

一种基于均值偏移的自动运动分割方法

蒋鹏 秦娜 周艳 唐鹏 金炜东

(西南交通大学电气工程学院 成都 610031)

摘要 运动分割是计算机视觉领域研究的重要内容。提出一种基于均值偏移的自动运动分割算法。该方法首先用特征点匹配关系获得特征点的运动轨迹,并以轨迹的运动向量作为特征,再用均值偏移算法对轨迹的运动向量进行聚类。均值偏移缩小相似的运动向量之间的差别,同时扩大不同运动的运动向量之间的差距。为了自动获得运动分类数,还提出了一种基于非参数核密度的自动分类方法,该方法通过估计运动向量的密度分布,用核密度图自动确定运动分类数。实验结果表明,该算法分割精度高、鲁棒性好,能够自动确定运动分类数。

关键词 均值偏移,运动分割,核密度

中图法分类号 TP391 **文献标识码** A

Automatic Motion Segmentation of Sparse Feature Points with Mean Shift

JIANG Peng QIN Na ZHOU Yan TANG Peng JIN Wei-dong

(College of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

Abstract We proposed an automatic motion segmentation operating on sparse feature points. Feature points are detected and tracked throughout an image sequence, and feature points are grouped using a mean shift algorithm. The motion segmentation is driven by the density of the motion vector in feature space. The kernel density estimation is performed on the mean-shifted motion vector and the number of motion present is estimated by the number of peaks in the kernel density curve. Experimental results on a number of challenging image sequences demonstrate the effectiveness and robustness of the technique.

Keywords Mean shift, Motion segmentation, Kernel density estimation

1 引言

运动分割是计算机视觉领域一个重要的研究内容。有效的运动分割往往是视频对象分割和视频目标跟踪等应用中一个重要的步骤,因此引起了国内外大量研究者的关注^[1-3]。

运动分割实质是利用不同类型的运动在边界上的不连续性将不同种运动区分开来。其中利用运动向量场对运动进行分割是目前一种常见的做法。该方法利用运动在边界上是不连续的特点,构成运动边界。基于运动向量的运动分割是以待分割的区域的运动特征的差异作为特征,因而这种分割方法与颜色纹理无关,只要像素或者特征点的运动向量相似,就可以被划分成同一个区域。当同一个运动物体是由多种颜色区域组成时,这种方法依然可以较好地运动分割。

如何有效提取运动向量以及如何减少运动向量包含的噪声是目前运动分割领域主要的研究内容^[4-7]。提取视频序列的运动向量实际上就是查找对应物体在连续视频帧中相对位置的变化。块匹配、光流场和特征点匹配等是目前比较常见的运动向量提取方法。由于块匹配和光流场在低纹理区域产生的噪声较多,如何有效地减少这些噪声是目前研究的重点。特征点匹配是通过跟踪角点等形状特征获得特征点在多帧之

间的匹配关系。这种方法可以利用运动的连续性,获得特征点在多帧的运动轨迹,因此相对块匹配和光流场等方法,特征点匹配方法稳定性较好^[4]。

由于特征匹配错误可能产生异常运动向量,如何减少异常运动向量引入的噪声是提高分割精度的关键。例如文献[5]利用参数模型对运动进行建模,该方法假设块匹配算法的误差是服从高斯分布的,因此可以通过最小二乘估计将异常数据剔除,并估算出仿射运动模型的6个参数。文献[6]提出利用小波分解等方法减少异常运动向量对特征点轨迹聚类的影响。而Faisal等人^[7]则利用主成分分析PCA来获得主要运动向量,将与主要运动向量不一致的运动向量作为异常运动向量剔除。

在获得运动向量场后,对运动向量场进行分割是一个重要的研究内容。过去的方法常常假设运动物体的数目是给定的。但在现实中,运动物体是根据场景的变化而动态变化的。因此如何动态地确定运动物体数目是目前一个亟待解决的问题。其中,文献[8]直接利用K均值聚类对运动向量进行分割,但该方法的聚类数是事先给定的,无法动态调整。文献[9]提出了一种自适应确定运动物体数目的方法,但该方法利用光流场,计算量很大。

到稿日期:2012-10-22 返修日期:2013-02-24 本文受中央高校基本科研业务费专项资金资助项目(SWJTU12CX027),国家自然科学基金(60971103,61134002)资助。

蒋鹏 男,博士,讲师,主要研究方向为计算机视觉、多媒体传感器网络。

为了解决分割中的运动向量噪声问题并实现自动分割,本文提出一种基于均值偏移的自动运动分割方法。该算法首先利用多帧的特征点匹配关系获得特征点的运动向量,再用均值偏移算法对运动向量进行去噪,最后采用一种基于核密度估计的自动运动分类方法获得最终的运动分类。

本文方法的特点如下:

1) 本文采用多帧特征点匹配关系形成的运动向量。这种运动向量本身具有一定的统计性,相对两帧之间的运动向量而言,这种运动向量可以较好地解决运动暂停等问题。

2) 为了扩大不同运动之间的差距、缩小相同运动之间的差别,用基于均值偏移算法对运动向量进行处理,可以使得与大多数运动向量不同的异常数据向主要运动进行偏移。这样可以有效地减少异常数据的干扰。

3) 提出的非参数核密度估计方法,可以自动确定运动类别数,无需用户干预,就可以根据场景自动完成运动分割。

2 基于均值偏移的运动分割

2.1 运动特征提取

根据运动连续性的特点,本文采用多帧特征点匹配的方法获得运动特征。首先采用 Harris 特征点检测算法^[10]来检测视频图像中的特征点,并用金字塔光流法^[11]跟踪特征点下一帧中的位置。设一帧图像中有 N 个特征点,则在时间 t 时第 i 个特征点的描述符为: $p_i^t(x_i^t, y_i^t)$, $i \in N$, 其中 x_i^t, y_i^t 为特征点的位置, $x_i^{t'}, y_i^{t'}$ 为该点在下一帧的位置, dx_i, dy_i 为特征点的运动向量, 其中 $dx_i = x_i - x_i', dy_i = y_i - y_i'$ 。当运动向量的幅值过大时,则认为是由特征点匹配出错造成的,将其去除。继而利用文献^[12]提出的基于路径连通(path coherence)的概念的迭代特征点轨迹判决算法获得每个特征点的 M 帧的轨迹。本文用 $x_i = \{p_i^1, p_i^2, \dots, p_i^M\}$ 来表示第 i 个特征点的轨迹。

本文将轨迹的运动向量作为轨迹特征。首先定义一个轨迹 x_i 所包含的运动向量 $v_i(\Delta x, \Delta y)$ 为:

$$\begin{aligned} \Delta x &= \sum_{k=1}^M |p_i^k(x) - p_i^{k-1}(y)| \\ \Delta y &= \sum_{k=1}^M |p_i^k(y) - p_i^{k-1}(y)| \end{aligned} \quad (1)$$

对于运动向量而言,同一个物体的运动轨迹虽然趋近相似,但也有微小差别。特别是在非刚体运动时,非刚体运动造成的局部位移使得同一个运动物体的轨迹具有相似的趋势但是又有细微差别。这些运动轨迹的细微差别对自动获得运动物体数目提出了挑战。为了解决该问题,本文提出利用均值偏移方法将同一个物体的运动轨迹的细微差别向中心值进行修正,以提高自动运动物体个数判定的准确度。

2.2 均值偏移处理

均值偏移算法的核心是通过核函数对数据点的加权,使得数据向密度较大的点靠拢。均值偏移算法通过核函数加权的方法实现局部最优,因此可以有效地缩小同一个运动物体轨迹的差别^[13]。由于同一个物体的运动向量可能存在一定的差异性,因此通过均值偏移,可以减少这种差异性,将相似的运动向量归为一类。针对运动向量而言,由于物体并非以完全刚体的形式运动,因此会造成部分的运动向量与主要的

运动向量不一致的现象。如果采用传统聚类技术,这部分将与主要运动向量不一致的少量运动向量可能会被分割成一个新的物体。这显然是不合理的,因为这部分少量运动向量只是与主要运动向量之间存在少许差别,通过均值偏移方法,可以将这小部分运动向量向主要运动向量靠近,从而加强了主要运动向量,便于判断运动物体数目。

设有 n 个数据点,每个数据点的运动向量为: $v_i(\Delta x, \Delta y)$, $1 \leq i \leq n$ 。

首先采用核密度方法估计每个运动向量的密度:

$$f(v) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{|v - v_i|^2}{h^2}\right) \quad (2)$$

其中 $K(\cdot)$ 为核函数,本文采用高斯核函数,则式(2)改写为:

$$f(v) = \frac{1}{n} \sum_{i=1}^n \exp\left(-\frac{|v - v_i|^2}{h^2}\right) \quad (3)$$

根据均值偏移算法,将均值偏移向量定义为:

$$m(v) = \frac{\sum_{i=1}^n v_i \cdot \exp\left(-\frac{|v - v_i|^2}{h^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{|v - v_i|^2}{h^2}\right)} - v \quad (4)$$

均值偏移向量实际上是用核函数求梯度来获得密度最大的运动向量。均值偏移是一种迭代算法,当 $m(v)$ 的值小于一个阈值时(本文设为 0.01),算法退出。这样可以根据 n 个运动向量,估算出 n 个经过均值偏移处理过的运动向量为: $u_i(\Delta x, \Delta y)$, $1 \leq i \leq n$ 。在实现过程中,为了避免每次都重新计算核密度函数,本文采用查表的方式提高计算速度。

图 1 为均值偏移对运动向量聚类前后的示意图。图 1 中左图为均值偏移处理前的运动向量,而右图为处理后的运动向量。从图 1 可以看出,在经过均值偏移处理以后,具有相似运动的向量向其中心值进行了偏移,缩小了相同类型运动向量的差别,同时扩大了不同类型数据之间的差距。由于加大了不同类型运动的差别,因此更有利于进行运动分割。同时,由于噪声等因素产生的少数与主要运动不相符的运动向量,通过均值偏移处理后,可以起到有效的降噪作用,这也为更精确的运动分割提供了基础。

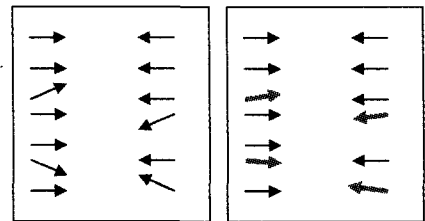


图 1 运动向量的均值偏移处理示意图

3 基于核密度估计的自动运动分割

在获得均值偏移处理过的运动向量 $u_i(\Delta x, \Delta y)$ 后,需要根据 $u_i(\Delta x, \Delta y)$ 获得运动分割结果。由于运动的分类数是未知的,因此传统的聚类方法 K 均值或者模糊 C 均值无法直接使用。为此本文提出一种基于非参数核密度估计的方法来自动获取运动类型分类数和进行聚类。

核密度估计的本质是通过加权平均中心点位于采样值的局部函数来估计未知的密度分布。本文通过核密度估计获得 n 个运动向量 $\{u_1, u_2, \dots, u_n\}$ 的分布。对于 n 个运动向量的观

测值 $\{u_1, u_2, \dots, u_n\}$, 检测值 u_i 的核密度估计用 $p(u_i)$ 来估计。本文依然采用高斯核函数, 则核密度 $p(u_i)$ 表示为:

$$p(u_i) = \frac{1}{n} \sum_{i=1}^n \exp\left(-\frac{|u_i - u_i|^2}{h^2}\right) \quad (5)$$

由于核密度估计是利用核函数加权平均中心值, 因此可以获得连续和平滑的密度分布。这种核密度估计方法相当于对概率直方图进行了平滑处理, 可以有效地避免噪声数据的干扰。

图2为以 coastguard 为例的运动向量幅值的核密度估计图, 其中 X 轴为运动向量的幅度, Y 轴为核密度。从图2可以看出, 在主要运动(principle motion)处, 经过高斯函数的加权, 形成多个波峰, 其中每种运动的运动向量在核密度图中表现为一个波峰。因此通过分析运动向量的核密度图就可以获得波峰数, 从而进一步获得运动分类数。

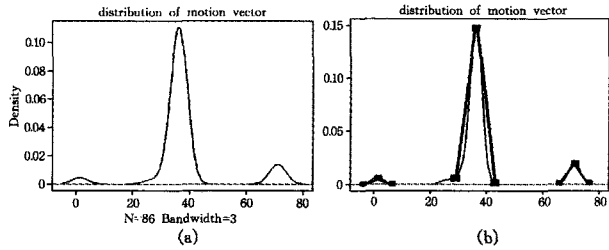


图2 运动向量的核密度分布

虽然通过均值偏移扩大了不同运动的差别, 但相似的运动表现在核密度图中可能会形成多个局部波峰。为了避免将局部波峰作为一个聚类, 本文先对核密度图进行平滑处理, 再计算波峰数。如果将核密度图看成是一个直方图 H , 则第 i 个 bin 处的密度值为 $H(i)$ 。可以通过直方图峰值分析方法对核密度进行处理。取相邻的 3 个 bin 的均值作为当前 bin 的值进行平滑处理获得 $H_S(i)$ 。通过平滑处理可以消除一部分局部波峰。由于噪声等干扰, 也会产生少量的在方向和幅值都与大多数运动向量不同的噪声运动向量, 这些噪声运动向量在 KDE 图中往往表现成密度值很小的波峰, 这类波峰也需要消除。为此, 本文采用一种根据波峰所代表的密度将局部峰值消除的方法。具体做法如下: 首先通过式(6)获得 $H_S(i)$ 中所有的峰值处位置:

$$\begin{aligned} H_S(i) - H_S(i-1) &> 0 \\ H_S(i) - H_S(i+1) &> 0 \end{aligned} \quad (6)$$

在获得所有峰值后, 利用式(7)分别求出每个峰值左右两边由陡变缓的转折点:

$$P_i = \arg \max_i \left| \frac{H_S(i-1) - H_S(i)}{H_S(i) - H_S(i+1)} \right| \quad (7)$$

因此由峰值点和左右两个转折点可以构成一个三角区域。由图2(b)所示, 由每个峰值和左右两边的转折点构成了三角区域。当一个峰值所表示的密度值越大时, 三角区域的面积相应越大。因此本文采用三角区域的面积作为一个波峰所代表的密度值。设波峰处到左右两个转折点的距离分别为 D_L 和 D_R , 则一个波峰所代表的密度值 S_i 为由 D_L 和 D_R 作为两条边围成的三角形面积:

$$S_i = D_L \cdot D_R \cdot \cos(\omega) \quad (8)$$

当 S_i 的值小于一个阈值时, 表示该处可能是由异常运动

向量构成的波峰, 因此需要剔除。在剔除局部峰值后, 利用剩余的峰值数作为分类数, 并用峰值处的 bin 值作为聚类中心。

至此, 完成运动向量的聚类。通过运动向量可以实现对应的特征点聚类。

4 实验结果及分析

为了验证本文算法的有效性和鲁棒性, 对多种标准测试视频进行了实验。

正确地获得分类数是进行运动分割的关键。因此首先对本文提出的自动运动数获取方法进行了验证。图3为针对测试视频 coastguard 和 mobile 自动获取分类数的实验结果, 其中第一行为原图, 第二行为相应的 KDE 图, 第三行为获得的分类数。coastguard 测试序列为摄像机跟踪拍摄多个运动物体, 摄像机随着物体运动而运动。图3中第一列为对 coastguard 前 50 帧的测试结果。在前 50 帧有两个运动物体出现。从 KDE 图中, 可以看出由这两个运动物体以及摄像机的运动构成的运动向量组成了 3 个波峰。根据本文第 3 节的核密度估计方法, 将每个波峰作为一个运动类型。因此前 50 帧中, 本文获得的分类数为 3。在 50 帧后, 向左运动的摩托艇消失。可以从图3第 2 列的 KDE 图中看出, 运动向量组成了两个主要波峰。但是由于噪声干扰, 在两个波峰之间出现了由较小密度值构成的局部峰值。根据本文提出的波峰所围成的面积计算出该局部波峰的面积很小, 因此剔除该局部波峰。根据主要波峰数可以获得运动分类数为 2。

Mobile 测试序列包含多种复杂运动。常规方法在这种复杂场景中正确地获得运动分类数比较困难。从 mobile 的 KDE 图中可以看出, 各种运动构成了多个波峰, 在主要波峰之间也包含有多个局部波峰。根据本文的局部波峰剔除策略, 成功地将一些峰刺剔除。最后根据主要波峰数获得的运动分类数为 5。从该实验可以看出本文提出的自动运动分割方法能够根据场景的变化有效地检测出实际的运动类型数。

图4为对 coastguard 和 walking man 的运动分割结果。其中第一行为对 coastguard 的分割结果。在正确地获得了分类数的同时也完成了运动向量对应特征点的聚类。从图4第一行可以看出, 不同的运动物体和背景点之间成功地进行了分离。该实验证明了本文提出的自动运动分割算法能够在摄像机运动的复杂场景下根据场景的变化自动进行运动分割。

第二行为对摄像机跟踪拍摄的一个步行者的视频片段(walking man)。由于人物的运动是非刚体的, 因此传统的参数运动模型无法对人物的运动进行建模, 文献[6, 7]中的方法都会失效。图4第2行为本文方法对该视频的测试结果。根据本文第3节的自动分割方法, 运动向量被分为两类。从图3(e)-(h)中可以看出, 属于人物的特征点被聚成一类, 这是因为本文方法采用了均值偏移处理技术, 将由非刚体运动产生的异常运动向量向主要运动进行了偏移。这种处理扩大了人体的主要运动, 同时减小了腿部区域等非刚体运动带来的影响。从实验结果可以看出, 本文方法对刚体或非刚体运动具有较好的自适应性。从连续 4 帧的测试结果, 可以看出本文方法的鲁棒性。

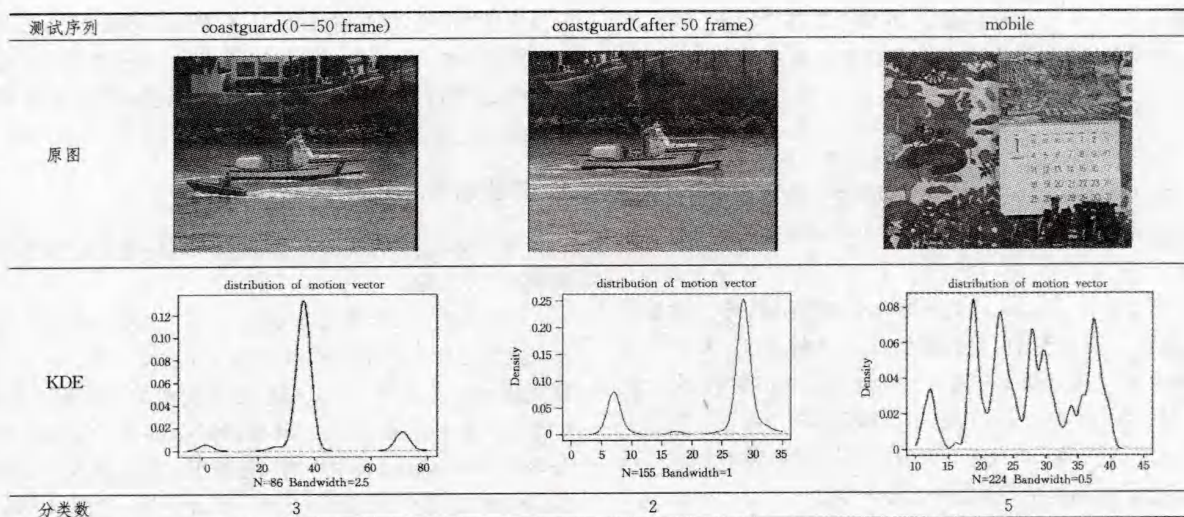


图3 运动分割实验结果图

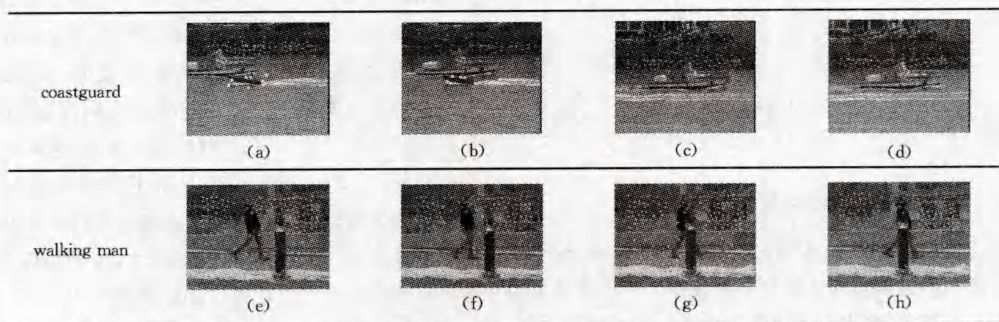


图4 运动分割实验结果图

结束语 本文提出了一种自适应的运动分割方法。该方法利用前后若干帧特征点的对应关系获得特征点运动向量。通过均值偏移方法修正了运动向量中与主运动不相符的异常运动向量后,再通过一种专用的聚类算法将特征点进行分类。本文方法是根据特征点的对应关系获得运动信息,其含有的噪声较少,而且不依赖特定的运动模型,运用了均值偏移对异常运动向量向主要运动成分偏移,减小了异常运动向量对运动分割的影响。实验结果证明了本文方法的有效性和鲁棒性。

参考文献

- [1] Sajid G. Motion segmentation in videos from time of flight cameras[C]//2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP). Vienna, Austria, April 2012
- [2] Gao Ji-zhou, Yang Rui-gang, Gong Ming-lun. Video Stereolization:Combining Motion Analysis with User Interaction[J]. IEEE Transactions on Visualization and Computer Graphics, 2012,7(18):1079-1088
- [3] Jang J, Lee H. Enhanced motion estimation algorithm with pre-filtering in video compression[J]. Optical Engineering, 2012, 51(3):37002-37009
- [4] Karavasilis V, Blekas K, Nikou C. Motion Segmentation by Model-Based Clustering of Incomplete Trajectories[J]. Lecture Notes in Computer Science, 2011, 6912:146-161
- [5] Wang Xiao-gang, Tieu K. Learning Semantic Scene Models by Trajectory Analysis[C]//Proceedings of European Conference on Computer Vision(ECCV). 2006
- [6] Antonini G, Thiran J P. Counting Pedestrians in Video Sequences Using Trajectory Clustering[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(8):1008-1020
- [7] Bashir F I, Khokhar A A. Real-Time Motion Trajectory-Based Indexing and Retrieval of Video Sequences[J]. IEEE Transactions on Multimedia, 2007, 9(1):58-65
- [8] Leiva L A, Vidal E. Revisiting the k-means algorithm for fast trajectory segmentation[C]//ACM SIGGRAPH 2011. New York, NY, USA, 2011
- [9] Sundberg P, Brox T. Occlusion boundary detection and figure ground assignment from optical flow[C]//2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI, 2011
- [10] Harris C G, Stephens M J. A combined corner and edge detector [A]//Proceedings Fourth Alvey Vision Conference[C]. Manchester, UK, 1988:147-151
- [11] Bradski G R, Pisarevsky V. Application in calibration, stereo, segmentation, tracking, gesture, face and object recognition[C]//IEEE Conference on computer vision and pattern recognition. SC, USA, 2000:796-797
- [12] Pundlik S J. Real-Time Motion Segmentation of Sparse Feature Points at Any Speed[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2008, 38(3):731-742
- [13] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002;24(5):603-619