

文章编号 1004-924X(2011)01-0183-09

复杂场景中基于变块差分的运动目标检测

朱娟娟, 郭宝龙

(西安电子科技大学 机电工程学院, 陕西 西安 710071)

摘要:针对复杂场景中包含的摄像机扫描运动、随机抖动和目标运动,提出一种基于帧间可变块差分的运动目标检测算法。首先,利用全局特征点估计运动参数对帧间背景进行补偿,提取图像的全局特征点并匹配,以特征点集的最小位置误差和作为目标进行迭代,获取误差不大于 0.5 pixel 的全局运动参数,并精确补偿当前帧实现背景校正。然后,利用可变块差分实现运动目标的检测。先用大尺寸对差分图像进行分割,将整幅图像粗略区分为背景区域、运动目标区域和边界区域,通过阈值判定来进一步细分,最后对运动目标区域进行准确标定。这种由粗到细的检测步骤,在保证精度的同时能够提高检测速度。实验结果表明,该算法能够准确检测含摄像机扫描和抖动的复杂运动场景中的前景运动目标,且处理速度达到 25 frame/s。

关键词:视频图像;运动目标检测;可变块差分;全局运动估计;帧间补偿

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/OPE.20111901.0183

Moving object detection based on variant block difference in complex scenes

ZHU Juan-juan, GUO Bao-long

(School of Electromechanical Engineering, Xidian University, Xi'an 710071, China)

Abstract: A fast moving-object detection algorithm based on inter-frame Variant Compensated Blocks Difference (VCBD) is presented to deal with the complex scenes with camera scan, dithering and object moving. Firstly, the global motion estimation is performed based on feature points to compensate the inter-frame background. The global feature points in a reference image are selected and matched in a current image. Then, the iteration is applied to realize the minimum sum of position errors of all matched points and to obtain the global motion parameter with the accuracy less than 0.5 pixel. Accordingly, the current frame is compensated to match the background area. Finally, the adaptive variant block difference is proposed to detect moving objects. The whole image is classified into background, foreground and boundary areas and the block is then judged with the threshold and divided into four blocks. These coarse-to-fine steps can greatly improve the velocity and veracity of detection. Experimental results show that the algorithm can detect moving objects in camera scan and dithering sequences and the processing speed achieves 25 frame/s.

收稿日期:2010-04-01;修订日期:2010-06-25.

基金项目:国家自然科学基金资助项目(No. 61003196);高校基本科研业务费专项资助项目(No. K50510040004)

Key words: video image; moving object detection; variant block difference; global motion estimation; inter-frame compensation

1 引言

运动目标检测是指将视频序列图像中的变化区域从背景中分割出来,其中以动态场景下的运动目标实时检测为难点。由于摄像载体的跟踪拍摄会导致运动目标和背景的相应运动;而当载体存在不规则抖动时,视频序列也会出现晃动甚至模糊现象,无法准确提取运动目标区域,因此,对于摄像机运动-目标运动的复杂场景所获得的视频序列不仅存在正常的扫描运动,还存在无规则的抖动,所以对其运动目标的实时检测十分重要。

传统运动目标检测算法有光流法^[1],背景差法^[2]和帧间差法^[3-4],其中光流法可应用于摄像机运动场景,但其光流场分布精度难以保证,且不能实时处理;背景差法实现最简单,但需要对背景进行建模^[5],且不适应于动态场景;帧间差法有较好的适应性,不足之处是易受背景运动干扰,不利于保留前景运动,且像素点差分运算的抗噪声干扰性能差。近年出现了基于曲线演化^[6]和不变矩匹配^[7]等新颖算法,但运算均较为复杂,无法实现实时检测。

通常,分析人眼对于运动判断的特性,人们总是将背景作为参照^[8],将相对于背景的运动感知出来就是前景运动目标。因而若去除了由相机运动而造成的背景相对运动,序列帧间的差别就仅与前景的绝对运动有关。因此,本文结合人眼的视觉特性提出了一种基于帧间补偿的可变块差分算法(Variant Compensated Block's Difference, V CBD)。该算法首先对相邻帧进行全局运动估计^[9],补偿当前帧以实现背景校正。然后,做与前一帧的基于可变块的差分运算:对差分图先用大尺寸进行分块处理,通过阈值判定来决定是否进行更细划分,将所有误差较大块在原始图像中进行标定,实现快速运动目标检测。实验结果充分证明,对于摄像机扫描和存在抖动的视频序列,V CBD算法均可以实时准确地检测运动目标,从而为后续目标精细边缘提取^[10-11]、实时跟踪^[12]等环节提供感兴趣区域。

2 V CBD 的帧间背景校正

帧间背景校正的核心问题是对摄像机的运动进行全局运动估计^[13],然后对当前帧进行补偿实现背景校正。在实际应用中,在相邻两帧的时间间隔内,摄像机的运动可近似认为是背景的平移和小角度的旋转(不可忽略);另一方面,假定运动目标相对于背景是一个小目标(即目标在整个图像中所占面积不大)。此时,可以通过全局运动估计算法,检测相邻帧间的背景相对运动,并补偿当前帧以校正背景,从而保留前景目标的绝对运动。

V CBD 算法的帧间背景校正流程如图 1 所示,具体原理是:利用特征点提取技术检测参考帧(即前一帧)中的特征点;然后构造特征窗进行匹配,确定当前帧中的对应特征点;再将所有特征点对代入运动模型,进行最小二乘解的迭代运算,获取帧间全局运动参数矩阵;最后对当前帧进行运动补偿,实现背景校正。其中的关键技术就是快速精确地获取全局运动矢量,从而保证背景运动的正确校正。

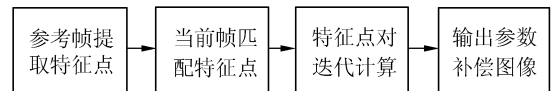


图 1 帧间背景校正流程图

Fig. 1 Flow diagram of inter-frame background validation

2.1 特征点提取

采用经典的 Harris 算子提取参考帧的特征点,其检测步骤如下:

(1) 计算图像亮度 $I(x, y)$ 在点 (x, y) 处的梯度:

$$\begin{cases} X = I \otimes (-1, 0, 1) = \partial I / \partial X \\ Y = I \otimes (-1, 0, 1)^T = \partial I / \partial Y \end{cases}, \quad (1)$$

(2) 构造自相关矩阵:

$$\mathbf{A} = X^2 \otimes \omega, \quad \mathbf{B} = Y^2 \otimes \omega, \quad \mathbf{C} = (XY) \otimes \omega, \quad (2)$$

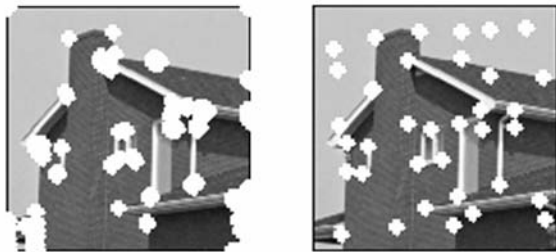
式中 \otimes 表示卷积算子, $\omega = \exp(-(x^2 + y^2)/2\delta^2)$ 是高斯窗平滑函数。由式(2)可以得到一个二阶实对称矩阵 $\mathbf{M} = [\mathbf{A}, \mathbf{C}; \mathbf{C}, \mathbf{B}]$, 必然存在两个特征值 λ_1 和 λ_2 , 代表自相关函数的主曲率。

(3) 提取特征点:如果特征值 λ_1 和 λ_2 是极大值时,则点 (x, y) 是一个特征点。即满足:

$$R = \det(\mathbf{M}) - k \cdot \text{trace}^2(\mathbf{M}) > T, \quad (3)$$

其中, $\det \mathbf{M} = \lambda_1 \times \lambda_2$, $\text{trace} \mathbf{M} = \lambda_1 + \lambda_2$ 。这里是一个固定的阈值,参数 $k = 0.04 \sim 0.06$ 。式(3)减号后面的部分是修正,目的是抑制边缘点的误检。

为了使特征点在图像中分布均匀,避免分布过于集中,本文采用改进的分区提取 Harris 特征点的方法,即:将图像分成互不重叠的区域,在各区域内选取响应 R 最大的点作为特征点。如图 2 所示是对 House 图像进行 Harris 角点检测的结果,图 2(a)是阈值 $T = 5\ 000$ 时角点检测结果,图 2(b)为改进的分区提取结果(6×6 个特征点),均匀分布在整幅图像中,绝大多数位于物体边缘处,具有较好的独特性,且信息量丰富。



(a) 角点检测结果 (b) 分区检测结果
(a) Feature point selection (b) Results of sub areas

图 2 特征点检测示意图

Fig. 2 Results of feature point selection

2.2 特征点匹配

所谓特征点匹配,就是在当前帧找到参考帧中每一特征点的唯一匹配点。本文采用特征窗区域的相似度测量法,以参考图像中的特征点 p_k 为例,其特征窗匹配步骤如下:

Step1:以特征点 p_k 为中心,将其邻域构造一个尺寸为 $M \times N$ 的模板图像 I_1 作为特征窗;

Step2:在当前图像中对应位置处,确定 $(M + 2p) \times (N + 2p)$ 的搜索区域 I_2 ,其中 p 表示设定的水平和垂直方向最大偏移量;

Step3:从搜索区域的中心开始搜索,搜索路径采用菱形法^[14],计算每点的绝对误差和 SAD (Sum of Absolute Difference) 值,其计算如公式(4)所示:

$$SAD(i, j) = \sum_{m=1}^M \sum_{n=1}^N |I_1(m, n) - I_2(m+i+p, n+j+p)|, \quad i, j = -p, \dots, p \quad (4)$$

Step4:找出最小 SAD 值,则其对应位置为匹配特征窗,其中心点即为匹配特征点。

为了进一步提高搜索速度,每次将阈值 T 指向已搜索点的 SAD 值的最小值,即阈值序列是单调递减的,使得非匹配点在更少的计算过程中就达到阈值而被丢弃,从而提高算法效率。

2.3 全局运动参数计算

假定参考帧中点 $[x, y]^T$ 运动到当前帧的位置 $[\hat{x}, \hat{y}]^T$,则对应的坐标关系为:

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} m_0 & m_1 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} m_2 \\ m_5 \end{bmatrix}, \quad (5)$$

其中 m_2 和 m_5 表示两图的平移量, m_0, m_1, m_3, m_4 表示尺度和旋转量,即为图像仿射变换模型。

将特征点对代入运动模型(5),求最小二乘解,求解的目标是使得各特征点的位置误差达到最小。因此,本文提出了基于求最小二乘解的迭代步骤:

Step1:计算 c 对特征点的坐标差 $\Delta x_i = x_i - \hat{x}_i, \Delta y_i = y_i - \hat{y}_i$,作为初始数据集合 S ;

Step2:将 S 中对应的特征点对代入仿射运动模型(5),建立仿射方程: $\mathbf{B} = \mathbf{A} \mathbf{M}_{\text{affine}}$;

$$\mathbf{B} = \begin{bmatrix} \hat{x}_1 & \hat{y}_1 \\ \hat{x}_2 & \hat{y}_2 \\ \vdots & \vdots \\ \hat{x}_c & \hat{y}_c \end{bmatrix}, \mathbf{A} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_c & y_c & 1 \end{bmatrix}, \mathbf{M}_{\text{affine}} = \begin{bmatrix} m_0 & m_1 \\ m_3 & m_4 \\ m_2 & m_5 \end{bmatrix}$$

Step3:计算仿射矩阵 $\mathbf{M}_{\text{affine}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B}$;

Step4:利用所求的参数矩阵 $\mathbf{M}_{\text{affine}}$,代入模型(5)计算参考帧中各特征点映射后的坐标 (\bar{x}_i, \bar{y}_i) ,并计算其与参考帧特征点的坐标差 $\Delta \bar{x}_i = x_i - \bar{x}_i, \Delta \bar{y}_i = y_i - \bar{y}_i$;

Step5:计算集合 S 中 $(\Delta x_i, \Delta y_i)$ 与 $(\Delta \bar{x}_i, \Delta \bar{y}_i)$ 的位置误差 E ;

$$E = \sqrt{(\Delta x_i - \Delta \bar{x}_i)^2 + (\Delta y_i - \Delta \bar{y}_i)^2}. \quad (6)$$

Step6:将 S 中位置误差大于阈值的对应特征点对删除,更新集合 S 并转至 step2;若误差小于设定阈值,则保留的全局特征点集合不再发生变化,即停止迭代,输出 $\mathbf{M}_{\text{affine}}$ 作为最终的全局运动参数。

2.4 帧间图像补偿

全局运动参数矩阵 $\mathbf{M}_{\text{affine}}$ 代表的是背景的运动,即摄像系统的旋转,水平和垂直偏移运动。以

此运动参数对当前帧进行补偿,即是对当前帧的像素坐标利用 M_{affine} 矩阵进行变换,得到新的坐标,并将原图中像素值赋给该新坐标,即可获取补偿后的结果图像,实现对帧间背景运动的校正。

在坐标转换时,变换后的坐标不一定在整数格上,选用双线性插值法进行插值,插值原理为:对于一个目的像素,设置坐标通过变换得到的浮点坐标为 $(i+p, j+q)$,其中 i 和 j 均为非负整数, p 和 q 为 $[0, 1)$ 区间的浮点数,则这个坐标的像素值 $f(i+p, j+q)$ 可由原图像中坐标为 (i, j) 、 $(i+1, j)$ 、 $(i, j+1)$ 和 $(i+1, j+1)$ 所对应的 4 个像素的值决定,即:

$$f(i+p, j+q) = (1-p) \cdot (1-q) \cdot f(i, j) + (1-p) \cdot q \cdot f(i, j+1) + p \cdot (1-q) \cdot f(i+1, j) + p \cdot q \cdot f(i+1, j+1). \quad (7)$$

3 V CBD 的帧间可变块差分

3.1 可变块差分原理

将相邻帧进行背景配准之后,两幅图像之间的差异就仅是由于前景目标的运动引起的,可以通过帧间差分来检测这种差异。传统的帧间差分法^[5]是对像素点进行灰度求差再二值化。由于帧间光照的变化,纹理的变化以及噪声等影响,单像素点作为差分对象容易出现两种问题:目标内部像素点由于颜色均匀会被漏检;背景中部分像素点由于差异较大会被误检。而且从统计的角度来看,单个像素点相对于图像块来说,其受噪声影响更大,个体差异也更突出。

因此,本文采用基于块的帧差法,其关键在于合适选取块的尺寸:块过大则块内同时含有背景和前景;块过小又会导致计算量过大。经分析,本文提出的可变块差分原理为:对于背景块,采用大尺寸进行快速判定;对于前景块,采用小尺寸;对于背景和前景交界块,进行细致划分。这样,图像中的块可以粗略分为 3 类:背景块,前景块和交界块。

3.2 可变块差分流程

通过以上分析,VCBD 算法采用基于可变块的差分方法,算法流程如下:对于背景配准后的相邻两帧做帧间差分运算,得到差分图;对差分图首

先利用大尺寸进行划分,判断每块内的平均绝对误差(Mean Absolute Difference, MAD)是否大于设定阈值(取所有子块误差的平均值),来分析该块内是否有属于运动目标的部分,从而来决定是否进行更细划分,直到划分到块的最小尺寸为止;将所有 MAD 大于设定阈值(将误差从大到小排列,取前 10% 的平均值)的块在原始图像中标定,结果以运动目标的外接矩形表示,实现运动目标的检测,如图 3 所示。

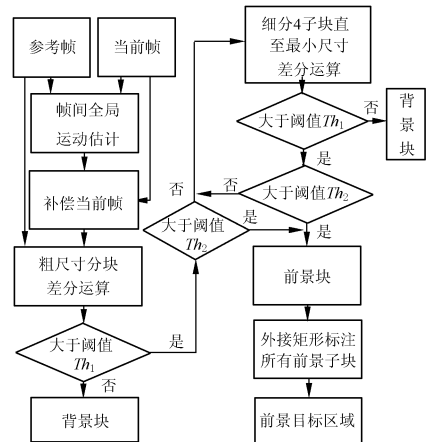


图 3 VCBD 算法流程图

Fig. 3 Flow diagram of VCBD

3.3 可变块差分实现步骤

(1) 对补偿后图像 I_k 与前一帧图像 I_{k-1} 做差分运算,得到差分图像 D ,其每一点的像素值为:

$$D(i, j) = \text{abs}(I_k(i, j) - I_{k-1}(i, j)). \quad (8)$$

(2) 将 D 分为互不重叠的子块 $B(64 \text{ pixel} \times 64 \text{ pixel})$;

在对差分图像进行划分时,对若干行列的边界信息不做处理。一方面由于补偿图像的边界信息丢失,导致差分不够精确;另一方面运动目标一般在场景中间,边界位置属于非感兴趣区域。

(3) 计算各子块 B 内的平均绝对误差 MAD;

$$MAD = \frac{1}{N \times N} \sum_{i=1}^N \sum_{j=1}^N B(i, j), \quad (9)$$

式中, N 为该子块的尺寸, $B(i, j)$ 表示该子块内每一点的像素值。

(4) 对 MAD 作如下判断:

如果 $MAD \leq Th_1$, 则该块属于背景块,则停止对该块的处理;

如果 $Th_1 < MAD < Th_2$, 则表明该块的残差

较大,属于背景和前景的交界区域,转步骤(5);

如果 $MAD \geq Th_2$, 则表明该块的残差很大,该块属于前景运动块,直接对该块进行标记;

(5)对子块继续分割成 4 个均匀子块,转步骤(3),直至最终的子块尺寸为最小尺寸 8×8 为止;

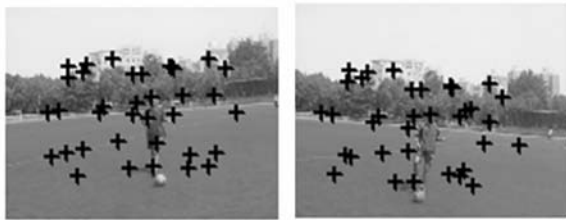
(6)统计所有标定的前景目标块,显示前景目标;或者以外接矩形的形式标记运动目标。

4 实验结果及分析

实验视频为:(1)标准视频序列 Coastguard 序列;(2)手持摄像机边走边拍的 Playground 序列;(3)高处拍摄的 Car 序列,摄像机被安放在移动平台上,存在慢速水平扫描和随机抖动,利用 VC++ 编程进行同步实时检测动态场景中的运动目标。

4.1 全局运动估计结果及分析

对于 Playground 序列,其背景和前景均存在较大运动,图 4 为序列中两帧图像,对参考帧采用分区提取 Harris 特征点,其结果如图 4(a),并在当前帧确定其匹配特征点,结果如图 4(b)所示。可以观察到,背景中多半数特征点获得了正确的匹配。

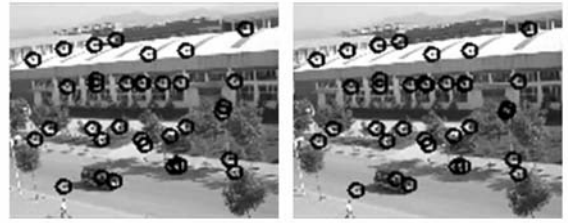


(a) 参考帧提取特征点 (b) 当前帧匹配特征点
(a) Detected points (b) Matched points

图 4 Playground 序列中特征点提取和匹配结果图
Fig. 4 Feature selection and matching of Playground sequence

对于 Car 序列,其背景较为复杂且存在垂直方向低幅抖动,如图 5 所示,图 5(a)为参考帧中均匀提取的特征点结果,图 5(b)为当前帧中匹配的特征点。可以观察到,所有特征点都得到了正确的匹配。

对该两组图像求全局运动参数时,对所有特征点代入模型(5)直接计算,和特征点迭代计算进行比较;另外,为了验证运动估计的精度,以设定参数 $[0.992\ 4, 0.034\ 8; -0.018\ 7, 0.998\ 5;$



(a) 参考帧提取特征点 (b) 当前帧匹配特征点
(a) Detected points (b) Matched points

图 5 Car 序列中特征点提取和匹配结果图
Fig. 5 Feature selection and matching of Car sequence

$14.120\ 8, -3.9567]$ 进行人工合成得到两幅图像作为第三组,表 1 为对这三组图像进行全局运动估计的结果。可以看出,特征点迭代计算的全局运动参数与设定参数之间的误差在很小范围内(水平垂直方向均小于 0.5 pixel)。

表 1 全局运动估计实验结果 1

Tab. 1 Results 1 of global motion estimation

实验图片	直接计算参数矩阵	迭代估计参数矩阵
No. 1	$\begin{bmatrix} 1.000 & -0.054 \\ 0.050\ 8 & 0.982\ 2 \\ -20.036 & 11.034\ 1 \end{bmatrix}$	$\begin{bmatrix} 1.013\ 8 & -0.044 \\ 0.061\ 0 & 0.980\ 9 \\ -19.584\ 0 & 11.6860 \end{bmatrix}$
No. 2	$\begin{bmatrix} 1.001\ 2 & -0.041 \\ 0.001\ 1 & 1.020\ 0 \\ -0.022\ 3 & 1.003\ 5 \end{bmatrix}$	$\begin{bmatrix} 0.999\ 3 & -0.002\ 7 \\ 0.000\ 4 & 1.000\ 0 \\ -1.019\ 0 & 0.192\ 7 \end{bmatrix}$
合成图片	$\begin{bmatrix} 0.981\ 1 & 0.023\ 5 \\ -0.008\ 4 & 0.967\ 4 \\ 13.642\ 9 & -3.427\ 5 \end{bmatrix}$	$\begin{bmatrix} 0.990\ 9 & 0.034\ 4 \\ -0.019\ 9 & 0.997\ 4 \\ 14.287\ 0 & -3.871\ 6 \end{bmatrix}$

另外,对当前帧以计算出的全局运动参数进行补偿,即沿着运动的反方向进行纠偏,得到补偿后的当前图像,计算其与参考图像之间的 PSNR (Peak Signal-to-Noise Ratio) 值^[14];并与原始的两图像之间的 PSNR 值进行比较(PSNR 值是在剪切若干边界的图像间进行计算的)。由表 2 的实验结果,发现 PSNR 值有很明显的提高,说明补偿后两幅图像之间的相似度较高,从而也进一步说明 V CBD 算法的全局运动估计具有较高精度,而且运算时间也能满足 25 frame/s 的要求。

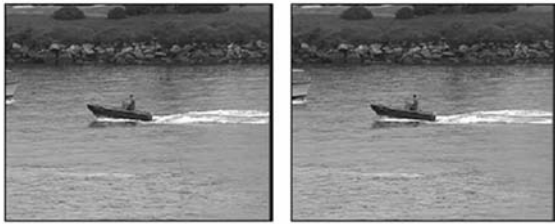
表 2 全局运动估计实验结果 2

Tab. 2 Results 2 of global motion estimation

实验图片	时间/ms	原始 PSNR/dB	补偿 PSNR/dB
No. 1	28.45	19.87	29.32
No. 2	21.89	22.96	33.29
合成图片	25.12	19.64	38.34

4.2 帧间差分结果比较与分析

实验中利用 Coastguard 序列,任意选取其中两帧进行实验,如图 6。序列中前景小船在缓速向左行驶,摄像机跟随拍摄且保持小船处于场景中间;同时,船行驶导致水面的波纹也有明显起伏晃动。



(a) 参考帧

(b) 当前帧

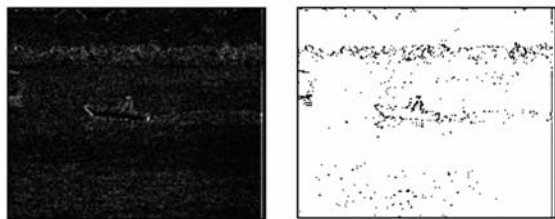
(a) Reference image (b) Current image

图 6 Coastguard 序列中相邻帧

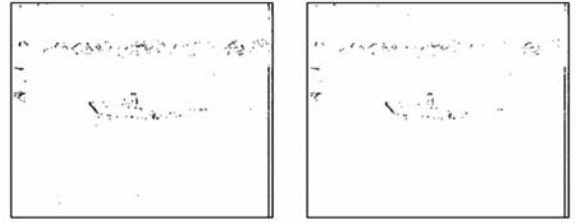
Fig. 6 Adjacent frames in Coastguard sequence

图 7 是对图 6 进行直接差分^[4]和二值化的结果,其中(b)~(d)图像为反色显示,即差值小于阈值的像素为白点。当阈值增加,背景像素点逐渐被剔除,但前景像素点同时被去除;当阈值过大时,如图 7(d)中前景目标已经很不完整,但是背景中的河岸却并没有完全去除,无法准确提取运动目标的位置。

直接差分法失败的原因在于摄像机运动导致背景区域的变化,因而需要消除其不利影响。此时,用 $M_{\text{affine}} = [1.00, 0.00; 0.01, 0.99; 2.00, -1.01]$ 全局参数矩阵对当前帧进行补偿,以实现



(a) 帧间直接差分结果 (b) 二值化结果, 阈值=20

(a) Image difference (b) Binarization result in $Th=20$ 

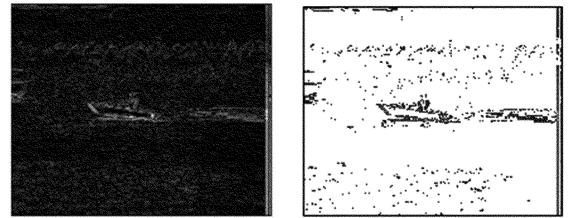
(c) 二值化结果, 阈值=50 (d) 二值化结果, 阈值=80

(c) Binarization result in $Th=50$ (d) Binarization result in $Th=80$

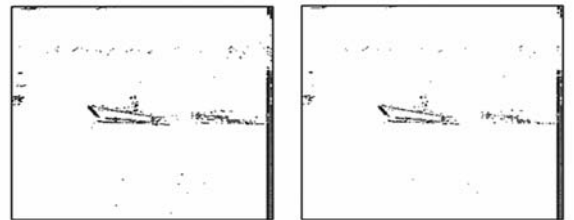
图 7 帧间直接差分和二值化处理结果比较

Fig. 7 Results of image difference and binarization

背景校正,再进行差分运算,实验结果如图 8 所示。与图 7 的直接差分法^[4,13]比较可知,补偿后差分在发生运动的位置更为突出,如图 8(a),因此对目标的保持更好;即使二值化设定阈值较大,如图 8(d),其运动目标的轮廓依然很完整,且背景也去除得较干净。



(a) 帧间补偿后差分结果 (b) 二值化结果, 阈值=20

(a) Compensated image difference (b) Binarization result in $Th=20$ 

(c) 二值化结果, 阈值=50 (d) 二值化结果, 阈值=80

(c) Binarization result in $Th=50$ (d) Binarization result in $Th=80$

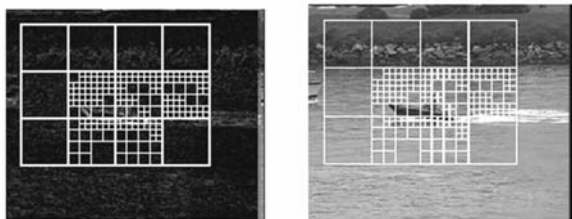
图 8 帧间补偿差分和二值化处理结果比较

Fig. 8 Results of compensated difference and binarization

4.3 运动目标检测结果与分析

图 9 是 Coastguard 序列中目标的检测结果图,其中图 9(a)是差分图像分块结果,为了更为直观地观察细化结果,这里将分割结果显示在原始图像的相同位置,如图 9(b)。可以看出,对于背景中相对变化较小的河岸,不需要进行细致划分,而对前景中的行驶船只和波动的水面,是需要进行细致划分的。图 9(c)显示的是通过阈值判

定后,对块内残差较大的子块进行标定的结果;图 9(d)显示的是检测出的前景目标,较为完整的包含了运动船只,同时船行驶导致的运动波浪也被检测出一部分。



(a) 差分图像分块结果 (b) 原始图中分块示意
(a) Blocks in difference image (b) Blocks in original image

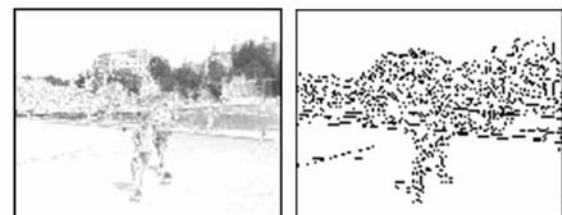


(c) 运动前景子块 (d) 运动目标区域
(c) Blocks in foreground (d) Moving object detection

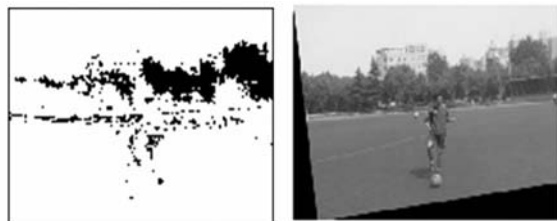
图 9 V CBD 算法检测运动目标示例

Fig. 9 Results of moving object detection

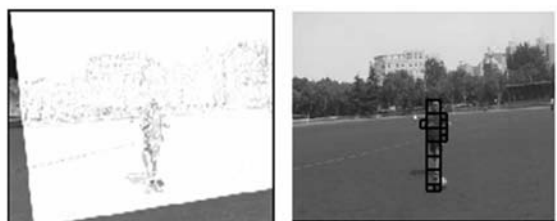
对 Playground 序列中的实验图片(图 4),图 10(a)为直接差分结果^[4],图 10(b)为运动目标边缘图^[6],图 10(c)为二值化后的运动目标模板^[13]。可以看出,相机运动导致差异较大的背景被误检,而前景的拖影也增大目标尺寸。本算法获取的全局运动矩阵为 $M_{\text{affine}} = [0.97, 0.15, -1.05; -0.12, 0.99, -1.28]$,对当前帧补偿结果如图 10(d)所示,计算其与参考帧的差分,其结果如图 10(e),可见差分结果保留了完整的前景运动,且去除了背景运动的不利影响。图 10(f)是通过阈值判定后,对块内残差较大的子块进行标定的结果。



(a) 直接差分结果(反色) (b) 目标边缘检测
(a) Original difference (b) Object edge detection



(c) 二值检测目标结果 (d) 当前帧补偿结果
(c) Binarization result (d) Compensated image



(e) 补偿后差分结果(反色) (f) 运动目标区域
(e) Compensated difference (f) Moving object detection

图 10 扫描视频运动目标检测结果

Fig. 10 Results of moving object detection in scan video

对抖动平台实际拍摄的 Car 序列如图 11 所示,是为对实验图片(如图 5)差分及其运动目标检测结果。其中,相邻两帧之间的全局运动估计参数矩阵为 $M_{\text{affine}} = [1.00, 0.00, 1.00; 0.00, 0.99, -0.01]$,对当前帧补偿结果如图 11(a),与



(a) 当前帧补偿结果 (b) 差分结果
(a) Compensated image (b) Image difference



(c) 运动前景子块 (d) 运动目标区域
(c) Blocks in foreground (d) Moving object detection

图 11 抖动视频运动目标检测结果

Fig. 11 Results of moving object detection in dithering video

参考帧差分结果如图 11(b)。图 11(c)是通过由粗到细地阈值判定后,对块内残差较大的子块进行标定的结果;图 11(d)显示的是检测出的前景目标,以外接矩形的形式进行标定,完整地包含了前景运动的车辆。

对该抖动视频序列通过全局运动估计后进行帧间补偿,得到稳定平滑的视频序列,然后隔 10 frame 进行运动目标检测,其结果如图 12 所示。可以看出,在稳定视频序列中,帧间背景的差异几乎被完全去除,因此运动车辆的检测结果较为精



图 12 连续多帧运动目标检测结果

Fig. 12 Moving object detection in consecutive frames

确。不过由于车辆影子的影响,使得将阴影也当成前景目标,因此造成检测的目标会比实际目标稍大。

5 结 论

本文针对摄像机运动-目标运动的复杂场景,提出一种基于帧间补偿的可变块差分法 V CBD 来检测运动目标。其优势在于:基于人眼视觉特性,对图像先利用全局运动估计和补偿来实现背景校准;对补偿后图像与参考图像进行差分运算时,改变基于点差异的检测为基于块差异的检测,在一定程度上消除了噪声的影响;对差分图像没有进行传统的二值化处理,而是直接进行由粗到细分块,粗分割快速判断背景块,细分分割逐渐检测前景运动目标,从而提高了运行速度,处理能力达到 25 frame/s。

V CBD 算法适于对摄像机扫描和抖动视频序列中的小尺寸单目标进行实时检测。其局限性在于:分块处理使得目标边缘有锯齿现象;若用外接矩形标定,则检测目标比实际目标稍大。后续研究将针对边缘精细分割,拟采用形态学滤波和边缘检测处理方法进行;同时,将进一步针对视频序列中多运动目标研究新的检测算法。

参考文献:

- [1] SCHMUDDERICH J, WILLERT V. Estimating object proper motion using optical flow, kinematics and depth information [J]. *IEEE Trans. On Systems, Man, and Cybernetics*, 2008, 38(4): 1139-1151.
- [2] HARITAOGLU I, HARWOOD D, DAVIS L. W4: real-time surveillance of people and their activities [J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 2000, 22(8): 809-830.
- [3] MEIER T, NGUN K N. Video segmentation for content-based coding [J]. *IEEE Trans. on Circuits and Systems for Video Technology*, 1999, 9(8): 1190-1203.
- [4] EBISAWA Y. Robust pupil detection by image difference with positional compensation[C]. *VEE-IMS*, 2009, 11-13: 143-148.
- [5] 姜永林, 屈楦深, 王常虹. 基于纹理及统计特征的视频背景提取 [J]. *光学 精密工程*, 2008, 16(1): 172-177.
JIANG Y L, QU ZH SH, WANG CH H. Video background extraction based on textural and statistical features [J]. *Opt. Precision Eng.*, 2008, 16(1): 172-177. (in Chinese)
- [6] CHENG J R, FOO S W. Dynamic directional gradient vector flow for snakes [J]. *IEEE Trans. on Image Processing*, 2006, 15(6): 1563-1571.
- [7] WU G, EU CH F, HOU Q Y, et al.. Target location method based on invariable moment feature

- matching [J]. *Opt. Precision Eng.*, 2009, 17(2): 230-238.
- [8] WEI ZH H, TIAN J, LIU CH G, *et al.*. An updating algorithm of self-adaptive background based on energy method[C]. *IEEE International Conference on Automation and Logistics*, 2008; 848-852.
- [9] ZHU J J, GUO B L. Electronic image stabilization system based on global feature tracking[J]. *Journal of Systems Engineering and Electronics*, 2008, 19(2): 228-233.
- [10] PHAM T D. Image segmentation using probabilistic fuzzy C-means clustering [C]. *Proceeding of International Conference on Image processing*, 2001, 21(3): 722-725.
- [11] FU Y, ERDEM A T, MURAT TEKALP A. Tracking visible boundary of objects using occlusion adaptive motion snake [J]. *IEEE Trans. on Image Processing*, 2000, 9(12): 2051-2060.
- [12] 薛陈, 朱明, 陈爱华. 鲁棒的基于改进 Mean-shift 的目标跟踪[J]. *光学精密工程*, 2010, 18(1): 234-239.
- XUE CH, ZHU M, CHEN A H. Robust object tracking based on improved Mean-shift algorithm [J]. *Opt. Precision Eng.*, 2010, 18(1): 234-239. (in Chinese)
- [13] MADDALENA L, PETROSINO A. Moving object detection for real-time applications[C]. *International Conference on Image Analysis and Processing*, 2007; 542-547.
- [14] 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.

作者简介:



朱娟娟(1980—),女,江苏南通人,博士,副教授,2005年、2009年于西安电子科技大学分别获得硕士、博士学位,主要从事图像工程及人工视觉方面的研究。E-mail: zhujoo@126.com



郭宝龙(1962—),男,陕西西安人,教授,博士生导师,1988年、1995年于西安电子科技大学分别获得硕士、博士学位,主要从事模式识别与智能系统,视频处理及通信等方面的研究。E-mail: blguo1199@126.com

(本栏目编辑:白雨虹)