

文章编号:1671-8860(2005)08-0723-05

文献标志码:A

一种新的运动检测及轮廓追踪方法

吕常魁¹ 姜澄宇¹ 王宁生¹

(1 南京航空航天大学 CIMS 工程研究中心,南京市御道街 29 号, 210016)

摘要:构建了基于块区域二阶矩比的三帧差分运动检测模型,应用文献[1]中的核密度估计模型来滤除其中的非运动区域,并采用支持向量聚类实现多目标检测。给出了运动物体轮廓快速追踪的 RW 算法,该算法不必考虑检测区域内部的细节问题,能够充分利用所检测到的边缘信息,比较符合人眼对轮廓的搜索习惯。最后给出了实验结果。

关键词:运动检测;轮廓追踪;时间差分;颜色矩;滚轮算法

中图法分类号:TN941.1;P237.3

运动检测是计算机视觉应用中一个重要环节。运动检测的主要目的是根据序列图像将运动物体从背景图像中提取出来。对运动区域大多利用连续图像的时间、空间信息进行分割,常用的方法如背景减除^[1~5]、时间差分^[6,7]及光流^[8]等。光流算法复杂,抗噪性能差,且如果没有特殊的硬件设施,几乎不可能被用于视频流的实时处理;背景减除一般能够提供最完全的特征数据,但对由于光照及其他外界扰动引起的场景变化过于敏感;时间差分运动检测方法对于动态环境具有较强的自适应性,但一般不能完全提取出所有相关的特征像素点,容易产生实体内部的空洞及边缘尾迹现象。

本文充分利用时间差分的自适应性,构建了基于局部区域二阶统计矩比的三帧差分统计模型,同时应用文献[1]中的背景减除模型来滤除其中的非运动区域,并采用支持向量聚类实现多目标检测。文章给出了运动物体快速轮廓追踪的 RW 算法,该算法不必考虑检测区域内部的细节问题,能够充分利用所检测到的边缘信息,比较符合人眼对轮廓的搜索习惯。

1 基本模型

时间差分在连续图像序列中的相邻帧间采用基于像素的时间差分来提取出图像中的运动区

域。Lipton 等^[6]利用两帧差分方法从实际视频图像中检测出运动目标,进而用于目标的分类与跟踪;改进的方法是利用三帧差分代替两帧差分来判断当前像素点 $I_n(x, y)$ 是否运动^[7],对于阈值 $T_n(x, y)$,如果满足:

$$|I_n(x, y) - I_{n-1}(x, y)| > T_n(x, y)$$

$$\text{并且 } |I_n(x, y) - I_{n-2}(x, y)| > T_n(x, y)$$

则判定该点为移动点。时间差分运动检测方法对于动态环境具有较强的自适应性,但也存在较大的缺陷。

鉴于此,本文利用统计分析的优点,与文献[1]中的核密度估算背景减除模型相结合,构建了具有较好鲁棒性的基于区域级统计分析的三帧差分检测模型。基本模型如下。

将所检测场景图像等分为 $M \times N$ 个大小均为 $n \times n$ (像素) 的子窗口,这里采用 Stricker 和 Orengo 所提出的颜色矩(color moments)^[9] 分别对每个窗口进行颜色空间的统计描述。由于三阶矩(skewness)对噪声过于敏感,这里只采用其中的一阶矩(mean)和二阶矩(variance):

$$\mu_k = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n P_k(i, j)$$

$$\sigma_k = \left(\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (P_k(i, j) - \mu_k)^2 \right)^{\frac{1}{2}}$$

其中, $P_k(i, j)$ 为 $n \times n$ 子窗口中坐标为 (i, j) 的像素的第 k 个颜色分量。对于颜色空间,这里采用

对阴影具有较好鲁棒性的(r, g, s)色彩模型^[1]:

$$\left. \begin{array}{l} r = \frac{R}{R+G+B} \\ g = \frac{G}{R+G+B} \\ s = \frac{R+G+B}{3} \end{array} \right\} \begin{array}{l} \text{色变} \\ \text{光强度} \end{array}$$

为便于说明,本文称被移动物体全部或部分遮挡的子窗口为覆盖区域(covered regions),称所有被模型所检测到的疑似覆盖区域的窗口为反常区域(anomalous regions)。实验表明,覆盖区域对于窗口部分被遮挡(即物体边缘进入或移出窗口)的情况, s 分量颜色矩的变化幅度最为明显,其中二阶矩 σ_s 的变化尤其具有代表性,如图1所示。图1中,假设物体移入部分的像素值高于背景,从移动物体边缘进入窗口直到完全覆盖, σ_s 的值经历了一个马鞍形的变化。这说明与一阶矩 μ_s 相比, σ_s 对物体边缘进入窗口的反应更为敏感。并且实验证明, σ_s 对于光照的均匀变化、阴影遮挡甚至背景物体小幅度扰动等情况都具有较强的鲁棒性,并对噪声有一定的抑制作用。基于此,选取相邻帧对应子窗口 σ_s 的比值来度量窗口的变化。设全局阈值 T_1 ,令 $\sigma_s(n)$ 为第 n 帧子窗口 W 的 s 二阶矩,构造三帧差分模型如下。

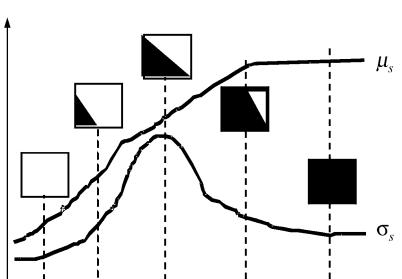


图1 覆盖区域 μ_s 和 σ_s 的变化示意

Fig. 1 Schematic Chart of the Variance of μ_s and σ_s in a Covered Region

W 为反常区域:

$$\text{s. t. } \left\{ \begin{array}{l} \frac{\sigma_s(n)}{\sigma_s(n-1)} > T_1 \\ \frac{\sigma_s(n)}{\sigma_s(n-2)} > T_1 \end{array} \right.$$

满足上述条件的 W 包含了比较完整的移动物体轮廓及内部变化的信息,但也可能包含帧间物体与背景分离区中比较明显的背景边缘以及其他外界干扰。图2为算法的检测结果示例,其中阈值 $T_1 = 2.5$,子窗口大小为 $5 \text{ 像素} \times 5 \text{ 像素}$ 。可以看出,二阶矩比是一个敏感的活动边缘探测器,行人的轮廓被完整地检测到,但由于两帧背景

分离区的部分区域存在明显的墙壁边缘,导致算法将该部分区域也误判断为反常区域。

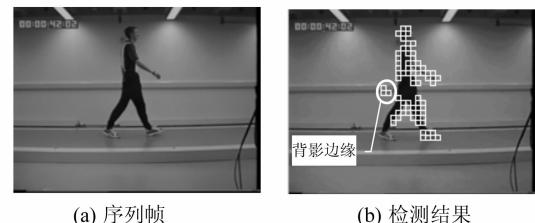


图2 基于 σ_s 比的三帧差分检测结果示例

Fig. 2 Experimental Result of the Three-frame Differencing Model Based on the Ratio of σ_s Between Neighbored Frames

借助目前比较流行的基于核密度估算的背景差分模型^[1]来滤除反常区域中的非覆盖区域。设定全局阈值 T_2 ,取颜色空间中 r, g 两分量的一阶矩作为输入参数,如果第 n 帧子窗口 W 为反常区域,则有:

W 为覆盖区域:

$$\text{s. t. } \Pr(W, r, g) = \frac{1}{m} \sum_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_r'^2}} e^{-\frac{(\mu_r(n)-\mu_r(i))^2}{\sigma_r'^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_g'^2}} e^{-\frac{(\mu_g(n)-\mu_g(i))^2}{\sigma_g'^2}} < T_2$$

其中, m 为连续采样的帧的数目,为避免与二阶矩混淆,以 σ_r' 、 σ_g' 表示核函数在 r, g 分量的带宽。同时假设一阶矩沿时间轴的局部分布满足正态分布 $N(\mu, \sigma^2)$,则根据文献[1],任一颜色分量对应的核函数带宽可估算为:

$$\sigma' = \frac{\Delta_{\text{mid}}}{0.68\sqrt{2}}$$

其中, Δ_{mid} 为取样帧中相应颜色分量所有相邻帧一阶矩绝对差值的中值。

设置0-1标识变量 $w(i, j)$ ($0 \leq i \leq M-1, 0 \leq j \leq N-1$, i, j 为对应窗口在窗口矩阵中的行标和列标),如果对应窗口为覆盖区域,则 $w(i, j) = 1$,否则为0。对于覆盖区域的每个窗口,构造四维向量 $\mathbf{X}_{ij} = (i, j, \mu_r, \mu_g)$ 。实际应用中,采用文献[10]中的邻近图法对集合 $\{\mathbf{X}_{ij}\}$ 进行快速支持向量聚类(support vector clustering, SVC)以实现实时的多目标检测,其中惩罚项系数 C 设置为1,即不允许存在野值点。

核密度估算的主要缺点是耗时的计算及其庞大的数据存储量需求,对硬件系统的配置要求比较高,难以满足实时性的要求。在本模型中,只存储每个子窗口的 μ_r, μ_g ,使得需要预先存储的数据量大大减少;而只对反常区域进行概率估计又

大幅度降低了计算的消耗,从而使得该模型能够在普通的 PC 机上实时运行。实验表明,以 (r, g, s) 模型的 r, g 分量作为输入参数,不但能够有效去除背景边缘的干扰,而且对由于阴影引起的光强度变化也很不敏感,这进一步增强了系统的鲁棒性。

2 轮廓追踪的 RW 算法

实际所检测到的覆盖区域往往形成一些互不连通的区域。在物体的最小边界矩形(minimum bounding rectangle, MBR)已知的前提下,物体轮廓的获得就成了不连续边界的轮廓追踪问题。常用的轮廓追踪算法如 Square 跟踪、Moore 近邻跟踪、Radial Sweep 及 Pavlidis 算法^[11]等对于不连续边界的追踪无能为力。传统的方法^[12]因为要计算整个物体所在区域,计算的耗费很大。目前大量文献采用活跃轮廓模型(active contour model)^[13]来实现基于轮廓的跟踪^[14],但活跃轮廓模型的以下局限性使得其应用受到一定限制:能量最小化过程收敛速度缓慢,实时性差。为避免受到噪声、遮挡等外界干扰,初始化轮廓需尽可能贴近物体的真实轮廓。

鉴于此,笔者创建了“滚轮算法”(roller wheel algorithm, RW 算法),同时实现了不连通边界区域的连接与边界追踪,而不必考虑区域内部的细节问题。RW 算法的基本原理是利用假想的单位滚轮按照一定的规则紧贴轮廓外沿滚动一周,所有与滚轮相接触的点构成的封闭轮廓即为移动物体边缘。该算法充分利用了边界区域的信息,比较符合人眼对轮廓的搜索习惯。

2.1 滚动规则与滚动向量

将 $w(i, j)$ 映射为 1×1 单元格空间,由 $\{w(i, j)\}$ 组成的矩阵 $H_{M \times N}$ 映射为由这些单元格组成的阵列;滚轮直径为 1, 占据 1 单元格。沿 $\{w(i, j) | w(i, j) = 1\}$ 外沿滚动(图 3), 定义滚动规则如下。

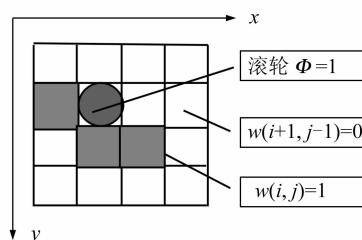


图 3 矩阵 H 的映射网格及滚轮示意图

Fig. 3 A Schematic View of the Roller Wheel and the Mapping from the Matrix H to the Grid Units

- 1) 只能向满足 $w(i, j) = 0$ 的单元格滚动;
- 2) 滚动步长为 1(即沿坐标轴四个方向滚动);
- 3) 不能回滚;
- 4) 下一滚动位置的 8 近邻至少包含一个当前与滚轮相切的满足 $w(i, j)$ 的单元格。

为便于说明,首先给出两个定义。

定义 1 滚动自由度。滚轮在某一位置按照滚动规则所能滚动的方向的总数,用 f 表示。上述规则将 f 限制为 1, 2, 3。

定义 2 滚动向量。设滚轮作一次滚动,前后所处单元格分别为 $w(m_1, n_1), w(m_2, n_2)$, 则定义滚动向量 $r = (m_2 - m_1, n_2 - n_1)$ 。根据滚动规则, r 只能为 $(0, \pm 1)$ 或 $(\pm 1, 0)$ 。

设当前位置为 $w(i, j)$, 下一滚动向量为 r , 则下一滚动位置为 $w(i + r_x, j + r_y)$ 。自由度 f 实际上确定了滚轮的位置状态。基本的位置状态有三种。不同位置状态的相邻滚动向量间存在一定的角度旋转关系。如图 4 所示(设滚轮作顺时针滚动,以下同),图 4 中, θ 为相邻滚动向量的夹角。

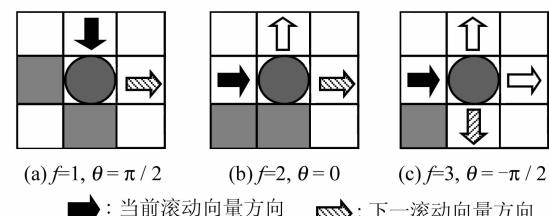


图 4 f 所确定的滚轮的三种位置状态及相应的滚动向量方向

Fig. 4 The Three Different States Describing the Position of the Roller Wheel Defined by f

为此,定义平面旋转矩阵 R_θ :

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

设当前滚动向量 r' , 则下一滚动向量 $r = r' \times R_\theta$ 。

2.2 搜索填充规则

滚轮在正式滚动之前要作预滚动,预滚动的目的是为了进行搜索填充。所谓搜索填充是指在预滚动位置沿搜索向量方向搜索滚轮所未“滚过”的满足 $w(i, j) = 1$ 的单元格,而且一旦搜索到满足条件的单元格,则沿搜索轨迹将所有位于预滚动点(含预滚动点)与该单元格之间的单元格的 w 置 1(即填充),然后再回到预滚动的初始位置作正式滚动,其目的是最大范围地利用并连接边缘覆盖区域片断。预滚动和正式滚动均按滚动规则进行。完整的预滚动与正式滚动过程如图 5 所

示。图5(b)中,空心圆为预滚动位置,箭头为搜索向量方向。

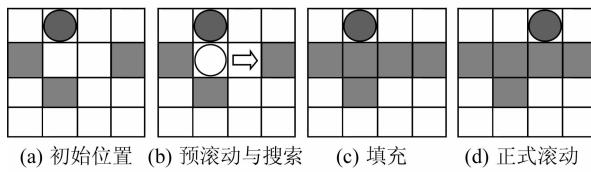


图5 搜索填充过程示意图

Fig. 5 Demonstration of the Search-and-Filling Process

设预滚动后的滚动向量为 r ,则无论当前预滚动位置处于何种状态,搜索向量均可由 r 旋转 $\pi/2$ 角度获得。设搜索向量为 p ,则 $p = r \times R_{\pi/2}$ 。由此,定义搜索填充规则如下。

- 1) 沿 p 方向搜索第一个满足 $w(i, j) = 1$ 的单元格(直到超出规定的最大边界);
- 2) 单元格 $w(i - p_x, j - p_y)$ 应该未曾被正式滚动的滚轮所占据过,即正式滚动的滚轮轨迹不包含 $w(i - p_x, j - p_y)$;
- 3) 如果存在单元格 $w(i, j)$ 同时满足以上两个条件,则沿 p 方向将预滚动点至 $w(i - p_x, j - p_y)$ 之间单元格的 w 置 1。

2.3 轮廓追踪的实现

具体实现中的关键是滚轮位置及第一步滚动的初始化。初始化过程如下。

- 1) 确定所处理的聚类区域的最大边界,周边各扩充一单元格以搁置滚轮;
- 2) 由边界中具有最小列标的单元格中任取其一,设为 $w(i, j)$,则初始化滚轮位置为 $w(i, j - 1)$;
- 3) 如果为顺时针滚动,滚轮沿 x 轴正向滚出第一步,则当前滚动向量为 $(0, 1)$ 。

设空集 A ,轮廓追踪的实现过程描述如下。

- 1) 按滚动规则确定当前滚轮的自由度,根据自由度及目前滚动向量确定预滚动向量;
- 2) 根据预滚动向量按照搜索填充规则确定搜索向量,并进行单元格的搜索填充;
- 3) 如果填充数目大于 0,则由步骤 2) 循环,否则进行如下操作;
- 4) 将滚轮当前位置四近邻中满足 $w(i, j) = 1$ 的点加入 A ,并按照滚动规则正式滚动;
- 5) 刷新当前滚动向量;
- 6) 如果到达起始点,将 A 映射回像素空间得到该聚类区域边缘,否则由步骤 1) 循环。

3 实验结果及结论

与传统的时间差分模型相比,采用块区域二阶

矩比的形式,能够更加准确、完整地检测出运动物体的边缘,对光照变化、阴影及背景物体小幅度扰动等的抗干扰能力增强,有较强的自适应性,并有效消除了传统时间差分算法中的尾迹现象,能够实现实时检测。鉴于物体特征抽取等后续处理的需求,有必要获得相对完整的运动物体轮廓或者区域。基于只提取活动物体边缘的思想,本文创建了 RW 算法,实现了快速、完整的运动物体的轮廓追踪。与传统的基于形态学的方法相比,该算法能够充分利用所检测到的边缘区域信息,同时实现了边缘运动区域的连通与整体轮廓追踪,而不必考虑检测区域内部的细节,比较符合人眼对轮廓的搜索习惯。该算法同样适用于像素级处理。

图6示例了二阶矩比模型与基于核密度估计的背景减除模型的联合检测结果以及 RW 算法的轮廓追踪效果,其中子窗口大小为 5 像素 \times 5 像素。可以看出,在图像对比度较小的情况下,二阶矩比模型仍然能够相对完整地检测到运动物体的边缘区域;RW 算法充分利用了所检测到的运动区域(窗口),完整地描述了物体的轮廓。甚至对于分布比较离散的边缘区域,RW 算法仍能实现合理的连接。根据实验结果,大多数情况下, T_1 的取值一般为 1.5~3;虽然随着子窗口的增大,模型的抗干扰能力增强,但同时会丢失活动边缘的一些细节信息,给随后的轮廓追踪工作带来困难,窗口大小一般为 5 像素 \times 5 像素 \sim 15 像素 \times 15 像素。

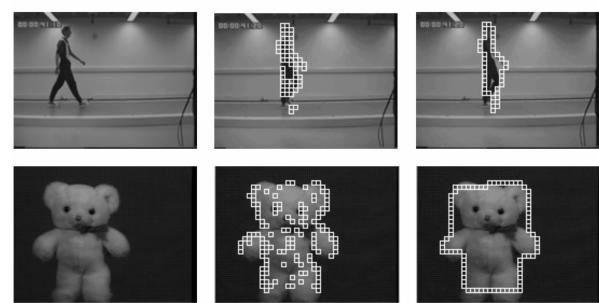


图6 动检测模型及 RW 算法的结果示例

Fig. 6 Experimental Results of Our Motion Detection Model and the RW Algorithm

参 考 文 献

- 1 Elgammal A, Duraiswami R, Harwood D, et al. Background and Foreground Modeling Using Non-parametric Kernel Density Estimation for Visual Surveillance. Proceedings of the IEEE, 2002
- 2 Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-time Tracking. IEEE Conference

- on Computer Vision and Pattern Recognition, Fort Collins, Colorado, 1999
- 3 Haritaoglu I, Harwood D, Davis L S. W4: Real-time Surveillance of People and Their Activities. *IEEE Trans. on Pattern Anal. Mach. Intell.*, 2000, 22 (8): 809~830
- 4 Takashi M, Takashi O, Hitoshi H. Background Subtraction for Non-Stationary Scenes. The 4th Asian Conference on Computer Vision, Taipei, 2000
- 5 Ma Y, Zhang H J. Detecting Motion Object by Spatio-temporal Entropy. *IEEE International Conference on Multimedia and Expo. (ICME 2001)*, Tokyo, 2001
- 6 Lipton A, Fujiyoshi H, Patil R. Moving Target Classification and Tracking from Real-time Video. *IEEE Workshop on Applications of Computer Vision*, Princeton, 1998
- 7 Collins R, Lipton A, Fujiyoshi H, et al. A System for Video Surveillance and Monitoring. American Nuclear Society (ANS) Eighth International Topical Meeting on Robotics and Remote Systems, Pittsburgh, PA, Nisan, 1999
- 8 Barron J, Fleet D, Beauchemin S. Performance of Optical Flow Techniques. *International Journal of Computer Vision*, 1994, 12 (1): 42~77
- 9 Stricker M, Orengo M. Similarity of Color Images.
- SPIE Storage and Retrieval for Image and Video Databases III, San Jose, California, 1995
- 10 Yang J H, Estivill-Castro V, Chalup S K. Support Vector Clustering Through Proximity Graph Modeling. The 9th International Conference on Neural Information Processing (ICONIP02), Singapore, 2002
- 11 Pavlidis T. Algorithms for Graphics and Image Processing. Rockville, Maryland: Computer Science Press, 1982
- 12 Wren C, Azabayejani A, Darrell T, et al. Pfinder: Real-time Tracking of The Human Body. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997, 19:780~785
- 13 Kass M, Witkin A, Terzopoulos D. Snakes: Active Contour Models. International Conference on Computer Vision, London, 1987
- 14 Tsechpenakis G, Rapantzikos K, Tsapatsoulis N, et al. A Snake Model for Object Tracking in Natural Sequences. *Signal Processing: Image Communication*, 2004, 19(3):219~238

第一作者简介:吕常魁,博士生,研究方向为计算机集成制造、计算机视觉及工业应用。

E-mail: mailck@yahoo.com.cn

A Novel Statistical Model for Motion Detection and Contour Tracing

LÜ Changkui¹ JIANG Chengyu¹ WANG Ningsheng¹

(1) CIMS Research Center, Nanjing University of Aeronautics and Astronautics, 29 Yudao Street, Nanjing 210016, China)

Abstract: In the first section, according to the assumption that only the silhouette could be extracted to describe the motion of a moving object, a three-frame differencing model is constructed based on the variance ratio of successive blocks of different frames, simultaneously the Gaussian kernel based density estimation model in [1] is used as a filter to suppress false positives, and the SVC algorithm in [10] is also employed to realize real-time multi-object motion detection. A novel contour tracing algorithm, namely the RW (roller wheel) method, is proposed in the next section. Edge information detected by the forgoing model can be fully used by this algorithm, while the interior details will be skipped. This algorithm works in a similar way that human vision system works. Experimental results are given in the end.

Key words: motion detection; contour tracing; temporal differencing; color moments; roller wheel algorithm