

一种 DWT 与背景重构相结合的运动目标分割方法

任应军 范九伦

(西安邮电学院通信与信息工程学院 西安 710061)

摘 要 针对视频处理中数据量大、消耗时间长的缺点,在假定“背景像素总是以较高的频率在图像序列中出现”的前提下,提出了一种新的运动目标分割方法。首先将待处理的视频图像序列经过 DWT 变换(Discrete Wavelet Transform)提取其近似分量,然后利用像素点聚类方法,结合双阈值和相似类合并,选择一段时间内频率出现较高的像素值来重构背景,最后借鉴图像匹配的评价标准来验证重构背景的准确性。实验结果表明,该方法能提取出较好的背景,从而实现对运动目标的高效且完整分割。

关键词 DWT,背景相减,像素聚类,背景重构,运动分割

中图分类号 TP391.4 **文献标识码** A

Moving Objects Segmentation Method Combining DWT and Background Reconstruction

REN Ying-jun FAN Jiu-lun

(School of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710061, China)

Abstract Due to the disadvantage of time-consuming and large volumes data of video processing, based on the hypothesis that the background pixel intensity always appears in image sequence with maximum probability, a novel method of moving objects segmentation was presented. This approach extracts the approximate components of video images sequence on hand firstly using the DWT (discrete wavelet transform), and then selects the higher frequency pixel values of an interval to reconstruct the background according to the pixels clustering combining dual-threshold with similar classes merged method, at last verifies background reconstruction accuracy by quoting the evaluation criteria of image matches. Experimental results demonstrate that the proposed method can extract better background to achieve efficient and integrated moving object segmentation.

Keywords DWT, Background subtraction, Pixels clustering, Background reconstruction, Moving segmentation

1 引言

视频序列中运动目标的检测与分割被广泛用于安全防范和工农业生产等各大领域。在摄像机固定条件下,图像中背景区域固定不动,因此可以利用光流法^[1]、帧间差分法^[2]和背景差分法^[3,4]来提取运动目标。

光流法是通过给视频图像中的每个像素点赋予一个运动矢量,形成一个运动场,根据每个像素点的速度矢量特征对视频图像进行动态分析。当视频中有运动物体运动时,运动物体所形成的速度矢量与邻域的背景速度矢量不同,据此检测运动物体及位置。该方法能够检测独立的运动对象,不需要预先知道场景的任何信息,但其运算公式复杂,计算量大,在没有特殊硬件支持的条件下很难达到实时要求。帧间差分法的基本原理是:对视频图像中相邻的两帧(或三帧)作差来获得运动目标轮廓。该方法只检测相对运动的物体,受光线变化影响小,检测有效稳定,但容易形成空洞,检测出的物体位置不精确,加之帧间差分法的分割效果受阈值的影响较大,若阈值选取不当,则分割效果不理想。背景差分法简单有效,易于实现,是实际应用中的首选,不过作为背景差分法核心的背

景图像必须准确重构。

近年来,国内外的研究人员针对背景图像的重构与更新开展了大量的工作^[5-7],这些工作大体上分为两类:第 1 类是通过概率密度函数建立统计模型,包括非参数模型法^[8]、混合高斯模型法^[9,10]、光流场等,这些模型在不含运动目标时才能正确地对模型初始化,加之参数靠经验获取,时间复杂度较高;第 2 类是计算背景像素值的方法,包括典型的时间平均法,即通过对一段时间的视频图像序列求和再平均来获得一帧近似的背景,此类方法计算简单,但对处理设备的内存要求较高,此外,提取的背景中常常有“鬼影”出现。

文献[11]假设背景在图像序列中总是被经常观察到,提出把出现频率最高的亮度值作为背景亮度值。在此基础上,文献[12]给出基于在线聚类的背景减法。虽然用以上方法都能重构完整的背景,但提取的目标均不是很理想。为此本文提出一种新的基于像素聚类的背景差分目标分割方法。

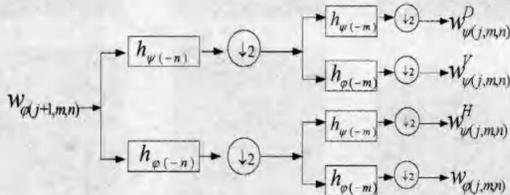
2 视频数据预处理

通常视频数据中存在冗余,采集卡采集到的数据量庞大,进行数据预处理将成为提高目标提取速度的首选途径。

到稿日期:2011-12-13 返修日期:2012-03-27 本文受国家自然科学基金项目(61102095)和陕西省教育厅专项科研计划(11JK1008)资助。

任应军(1985-),男,硕士生,主要研究方向为图像与视频通信, E-mail: ryj138@126.com; 范九伦(1964-),男,教授,博士生导师, CCF 高级会员,主要研究方向为模式识别、信息安全。

DWT 变换广泛用于信号分析、图像处理中,通过 DWT 变换可以将输入信号分解为 4 个较低分辨率(或低尺度)的分量,其分别为近似值($w_{\psi(j,m,n)}$)、水平细节($w_{\psi^H(j,m,n)}$)、垂直细节($w_{\psi^V(j,m,n)}$)、对角细节($w_{\psi^D(j,m,n)}$),如图 1 所示,其中在第一次迭代中 $w_{\psi(j+1,m,n)} = f(x,y)$, $h_{\psi(-n)}$, $h_{\psi(-n)}$ 分别为高通滤波器和低通滤波器。



DWT 变换后的近似值只占原来图像的 1/4,因此,将其作为下一步重构背景的数据会大大提高处理速度。

3 背景重构

在文献[11]的基础上,文献[12]采用在线聚类算法,其思路清晰,效率较高,但存在以下不足:①采用单一阈值,如果阈值选取不当,重构效果不佳;②经过聚类以后,可能出现两个类的数据接近的情况,这时把它们进行合并是最理想的,对此文献[12]的方法在处理时存在困难;③以均值作为聚类中心会导致类中的数据在阈值范围内较为分散,若以固定值为聚类中心则数据的“紧致性”会更好。为此,本文对其进行改进。

3.1 像素聚类法提取背景的原理

图 2、图 3 为同一视频图像序列中不同两点在 100 帧中的直方图统计和灰度变化曲线。

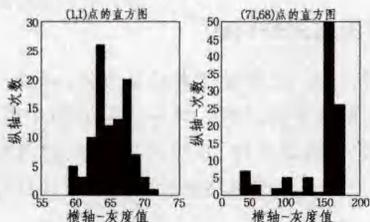


图 2 同一视频序列中不同两点的直方图

其中(1,1)点是没有受到前景目标影响的点,(71,68)是受到前景运动目标影响的点,从直方图统计结果和点运动曲线来看,“背景点总是以较大频率出现”的假设是完全合理的。(1,1)点的像素值总是在 60~72 之间变化,(71,68)点的像素值当受到运动目标干扰后会迅速变化,当运动目标离开时又回到原来的像素值附近。基于此通过对一段时间内像素值的统计聚类,可求取出现频率最高的像素值,并将其作为背景点像素值来重构背景,提取运动目标。

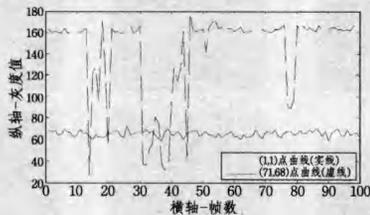


图 3 同一视频序列中不同两点变化曲线图

3.2 像素聚类下的背景重构

像素聚类的思想是:首先将输入的第一帧中的像素灰度

值作为第一个类,当新的数据到来时与该类中的首元素相比较,如果该数据与首元素的绝对值的差大于设定的阈值 $thre1$,则重新分类。反之,将该数据归入该类,直到所有数据处理完毕。

本文提出的背景重构步骤如下:

1) 以 $samp_per$ 为固定采样周期采样输入视频序列 (L_1, L_2, \dots, L_M) (本文的 $samp_per$ 为 3), 采样后的序列记为 (L_1', L_2', \dots, L_M'), 对序列 (L_1', L_2', \dots, L_M') 进行一级离散小波变化, 取其近似分量部分得到的序列记为 (f_0, f_1, \dots, f_N)。这里 $f_k(x, y)$ 表示第 k 帧中 (x, y) 像素点的灰度值 ($k=0, 1, 2, \dots, N$)。

2) 像素聚类

(1) 创建一个类 C_1 , 把第一帧的元素 $f_0(x, y)$ 放入类中。

(2) 从第二帧开始将位于同一坐标处的像素 $f_k(x, y)$ ($k=1, 2, \dots, N$) 与前面生成类的首元素 $C_j(1)$ 相比较, 计算其距离 $d(f_k(x, y), C_j(1)) = |f_k(x, y) - C_j(1)|$, 取 $j(j=1, 2, \dots, k)$ 个类中距离最小的类。如果距离 $d(f_k(x, y), C_j(1))$ 小于阈值 $thre1$, 则将元素 $f_k(x, y)$ 放入该类中。反之, 创建新类, 并将此元素放入新类中, 直到该坐标处的元素全部归类完毕。

3) 统计各个类中数据的平均值。设 C_m 类中有 n 个元素, 则该类的灰度平均值通过式(1)计算:

$$\bar{C}_m = \frac{\sum_{l=1}^n C_m(L)}{n} \quad (1)$$

4) 如果两个类的灰度平均值比较接近, 即 $|\bar{C}_m - \bar{C}_n| < thre2$, 则合并该类。

5) 遍历所有的分类数, 找出合并后类中数据元素最多的类的数据中值来充当该点的像素灰度值。

6) 遍历所有的像素点以实现背景的重构。

4 实验结果和分析

为验证本文算法的合理性和有效性, 使用多个视频图像序列用 matlab 进行仿真实验, 并与文献[12]算法进行比较, 所有检测结果没有经过任何后处理。实验结果如下。

视频 1 为高速公路隧道监控视频, 本文使用视频的前 60 帧作为训练序列进行背景重构。重构背景如图 4(b) 所示, 为了对重构背景进行对比需先得到一个准确的参考模板, MPEG-4 并没有提供相关序列的参考模板, 因此, 我们用常用的图像处理软件 photoshop 7.0 手动提取了背景[13], 即本文以下所简称的“理想背景”, 如图 4(a) 所示。

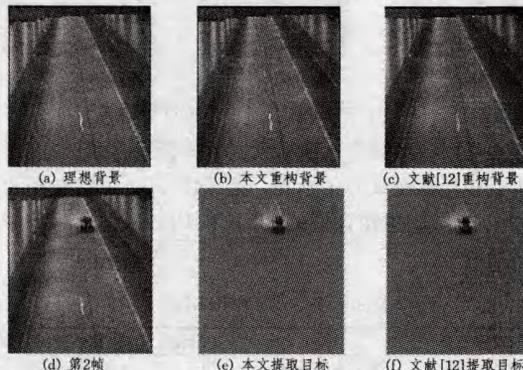


图 4 测试序列 tunnel video

为了验证重构背景的效果, 本文借鉴图像匹配的思想对

重构的背景进行量化评价。图像质量客观评价方法有均方误差 MSE(Mean Squared Error)和峰值信噪比 PSNR(Peak Signal to Noise Ratio)。该类方法具有计算复杂度低、物理意义清晰等优点,MSE 和 PSNR 成为多年来应用最广泛的客观评价标准。

图像的均方误差值:令 I_1 和 I_2 分别表示理想图像与重构图像,它们的均方误差如式(2)所示:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [I_1(i, j) - I_2(i, j)]^2 \quad (2)$$

式中, M 和 N 分别表示图像的行数和列数。MSE 的值越小,说明这两幅图像相似度越高,即重构效果越好。

图像的峰值性噪比——PSNR 用式(3)计算:

$$PSNR = 10 * \log_{10}(255^2 / MSE) \quad (3)$$

PSNR 值越大,就代表图像匹配的失真越少;相反 PSNR 值越小,就代表图像匹配的失真越多。

表 1 是用理想背景分别与本文重构背景、文献[12]重构背景进行相似性测试的数据。

表 1 MES&PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	8.3811	38.8978
理想背景	文献[12]重构背景	28.5946	33.5679

从图 4(c)可知,文献[12]重构的背景中噪音严重,图 4(f)提取的目标后面存在严重的拖影现象。如表 1 所列,本文算法的 MSE 和 PSNR 均优于文献[12]重构背景的值,从数据和提取的效果来看,本文重构的背景优于文献[12]重构的背景。

图 5 的视频 2 为 PETS2000 标准测试视频,该视频中目标从停车处运动到目的地再返回,使用视频的 100 帧作为训练序列进行背景重构。由于文献[12]采用单一的阈值,因此利用重构的背景所提取的目标中混杂现象比较严重,如图 5(f)所示。本文提取的背景较为理想。

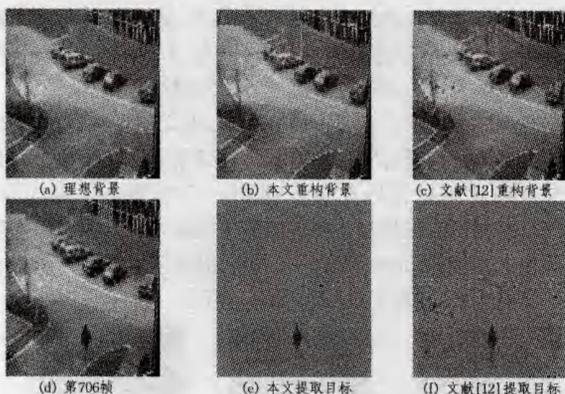


图 5 PETS2000 标准测试视频

表 2 是用理想背景分别与本文重构背景、文献[12]重构背景进行相似性测试后的数据。由于图 5(c)中混杂现象严重,因此文献[12]重构背景的 MSE 和 PSNR 值远远差于本文算法的值。

表 2 MES & PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	29.7412	38.8978
理想背景	文献[12]重构背景	107.5567	27.8144

图 6 的视频 3 为 PETS2000 标准测试视频。在该视频中

出现多运动目标,使用视频的 100 帧作为训练序列进行背景重构,通过提取多运动目标来验证本文算法重构效果的优越性。

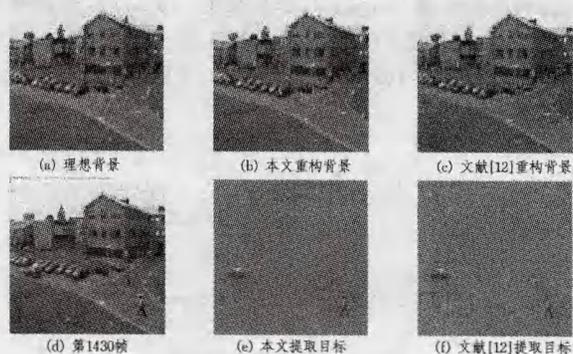


图 6 PETS2000 标准测试视频

表 3 是用理想背景分别与本文重构背景、文献[12]重构背景进行相似性测试的数据。从图 6(f)可知,文献[12]提取的目标中有路沿留下的拖影。从表 3 也可以看出本文算法的 MSE 和 PSNR 值均优于文献[12]重构背景的值,本文提取的目标效果较好。

表 3 MES & PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	24.861	34.1756
理想背景	文献[12]重构背景	39.5509	32.1592

通过以上重构的背景和提取的目标结果可以看出,本文提出的算法均能明确地提取出目标,文献[12]提取的目标混杂现象比较严重。从提取目标的效果与各表中统计出来的数据来看,本文提出的算法优于文献[12]提出的算法。

5 本文算法适用性讨论

本文算法是基于“背景像素总是以较高的频率在图像序列中出现”的前提下提出的。但是,实际场景中存在的一些运动场景不满足此假设条件,如摆动的树叶、起涟漪的水面等,在此类场景下本文算法虽然也能提取出运动目标,但是存在混合现象。

图 7 为 PETS2000 标准测试视频。由于在场景中两车相遇会车所需,右侧车辆从第 900 帧开始停在原地,直到 1229 帧才开始运动。基于本文算法的假设,如采用此段帧间的图像重构背景,则将右侧车辆视为背景重构出来,通过背景差分提取的目标如图 7(e),(f)所示,图中右侧目标车辆前方的浅白色部分,就是由于不符合假设前提所得到的结果,虽然本文算法也能提取出清晰的运动目标,但是存在瑕疵。

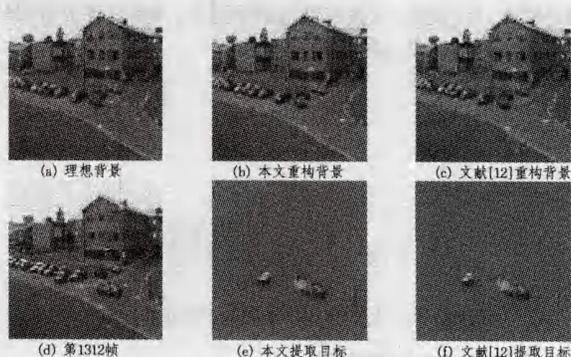


图 7 PETS2000 标准测试视频

结束语 针对视频处理中数据量大的特点,本文结合 DWT 变换和背景重构提出一种有效的运动目标分割方法。该算法无需对场景中的背景和目标建立模型,在像素级上结合改进的聚类算法并利用双阈值与类合并,使得类间相似度较小且类内相似度较大,从而实现背景重构。我们通过引用图像匹配的参数量化背景,验证了本文方法的有效性,该方法能完成对运动目标的完整分割。但是,当视频序列不满足本文算法的假设时,会出现混合现象,如何消除该现象是我们下一步的工作。

参 考 文 献

[1] Silva L S, Scharcanski J. Video segmentation based on motion coherence of particles in a video sequence[J]. *Image Processing, IEEE Transactions*, 2010, 19(4): 1036-1049

[2] Chien S-Y, Ma S-Y, Chen L-G. Fast video segmentation algorithm with shadow cancellation, global motion compensation, and adaptive threshold techniques[J]. *Multimedia, IEEE Transactions*, 2004, 6(5): 732-748

[3] Venkateswaran S, Desai U B. DWT based hierarchical video segmentation[C]// *Circuits and Systems, 2002, IEEE International Symposium*. 2002: 815-818

[4] Cevher V, Sankaranarayanan A, Duarte M, et al. Compressive sensing for background subtraction[C]// *Computer Vision-EC-*

CV 2008. 2008: 155-168

[5] Mandellos N A, Keramitsoglou I, Kiranoudis C T. A background subtraction algorithm for detecting and tracking vehicles[J]. *Expert Systems with Applications*, 2011, 38(3): 1619-1631

[6] 郑锦, 李波. 面向室外视频监控的背景重构算法[J]. *电子学报*, 2009, 37(8): 1854-1859

[7] Lu Hong, Li Hong-sheng, Liu Lan-ying, et al. Background Reconstruction Using DWT and Grayscale Classification[C]// *Artificial Intelligence and Computational Intelligence (AICI), 2010 International Conference*. 2010: 57-61

[8] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction [J]. *Computer Vision—ECCV 2000*, 2000: 751-767

[9] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]// *Computer Vision and Pattern Recognition, 1999, IEEE Computer Society Conference*. 1999, 2: 246-252

[10] 林庆, 徐柱, 王士同, 等. HSV 自适应混合高斯模型的运动目标检测[J]. *计算机科学*, 2010, 37(10): 254-256

[11] 侯志强, 韩崇昭. 基于像素灰度归类的背景重构算法[J]. *软件学报*, 2005, 16(9): 1568-1576

[12] 肖梅, 韩崇昭. 基于在线聚类的背景减法[J]. *模式识别与人工智能*, 2007, 20(1): 35-41

[13] 包红强, 张兆扬. 一种基于区域 Gibbs 势能函数的视频运动对象分割算法[J]. *通信学报*, 2005, 6(6): 57-61

(上接第 277 页)

法的自动生成,即产生式的自动生成,以及图文法在模式识别领域的更多应用。

参 考 文 献

[1] 许红霞, 张莉. 可视化语言文法形式化描述综述[J]. *计算机科学*, 2005, 32(4): 201-204

[2] Pfaltz J. Web Grammars and Picture Description[J]. *Computer Graphics and Image Processing*, 1972, 1(2): 193-220

[3] Bunke H. Attributed Programmed Graph Grammars and Their Application to Schematic Diagram Interpretation[J]. *IEEE Pattern Analysis and Machine Intelligence*, 1982, 4(6): 574-582

[4] 王莉, 钟春香. 基于规则的网状结构目标识别研究[J]. *华中理工大学学报*, 1996, 24(2): 15-18

[5] Fahmy H, Blostein D. A Graph-Grammar Programming Style for Recognition of Music Notation[J]. *Machine Vision and Applications*, 1993, 6(2/3): 83-88

[6] Grbavec A, Blostein D. Mathematics Recognition Using Graph Rewriting[C]// *Third Int. Conf. on Document Analysis and Recognition*. Montreal, Canada, Aug 1995: 417-421

[7] Schiirra B D. Visual Modeling and Programming with Graph Transformations[C]// *Tutorial at 14th IEEE Symp on Visual*

Languages. Halifax, NS, Canada, 1998

[8] 冉平, 石兵, 马晓星, 等. 上下文相关图文法分析及其应用初探[J]. *计算机科学*, 2006, 33(3): 255-260

[9] Rekers J, Schürr A. Defining and Parsing Visual Languages with Layered Graph Grammars[J]. *Journal of Visual Languages and Computing*, 1997, 8(1): 27-55

[10] Zhang Da-qian, Zhang Kang, Cao Jian-nong, et al. A Context-sensitive Graph Grammar Formalism for the Specification of Visual Languages[J]. *The Computer Journal*, 2001, 44(3): 187-200

[11] Kong Jun, Zhang Kang. On a Spatial Graph Grammar Formalism [C]// *Proc of the 2004 IEEE Symposium on Visual Languages and Human Centric Computing*. IEEE, 2004: 102-104

[12] Zeng Xiao-qin, Zhang Kang, Kong Jun, et al. RGG+: An Enhancement to the Reserved Graph Grammar Formalism[C]// *Proc of the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing*. IEEE, 2005: 272-274

[13] 曾晓勤, 韩秀清, 邹阳. 一种基于边的上下文相关图文法形式化框架[J]. *软件学报*, 2008, 19(8): 1892-1901

[14] 谭浩强. C 语言程序设计(第二版)[M]. 北京: 清华大学出版社, 1999: 21-28