

高速密集视频目标场景下的运动分析

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: 本文首先分析了高速、密集图像目标场景下邻帧差分运动估计检测法的缺陷. 然后综合图像能量、高阶统计量预测理论以及块处理技术, 对高速密集图像序列进行快速背景重建、运动区域提取与运动分析. 基于 DSP 的平台测试表明: 本文背景重建算法速度快, 运动区域检测精度高, 目标跟踪准确. 为密集、高速的图像目标跟踪系统的研制提供了新的可借鉴方法.

关键词: 高速运动; 背景重建; 密集图像目标场景; 跟踪与识别

中图分类号: TN941.1 **文献标识码:** A **文章编号:** 0372-2112 (2000) 10-0114-04

Motion Analyses under High Speed Dense Visual-Target Scenes

ZHANG Wen-tao, LI Xiao-feng, LI Zai-ming

(Dept. of Communication and Information, UESTC, Chengdu 610054, China)

Abstract: The paper discusses shortcomings of adjacent frame difference. Based on the image energy, high order statistic (HOS) prediction theory and block processing techniques, background of high speed dense visual target motion sequences is reconstructed, and motion states are studied. Our DSP based platform tests indicate the provided algorithm performs excellently in real visual sequences motion analyses.

Key words: high speed motion; background reconstruction; dense visual target scene; recognition and tracking

1 引言

运动目标检测是图像跟踪识别系统中的一个重要环节, 算法的不精确常会导致估计参数误差的增大, 甚至跟踪的失败. 目前多数有关运动目标跟踪、识别的文献^[1-3]中, 对运动目标区域提取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, 该方法检测误差不大; 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. 本文利用图像能量、高阶统计量理论及块处理技术, 首先对图像序列场景进行快速背景重建, 然后提出基于该背景的高速密集目标运动分析方法. 基于本文算法的实时 DSP 运动分析系统测试表明: 基于块处理技术的背景重建算法所需帧数少, 重建效果理想; 将重建背景用于运动分析时, 取得了理想的效果.

2 运动区域组成划分

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

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

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

其中 $b_{s,s+1}(x, y)$ 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; $T_i(x, y)$ 、 $T_i(x + x, y + y)$ 对应目标在 s 、 $s+1$ 帧的运动位置 ((x, y) 为目标形心位移矢量), 噪声 $m_s(x, y)$ 主要由 CCD 充电耦合器件的电子噪声、光电子噪声和感光胶片颗粒噪声组成^[4]; 定义 $n_{s,s+1}(x, y) = m_{s+1}(x, y) - m_s(x, y)$ 为 $s+1$ 帧场景对 s 帧的相对干扰噪声. 对式 (1)、式 (2) 相减, 则邻帧差分结果为

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

式 (3) 中的运动区由 $v_{s+1}(x, y)$ 、 $u_s(x, y)$ 和 $T_i(x + x, y + y) - T_i(x, y)$ 三者相并而成. 图 1 为椭圆形运动目标邻帧差分后各区域间关系: 白色区域 R_4 为共同背景区 $b_{s,s+1}(x, y)$, R_2 为遮挡区 $u_s(x, y)$, R_1 为恢复区 $v_{s+1}(x, y)$, R_3 对应 $T_i(x + x, y + y) - T_i(x, y)$. 若定义运动区域估计相对误差 $Error$ 为

$$Error = (R_1 + R_2 + R_3 - R_4) / R_4 \quad (4)$$

其中 R_4 为目标 T_i 的实际运动区域, 得到邻帧差分法的 $Error$ 有以下结论:

结论 1 若 $\sqrt{x^2 + y^2} \rightarrow 0$, 则 $R_1 \rightarrow 0, R_2 \rightarrow 0, R_3 \rightarrow R_{T_i}$, $Error \rightarrow 0$.

结论 2 若 $\sqrt{x^2 + y^2} \uparrow$, 则 $R_1 \uparrow, R_2 \uparrow, R_3 \downarrow, Error \uparrow$, 直至 $Error = 1$.

结论 2 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; 密集目标场景下, 重叠现象不仅来自同一目标, 还来自相互间靠近的其它目标. 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.

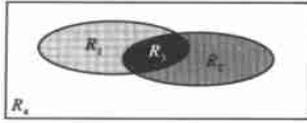


图 1 邻帧差分法原理

3 背景重建理论模型

基于上述原因, 邻帧差分法不适用于密集、高速运动目标区的估计. 但假如知道无任何运动目标干扰时的背景 $B(x, y)$, 则将邻帧差分变为各帧与 $B(x, y)$ 差分, 则显然不存在目标重叠现象. 由于本文中 $B(x, y)$ 代表无运动目标干扰(但包含电子噪声、光电子噪声和感光胶片颗粒噪声)的真实观测视频场景, 因此称 $B(x, y)$ 为理想背景或无运动目标干扰背景. 下面讨论如何利用图像序列 $\{f_s(x, y)\}_{s=0}^N$ 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.

由式(1)和式(2)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 $b_{s+1, s+2}(x, y)$, ..., $N-1$ 帧和 N 帧的共同背景 $b_{N-1, N}(x, y)$, 则可利用

$$b(x, y) = \sum_{s=0}^{N-1} b_{s, s+1}(x, y) \quad (5)$$

来逼近、重建理想背景 $B(x, y)$. 其中“ \sum ”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), 如图 2 所示. 将差分场景 $d_s(x, y)$ 按同样方式分成子块系列 $\{Dbk(s, j)\}_{j=0}^M$.

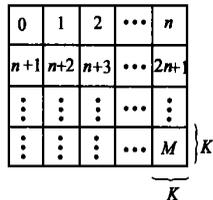


图 2 子块划分

考虑式(1)、(2), 由于噪声 $m_s(x, y)$ 由电子运动产生的, 因此相对由被拍摄目标反射而形成的非噪声图像灰度值相比, $n_{s, s+1}(x, y)$ 幅度较小. 该现象可具体体现在对相邻时刻同一场景的两幅图像进行差分, 获得图像的灰度值几乎为零, 人眼无法发现的实事基础之上; 另外, 文献[4]还指出, 热噪声、光电子噪声和感光颗粒噪声随机过程分布在理论上属平稳随机过程且呈遍历性, 均可采用高斯分布的数学模型对其分析、研究. 因此, 对于差分场景 $d_s(x, y)$, 可导出子块 $Dbk(s, j)$ 内序列点 $\{n_{s, s+1}$

$(x, y)\}$ 具有以下两个性质:

性质 1 若 $Bk(s, j) \subset b_{s, s+1}(x, y)$, 则概率密度函数 $p(n_{s, s+1}(x, y))$ 分布近似呈高斯性; 否则 $p(n_{s, s+1}(x, y))$ 不具有高斯性.

性质 2 若 $Bk(s, j) \subset b_{s, s+1}(x, y)$, 则子块 $Dbk(s, j)$ 能量较小; 否则能量较大.

由于高斯分布随机变量一个重要属性是 4 阶统计量为零^[5], 因此由性质 1 知, 区域 $b_{s, s+1}(x, y)$ 对应子块 $Dbk(s, j)$ 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], 按式(6)~(9)方便求解, 即

$$HOS_4 = m_4 - 3m_2^2 \quad (6)$$

$$m_4 = \frac{1}{K^2} \int_{(x, y) \in Dbk(s, j)} (d_s(x, y) - m_0)^4 \quad (7)$$

$$m_2 = \frac{1}{K^2} \int_{(x, y) \in Dbk(s, j)} (d_s(x, y) - m_0)^2 \quad (8)$$

$$m_0 = \frac{1}{K^2} \int_{(x, y) \in Dbk(s, j)} d_s(x, y) \quad (9)$$

同时定义子块平均能量为

$$Energy(Dbk(s, j)) = \frac{1}{K^2} \int_{(x, y) \in Dbk(s, j)} d_s(x, y)^2 \quad (10)$$

综合性质 1、2 知, 对于场景 $f_s(x, y)$ 中子块 $Bk(s, j)$, 归属 $b_{s, s+1}(x, y)$ 的判定函数可定义为

$$Cons(Bk(s, j)) = \frac{1}{4} (\text{Sig}(T_0 - HOS_4(Dbk(s, j))) + 1) (\text{Sig}(T_1 - Energy(Dbk(s, j))) + 1) \quad (11)$$

其中符号函数 $\text{Sig}(\cdot)$ 为

$$\text{Sig}(x) = \{ p \mid \text{if } x \geq 0, p = 1; \text{ else } p = -1 \} \quad (12)$$

式(11)中 T_0 代表从高斯分布角度来衡量子块 $Bk(s, j)$ 是否属于背景的最大可能检测门限. 其物理含义为: T_0 越大, 则对子块灰度集合是否满足高斯性的检测标准要求越低, 否则越高. 当 $T_0 = 0$ 时最为苛刻, 对应子块灰度分布为理想的高斯分布. 尽管此种情况实际中并不存在, 但研究表明实际情况下目标子块与背景子块的 $HOS_4(\cdot)$ 值相差数十倍, 因此 T_0 比较容易确定; 同时, 参数 T_1 从子块内总体灰度能量的角度来判别子块的前背景归属, 其值越小, 要求越高. 通常当子块内存在目标时, 其差分后仍可发现一定的目标轮廓, 说明该子块的能量较大. 而无目标的静止背景差分后的能量则很小, 这两种情况下的 $Energy(\cdot)$ 值同样相差悬殊. 同样很容易通过实验确定 T_1 的判别门限.

综上所述讨论, 若同时考虑子块的高斯性及能量两种判别条件, 则如果

$$Cons(Bk(s, j)) = 1 \quad (13)$$

则 $Bk(s, j) \subset b_{s, s+1}(x, y)$, 否则 $Bk(s, j) \not\subset b_{s, s+1}(x, y)$. 显然, 当推广到密集多运动目标情况时, 作为共同背景判别条件的式(13)仍然成立. 这样, 帧 $s, s+1$ 共同背景 $b_{s, s+1}(x, y)$ 为

$$b_{s, s+1}(x, y) = \{ Bk(s, j) \mid Cons(Bk(s, j)) = 1, j = 0, 1, \dots, M-1 \} \quad (14)$$

3.2 重建失真性讨论

由上述讨论知,针对图像序列 $\{f_s(x, y)\}_{s=0}^N$,可由式(5)和式(14)重建背景 $b(x, y)$.若无失真重建,由式(5)知,应满足

$$b(x, y) = B(x, y) \tag{15}$$

式(15)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.这是背景无失真重建的充要条件.若提供的帧长 N 较短,以至使 j 标号子块在各 $f_s(x, y)$ 帧中均不满足式(13)约束条件,则分别沿上、下、左、右四个方向搜索四个最近的无失真重建子块,标号为 u 、 d 、 l 、 r ,如图3所示.这样,对于 j 标号子块内任意位置点 (x, y) ,其灰度值可由式(16)的线性预测器得到

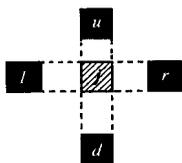


图3 子块搜索

$$g(x, y) = \frac{\prod_{i=0}^3 p_{h_i}(x, y)}{\prod_{i=0}^3 h_i} \tag{16}$$

其中 $h_i, i=0, 1, 2, 3$ 代表子块 j 内点 (x, y) 分别距离这四个子块内所有像素的最近距离值, p_{h_i} 则代表与上述各距离所对应的四个像素点灰度值.

4 运动分析

当按上述块处理技术完成对序列图像场景背景的重建

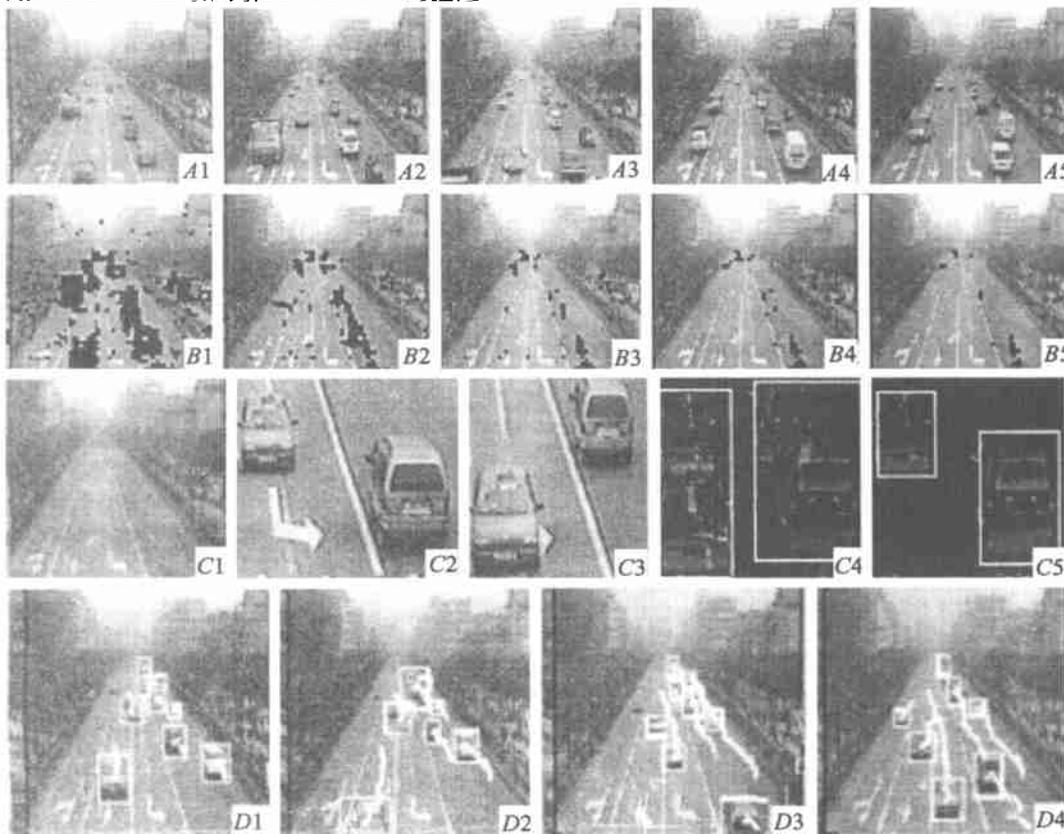


图4 序列场景背景重建、运动分析结果

5 测试结果与分析

由上述算法,作者开发出一个基于 DSP 平台,用于研究

后,可精确提取帧 $f_s(x, y)$ 中的系列运动目标区为

$$R(k, x, y) = \text{close}(\text{structure}, \text{differ}(k, x, y) = \bigcup_{i=1}^{n_k} r_i(k) \tag{17}$$

其中函数 $\text{close}(\cdot)$ 代表形态学中的闭合算子^[8], structure 为结构元素, $\{r_i(k)\}_{i=1}^{n_k}$ 代表第 k 帧所检测的运动目标集合,对于二值化门限 S_0 , $\text{differ}(k)$ 为

$$\text{differ}(k) = \{(x, y) \mid |f(k, x, y) - B(x, y)| \geq S_0\} \tag{18}$$

对于目标为 $r_i(k)$, 设其形心坐标为 $z_i(k)$, $k+1$ 帧时该目标形心测量值为 $\tilde{z}_i(k+1)$, 则该目标在 $k+1$ 帧场景中位置、速度 - 滤波器为

$$\hat{z}_i(k+1/k) = \hat{z}_i(k) + T\hat{z}_i(k) \tag{19}$$

$$\hat{z}_i(k+1/k) = \hat{z}_i(k) \tag{20}$$

$$\hat{z}_i(k+1) = \hat{z}_i(k+1/k) + [\tilde{z}_i(k+1) - \hat{z}_i(k+1/k)] \tag{21}$$

$$\hat{z}_i(k+1) = \hat{z}_i(k+1/k) + \frac{1}{T}[\tilde{z}_i(k+1) - \hat{z}_i(k+1/k)] \tag{22}$$

其中 $\hat{z}_i(k)$ 、 $\hat{z}_i(k)$ 为目标 $r_i(k)$ 的形心位置、速度估值, α 、 T 为滤波器系数. 滤波器的优点在于形式简单,易于工程实现,每次滤波循环迭代时间比 Kalman 滤波器节约 70% 左右^[7].

车、船等高速运动目标运动状况的运动分析系统.该系统由背景重建和运动分析两个模块组成.系列实地场景测试取得了理想效果.图4为一典型多运动目标有失真背景重建、目标形

心运动轨迹分析结果,各参数为: $N = 30$ 、 $K = 12$ 、 $T_0 = 40$ 、 $T_1 = 70$ 。图中 $A1$ 至 $A5$ 标号图像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; $C1$ 为按式(16)预测、补齐有失真子块后最终重建场景结果; $C2$ 、 $C3$ 为实际两相邻采样运动场景图像, $C4$ 是采用传统邻帧差分的运动区域检测结果。作为对比,图 $C5$ 给出了本文的基于重建背景的运动差分检测结果;最后,视频序列 $D1$ 至 $D4$ 是相同场景下某段时间内总体算法的运动检测、分析结果。

参考文献:

- [1] Thomas Meier. Automatic segmentation of moving objects for video object plane generation [J]. IEEE Trans. on Circuits and Systems for Video Technology, 1998, 8(5): 525 - 538.
- [2] S. T. Comparing features for target tracking in traffic scenes [J]. Pattern Recognition, 1996, 29(8): 1285 - 1296.
- [3] Me. A. J. Vehicle Segmentation and classification using deformable templates [J]. IEEE Trans. On PAMI, 1996, 18(3): 293 - 308.
- [4] Kenneth. R. Castleman 著,朱志刚等译. 数字图像处理 [M]. 北京:电子工业出版社, 1998: 359 - 361.
- [5] 张贤达. 现代信号处理 [M]. 清华大学出版社, 1994.
- [6] 王宏禹. 现代谱估计 [M]. 东南大学出版社, 1991.
- [7] 周宏仁, 敬忠良, 王培德. 机动目标跟踪 [M]. 国防工业出版社.
- [8] L. D. and C. G. Analysis and tracking of destructive tests on working turbfans involving image processing techniques [A]. SPIE 1996 [C], 2785: 142 - 149.

作者简介:



张文涛 博士生,研究方向:图像目标的识别、跟踪与基于图像内容的网络检索。

李晓峰 副教授,硕士生导师,从事通信信息处理与 DSP 方面研究。

李在铭 教授,博士生导师,从事通信与多媒体技术研究。

(上接第 129 页)

参考文献:

- [1] 杨福生. 小波变换的工程分析与应用 [M]. 北京:清华大学出版社, 1996.
- [2] IEEE Transactions on Pattern Analysis and Machine Intelligence [J]. July 1992, 14(7): 710 - 732.
- [3] 高振明, 陈献民. 窄带无线通信中数字信号的频谱编码传输方法 [J]. 电子学报, 1997, 25(10): 82 - 84.
- [4] [美] A. V. 奥本海姆. 信号与系统 [M]. 西安交通大学出版社, 1990.
- [5] [美] Mathworks 公司. MATLAB wavelet toolbox user's guide [Z], V5. 12
- [6] Gossmann A. Wavelet transform and edge detection [A]. In: Hazewinkel M. ed. Stochastic Processing in Physics and Engineering [A]. Dordrecht: Reidel, 1986.
- [7] 数字信号处理原理及应用 [M]. 电子工业出版社, 1996.