

基于改进型 C-1BT 变换的快速全局运动估计*

席志红, 初守艳

(哈尔滨工程大学 信息与通信工程学院, 哈尔滨 150001)

摘要: 针对存在单一运动目标视频序列的全局运动估计问题, 提出一种结合 C-1BT 变换和自适应十字搜索法 (ARPS) 的快速块匹配运动估计方法。采用简单的预处理, 提高灰度突然变化和存在局部运动目标时的全局运动估计的鲁棒性; 简化传统 C-1BT 变换算法中的核函数, 降低计算负担; 考虑到实际视频中存在大量的零运动矢量 (ZMV) 区域和视频的空间相关性, 改进原有的 ARPS 搜索法。MATLAB 仿真结果证明, 提出的算法在保证精度的同时极大地减少了搜索点数。

关键词: 全局运动估计; C-1BT 变换; ARPS; 灰度归一化; 投影剪切

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 1001-3695(2012)09-3554-04

doi: 10.3969/j.issn.1001-3695.2012.09.094

Fast global motion estimation based on improved C-1BT transformation

XI Zhi-hong, CHU Shou-yan

(Dept. of Information & Communication Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Considering the video sequences with a single moving object, this paper proposed a new fast block matching motion estimation approach, which combining C-1BT transformation and the ARPS approach. The proposed pre-processing step was simple, and could increase the robustness of the global motion estimation in the intensity suddenly changed and existence of local motion of objects. Then, it used a simplified C-1BT transformation kernel function to reduce the computational burden. In the search process, improved ARPS algorithm not only used spatial correlation, but also used a lot of early termination of the search methods. MATLAB simulation results show that the proposed approach can greatly reduce the search points while ensuring accuracy.

Key words: global motion estimation; C-1BT; ARPS; gray-scale normalization; projection clip

全局运动估计 (GMV) 技术已经广泛应用在视频编码、抖动视频序列的稳定、图像拼接、全景图生成等领域^[1], 一般地, 主要是由摄像机运动形成的^[2], 以相邻帧间的平移运动类型为常见研究对象。检测平移运动的方法有很多种, 最常用的简单方法就是块匹配法^[3], 其结合平均绝对误差 (MAD) 准则的穷尽搜索法 (ES) 能取得很高的精度, 经常作为检测其他方法的标准。但是由于其较大计算量而难以实时处理, 因此, 人们提出多种改进方法。采用减少搜索位置的方法, 常用的如三步法^[4] (three step search, TSS)、菱形法^[5] (diamond search, DS), 以及最近提出的自适应十字搜索法 (ARPS)^[6,7] 等快速搜索方法。另一种加快运动估计速度的方法就是减少表示每个灰度值的比特位数, 如 1BT 变换^[8]、C-1BT 变换^[9], 改变了匹配准则, 采用简单的 BOOL 运算, 易于硬件实现, 极大地提高了运算速度。

在深入研究上述两类方法的基础上, 提出一种针对特定视频序列的快速平移运动估计方法 (PARPS-CBT)。首先预处理方法的提出可以增加方法的鲁棒性, 提高全局估计的精度; 并对变换中的滤波核函数进行简化, 在不影响系统性能的前提下进一步减小运算复杂度; 同时结合提前预测方法来降低搜索陷入局部最优的风险; 考虑零运动矢量的存在, 增

加搜索提前终止的条件, 进一步减少局部运动的搜索步数。最后通过运动矢量场的统计特性得到图像的全局运动估计矢量。

1 C-1BT 变换原理

在文献[8]中使用一个 17×17 的多带通滤波核函数, 其形式为

$$K(i, j) = \begin{cases} 1/25 & \text{if } i, j \in [0, 4, 8, 12, 16] \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

使用上述核函数对每帧图像进行卷积滤波, 由于它的多带通性质, 通过将原图像与滤波后图像逐点对比来构造二值图像 B 。其构造式为

$$B(i, j) = \begin{cases} 1 & \text{if } I(i, j) \geq I_f(i, j) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

其中: I 和 I_f 分别代表原图像和滤波后图像。

从式(1)可以看出, 在滤波的归一化过程中需要用到复杂的浮点数运算。而文献[9]提出的 C-1BT 采用 16 非零元素的核函数, 每次运算只需 16 次整数加法和 1 次移位运算就可完成。同时, 为提高阈值附近像素匹配的准确性, 引进限制模板 CM 到原 1BT 匹配准则中来增加估计精度。CM 表示为

收稿日期: 2012-01-03; **修回日期:** 2012-02-13 **基金项目:** 国家自然科学基金资助项目 (60875025/030410); 中央高校基本科研业务费专项资金资助项目

作者简介: 席志红 (1965-), 女, 黑龙江哈尔滨人, 教授, 博士, 主要研究方向为图像处理、视频通信 (xizhihong@hrbeu.edu.cn); 初守艳 (1987-), 女, 博士研究生, 主要研究方向为视频稳定。

$$CM'(i,j) = \begin{cases} 1 & \text{if } |I'(i,j) - I'_j(i,j)| \geq D \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

CM 模板可以矫正由于门限的设定导致的相近灰度值的像素被分到二值图像两侧的匹配错误。C-1BT 变换中,只有当图像的灰度值远离门限大于距离 D 时,CM 模板才为 1。改进的匹配准则 CNNMP 为

$$CNNMP(m,n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{ [CM'(i,j) \parallel CM'^{-1}(i+m,j+n)] \odot [B^t(i,j) \oplus B'^{-1}(i+m,j+n)] \} \\ -s \leq m, n \leq s \quad (4)$$

其中:符号 $\parallel$ 、 $\oplus$ 、 $\odot$ 分别表示 OR、XOR、AND。当 CM 为 0 时,该像素被认为不可靠,需要舍弃。增加的判断过程改善了原 1BT 变换中由于门限作用造成的准确性下降的问题。

2 ARPS 算法

ARPS 是目前存在的一种非常优秀的搜索方法,其搜索精度与 DS 算法接近,但搜索步数仅为 DS 的一半。ARPS 算法利用当前宏块的左侧邻块来预测初始运动矢量 (Pre-MV)。算法包括两个步骤:a) 自适应的初步搜索;b) 精确局部搜索。ARPS 的初始臂长 Γ 由左侧邻块的运动矢量决定

$$\Gamma = \max(MV_{\text{left}}(x), MV_{\text{left}}(y))$$

如图 1,左侧运动矢量为 $(2, -1)$,初始臂长即为 2。如果该子块处于第一列,设臂长为 2。对于每一个宏块,初始搜索只执行一次,计算 4 或 5 个点,为以后的精确搜索找到一个好的起始点。第二步,使用固定的小菱形模板进行搜索,直到找到最后的运动矢量。

ARPS 算法初始预测时只用到空间中的左侧相邻宏块。得到的预测值不够精准,致使最终的搜索精度不高,同时,预测臂长采用水平和垂直位移绝对值中的较大值,这样会导致在取值较小的方向搜索更多的步骤。所以,ARPS 算法的性能还有提升的空间。最后,由于匹配准则的改变,ZMV 判断的门限值需要实验重新讨论求解。

3 PARPS-CBT 方法

针对存在单一局部运动物体的视频 PARPS-CBT 方法首次将 C-1BT 变换和 ARPS 算法相结合,用灰度投影剪切法去掉局部运动物体;同时灰度归一化的执行保证了估计结果对光照变化的鲁棒性。其次,对 ARPS 进行一定的改进,以更少的搜索步数获得令人满意的结果。最后,简化传统 1BT 变换的滤波核函数,以提高 GMV 求解的精度,降低运算负担。

3.1 预处理

预处理包括灰度归一化和灰度投影剪切法两个步骤。前者可以极大地减小光照变化对 GMV 求取的不利影响,而后者能部分解决在视频图像中单独局部运动物体存在导致的 GMV 求解精度下降的问题。具体步骤如下:

a) 滤波处理。采用均值滤波器进行滤波,减少图像中噪声的影响。

b) 灰度归一化。采用视频相邻两帧的灰度值均值比作为归一化系数,对当前图像进行归一化处理。归一化系数定义为 MI'^{-1}/MI' ,上标表示时间顺序。如果只有局部区域发生灰度突变,可以只对该块区域进行归一化处理。

c) 投影处理。对归一化后的两帧图像进行差值处理,对

差值图像分别进行行、列投影。

d) 剪切处理。通过统计上述投影曲线中大于零的像素数量,来粗略确定运动物体范围。如运动范围为 (a, b) ,则剪切掉此部分,并将图像 $(b+1, \dots, n)$ 复制到 $(a+1, \dots, n-b+a)$ 。如图 2 所示。

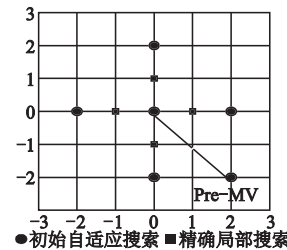


图1 自适应十字搜索模式

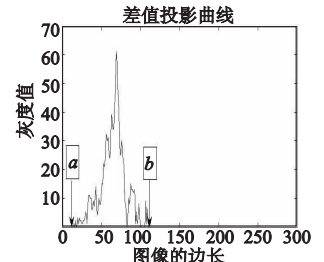


图2 投影剪切示意图

3.2 改进的 ARPS

为弥补 ARPS 算法的不足,根据 C-1BT 变换的特点,提出了如下改进:

a) ZMV 的判定阈值。文献[10]中在 MAD 准则下采用阈值 2 可以得到很好的效果。通过大量的实验得知,C-1BT 变换的匹配准则 CNNMP 与 MAD 的值近似成正比例,同时随着模板阈值 D 的增大,略微减小,所以 ZMV 的阈值采用 $D=12$ 时的准则值进行估算。文中选择为 0.016。

b) Pre-MV 的选择。宏块顶点处于图像首列,则利用上部邻块进行预测;首行的宏块则采用左邻块进行预测;其他位置选用两者中 Pre-MV 绝对值较大的宏块。

c) 臂长的设定。第一个宏块的臂长设为 2;其他位置, X 轴和 Y 轴的搜索臂长通过预测值分别定义为

$$\Gamma_x = \text{ABS}(MV_{\text{predict}}(x)) \\ \Gamma_y = \text{ABS}(MV_{\text{predict}}(y)) \quad (5)$$

实际图像序列中存在大量的静止或者运动幅度极小的区域,步骤 a) 采用使得只需搜索极少的点数即可达到最佳匹配,得到较高峰值信噪比。步骤 b) c) 则更好地利用了空间相关性,增加算法的自适应性,减小陷入局部最优风险的同时缩减搜索时间。

3.3 新的核函数

在文献[9]中, C-1BT 变换中采用 16 个非零元素的核函数,每次滤波计算仅需 16 次加法和一次移位运算。基于对原核函数的分析,发现非零元素间距离的变化只改变滤出元素的多少,并不影响函数本身的多带通性质。所以可以进一步简化该函数,提出仅有 4 个非零元素的核函数,这样每次滤波处理的加法运算只需 4 次即可,极大地加快了运算速度,新核函数的表达式为

$$h(x,y) = \frac{1}{4} \times \begin{cases} 1 & \begin{cases} x=1, y=5; x=5, y=1 \\ x=9, y=5; x=5, y=9 \end{cases} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

为了直观地看出核函数的多带通性质和非零元素间距离变化的关系,选择文献中 16 元素的滤波器与提出的 4 元素的滤波器的频率响应进行比较,如图 3 所示。

通过对图 3 频响图像的分析可以看出,简化的滤波器依旧保留了多带通性质,只是由于非零元素间距离的增大导致滤出的元素数量略微改变。但是采用提出的 PARPS-CBT 法,能在保证匹配精度的同时极大地减少搜索点数,通过实验证实,PSNR 增加约 0.5 dB,同时搜索点数也极大减少,证实了改进

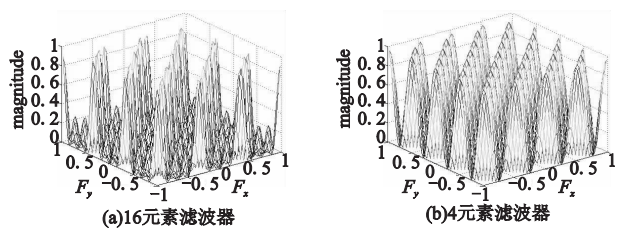


图3 不同滤波器的频响图

的合理性。

选取 foreman 序列图像中的任一帧,进行均值滤波,分别采用上述两种滤波器做 C-1BT 变换得到的模板图像和二值图像,如图 4 所示(模板图像的获取中, D 取 10;CM 表示模板图像,B 表示二值图像)。



图4 两种滤波器的C-1BT变换结果

从变换后图像可以更直观、清晰地看出,新核函数仍可提取出足够的信息用来匹配,保证了运动估计的精度。

3.4 求解 GMV

针对局部运动物体较小的视频序列进行稳像时,需要补偿的是代表背景运动的全局运动矢量,而前景目标的运动是局部运动,因此需要对局部运动矢量进行统计分析,以提取出全局运动矢量。由于局部运动的个体差异较大以及不规律性,导致局部运动矢量表现出高度的不一致,而处于背景中的全局运动在视频中占有较大比例,具有良好的一致性,因此采用统计概率的方法来求取全局运动矢量。首先将检测到的所有子块的水平和垂直运动矢量作为两个样本集合;然后对集合中每一样本统计其出现个数;最后找到概率最大的样本对,认为该样本对即为最终的全局运动矢量。

4 实验仿真分析

4.1 局部运动估计算法性能的比较

采用 football、foreman、silent、stefan (各 30 帧)运动程度不同的序列作为实验对象,宏块尺寸为 16×16 ,搜索范围为 $(-16, 16)$ 。由于文献[11,12]已对 NMMP 准则和 MAD 准则作过比较,因此,本文只针对统一的 CNMMP 准则下不同搜索方法 & 不同核函数的结果进行对比。

准确性和实时性是评价运动估计的基本要求,分别采用峰值信噪比 (PSNR) 和平均搜索点数作为比较的依据,采用文中提出的核函数结合 PARPS-CBT 法与文献[9]中的 16 元素核函数结合 ES、DS 和 ARPS 三种搜索法分别进行对比验证。

从表 1、2 的对比可以看出,除了大运动序列 football 的峰值信噪比比较 ES 和 DS 算法略有下降,其他序列,采用提出的 PARPS-CBT 法,在保证了较高的峰值信噪比的同时,搜索点数大幅下降,极大地减轻了运算负担。

表 1 搜索点数对比

序列	football	foreman	silent	stefan
ES-16	984.92	984.92	984.92	984.92
DS-16	23.96	25.00	19.67	22.96
ARPS-16	14.81	16.72	11.61	12.04
PARPS-4	10.45	4.30	2.38	4.85

表 2 峰值信噪比对比

序列	football	foreman	silent	stefan
ES-16	23.96	29.25	28.54	25.65
DS-16	22.22	29.94	28.90	25.44
ARPS-16	22.84	29.81	29.21	25.77
PARPS-4	21.46	30.49	31.91	25.89

为进一步证实文中方法的普遍性,采用实际拍摄的视频进行验证,选取 30 帧,图像大小为 672×512 ,块大小为 32×32 ,搜索范围为 $(-32, 32)$ 。仍然采用上述对比方法进行比较。比较结果如图 5 所示(其中,ES 搜索法的点数为 3779,未在图中标出)。

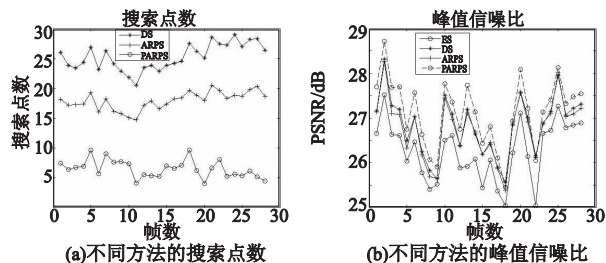


图5 仿真结果比较

通过观察图 5,可清楚地看出由于加入了 ZMV 的预判断和自适应的臂长确定过程,提高了初始点预测的准确性,减少了陷入局部最优的风险,尤其对于运动不太剧烈的图像,无论是搜索点数还是精度都有很大的提高。

4.2 全局运动估计算法性能验证

为了验证提出的预处理法,本文选取两幅具有不同运动强度的图像作为测试对象,并表示出运动矢量,如图 6 所示。

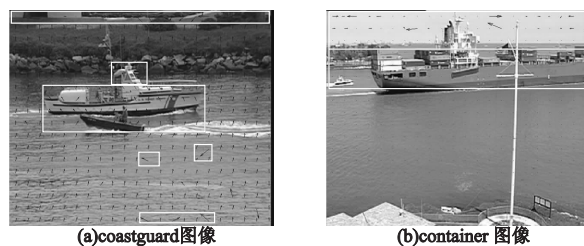


图6 视频序列中的任意两帧

其中矩形区域为局部运动物体或误匹配的区域,对上述两帧图像进行灰度归一化及剪切处理后得到的图像如图 7 所示。

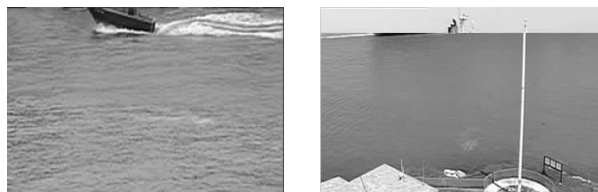


图7 预处理后的两幅图像

以 coastguard 图像为例,计算剪切前后得到的正确全局运动矢量的概率,实验仿真结果如表 3 所示。

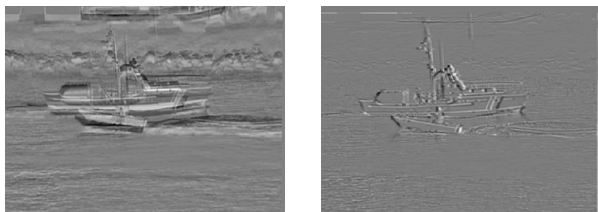
表 3 剪切前后的 GMV 准确率

准确率	正确数	错误数	百分比
剪切前	215	181	0.5429
剪切后	147	62	0.7033

对比图 6 和 7 可以看到,由于去掉了局部运动物体,图像的整体尺寸变小,可以减少搜索点数。参考表 3 分析可得,由于在全局性相对一致的区域进行搜索,能在一定程度上提高全局运动估计的准确性。同时对其他类型的视频序列进行了仿

真实验,由于局部运动物体的形式不同,实验效果不尽相同,但是剪切处理都能使得全局运动准确性有不同程度的改善。但是本文提出的方法提取局部运动物体过于粗糙,只适用于存在单一物体的情况,下一步打算结合聚类等方法更精细地提取局部运动物体。

为了更直观地看出补偿的效果,仍以 coastguard 序列中的两帧为例,估计出的平移运动矢量为 $(-9, 1)$, 补偿前后的差值图像如图 8 所示。



(a)原差值图像 (b)补偿后差值图像

图8 补偿前后的差值图像

补偿图像的无定义区采用参考帧相应位置的灰度值代替,原相邻图像的 PSNR 为 17.2122 dB,补偿后图像的 PSNR 为 24.0510 dB。从图 8 可以看出,补偿后的图像除局部运动物体外还有水面波动造成的残差图像,如果除掉抖动,信噪比提高得更多。

5 结束语

PARPS-CBT 全局运动估计法采用预处理过程减少由于灰度突变和存在局部运动物体导致的匹配错误,增强全局运动估计的鲁棒性,提高估计精度。同时,将 C-1BT 变换和改进的 ARPS 法相结合来实现局部运动矢量的估计。简化 C-1BT 变换的核函数,在整个图像范围内均匀提取特征点,形成 1 bit 的单一位平面。根据其特点,匹配准则转换成适合 BOOL 运算的形式,利于硬件实现。最后,由于改进的 ARPS 能更好地利用运动矢量的空间相关性来进行预测估计,使搜索起点更接近最优值附近,并且综合 ZMV 提前中止条件,能快速结束搜索。

但是本文提出的方法仅适用于具有单一局部运动物体的小范围平移运动。下一步的目标分为两部分:对预处理部分进

行改进和完善,增加提取局部运动物体的精度;考虑图像帧间存在旋转运动的情况。

参考文献:

- [1] 刘磊,王志良,刘冀伟,等. 一种改进的快速全局运动估计算法[J]. 计算机工程,2010,36(20):28-31.
- [2] 樊玉平,田裕鹏. 一种基于皇后模板抽样的全局运动估计方法[J]. 中国图象图形学报 A,2009,14(12):2497-2502.
- [3] SAHA A, MUKHERJEE A, SURAL S. New pixel-decimation patterns for block matching in motion estimation[J]. Signal Processing: Image Communication, 2008,23(10):725-738.
- [4] DUANMU C J, AHMAD M O, SWAMY M N S. A fast three-step search algorithm by the utilization of multilevel vector partial sums [C]//Proc of IEEE CCECE. 2003:1981-1984.
- [5] ZHU S, MA K K. A new diamond search algorithm for fast block-matching motion estimation[J]. IEEE Trans on Image Processing, 2000,9(2):287-290.
- [6] NIE Yao, MA Kai-kuang. Adaptive rood pattern search for fast block-matching motion estimation[J]. IEEE Trans on Image Processing, 2002,11(12):1442-1449.
- [7] 李会宗,陈雷霆,卢光辉,等. 基于起点预测的不连续十字形快速搜索算法[J]. 计算机应用研究,2008,25(10):2929-2931.
- [8] URHAN O, ERTURK S. Constrained one-bit transform for low-complexity block motion estimation[J]. IEEE Trans on Circuits and Systems for Video Technology, 2007,17(4):478-482.
- [9] NATARAJAN B, BHASKARAN V, KONSTANTINIDIS K. Low-complexity block-based motion estimation via one-bit transforms[J]. IEEE Trans on Circuits and Systems for Video Technology, 1997,7(4):702-706.
- [10] 梁燕,刘文耀. 基于起点预测的自适应交叉—准菱形运动估计算法[J]. 光学精密工程,2005,13(2):237-246.
- [11] CELEBI A, AKBULUT O, URHAN O, et al. Truncated gray-coded bit-plane matching based motion estimation and its hardware architecture[J]. IEEE Trans on Consumer Electronics, 2009,55(3):1530-1536.
- [12] KO S J, LEE S H, JEON S W, et al. Fast digital image stabilizer based on gray-coded bit-plane matching[J]. IEEE Trans on Consumer Electronics, 1999,45(3):598-603.