

基于光流技术的运动目标检测和跟踪方法研究

戴 斌, 方宇强, 孙振平, 王 亮

国防科学技术大学机电工程与自动化学院, 长沙 410073

摘要 光流技术作为一种重要的二维运动估计技术, 在运动目标检测和跟踪中有着重要的作用。为了更好地将光流技术运用到实时的运动目标检测和跟踪系统中, 针对微分光流信息丰富但计算量大、特征光流计算量小但信息量不足的现状, 提出了一种基于最优估计的点匹配技术和光流均匀采样策略的光流场计算方法, 并通过对灰度化后的光流场进行自适应阈值分割、形态学滤波等处理, 实现了实时的运动目标检测和跟踪。通过对图像序列取 700 个样本点的仿真实验表明, 该方法帧间处理时间基本小于 100 ms, 同时基本解决了信息量和计算量的矛盾。

关键词 微分光流; 特征光流; 运动目标检测; 运动目标跟踪

中图分类号 TP391.4

文献标识码 A

文章编号 1000-7857(2009)12-0055-06

The Detection and Tracking of Moving Objects by Using Optical Flow Technique

DAI Bin, FANG Yuqiang, SUN Zhenping,
WANG Liang

*College of Electromechanical Engineering and Automation,
National University of Defense Technology, Changsha 410073,
China*

Abstract The Optical Flow Technique (OFT), as an important technique for estimating a two dimensional movement, is widely used in Moving Object Detecting and Tracking (MODT). However, the OFT method usually suffers from a heavy calculation load and inadequacy in information. In this paper, we propose an OFT method which is based on the technology of point matching and the strategy of uniform sampling and can be used in MODT in a real time through graying optical flow field. The simulation experiment with samples of 700 points per image indicates that the frame processing time is less than 100 ms, which means an adequate amount of information handled with reasonable computation cost.

Keywords derivative optical flow; feature optical flow; moving object detecting; moving object tracking

0 引言

运动目标检测和跟踪是动态图像分析的重要组成部分, 具有广泛的应用价值和发展前景。从作用和地位来看, 它是图像理解 and 行为分析的前提, 也是计算机视觉要解决的基础性工作; 从应用和影响来看, 它已经深入到了人们生活的各个部分, 正在各行各业发挥着不可替代的作用。从学科的发展来看, 它涉及图像处理、模式识别、人工智能和自动控制等多个领域, 有利于促进多学科共同发展。目前它已在军事、工业、交通、医学、教育、商业、体育等领域中有着广泛的应用。

光流技术是一种重要的二维运动估计技术, 它不仅可以进行运动目标检测, 也可以进行运动目标跟踪。这是由光流技术的内在特性所决定的, 也是光流技术的优势所在。光流是图像灰度模式的运动速度, 只要图像中有运动存在, 就会有光流产生。运动一致的区域, 光流表现出一致性; 运动不一致的区域, 光流也就表现出不一致。通过对光流场进行分析, 可以实现运动目标的检测和跟踪。例如, Liu^[1]和 Zinbi^[2]等都利用光流场及活动轮廓模型来进行运动目标的检测, Toshio 等^[3]也结合了光流场和图像的边缘信息进行目标检测, Shin 等^[4]利用基于特征的光流技术进行跟踪。与此同时, 国外已经有了很多基于光流技术的目标检测和跟踪的系统, 例如英国牛津大学的 ASSET-2 系统^[5]是个典型的光流分析系统, 利用特征光流技术实现运动目标的实时检测和跟踪。

可以看出, 以往的基于光流的目标检测方法主要分为两类: ① 利用微分光流技术如文献[1-3]所述, 即利用光流的基本方程, 附加一定的约束, 得到致密的光流场, 再提取运动目标, 此方法的不足在于计算量较大, 实时性不强。② 像文献

收稿日期: 2008-12-09

作者简介: 戴斌, 副教授, 研究方向为模式识别与智能系统, 电子信箱: 13807318768@139.com

[4-5]所述,利用特征光流技术,在图像中寻找特征点进行匹配,得到稀疏光流场,然后提取目标,这种方法的实时性得到了提高,但信息量不足,容易造成目标的漏检。而在目标跟踪方面,以往的做法常常将其分离开来,在实现检测后,再基于目标的特征进行跟踪。这样做便增加了算法处理的复杂度,在目标的进入和退出时带来复杂的处理过程。

针对这种情况,本研究提出了一种基于光流均匀采样策略的运动目标检测和跟踪方法。该方法通过采样处理,在点匹配的基础上实现了目标的检测,一定程度上解决了信息量和计算量之间的矛盾,取得了较好的效果。同时可以将检测和跟踪融为一体,降低算法的复杂程度。本文首先介绍了基于最优估计的点匹配技术和均匀采样策略的光流场计算,然后说明了如何在光流场实时地提取运动目标,最后对整个运动目标检测和跟踪的算法过程加以总结,并给出实验结果及分析。

1 光流场的计算

要利用光流来进行运动目标的检测和跟踪,首先要做的就是光流场的计算,这是实现目标检测和跟踪的基础。微分光流技术如 Horn-Schunck^[6]方法、Lucas-Kanade^[7]方法等,虽然信息量大,但同时由于其计算量较大,无法满足实时性的要求,所以本文采用基于点匹配技术和均匀采样的策略来实现光流场的计算。

1.1 基于最优估计的点匹配技术

本文提出了一种改进的点匹配方法。该方法在最优估计的引导下进行局部区域搜索,在保持精度的同时大大提高了匹配速度。基于最优估计的点匹配是通过将2个点特征之间的某种相似性度量进行最优估计,进而得到最佳匹配位置的一种方法,该方法无需进行搜索,因而时间是可控的。

平移模型是最简单的图像变换模型。在序列图像中,由于相邻两幅图像的时间间隔很小,图像之间的变化可用平移模型来近似。在该模型下,对应点之间有简单的几何关系。

设 $A(x, y)$ 、 $B(x, y)$ 代表前后两幅图像,对于第一幅图像上的点 $x=(x, y)^T$,匹配的目的就是寻求第二幅图像上的对应点: $(x', y')^T=(x+v_x, y+v_y)^T$,使得在点 $x=(x, y)^T$ 的 $w_x \times w_y$ 邻域内,下式所表示的灰度差平方和为最小

$$\varepsilon(v)=\varepsilon(v_x, v_y)=\sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} (A(x, y)-B(x+v_x, y+v_y))^2 \quad (1)$$

$\varepsilon(v)$ 最小值对应的就是最佳匹配位置 $v_{opt}=(v_x, v_y)^T$,采用优化法求解该问题,令 $\varepsilon(v)$ 对 v 的一阶导数为零: $\frac{\partial \varepsilon(v)}{\partial v} \Big|_{v=v_{opt}}=(0,0)$ 。

将这个式子展开以后,得到

$$\frac{\partial \varepsilon(v)}{\partial v}=-2 \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} (A(x, y)-B(x+v_x, y+v_y)) \cdot \left[\frac{\partial B}{\partial x} \quad \frac{\partial B}{\partial y} \right]$$

用 $B(x+v_x, y+v_y)$ 在点 $v=(0, 0)^T$ 的一阶泰勒展开式来代替,得

$$\frac{\partial \varepsilon(v)}{\partial v}=-2 \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} (A(x, y)-B(x, y)-\left[\frac{\partial B}{\partial x} \quad \frac{\partial B}{\partial y} \right] \cdot v) \cdot \left[\frac{\partial B}{\partial x} \quad \frac{\partial B}{\partial y} \right] \quad (2)$$

注意到 $A(x, y)-B(x, y)$ 可以看作是 $(x, y)^T$ 的一个时间方向上的导数,所以

$$I_t(x, y) \approx A(x, y)-B(x, y)$$

$$\text{令 } \nabla I = \begin{pmatrix} I_x(x, y) \\ I_y(x, y) \end{pmatrix} = \begin{pmatrix} \frac{\partial B}{\partial x} \\ \frac{\partial B}{\partial y} \end{pmatrix}^T$$

I_x 和 I_y 分别为 x 和 y 方向上的图像梯度,可根据任一幅图像中该点的 $w_x \times w_y$ 邻域信息求得。如果使用差分近似微分算子,则有

$$I_x(x, y) = \frac{\partial A(x, y)}{\partial x} = \frac{A(x+1, y)-A(x-1, y)}{2}$$

$$I_y(x, y) = \frac{\partial A(x, y)}{\partial y} = \frac{A(x, y+1)-A(x, y-1)}{2}$$

此时式(2)化简为

$$\frac{1}{2} \frac{\partial \varepsilon(v)}{\partial v} = \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} (\nabla I^T v - I_t) \cdot \nabla I^T$$

两边取转置后化简得

$$\frac{1}{2} \left[\frac{\partial \varepsilon(v)}{\partial v} \right]^T = \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} v - \begin{bmatrix} I_t \cdot I_x \\ I_t \cdot I_y \end{bmatrix}$$

令 $G = \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$ 和 $b = \sum_{-w_x}^{w_x} \sum_{-w_y}^{w_y} \begin{bmatrix} I_t \cdot I_x \\ I_t \cdot I_y \end{bmatrix}$, 则式(2)变成

$$\frac{1}{2} \left[\frac{\partial \varepsilon(v)}{\partial v} \right]^T = Gv - b$$

其中, G 为对称半正定矩阵,其特征值 $\lambda_1, \lambda_2 \geq 0$,二者可以有效地描述区域的纹理特性,当 λ_1, λ_2 都很小时,说明区域有大致恒定的灰度分布,而 λ_1, λ_2 大小差异大说明窗口有不定向的纹理模式,如果 λ_1, λ_2 都很大,则代表角点或椒盐纹理等,所以 G 可以看作是一个描述图像局部特征的结构张量。令上式为零,当 G 可逆时,则

$$v_{opt} = G^{-1}b \quad (3)$$

此时得到的 $v_{opt}=(v_x, v_y)^T$ 即为满足灰度差平方和最小的平移向量,对应第二幅图像上的匹配点就是 $(x', y')^T=(x+v_x, y+v_y)^T$ 。当 G 不可逆时,说明图像中的局部信息无法给出好的匹配结果,所以可以设 $v_{opt}=0$ 。这样,可以通过这种最优估计的点匹配技术寻求两幅图像中的匹配点,得到相应的光流值。这样做的好处在于匹配点是可选的,不必像微分光流技术通过求解约束方程计算图像中每一个点的光流值,所以计算时间是可控的。

1.2 光流的均匀采样策略

Smith^[5]首先在图像中寻求一些特征点,如角点、SIFT点^[8]等,再利用点匹配的技术来计算稀疏光流场,根据流场中不同光流来检测和跟踪运动目标。这种方法,虽然说由于特征点的数量较少,减少了计算量,同时由于特征点所具有的一些尺度、旋转不变的良好特性,具有较高的匹配精度。然而,因为图像中这些特征点的数量并不多,会造成信息量不足,特别是在弱纹理的区域中,无法像微分光流那样得到致密的

本文的光流灰度化处理针对光流的模值进行,统计光流场中光流模值的最大值和最小值,依次指定为 255 和 0,其余光流按其相对于最大和最小光流模值的位置确定其应当取的值,经过灰度化后的光流数据,就成了类似灰度图像的整型数据。

如图 3(a)所示,在对光流场进行灰度化处理之后,运动目标的灰度值往往较小,背景区域的灰度值较大。本文采用 Otsu 自适应阈值分割技术^[12],假设图像灰度值 i 的像素数为 n_i ,像素总数为 N ,某一阈值 T 将灰度值划分为 $0 \sim T-1$ 和 $T \sim 255$ 两组,两组的产生概率 w_i 和平均值 μ_i 分别为

$$w_0 = \sum_{i=0}^{T-1} \frac{n_i}{N}, \quad \mu_0 = \sum_{i=0}^{T-1} \frac{in_i}{Nw_0}$$

$$w_1 = 1 - w_0, \quad \mu_1 = 1 - \mu_0$$

总体的灰度平均值为

$$\mu = w_0 \mu_0 + w_1 \mu_1$$

两组间的方差为

$$\delta^2(T) = w_0(\mu_0 - \mu)^2 + w_1(\mu_1 - \mu)^2$$

从不同的灰度级中寻求 $\max \delta^2(T)$ 的值 T^* ,即为所求的最优阈值。

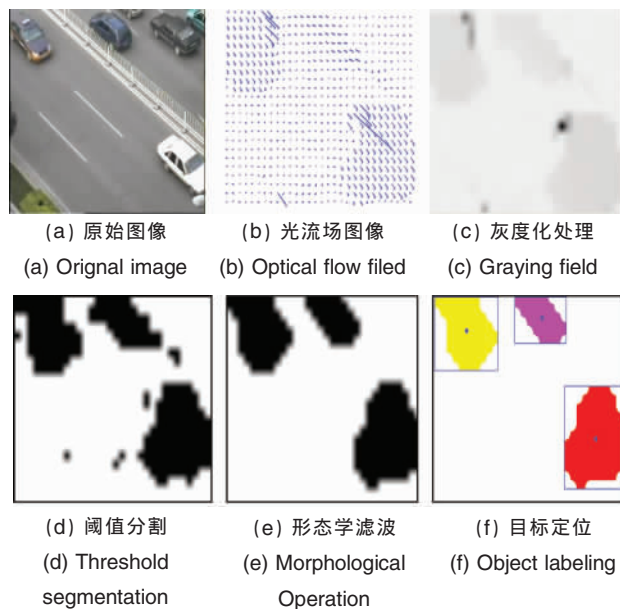


图 3 运动目标提取

Fig. 3 Moving Object Extraction

分割后的图像为二值图像,其中往往由于光流计算上的误差存在部分杂点,这时采用了形态学的开运算来完成滤波处理,如图 3(e)所示。最后在处理后的二值图像上,基于区域的连通性进行分类,采用 8 邻域的区域生长搜索策略寻求不同的目标区域。如图 3(f)所示,3 种颜色分别代表算法搜索到的具有 8 方向连通性质的 3 个不同的运动目标区域。

最后,统计各自区域的像素信息,计算出目标的质心,质心是实现目标跟踪的重要信息。假设目标区域的最小外界矩形大小为 $M \times N$,由于处理图像为二值图像,所以质心计算公

式为

$$(\bar{x}, \bar{y}) = \left(\frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x_{i,j}}{(M-1)(N-1)}, \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} y_{i,j}}{(M-1)(N-1)} \right)$$

其中 $x_{i,j}, y_{i,j}$ 分别代表目标区域像素的行列坐标。如图 3(f)所示,可以根据得到的最小外界矩形求出不同目标的区域的质心位置,定位不同的目标。

然而,这种方法的不足之处在于,分类时仅仅考虑了运动目标的空间位置关系,当目标接近时,容易造成目标的混叠。针对这种问题,本文利用运动区域中光流的大小和方向信息再分割,其中在光流场中用不同的颜色代表方向相近的光流。

对于图 4(a)中的同向车辆而言,其光流方向相近,但对于不同速度的邻近目标而言,光流值大小差异较大,可以根据运动区域内光流大小进行分割得到不同的运动目标,如图 4(c)。而对于图 4(d)中的相向而行的目标,可以看出其光流方向是不同的,如图 4(e)。那么根据区域内光流的方向就可以分割出两个邻近目标,如图 4(f)。如果目标始终距离接近,且运动速度相同,可以认为是一个运动整体来检测和跟踪。

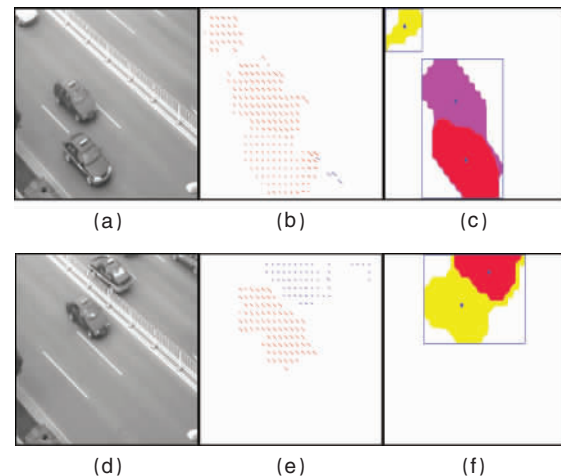


图 4 目标邻近情况处理

Fig. 4 Process of segmenting adjacent objects

最后,对本文的算法加以总结。算法的总体流程如图 5 所示,其中包括了光流场的计算和目标提取以及跟踪两个主要组成部分。而本文方法在处理运动目标检测和跟踪上有 3 点好处:① 光流的计算效率高,这是因样本的数目和计算方法确定的;② 能够成功检测到运动目标的进入和退出,场景的全局运动使得光流场产生整体流动,而目标的局部运动则使流场局部区域产生特殊的流动状态;③ 由于点匹配的技术本身就是目标跟踪的一种基本技术,以往的算法通常先进行目标的检测,然后提取目标区域,根据区域的特征利用匹配的方法进行跟踪,而本文采用均匀采样的点匹配技术计算光流场,本身就是一种对样本点的跟踪,使得运动目标的检测和跟踪合二为一,在序列图像中进行目标检测的同时,也实现了目标的跟踪。

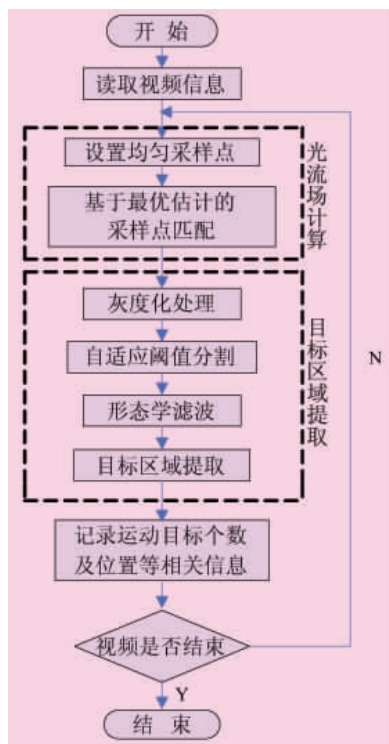


图 5 基于本文方法的运动目标检测和跟踪流程图

Fig. 5 Flowchart of MODT based on proposed method

3 实验及结果分析

本文实验数据来源于一段描述交通路口车辆行驶情况的视频,视频中每帧图像的大小为 320×240 ,实验利用本文描述的基于光流的技术方法对运动的车辆进行实时检测和跟踪,检验本文算法的可靠性。实验程序基于 Visual C++6.0 开发,同时利用了 OpenCV 1.0。图 6 为典型帧的效果示意图,分别为序列中的第 274、236、322、942 帧。每一帧依次列出了提取的目标区域轮廓与原图的叠加图、光流场图和灰度图进行自适应阈值分割等处理后的目标提取图。

实验图像序列中单幅图像大小为 320×240 ,实验中样本间隔设为 10 个像素,即最大样本点数为 32×24 。

图 6 中列出的各帧,帧间处理时间依次为:47、79、61、93 ms。可以看出,算法对运动目标的检测和跟踪有较好的鲁棒性,同时这个时间消耗还是可以接受的。如果采用硬件加速设备,加之合理的采样间隔,能够做到更好的运动目标实时检测和跟踪。

4 结论

运动目标检测和跟踪涉及计算机视觉、模式识别及图像处理等学科,是当今研究的重点和难点,具有广泛的应用前景。考虑到环境的复杂变化和摄像机的运动情况,光流法是一种比较有效的运动目标检测和跟踪方法。本文提出了一种基于最优估计的点匹配技术和光流均匀采样策略的运动目



图 6 运动目标检测和跟踪效果图

Fig. 6 Results of MODT based on proposed method

标检测和跟踪方法。该方法可以同时进行运动目标的检测和跟踪,并能在一定程度上解决光流技术中的计算量和信息量的矛盾。但同时也应看到,该方法的计算量依然很大,尤其在较大的图像、采样较密的情况下更是如此。因此,解决高精度光流计算的时间消耗问题,或者寻求一种更高效的光流算法,或是采用专门的硬件加速设备,均是下一步应涉及的研究工作。总之,计算效率问题是光流继续发展和得以实用的关键。

参考文献 (References)

- [1] Liu P R, Meng M Q H, Liu P X, *et al.* Optical flow and active contour for moving object segmentation and detection in monocular robot [C]// Proceedings 2006 IEEE International Conference on Robotics and Automation. 2006: 4075-4080.
- [2] Zinbi Y, Chahir Y, Elmoataz A. Moving object segmentation using optical flow with active contour model [C]//ICTTA 2008 3rd International Conference on Information and Communication Technologies: From

- Theory to Applications. 2008: 1-5.
- [3] Naito T, Ito T, Kaneda Y. The obstacle detection method using optical flow estimation at the edge image [C]//Proceedings 2007 IEEE Intelligent Vehicles Symposium. 2007: 817-822.
- [4] Shin J H, Kim S J, Kang S K, *et al.* Optical flow-based real-time object tracking using non-prior training active feature model [J]. *Real-Time Imaging*, 2005, 11: 204-218.
- [5] Smith S M. ASSET-2: Real-time motion segmentation and shape tracking [C]//Proceedings 1995 Fifth International Conference on Computer Vision. 1995: 237-244.
- [6] Horn B K P, Rhunck B G. Determining optical flow [J]. *Artificial Intelligence*, 1981, 17(1-3): 185-203.
- [7] Bouguet J Y. Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm[EB/OL]. [2008-12-12]. http://robots.stanford.edu/cs223b04/algo_tracking.pdf.
- [8] Lowe D G. Distinctive image features from scale invariant key points[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [9] Shi J, Tomasi C. Good features to track, TR93-1399 [R]. New York: Cornell University, 1993.
- [10] Bruhn A, Weickert J. Lucas-Kanade meets Horn-Schunck: Combining local and global optical flow methods [J]. *International Journal of Computer Vision*, 2005, 61(3): 211-231.
- [11] Fernando W S P, Udawatta L, Pathirana P. Identification of moving obstacles with Pyramidal Lucas Kanade optical flow and K means clustering [C]//ICIAFS 2007 Third International Conference on Information and Automation for Sustainability, December 4-6, 2007: 111-117.
- [12] Otsu N. Discriminant and least square threshold selection [C]//Proc 4IJCP, 1978: 592-596.

(责任编辑 赵佳)

第6届数字博物馆与文化自然遗产数字化及保护研讨会 暨第一届全国数字博物馆与数字科技馆高级论坛

时间: 2009年7月27-29日

地点: 山东·青岛

主办: 中国图像图形学会虚拟现实专业委员会
青岛文物局
浙江大学传媒与文化产业研究中心
山东省计算机学会

会议网址: <http://iec.qdu.edu.cn/conference/index.html>电子信箱: weiming1202@163.comqiuxiao@haivence.com

电 话: 0532-85953992, 88727167