种变形匹配误差函数称为基于全局均值 (Global Block Mean, GBM) 的匹配误差函数, 其计算公式为  $GBM\_MEF_{(x, y)}(u, v) = \zeta\left(\frac{1}{N_1 \times N_2} \sum_{i=-N_1/2}^{N_1/2} \sum_{j=-N_2/2}^{N_2/2} f_k(x+i, y+j), \frac{1}{N_1 \times N_2} \sum_{i=-N_1/2}^{N_1/2} \sum_{j=-N_2/2}^{N_2/2} f_{k+l}(x+i+u, y+j+v)\right)$ 。很显然, GBM\_MEF 只利用了匹配块的均值信息, 而忽略了匹配块的分布信息, 所得到的运动估算结果会有较大的误差。

### 1.2.3 基于子块均值的匹配误差函数

GBM\_MEF 作为匹配误差函数效果不佳, 主要是缺少匹配块的分布信息, 而这可以通过将匹配块均匀划分为多个子块, 计算匹配块中所有对应子块均值之间的距离, 然后逐个子块地对距离进行求和操作来解决。将匹配块划分为多个子块, 可以提供匹配块的分布信息; 子块均值的计算, 可以根据积分图像进行加速。只要对匹配块进行合理的子块划分, 这种基于子块均值的匹配误差函数一定可以在运动估算的质量和速度之间取得一个很好的平衡。

将每个匹配块均匀划分为  $N \times N$  ( $N \leq N_1$  且  $N \leq N_2$ ) 的子块, 每个子块的尺寸为  $\frac{N_1}{N} \times \frac{N_2}{N}$ , 那么基于子块均值 (Sub Block Mean, SBM) 的匹配误差函数的计算公式为  $SBM\_MEF_{(x, y)}^{N \times N}(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \zeta(SBM_{f_k}(X, Y), SBM_{f_{k+l}}(X+u, Y+v))$ , 其中  $X = x - N_1/2 + (1/2 + i)N_1/N$ ,  $Y = y - N_2/2 + (1/2 + j)N_2/N$ ,  $SBM_f(X, Y)$  是以像素  $(X, Y)$  为中心的子块均值  $SBM_f(X, Y) = \frac{1}{\frac{N_1}{N} \times \frac{N_2}{N}} \sum_{i=X-\frac{N_1}{2N}}^{X+\frac{N_1}{2N}} \sum_{j=Y-\frac{N_2}{2N}}^{Y+\frac{N_2}{2N}} f(i, j)$ , 它在

积分图像表示下的计算公式为  $SBM_f(X, Y) = \frac{1}{\frac{N_1}{N} \times \frac{N_2}{N}} \left( I_f\left(X + \frac{N_1}{2N}, Y + \frac{N_2}{2N}\right) - I_f\left(X - \frac{N_1}{2N} - 1, Y + \frac{N_2}{2N}\right) - I_f\left(X + \frac{N_1}{2N}, Y - \frac{N_2}{2N} - 1\right) + I_f\left(X - \frac{N_1}{2N} - 1, Y - \frac{N_2}{2N} - 1\right) \right)$ .

$$Y - \frac{N_2}{2N} - 1) + I_f\left(X - \frac{N_1}{2N} - 1, Y - \frac{N_2}{2N} - 1\right) \Bigg)。$$

如果  $N$  等于 1, 那么  $SBM\_MEF_{(x, y)}^{N \times N}(u, v) \big|_{N=1} = GBM\_MEF_{(x, y)}(u, v)$ 。也就是说, 当  $N$  等于 1 时, SBM\_MEF 就退化为 GBM\_MEF。  $N$  越大, SBM\_MEF 的运动估算质量越接近 FULL\_MEF; 而  $N$  越小, SBM\_MEF 的运动估算速度也就越快。能够使得在运动估算的质量和速度之间取得最佳平衡的  $N$  称为最佳子块划分。

## 2 实验分析

块匹配算法主要包括匹配误差函数、搜索过程和块尺寸的选择这三个方面。本文重点的研究对象是匹配误差函数, 通过与完全匹配误差函数比较进行基于子块均值匹配误差函数的分析, 实验中所采用的距离函数为平方误差函数。搜索过程选用二维对数搜索算法, 在块尺寸的选择上对多级块匹配进行实验。用峰值信噪比 (Peak Signal Noise Ratio, PSNR) 评价运动估算的质量, 用消耗时间 (单位为 ms) 评价运动估算的速度。实验环境为 Pentium III-900MHz 处理器、256M 内存、Windows 2000 + Service Pack 3 的 IBM ThinkPad 笔记本电脑, 编程工具为 Microsoft Visual C++ 6.0。

### 2.1 最佳子块划分的确定

最佳子块划分的确定是基于子块均值匹配误差函数的关键。基于子块均值的匹配误差函数属于局部匹配误差函数, 一般来说其运动估算质量要逊于完全匹配误差函数, 那么最佳子块划分就可以认为是, 相对完全匹配误差函数的运动估算质量损失在一定可接受范围之内, 而运动估算速度达到最大时的子块划分。MPEG 委员会使用 0.5dB PSNR 值作为非正式阈值来决定是否结合一种编码优化, 因为他们认为小于 0.5dB 幅度的改进将是不可见的。反之可以认为 0.5dB PSNR 值的运动估算质量损失是可以接受的, 最佳子块划分就是相对完全匹配误差函数的 PSNR 差值大于 -0.5dB 的最小子块划分。

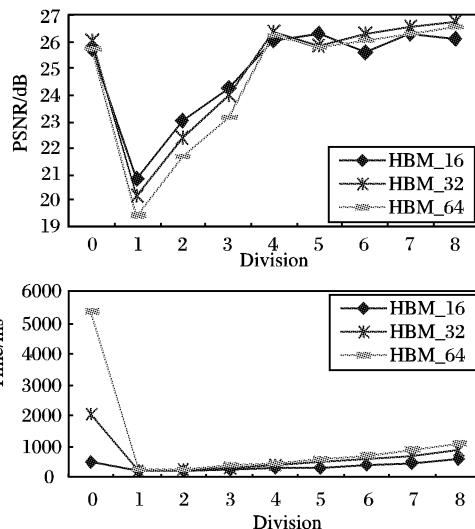


图1 用 Synthetic 标准视频序列进行基于子块均值的块匹配

我们绘制峰值信噪比/子块划分 (PSNR/Division) 曲线图和消耗时间/子块划分 (Time/Division) 曲线图, 其中子块划分从 0 到 8, 0-子块划分表示完全匹配误差函数。多级块匹配则采用 3 级、4 级和 5 级块匹配, 分别用 HBM\_16、HBM\_32 和

HBM\_64 来表示。用 Synthetic 标准视频序列进行基于子块均值的块匹配,图1为视频序列中所有相邻图像帧多级块匹配的结果平均。可以看出,1-子块划分具有最低的 PSNR 值和最少的消耗时间,随着子块划分的增加 PSNR 值和消耗时间都随之增加。在多级块匹配的情况下,到达4-子块划分的 PSNR 值与完全匹配误差函数的差值都大于  $-0.5\text{dB}$ ,4-子块划分后续的 PSNR 值都维持在与4-子块划分相当的水平上。大量实验表明,在多级块匹配的情况下4-子块划分符合最佳子块划分的要求,最佳子块划分所对应的消耗时间低于完全匹配误差函数。

## 2.2 多个匹配误差函数对比

用26个标准视频序列对基于不同匹配误差函数的多级(5级)块匹配进行比较,分别为完全匹配误差函数、基于子块均值的匹配误差函数和基于子抽样的匹配误差函数<sup>[2]</sup>。对于每一个标准视频序列,对视频序列中所有相邻图像帧的块匹配结果进行平均,作为视频序列对应的运动估算参数。完全匹配误差函数用 FM(Full Matching)来表示;基于子块均值的匹配误差函数用 SM(Sub Block Mean)来表示,采用4-子块划分(最佳子块划分);基于子抽样的匹配误差函数用 SS(Sub-Sampling)来表示,其每4个像素(X和Y轴各2个像素)进行一次子抽样。图2给出了相应的峰值信噪比/视频序列曲线图和消耗时间/视频序列曲线图。

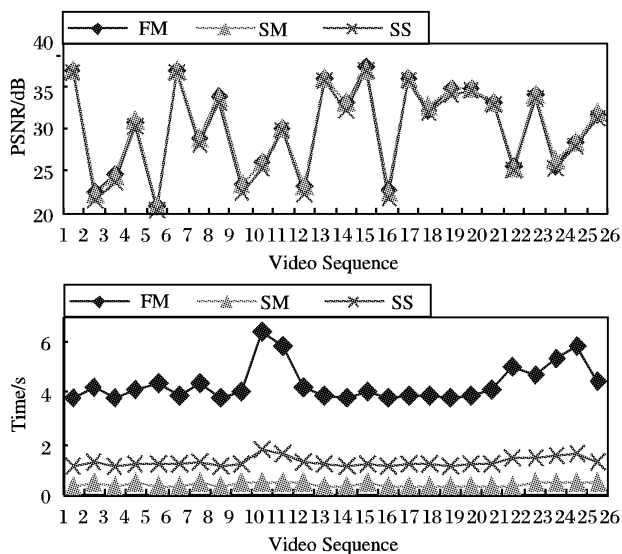


图2 用26个标准视频序列比较不同匹配误差函数的多级块匹配

根据图2可以得出,基于子块均值的匹配误差函数对应的所有 PSNR 值与完全匹配误差函数相当接近,符合最佳子块划分的  $0.5\text{dB}$  阈值要求,而基于子抽样的匹配误差函数对应的 PSNR 值与完全匹配误差函数的差距则大多超出了  $0.5\text{dB}$  阈值。在消耗时间方面两种局部匹配误差函数都快于完全匹配误差函数,其中基于子块均值的匹配误差函数要成

倍地快于基于子抽样的匹配误差函数。而且,基于子块均值的匹配误差函数相对于其他两种匹配误差函数还有一个优点,就是对于不同的匹配图像其所消耗的时间大致相同,这非常适合于实时视频压缩等应用。

为了进一步地说明,我们对使用相同匹配误差函数的所有标准视频序列的运动估算参数进行平均,并制成表格(见表1)。从表1可以得出,在  $0.5\text{dB}$  阈值的衡量下基于子块均值的匹配误差函数所对应的 PSNR 值与完全匹配误差函数相当,而基于子抽样的匹配误差函数所对应的 PSNR 值则不符合要求。基于子块均值的匹配误差函数所对应的消耗时间为完全误差函数的 9%,而相对基于子抽样的匹配误差函数则为 30%。实验表明,基于子块均值的匹配误差函数作为一种局部匹配误差函数,取得了与完全匹配误差函数相当的运动估算质量和快于完全匹配误差函数的运动估算速度;与基于子抽样的匹配误差函数相比具有更好的运动估算质量和更快的运动估算速度。

表1 基于不同匹配误差函数的多级块匹配性能比较

匹配误差函数	PSNR/dB	Time/ms
FM	30.06	4407.27
SM	30.09	389.23
SS	29.53	1284.31

本文进一步的研究工作有两点:一是将基于子块均值的匹配误差函数与局部距离技术结合起来,以提供更快的匹配速度;二是将子块划分的恒常性与子抽样进行结合,就是对子块抽取单个像素代替均值参与匹配,并与基于子块均值的匹配误差函数进行比较。

## 参考文献:

(上接第1589页)

- [7] 王姝华,李佐,蔡士杰. 基于最小二乘法的文档图像倾斜检测方法[J]. 计算机应用与软件, 2001, 18(9): 43-46.
- [8] PSTL W. Detection of Linear Oblique Structure and Skew Scan in Digitized Documents [A]. Proceeding of English International Conference on Pattern Recognition[C], 1986. 687-689.
- [9] YAN H. Skew Correction of Document Images Using Inner Line Cross-correlation [J]. Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing, 1993, 55(6): 538-543.
- [10] KAO C-H, DON H-S. Skew Detection of Document Images Using Line Structural Information [A]. Proceedings of the Third International Conference on Information Technology and Applications[C], 2005.
- [11] SHI Z, GOVINDARAJU V. Skew detection for complex document images using fuzzy runlength [A]. Proceedings of Seventh International Conference on Document Analysis and Recognition [C]. Edinburgh, Scotland, 2003. 715-719.
- [12] 陈宝林. 最优化理论与算法[M]. 北京: 清华大学出版社, 1989.