

一种改进的运动目标检测方法

杨会锋¹, 曹洁¹, 帅立国²

YANG Huifeng¹, CAO Jie¹, SHUAI Ligu²

1. 兰州理工大学 计算机与通信学院, 兰州 730050

2. 东南大学 机械工程学院, 南京 210096

1. School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

2. School of Mechanical Engineering, Southeast University, Nanjing 210096, China

YANG Huifeng, CAO Jie, SHUAI Ligu. Improved method for moving target detection. Computer Engineering and Applications, 2012, 48(20): 206-210.

Abstract: In order to overcome the drawbacks of the traditional background subtraction method, a moving target detection method based on the combination of edge features and improved K -means clustering is provided. Background model is established using improved K -means clustering method. The difference between foreground image and background image is calculated and the binarization process is employed. False targets are judged and rejected with the foreground edge information. Experimental results show that the method can effectively eliminate the false moving target resulted from object move-in or move-out and improve the accuracy of moving target detection.

Key words: K -means clustering; background subtraction; edge information; false moving target

摘要: 为了克服传统背景差分法所存在的不足, 提出了一种基于边缘特征和改进 K -均值聚类相结合运动目标检测方法。运用改进的 K -均值聚类方法建立背景模型, 将其与前景图像相差后得到的二值化图像, 通过前景边缘信息的鲁棒性来判别及去除存在的虚假目标。实验结果表明, 该方法可以有效去除背景中突然有物体移入或移出引起的虚假运动目标, 提高了运动目标检测的准确性。

关键词: K -均值聚类; 背景差; 边缘信息; 虚假运动目标

文章编号: 1002-8331(2012)20-0206-05 **文献标识码:** A **中图分类号:** TP391.4

如何从最底层的原始视频数据得到高层的语义信息并进行理解是计算机视觉领域中最根本的一个问题^[1]。目前, 国内外的研究大多是在对运动目标的视觉分析方面。而运动目标的视觉分析作为计算机视觉领域中最活跃的研究课题之一, 其重点是利用计算机视觉技术对运动目标进行检测、跟踪和识别后对其行为进行描述和理解。

运动目标检测作为计算机视觉领域最底层也是最基础的部分, 其结果直接关系到下一步目标识别与跟踪、行为理解等视频分析和处理应用的好坏^[2]。

目前常用的目标检测方法主要有以下几种^[3-4]: 光流法、帧差法及背景差法。其中, 光流法是基于运动目标随时间变化的特性, 它能够在不知道场景的任何信息的情况下, 检测出运动目标, 但其计算耗时, 无法满足实时性的要求。帧差法通过对两幅相邻帧图像的相减, 以滤除图像中的静止景物, 而仅保留运动物体, 该方法的优点是对环境的光线变化不敏感, 缺点是无法检测静止车辆。背景差法通过当前帧图像与背景图像的相减得到运动目标, 该算法原理简单、易于实现, 所以一直是研究的热点。

基金项目: 甘肃省自然科学基金项目(No.1010RJZA046); 甘肃省教育厅研究生导师基金项目(No.0914ZTB003)。

作者简介: 杨会锋(1984—), 男, 硕士研究生, 主要研究领域为智能交通系统; 曹洁(1966—), 女, 教授, 博导, 主要研究领域为信息融合理论与应用、智能交通系统等; 帅立国(1968—), 男, 教授, 博导, 主要研究领域为机器人触觉技术、制造业技术研究等。
E-mail: nongchaoer000@163.com

收稿日期: 2011-03-03 **修回日期:** 2011-05-11 **CNKI 出版日期:** 2011-08-04

DOI: 10.3778/j.issn.1002-8331.2012.20.043 <http://www.cnki.net/kcms/detail/11.2127.TP.20110804.1611.158.html>

在室外复杂场景中,往往存在着各种情况影响到前景和背景的有效分离,比如风吹树叶的晃动、摄像机的抖动等一系列的因素都会对检测结果产生影响^[5],进而影响到前景运动目标的精确获取。为此,Zhang H和Xu D^[6]提出了一种融合了颜色和纹理特征的检测方法,该方法有效地抑制了微小树木和摄像机抖动引起的影响;Omar Javed^[7]等提出了一种分级检测方法,算法包括像素级、区域级和帧级处理,解决了光照突变、背景中物体的移动等问题;Stauffer与Grimson^[8]通过对视频图像中的像素点进行混合高斯分类建模,并通过模型中每一个高斯分布的变化性和持续性,来判断哪个高斯分布最接近背景分布,其有效地解决了因光照变化、背景噪声的干扰等所产生的影响,但这种方法不能快速适应背景的运动,如果背景不能及时更新,突然移入或移出背景的物体就仍然被判为前景,形成虚假运动目标。

为了克服虚假目标造成的影响,本文提出了一种基于改进K-均值聚类和边缘特征相结合的检测方法。

1 基于K-均值聚类的运动目标检测方法

1.1 背景建模

在复杂的室外环境中,摄像机抖动、风吹树引起的晃动及水面的波纹等多方面的因素会给检测结果造成影响。鉴于此,本文用K-均值聚类的思想对K个子类进行描述。

K-均值聚类算法^[9]是根据事先给定的初始类别数来对样本进行划分的,但是在一般的实际应用场景当中,很难在背景建模之前就能确定各像素位置处的像素值的类别数目,这就使得K-均值聚类在应用时受到了一定的局限。为此,本文对K-均值聚类进行了一定的改进。首先在背景建模之前,假设初始的类别数 $K=1$,即各像素位置始终能够正确地表达出相应的背景信息,也就是说所有的像素值都属于背景。在接下来的背景建模过程中随着每个像素位置处的样本数目的增多,其样本之间的差异也越来越大。只要某一样本与该像素位置处某一类别中心之间的距离小于规定的阈值,则该样本就归入这个子类;反之,则认为新的样本值不属于已有的所有类,这时就给该像素的样本增加一个类别,使 $K=K+1$ 。如此按一定的规则对各个像素的观察值进行聚类,并在聚类过程中逐步确定观察值的类别。经过一段时间的背景建模学习之后,次数出现最多即类中样本数最多的类就构成了动态背景的模式。其中,样本数较多的子类就构成了静态背景,而其他子类则描述了背景干扰如晃动的树、摄像机的抖动等多模态分布情况。

根据以上思想,本文基于改进的K-均值聚类算法的背景学习过程如下:

步骤1 对每个像素,令K为聚类的初始类别数,且 $K=1$,初始聚类均值定义为采集到的第一帧图像像素值。

步骤2 采集一帧视频图像,对每个像素位置,确定当前像素值与已经存在的聚类中心的最小距离 D_{\min} ,若 D_{\min} 大于规定的阈值 T_a ,即 $D_{\min} > T_a$ 时,则认为当前像素值已经不属于已有的K个类,这时该像素的样本增加一个新的类别,即 $K=K+1$,并将该像素值作为第K类的聚类中心;反之,若 $D_{\min} < T_a$ 时,则当前像素值属于子类K,该子类样本数加1,并更新子类中心,完后转步骤3。

步骤3 对学习完的每个像素,按样本数目的多少由大到小对已存在的子类进行排序。

步骤4 按步骤2、3一直计算直到满足学习结束条件,比如达到指定的帧数,最后选择出样本数大于阈值 T_b 的子类作为背景模型。否则,返回步骤2,继续学习直到满足以上结束条件为止。

1.2 背景更新

外复杂环境下,由于光线和阴影等因素的影响,背景并不是静止不变的。如果提取出的背景图像不能及时跟上实际的背景图像的变化时,则检测出来的结果中可能会产生大量的噪声而影响到结果的准确性。因此,为了增强背景模型的适用性,需要根据实际情况对背景模型进行实时更新。目前背景更新的方法有很多,最常用的方法有多帧平均法和选择更新法。多帧平均法是一种非选择性算法,它按下式对背景图像进行动态更新:

$$B_k(x, y) = \alpha \cdot I_k(x, y) + (1 - \alpha) \cdot B_{k-1}(x, y) \quad (1)$$

其中, $B_k(x, y)$ 为当前更新的背景, $B_{k-1}(x, y)$ 为前一帧背景, $I_k(x, y)$ 为上一幅图像,系数 $\alpha \in [0, 1]$,表示背景模型的更新速度, α 越趋向于1,则新背景就越接近于上一幅背景。但其不足是对场景的改变过于敏感。研究表明,选择性更新算法更优于选择性更新,故本文中采用选择性更新来对背景进行更新。选择性更新法的基本思想是只将没有检测到运动的区域即真正的背景进行更新。其更新算法描述如下:

$$B_k(x, y) = \begin{cases} \alpha \cdot I_k(x, y) + (1 - \alpha) \cdot B_{k-1}(x, y), & (x, y) \text{ 为变化的像素} \\ B_{k-1}(x, y), & \text{otherwise} \end{cases} \quad (2)$$

其中 α 也是更新率,范围取值一般在0到1, α 越小,背景更新的速度越慢,运动目标对背景的影响越小; α 越大,背景的更新速度越快,运动目标的影响也就越大。在具体的实验中,可根据实际环境光线的变化来确定 α 的取值。判断当前帧图像中的像素位置

处 (x, y) 是否发生变化, 是否属于运动目标的像素。如果是则用式(2)对背景进行更新, 如果不是则 (x, y) 属于背景像素不需要更新。当背景按上述方法更新后, 就可以通过当前帧图像和更新后的背景的差值来检测和提取运动目标。

1.3 前景分割

图像的二值化作为图像处理中一个重要的环节, 其中的关键问题则是阈值的选取。目前, 常用的二值化方法主要有全局阈值法和局部阈值法^[10]。全局阈值法一般计算速度较快, 在物体和背景的灰度差别较明显时效果比较好; 局部阈值法速度一般比全局阈值法要慢, 但是在图像比较复杂或者背景存在噪声, 用单一的阈值很难区分物体和背景的情况下, 这种方法往往要优于全局阈值法。本文采用基于 Otsu 的非参数、无监督自适应阈值选取方法, 它将图像分为目标与背景两类, 通过搜索计算类间方差最大而动态地确定图像分割门限值, 且不需要其他先验知识, 其基本原理是:

设一幅图像的像素点数为 N , 它有 L 个灰度级 $(0, 1, \dots, L-1)$, 灰度级为 i 的像素点数为 n_i , 那么

$$N = \sum_{i=0}^{L-1} n_i \quad (3)$$

对图像直方图归一化, 有概率密度分布:

$$P_i = \frac{n_i}{N} \quad (4)$$

其中, $P_i \geq 0$, $\sum_{i=0}^{L-1} P_i = 1$, 假设阈值 T 将图像分成两类 C_o 和 C_b (物体和背景), 即 C_o 和 C_b 分别对应具有灰度级 $\{0, 1, \dots, t\}$ 和 $\{t+1, t+2, \dots, L-1\}$ 的像素。则 C_o 类和 C_b 类的发生概率分别为:

$$w_o = \sum_{i=0}^t P_i \quad (5)$$

$$w_b = \sum_{i=t+1}^{L-1} P_i \quad (6)$$

C_o 和 C_b 类的均值分别为:

$$\mu_o = \frac{1}{w_o} \sum_{i=0}^t iP_i \quad (7)$$

$$\mu_b = \frac{1}{w_b} \sum_{i=t+1}^{L-1} iP_i \quad (8)$$

图像和总均值为:

$$\mu = w_b \mu_b + w_o \mu_o \quad (9)$$

图像目标和背景两类的类间方差为:

$$\sigma^2 = w_o (\mu_o - \mu)^2 + w_b (\mu_b - \mu)^2 \quad (10)$$

按照最大类间方差的准则, 从 0 至 $L-1$ 改变 t , 最佳阈值 T 应使得类间方差最大, 即:

$$\sigma^2(T) = \sum_{0 \leq t \leq L-1} \max\{\sigma^2(t)\} \quad (11)$$

根据以上确定出的分割最佳阈值, 前景目标就被完整地检测出来。

2 基于边缘的目标检测方法

建立边缘作为图像的基本特征之一, 是图像中灰度发生急剧变化的区域边界。边缘检测的实质是采用某种算法来提取出图像中目标与背景间的交界线。本文采用 Canny 边缘检测方法, 其基本思想是^[11]: 首先选择一定的 Gauss 滤波器对图像进行平滑滤波; 然后采用非极值抑制 (Non-maxima Suppression) 技术, 对平滑后的图像处理, 得到最后的边缘图像。Canny 边缘检测算子的检测步骤如下:

步骤 1 将待处理图像用高斯滤波器进行平滑处理, 去除图像噪声, 高斯滤波函数为:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\delta^2}\right) \quad (12)$$

一般选取方差为 1.4 的高斯函数模板 $G(x, y)$ 和图像 $f(x, y)$ 进行卷积运算, 得到平滑图像 $s(x, y)$ 。

步骤 2 用一阶偏导数的有限差分计算梯度的幅度 $A(i, j)$ 和方向 $a(i, j)$ 。

步骤 3 对梯度幅值进行非极大值抑制。仅确定图像 $f(x, y)$ 上 (i, j) 点处的边缘强度 $A(i, j)$ 值的大小, 并不能确定该点是否为边缘点, 因此为确定边缘, 必须保留局部梯度最大的点, 而抑制非极大值。即如果图像 $f(x, y)$ 上 (i, j) 像素点的边缘强度 $A(i, j)$ 小于沿着梯度线方向的两个相邻像素点的边缘强度, 则认为该点为非边缘点。

步骤 4 图像在非极大值抑制后依然存在许多由于纹理和噪声引起的假边缘, 可以用双阈值算法来检测和连接边缘。通过高阈值 T_h 和低阈值 T_l 可以分别得到检测结果, 然后在高阈值 T_h 检测结果中连接边缘轮廓, 连接到端点时, 到低阈值 T_l 结果中寻找边缘点。

3 基于边缘和改进 K-均值聚类算法相结合的目标检测方法

一般情况下, 场景中总会出现运动的物体突然静止在某一位置而遮盖住了原先的背景位置, 或者原本属于背景中的物体突然在某一时刻可以运动, 使得原本被物体遮盖的背景显露出来。在这两种情况下背景都发生了变化, 基于模板的更新方法往往不能及时更新背景, 这就使得背景发生变化却实际并不存在物体的区域被检测为运动物体, 产生虚假运动目标。

为了解决上述问题, 本文将边缘信息和基于背

景差检测结果相结合。首先在时间域上通过 K -均值聚类构造出背景模型并及时维护,与前景差值后提取出目标的初始轮廓;然后在空间域上用canny边缘检测算子提取出前景目标的边缘轮廓,然后将其与前一步提取出的结果作“与”运算,对结果数学形态学处理之后并最终提取出运动目标的真实轮廓。本文算法的实现过程如图1所示。

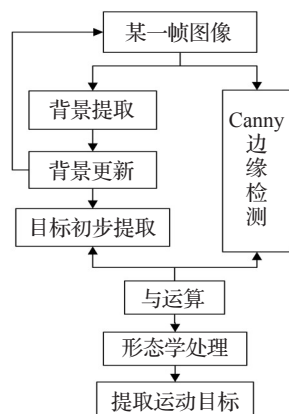


图1 运动目标检测过程图

4 实验结果与分析

为了验证算法的有效性,本程序采用MATLAB编程,并分别用一段室外高速公路视频和大学校园视频进行测试。其中两段视频格式都为PAL制式,帧率为25帧/s,图像分辨率为 320×240 像素,真彩色格式。视频中一直有运动目标存在,初始背景图像必须由多帧图像生成。

背景提取的对比实验分别用基于传统的 K 均值聚类算法和改进算法通过对同一视频序列进行。其中,图2(a)、图2(b)分别为视频中第78帧和第99帧含有运动目标的视频帧,图2(c)、图2(d)分别是基于

传统的 K 均值聚类算法和本文改进算法对99帧图像序列背景建模提取出的背景。其中,两种算法在同一视频序列中生成背景所耗费的时间分别为:本文算法为1.9 s,传统算法为1.7 s。

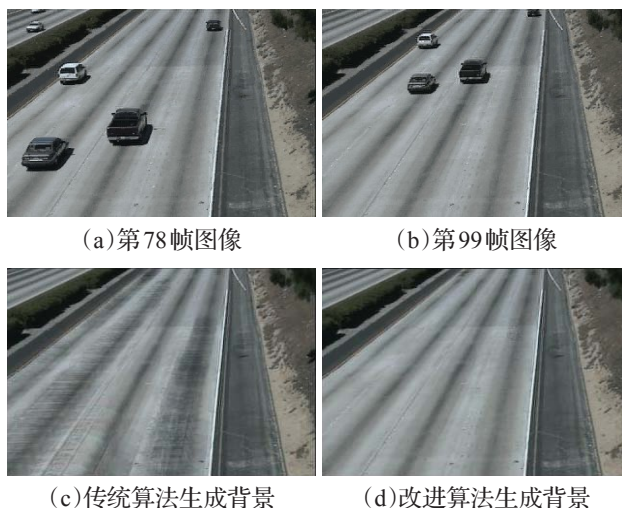


图2 背景提取

从图2中可以看出,由于传统 K -均值聚类算法是在给定类 K 的前提下进行的,所以其在聚类初始中心的选择上比改进的算法节省了时间。但由于其初始聚类中心点选择的随机性,使得聚类结果容易陷入局部最优解。如图2(c)所示,在含有各种噪声影响的室外环境中用传统算法提取出来的背景中出现了很深的车辆压痕,容易产生累积误差,这对于后续的运动目标检测会产生较大影响。通过比较发现,本文的改进算法能提取出更优的背景图像。

在图3中,图3(a)为第78帧图像,此时如图右下方深色轿车驶入场景中最终停留在图中的停车位,并成为背景的一部分。图3(a)为第120帧图像,此时

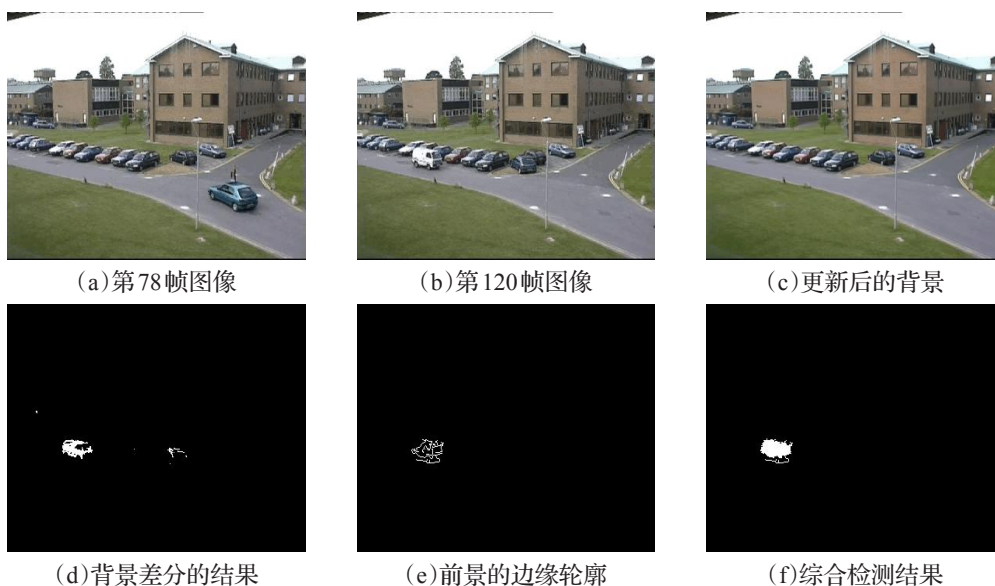


图3 目标分割

场景中的背景发生了变化,从右方驶入一辆白色的小面包车,由于背景模型不能及时更新,所以更新后的背景图像还是如图3(c)所示,其中并没有包括已经成为背景的小车,这时用背景差分方法会把真实背景误检测为前景运动目标(如图3(d)),图3(e)为提取出的第120帧图像中运动车辆的边缘轮廓。本文算法的检测结果如图3(f)所示,可以看出,边缘信息良好的鲁棒性保证了真实的前景目标被完整地检测出来,去除了由于背景不能及时更新而引起的虚假目标。

5 结论

本文提出了一种基于改进 K -均值聚类与边缘特征相结合的运动目标检测方法。该算法能够较好地提取出背景模型,有效地抑制了背景中物体突然移入或移出引起的虚假运动目标,提高了运动目标检测的精准性。但是,通过实验发现,该方法还不能有效地去除由于光照引起的与真实目标粘连的阴影的影响,这也是下一步工作要解决的问题。

参考文献:

- [1] 张治国,刘怀亮,马志辉,等.基于高层语义的视频检索研究[J].计算机工程与应用,2007,43(18):168-180.
- [2] James W D, Vinay S. Background-subtraction in thermal imagery using contour saliency[J]. International Journal of Computer Vision, 2007, 71(2): 161-181.
- [3] 张娟,毛晓波,陈铁军.运动目标跟踪算法研究综述[J].计算机应用研究,2009,26(12):4407-4410.
- [4] Motamed C. Motion detection and tracking using belief indicators for an automatic visual-surveillance system[J]. Image and Vision Computing, 2006, 24(11): 1192-1201.
- [5] Ding Y, Wenhui L. Robust moving object detection under complex background[J]. Computer Science and Information Systems, 2010, 7(1): 201-210.
- [6] Zhang H, Xu D. Fusing color and texture features for background model[C]//3rd Int Conf on Fuzzy Systems and Knowledge Discovery, 2006, 4233(7): 887-893.
- [7] Javed O, Shafique K, Shah M. A hierarchical approach to robust background subtraction using color and gradient information[C]//Proc of the IEEE Workshop on Motion and Video Computing, 2002: 22-27.
- [8] Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(8): 747-757.
- [9] 云廷进,郭永彩,高潮. K -均值聚类中心分析法实现红外人体目标分割[J].光电工程,2008,35(3):140-144.
- [10] 汪海洋,潘德炉,夏德深.二维Otsu自适应阈值选取算法的快速实现[J].自动化学报,2007,33(9):968-971.
- [11] 李牧,闫继红,李戈,等.自适应Canny算子边缘检测技术[J].哈尔滨工程大学学报,2007,28(9):1002-1007.
- [12] 施家栋,王建中.动态场景中的运动目标监测与跟踪[J].北京理工大学学报,2009,29(10):858-876.
- [13] Wang Z J, Zhan D C, Xu X F. A component retrieval method based on feature tree matching[J]. International Journal of Information Technology, 2006, 12(8): 60-72.
- [14] Zhou J T, Zhang S S, Wang M W, et al. XML-RDB driven semi-structure data management[J]. Journal of Information and Computing Science, 2006, 1(5): 266-274.
- [15] Nierman A, Jagadish H V. Evaluating structural similarity in XML documents[C]//Int'l Workshop on the Web and Database(WebDB), 2002: 61-66.
- [16] 聂卉,黄贵鹏.树编辑距离在Web信息抽取中的应用与实现[J].现代图书情报技术,2010,9(5):29-34.
- [17] Kim J, Kim H J. A partition index for XML and semi-structured data[J]. Data & Knowledge Engineering, 2004, 51: 349-368.
- [18] 郑伟.基于XML的树型结构编码及结构相似性匹配方法[D].长春:东北师范大学,2006.
- [19] Wang L, Cheung D W, Mamoulis N, et al. An efficient and scalable algorithm for clustering XML documents by structure[J]. IEEE Transactions on Knowledge and Data Engineering, 2004, 16(1): 82-96.
- [20] 鲁远,万常选.XML信息检索中结构相似度计算[J].情报杂志,2007(11):109-111.
- [21] 谌志群.XML文档相似度计算方法研究[J].情报学报,2009,28(1):48-57.

(上接181页)

- [10] Zhai Y, Liu B. Structured data extraction from the Web based on partial tree alignment[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(12): 1614-1628.
- [11] Ben Aouicha M, Boughanem M, Tmar M, et al. XML information retrieval based on tree matching[C]//Proceedings of the 15th Annual IEEE International Conference and Workshop on Engineering of Computer Based Systems, 2008: 499-500.
- [12] 王渊峰,薛云皎,张涌,等.刻画分类构件的匹配模型[J].软件学报,2003,14(3):401-408.
- [13] Wang Z J, Zhan D C, Xu X F. A component retrieval method based on feature tree matching[J]. International Journal of Information Technology, 2006, 12(8): 60-72.
- [14] Zhou J T, Zhang S S, Wang M W, et al. XML-RDB driven semi-structure data management[J]. Journal of Information and Computing Science, 2006, 1(5): 266-274.