

高速运动可视多目标精确检测研究¹

张文涛 李晓峰 李在铭

(电子科技大学通信信息学院 成都 610054)

摘 要 讨论了图像目标高速运动情况下邻帧差分运动检测法的严重缺陷。利用图像能量、高阶统计量理论,建立背景约束准则,运用块处理技术对序列场景进行快速背景重建,依靠重建背景进行运动区域检测。基于 DSP C54X 的平台测试表明:该文算法背景重建速度快,1 秒左右即可完成无失真重建;动目标检测精度高,与目标速度无关。具有重要的理论及应用价值。

关键词 运动检测,背景重建,图像能量,高阶统计量,高速动目标,块处理

中图分类号 TN911.74

1 引 言

运动目标检测技术在图像跟踪识别系统中具有重要地位,检测算法不精确会导致特征参数误差增大,甚至跟踪失败。在多数有关运动目标跟踪、识别的文献^[1-4]中,对动目标初始检测普遍采用邻帧差分检测法,认为差分后大于特定门限的像素点为运动目标点。目标非高速运动时,邻帧间位移较小,此种方法检测误差不大;目标高速运动时,邻帧间位移较大,此种方法将导致严重的目标区重叠现象,极大地影响了目标特征参数估计精度。本文利用图像能量、高阶统计量理论及块处理技术对序列场景进行快速背景重建,提出基于重建背景的背景差分多运动目标区检测方法。基于本文算法的实时运动目标检测系统系列测试表明:基于块处理技术的背景重建算法所需帧数少,通常情况下 $10 \leq N \leq 30$ (1s 左右)长度的序列帧即可无失真重建背景,即使对于目标较多情况(10 个以上),上述帧序范围内无失真重建区域也多达总背景 98% 以上;将重建背景用于动目标检测时,动目标区的检测精度高,取得了理想的效果。

2 图像动目标理论模型

设所研究图像序列为 $\{f_s(x, y)\}_{s=0}^N$, 假定目标 T_i 的帧序 s 存在范围为 $s \in [k_0, k_1]$, 则该范围内任意相邻两帧场景 $f_s(x, y)$, $f_{s+1}(x, y)$ 都可表示为

$$f_s(x, y) = b_{s,s+1}(x, y) + T_i(x, y) + u_s(x, y) \quad (1)$$

$$f_{s+1}(x, y) = b_{s,s+1}(x, y) + T_i(x + \Delta x, y + \Delta y) + v_{s+1}(x, y) + n_{s+1}(x, y) \quad (2)$$

式中 $b_{s,s+1}(x, y)$ 为场景 $f_s(x, y)$, $f_{s+1}(x, y)$ 的共同背景区, $u_s(x, y)$, $v_{s+1}(x, y)$ 分别为目标 T_i 在帧 s , $s+1$ 的遮挡区和恢复区, $T_i(x, y)$, $T_i(x + \Delta x, y + \Delta y)$ 对应目标在 s , $s+1$ 帧的运动位置 ($(\Delta x, \Delta y)$ 为目标形心位移矢量), $n_{s+1}(x, y)$ 为 $s+1$ 帧场景对 s 帧的相对干扰噪声。(1)式、(2)式相减,则邻帧差分法结果为

$$\begin{aligned} d_s(x, y) &= f_{s+1}(x, y) - f_s(x, y) \\ &= T_i(x + \Delta x, y + \Delta y) - T_i(x, y) + v_{s+1}(x, y) - u_s(x, y) + n_{s+1}(x, y) \end{aligned} \quad (3)$$

¹ 1999-06-24 收到, 1999-11-08 定稿

(3) 式中的运动区由 $v_{s+1}(x, y)$, $u_s(x, y)$ 和 $T_i(x + \Delta x, y + \Delta y) - T_i(x, y)$ 三者相并而成。图 1 为椭圆形动目标邻帧差分后各区域间关系: 白色区域 4 为共同背景区 $b_{s,s+1}(x, y)$, 区域 2 为遮挡区 $u_s(x, y)$, 区域 1 为恢复区 $v_{s+1}(x, y)$, 区域 3 对应 $T_i(x + \Delta x, y + \Delta y) - T_i(x, y)$ 。由图 1 知, 邻帧差分法所得运动区域具有以下特点:

(1) 目标非高速运动时, 形心位移矢量 $(\Delta x, \Delta y)$ 较小, 区域 1, 2 和 3 的并集与实际目标区差别不大, 此时区域 1 和区域 2 都比较小, 区域 3 接近实际目标区。运动区域估计误差较小。

(2) 随着目标速度提高, 形心位移 $(\Delta x, \Delta y)$ 逐渐增大。区域 1, 2 逐渐增大 (最大为实际目标区), 区域 3 逐渐减小, 直至为零。区域 1, 2 和 3 的并集与实际目标区差别增大, 目标重叠越来越明显, 最严重情况为 2 倍实际目标区。

当目标较多时, 重叠现象不仅来自同一目标, 还来源于相互间靠近的其它目标。因此, 检测高速运动的多目标时, 邻帧差分法不适用于运动目标区域的初始检测。

研究表明: 邻帧差分法缺陷在于其估计误差随目标运动速度增加而增大。因此只有采用与上述因素无关的检测算法才能从根本上提高目标区估计精度。

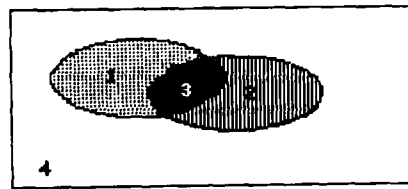


图 1 邻帧差分法原理示意图

3 背景重建与动目标区精确估计

基于上述原因, 邻帧差分法不适用于高速动目标区的初始检测。但如果知道理想背景 $B(x, y)$, 将邻帧差分变为各帧与背景差分, 显然无论运动速度多高, 都不存在目标重叠现象。因此, 获得精确动目标区关键是理想背景的重建。下面讨论如何在给定 $N + 1$ 帧图像序列 $\{f_s(x, y)\}_{s=0}^N$ 情况下, 来重建理想背景 $B(x, y)$, 以作为后续帧 $f_m(x, y)$, ($m > N$) 背景差分检测法中的参考背景。

由 (1) 式 and (2) 式知, 若利用某种合理约束准则, 估计出 $s, s+1$ 帧中共同背景 $b_{s,s+1}(x, y)$, $s+1, s+2$ 帧中共同背景 $b_{s+1,s+2}(x, y), \dots, N-1$ 帧和 N 帧的共同背景 $b_{N-1,N}(x, y)$, 则可利用式

$$b(x, y) = \cup_{s=0}^{N-1} b_{s,s+1}(x, y) \quad (4)$$

来重建理想背景 $B(x, y)$, 其中“ \cup ”为图像拼接算子, 表示将 $s, s+1$ 帧中共同背景 $b_{s,s+1}(x, y)$ 拼接到已估计背景 $b(x, y)$ 中。下面就共同背景约束准则建立、背景重建失真性进行讨论。

3.1 背景区约束准则建立

将场景 $f_s(x, y)$ 按从左到右, 从上到下空间顺序分成系列 $K \times K (K \geq 8)$ 大小的图像子块集合 $\{Bk(s, j)\}_{j=0}^M (M$ 为子块总数), 将差分场景 $d_s(x, y)$ 按同样方式分成子块系列 $\{Dbk(s, j)\}_{j=0}^M$ 。研究表明: 在 $d_s(x, y)$ 中, 空间位置对应于共同背景区 $b_{s,s+1}(x, y)$ 的子块 $Dbk(s, j)$ 内随机噪声点序列 $\{n_{s+1}(x, y), (x, y) \in Dbk(s, j)\}$ 通常呈高斯分布; 而对应于运动目标区的子块 $Dbk(s, w)$, 由于目标轮廓仍然存在, 其分布不具有高斯性。高斯分布随机变量一个重要属性是 4 阶统计量为零^[5], 因此区域 $b_{s,s+1}(x, y)$ 对应的各子块 $Dbk(s, j)$ 的 4

阶统计量近似为零或很小。通常直接估计随机分布变量 4 阶统计量比较繁琐,但对于任何随机分布,可以利用 4 阶统计量与 4 阶中心矩 m_4 , 方差 m_2 的关系^[6],按 (5) 式至 (8) 式方便求解,即

$$\text{HOS}_4 = m_4 - 3m_2^2 \quad (5)$$

$$m_4 = \frac{1}{K^2} \sum \sum_{(x,y) \in \text{Dbk}(s,j)} (d_s(x,y) - m_0)^4 \quad (6)$$

$$m_2 = \frac{1}{K^2} \sum \sum_{(x,y) \in \text{Dbk}(s,j)} (d_s(x,y) - m_0)^2 \quad (7)$$

$$m_0 = \frac{1}{K^2} \sum \sum_{(x,y) \in \text{Dbk}(s,j)} d_s(x,y) \quad (8)$$

各式中 $d_s(x,y)$ 为差分场景中点 (x,y) 处灰度值。并且,由于相对干扰噪声 $n_{s+1}(x,y)$ 幅度较小,空间位置上对应于共同背景区 $b_{s,s+1}(x,y)$ 的差分场景 $d_s(x,y)$ 子块 $\text{Dbk}(s,j)$ 内像素灰度值总体幅度很小,而对应于目标区的子块 $\text{Dbk}(s,w)$ 内像素灰度值总体幅度较大。因此,定义子块平均能量来判定其(运动/背景)归属。子块平均能量定义为

$$\text{Dbk}(s,p)_E = \frac{1}{K^2} \sum \sum_{(x,y) \in \text{Dbk}(s,p)} d_s(x,y)^2 \quad (9)$$

综合上述约束条件,针对场景 $f_s(x,y)$ 中子块 $Bk(s,j)$,其背景、动目标区归属判定由 $d_s(x,y)$ 中与其同位置(标号)子块 $\text{Dbk}(s,j)$ 约束条件决定。即

$$\text{Cons}(Bk(s,j)) = \lambda_1 \text{HOS}_4(\text{Dbk}(s,j)) + \lambda_2 \text{Dbk}(s,j)_E \quad (10)$$

其中正加权系数 λ_1, λ_2 满足条件 $\lambda_1 + \lambda_2 = 1$ 。这样,对于事先通过实验确定的门限 TB ,如满足条件:

$$\text{Cons}(Bk(s,j)) \leq TB \quad (11)$$

则判定子块 $Bk(s,j)$ 为帧 $s, s+1$ 的共同背景子块,否则为帧 $s, s+1$ 的动目标区子块。显然,当推广到多运动目标情况时,作为共同背景判别条件的 (11) 式仍然成立。这样,帧 $s, s+1$ 共同背景 $b_{s,s+1}(x,y)$ 为

$$b_{s,s+1}(x,y) = \{Bk(s,j) | \text{Cons}(Bk(s,j)) < TB, j = 0, 1, \dots, M-1\} \quad (12)$$

3.2 重建失真性讨论

由上述讨论知,针对图像序列 $\{f_s(x,y)\}_{s=0}^N$,可由 (4) 式和 (12) 式重建背景 $b(x,y)$ 。若无失真重建,由 (4) 式知,应满足

$$b(x,y) \geq B(x,y) \quad (13)$$

(13) 式物理意义是,图像场景任意子块在序列 $\{f_s(x,y)\}_{s=0}^N$ 中至少有一次在某相邻两帧中没有运动目标(全体或部分)存在。这是无背景失真重建的充要条件。若提供的帧列长 N 较短,以至使 j 标号子块在各 $f_s(x,y)$ 帧中均不满足 (11) 式约束条件,则取使 (11) 式达到最小的帧的 j 标号子块作为重建背景的 j 标号子块,即将 \hat{s} 帧

$$\hat{s} = \arg \min_s (\text{Cons}(Bk(s,j))), \quad s = 0, 1, \dots, N \quad (14)$$

中子块 $Bk(\hat{s}, j)$ 作为重建背景的 j 标号子块。因为目标高速运动时, 图 1 中区域 3 近似为零, 目标区 1, 2 都是由背景与动目标差分产生, 因此 (14) 式表示取各帧中目标区与背景差异最小的子块作为该标号处有失真重建背景的子块, 这显然是符合人的视觉特性的。

3.3 背景重建算法实现

依据上述思想, 建立基于 $N + 1$ 帧图像序列 $\{f_s(x, y)\}_{s=0}^N$ 背景重建 C 算法如下:

步骤 1 置子块标号数组 $\text{Label}[M] = \{0, 1, 2, \dots, M - 1\}$, 恢复背景 $b(x, y) = \Phi(\text{空集})$, 初始化满足 (14) 式的帧序数组 $\text{Frame}[M] = \{0, 0, \dots, 0\}$, 置各子块按 (10) 式取值数组 $N\text{Back}$ 为正无穷大, 即 $N\text{Back}[M] = \{\infty, \infty, \dots, \infty\}$ 。

步骤 2

```

For( $s = 0$ ;  $s < N$ ;  $s++$ )
  { $d_s(x, y) = f_{s+1}(x, y) - f_s(x, y)$ ;
  For( $i = 0$ ;  $i \leq M - 1$ ;  $i++$ )
    {If(( $\text{Cons}(Bk(s, i)) < \text{TB}$ )&&( $\text{Label}[i] \neq (-1)$ ))
       $\text{Label}[i] = -1$ ;  $b(x, y) = b(x, y) \cup Bk(s, i)$ ;
    else
      If( $N\text{Back}[i] > \text{Cons}(Bk(s, i))$ )
        { $N\text{Back}[i] = \text{Cons}(N.Bk(s, i))$ ;  $\text{Frame}[i] = s$ ; }
      {If( $b(x, y) \geq B(x, y)$ )
        {goto 步骤4: }
      }
    }
  }

```

步骤 3

```

For( $i = 0$ ;  $i \leq M - 1$ ;  $i++$ )
  {If( $\text{Label}[i] \neq (-1)$ )
     $b(x, y) = b(x, y) \cup Bk(\text{Frame}[i], i)$ ;
  }

```

步骤 4 程序结束。

算法原理如下: 首先取当前两邻帧场景 $s, s + 1$, 获得差分场景 $d_s(x, y)$; 然后从子块标号集合 Label 取出到帧 s 为止仍未被背景重建的 i 标号子块 (即 $\text{Label}[i] \neq (-1)$), 如该子块满足 (11) 式约束条件, 则将其拼接到重建背景上, 标注重建背景 i 标号子块已被找到, 以免再重建该子块 (即置 $\text{Label}[i] = (-1)$)。同时判断整个背景是否重建完成, 以决定是否结束无失真背景重建。当 $Bk(s, i)$ 不满足 (11) 式时, 通过 (14) 式、 $N\text{Back}[i]$, 保存前 s 帧中使 i 标号子块背景约束条件最小的帧序 \hat{s} 置于数组元素 $\text{Frame}[i]$ 中; 最后, 如果 N 帧信息利用完之后 $b(x, y)$ 仍不满足 (13) 式条件, 则将 Frame 中剩下的背景约束条件最小子块系列拼接到 $b(x, y)$ 中, 完成有失真背景重建。由于采用块处理技术, 并且每次只对未被重建子块进行评估, 因此运算负荷很小。

4 测试结果与分析

由上述算法, 我们开发出一个基于 DSP 平台, 用于实时精确检测车、船等高速运动目标的检测系统。该系统由背景重建和运动区域检测两个模块组成。系列实地场景测试取得了理想效果。背景重建结果表明: 本文算法重建时间短, 通常情况下 $10 \leq N \leq 30$ (1s 左右) 长度序列帧即可满足 (13) 式无失真背景重建条件。即使对于目标较多情况 (10 个以上), 上述帧序范围内无失真重建区域也多达 98% 以上, 且失真子块多呈分散状, 与背景差别不大, 对区域检测模块影响很小。图 2 为一典型多运动目标有失真背景重建、运动目标检测结果, 各参数为: $N=30$, $K=12$, $\lambda_1=0.4$, $\lambda_2=0.6$, $TB=120$ 。图中 A1 至 A5 标号图像

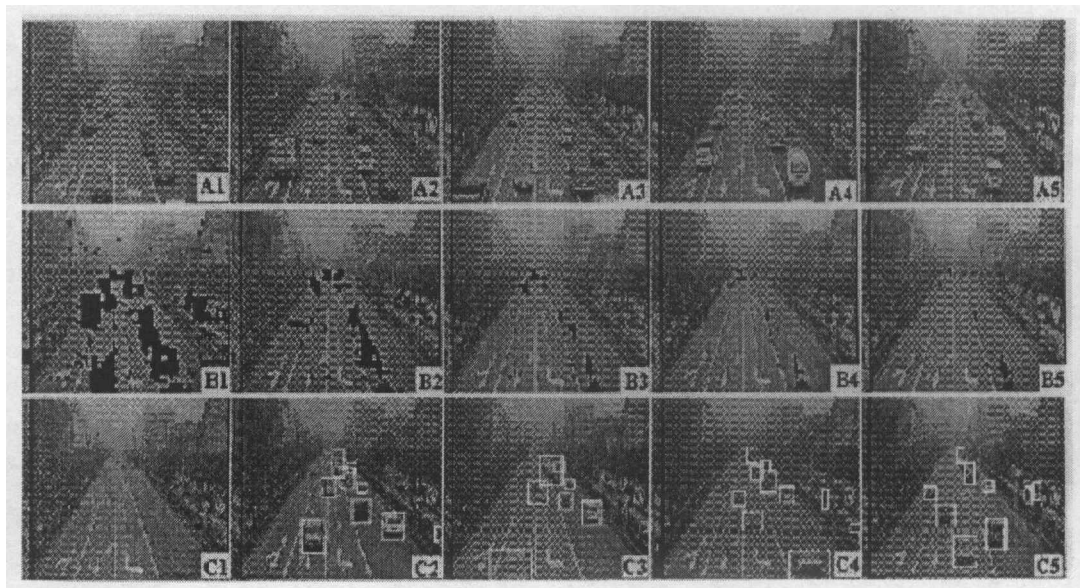


图 2 序列场景背景重建、运动目标典型检测结果

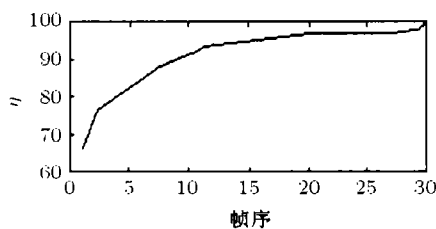


图 3 各帧场景时无失真背景重建百分比进度曲线

对应序列场景中的第 2, 10, 18, 24, 30 帧; B1 至 B5 标号图像则分别对应到 A1 至 A5 为止的无失真重建背景 $b(x, y)$ 结果; C1 为按 (14) 式补齐有失真子块后最终结果, C2 至 C5 为几帧基于重建背景的背景差分运动目标检测结果。各帧时背景重建百分比 $\eta = b(x, y)/B(x, y)$ 的进度如图 3 所示。

参 考 文 献

- [1] Meier Thomas, Automatic segmentation of moving objects for video object plane generation, IEEE Trans. on CAS VT, 1998, 8(5), 525-538.
- [2] J. G. Jeon, K. T. Park, Multi-target tracking system using texture, SPIE, Vol.3024, 1997, 229-236.
- [3] S. Gil, T. Pun, R. Milanese, Comparing features for target tracking in traffic scenes, Pattern Recognition, 1996, 29(8), 1285-1296.

- [4] M. A. Jain, Vehicle segmentation and classification using deformable templates, IEEE Trans. on PAMI, 1996, 18(3), 293-308.
- [5] 张贤达, 现代信号处理, 北京, 清华大学出版社, 1996, 373-380.
- [6] 王宏禹, 现代谱估计, 南京, 东南大学出版社, 1991, 304-320.

ACCURATE REGION DETECTION OF HIGH-SPEED MULTI-TARGET VISUAL SYSTEM

Zhang Wentao Li Xiaofeng Li Zaiming

(*Institute of Comm. and Info. Eng., UEST of China, Chengdu 610054, China*)

Abstract This paper discusses shortcomings of adjacent-frame difference. Based on the image energy and high order statistic (HOS) theory, background reconstruction constraint conditions are setup. Under the help of block- processing technology, background is reconstructed quickly. The DSP based platform tests indicate that the background can be recovered lossless in about one second, while moving regions estimations are not influenced by target speed.

Key words Motion detection, Background reconstruction, Image energy, HOS, High-speed target, Block processing

张文涛: 男, 1971 年生, 博士生, 研究领域: 视频序列多运动图像目标特征提取、识别、运动分析、基于内容功能的视频序列检索与过滤.

李晓峰: 男, 1963 年生, 副教授, 研究领域: 通信信息处理及 DSP 实现技术研究.

李在铭: 男, 1939 年生, 教授, 博士生导师, 研究领域: 多媒体技术、图像跟踪、编码与压缩技术研究.