

基于DWT的自适应在线聚类运动目标提取方法^{*}

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: 在假定背景像素以较高频率在图像序列中出现的前提下, 利用DWT变换的多分辨率优点, 提出一种自适应在线聚类的运动目标提取方法。首先将待处理的视频图像序列经过DWT变换提取近似分量; 然后利用像素点聚类方法, 结合自适应动态阈值和相似类合并来重构背景; 最后借鉴图像匹配的评价标准验证重构背景的准确性。实验结果表明, 该方法能够准确快速地提取运动目标, 并对环境变化、目标迂回移动、多目标运动情况具有较好的鲁棒性。

关键词: 离散小波变换; 像素聚类; 背景重构; 运动目标提取

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1001-3695(2012)09-3561-04

doi:10.3969/j.issn.1001-3695.2012.09.096

Adaptive on-line clustering moving objects extracting method based on DWT

REN Ying-jun, FAN Jiu-lun

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

Abstract: Based on the hypothesis that the background pixel intensity appears in image sequence with maximum probability, this paper proposed a novel method of moving objects extracting by utilizing the merits of multi-resolution of DWT. The method extracted the approximate components of video images sequence firstly using the DWT, and then reconstructed the background according to the pixels clustering combining adaptively dynamic threshold with similar classes merged, verified the accuracy of background reconstruction by quoting the evaluation criteria of image matches. Experimental results indicate that the proposed method can extract the motion objects correctly and fast, even in the cases of environments change, moving targets round-about, multi-targets moving.

Key words: DWT(discrete wavelet transform); pixels clustering; background reconstruction; moving targets extracting

0 引言

视频序列中运动目标提取是计算机视觉和模式识别领域中的一个重要课题, 被广泛应用于视频监控、智能导航、交通监测和分析中。对运动目标检测的研究是一项既有理论意义又有实用价值的课题, 目前运动目标检测的方法大致可分为:

a) 帧差法^[1]。它是一种通过对视频图像序列中相邻两帧(或三帧)作差分运算来获得运动目标轮廓的方法, 它可以很好地适用于存在多个运动目标和摄像机移动的情况。当监控场景中出現异常物体运动时, 帧与帧之间会出现较为明显的差别, 两帧相减得到两帧图像亮度差的绝对值, 判断它是否大于阈值来分析视频或图像序列的运动特性, 确定图像序列中是否有物体运动。帧差法的优点是: 算法实现简单, 程序设计复杂度低; 对光线等场景变化不太敏感, 能够适应各种动态环境, 稳定性较好。其缺点是: 不能提取出对象的完整区域, 只能提取出边界; 提取效果依赖于选择的帧间时间间隔。总体来说, 该方法不利于进一步的目标分析与识别。

b) 光流法^[2,3]。它检测运动物体的基本原理是: 给图像中的每一个像素点赋予一个速度矢量, 形成一个图像运动场, 在

运动的一个特定时刻, 图像上的点与三维物体上的点一一对应, 根据各个像素点的速度矢量特征, 可以对图像进行动态分析。当图像中有运动物体时, 目标和图像背景存在相对运动, 运动物体所形成的速度矢量必然和邻域背景速度矢量不同, 从而检测出运动物体及位置。光流法的优点在于光流不仅携带了运动物体的运动信息, 而且还携带了有关景物三维结构的丰富信息, 它能够在场景未知的情况下, 检测出运动对象。采用光流法进行运动物体检测的问题主要在于大多数光流法计算耗时, 实时性和实用性都较差。

c) 背景差分法^[4-6]。它是当前常用的一种运动目标检测方法, 尤其是在摄影机固定的情况下。在视频序列中, 如果得到场景的静态背景, 当目标进入场景时, 通过比较视频帧与静态背景在同位置的特征变化就可以把目标提取出来, 背景重构越精确则提取的目标越准确。因此, 背景重构既是背景差分法的核心, 也是当前研究的重点。

近年来, 国内外的研究人员对背景重构进行了大量而深入的研究, 提出了许多行之有效的方法。这些方法主要可分为两大类: a) 建立背景模型^[7]并采用自适应方法对模型参数进行调整, 从而获得新的背景图像, 包括非参数模型法^[8]、混合高

收稿日期: 2012-01-13; **修回日期:** 2012-02-25 **基金项目:** 国家自然科学基金资助项目(61102095); 陕西省教育厅专项科研计划(11JK1008)

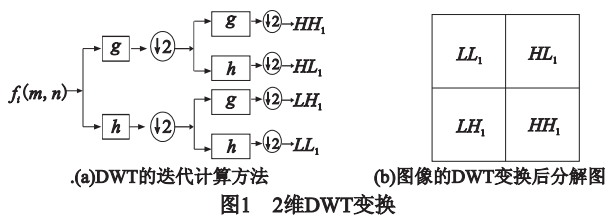
作者简介: 任应军(1985-), 男, 陕西扶风人, 硕士研究生, 主要研究方向为图像与视频通信(rj138@126.com); 范九伦(1964-), 男, 河南温县人, 教授, 博导, CCF高级会员, 主要研究方向为模式识别, 信息安全。

斯模型法^[9]、光流场等,此类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,计算量大;b)从过去的一组观测图像按照一定的假设选择像素灰度构建当前的背景图像的方法,包括典型的时间平均法、像素聚类法^[10]等。

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重构算法。文献[13]针对以上聚类中需要人工设定阈值的缺点,提出一种自适应在线聚类的背景提取方法。本文指出文献[13]中阈值设置存在的问题并进行改进,提出一种基于DWT的自适应在线聚类运动目标提取方法。

1 数据预处理

通常视频采集卡采集到的数据量庞大,视频图像数据之间往往存在各种信息的冗余,在同等的通信容量下,进行数据预处理将成为提高目标提取速度的首选途径^[14]。DWT变换因其小波基函数选择的灵活性,广泛使用在视频压缩、计算机视觉、信号分析、图像处理、生化医学中。图像的数据信息大多集中在低频部分,而高频部分的信息相对较弱,对人眼视觉的影响也较小。因此,对图像进行小波变换后,概貌信息大多集中在低频部分,如果只保留数据量1/4的低频部分,可极大地降低处理的复杂度。图1为2维DWT变换框图,在图1(a)中, $f_i(m,n)$ 表示第*i*帧图像上坐标为(*m*,*n*)的点,*g*表示高通滤波器,*h*表示低通滤波器, $\downarrow 2$ 表示下采样。 $f_i(m,n)$ 经过两次滤波和两次采样后被分解成 HH_1 、 HL_1 、 LH_1 、 LL_1 四个分量。 LL_1 是经过两次低通滤波器后的分量,被称为近似系数。一次2维DWT变换后输入的图片被分解成如图1(b)所示的4张子图,由于噪声保留在高频分量中,采用低频分量 LL_1 (近似分量)来提取目标,不但抑制了噪声,而且提高了处理速度。



2 基于像素值频次最高假设的背景重构

通过对不同视频中像素点的观察统计发现,没有运动目标经过的像素点数据分布比较均匀,有运动物体经过的像素点数据会有大幅度波动。图2、3分别为PETS标准视频图像序列hall中不同两点在100帧中的直方图统计和灰度变化曲线。其中:(1,1)点是没有受到前景目标影响的点,(71,68)点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。

2.1 阈值的改进

为了更好地重构背景,文献[13]对像素聚类中的两个关键参数,即聚类数目*k*和聚类阈值 α 进行详细分析,借鉴文献[12]的在线聚类方法,解决了预先设定聚类数目的问题。由于阈值的选择直接影响到聚类效果,当遇到变化的光照等环境因素时,固定阈值会对大量的运动像素误判,使得运动目标图

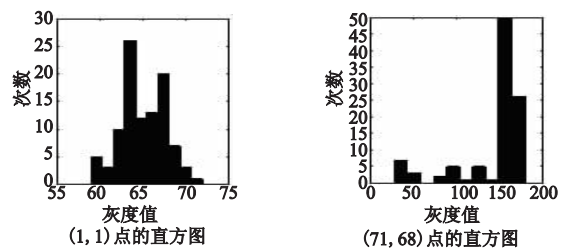


图2 同一视频序列中不同两点的直方图

像区域出现大面积的混合现象。文献[13]指出现有算法采用固定阈值的不足,提出了自适应 α 阈值的在线聚类算法。

文献[13]提出的自适应在线聚类算法思路清晰,运算速度快。但存在以下不足:a)经过一次聚类以后,可能出现两个类的数据接近的情况,这时把它们进行类合并是最理想的,但文献[13]在处理这种情况时存在困难;b)提出的自适应阈值不准确。图4是对PETS标准视频图像序列hall中(1,1)点、(71,68)点100帧相邻两帧做差之后绘制的点运动曲线。可以看出,(1,1)点运动曲线平缓,灰度差值在0~10之间分布;(71,68)点运动曲线跌宕起伏,数据分散,采用文献[13]的自适应阈值,明显会造成运动点的误分类。为此,本文对其进行改进。

在文献[13]中阈值 α 由式(1)确定:

$$\alpha = \arg\max(\arg\max |f_i(x,y) - f_{i-1}(x,y)|, \alpha) \quad (1)$$

其中: $f_i(x,y)$ 表示第*i*帧中坐标(*x*,*y*)点处的灰度值,阈值 α 的初始值为0。本文改进后的阈值 α 由式(2)求得:

$$\alpha = \beta \times \arg\max(|f_i(x,y) - f_{i-1}(x,y)|, \alpha) \quad (2)$$

在本文中,比例系数 β 的取值在0.1附近都能取得较好的重建效果。

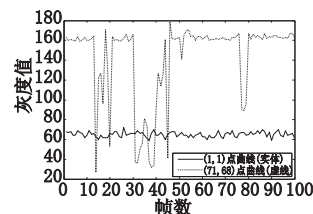


图3 同一视频序列中不同两点变化曲线

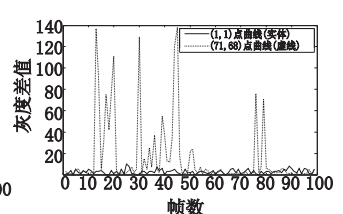


图4 同一视频序列中不同两点相邻帧差后的变化曲线

2.2 本文提出的像素聚类背景重构算法

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 α ,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:

a)对输入视频序列 $(L_1, L_2, \dots, L_M)$ 进行等间隔采样(本文的采样周期为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2维离散小波变换,取其近似分量部分 LL_1 得到序列记为 $(f_1, f_2, \dots, f_N)$ 。这里 $f_k(x,y)$ 表示第*k*帧像素点(*x*,*y*)的灰度值($k=1,2,\dots,N$)。

b)像素聚类:

(a)创建一个类 C_1 ,把第一帧的元素 $f_1(x,y)$ 赋值给该类。

(b)从第二帧开始计算同一坐标处的像素 $f_k(x,y)$ ($k=2, 3, \dots, N$)与已经创建类的首元素 $C_j(1)$ 之间的距离 $d(f_k(x,y), C_j(1)) = |f_k(x,y) - C_j(1)|$,并取 $j(j=1,2,\dots,k)$ 个类中距离最小的。如果距离 $d(f_k(x,y), C_j(1))$ 小于阈值 α ,则将元素 $f_k(x,y)$ 放入该类;反之,创建新类,将该元素放入新类中,直到所有的元素都归类完毕。

(c)统计各个类中元素的均值,设 C_m 类中有 n 个元素,通过式(3)计算其均值:

$$\bar{C}_m = \frac{\sum_{l=1}^n C_m(L)}{n} \quad (3)$$

(d)如果两个类的灰度值比较接近,即 $|\bar{C}_m - \bar{C}_n| \leq 0.5 \times \alpha$,则合并该类。

(e)遍历所有的分类,找出类合并后元素最多的类的均值作为该点的像素灰度值。

(f)遍历所有的像素点实现背景的重构。

3 重构背景效果评价方法

鉴于目前对重构背景效果评价的体系不够完善,为了验证重构背景的效果,本文借鉴图像配准的思想,对重构背景进行客观的评价。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), 此类方法具有计算形式简单,运算效率高,物理意义明显等优点, MSE 和 PSNR 已成为应用广泛的客观评价标准。

计算均方误差值:令 I_1 和 I_2 分别表示理想图像与重构图像,它们的均方误差按式(4)计算:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |I_1(i,j) - I_2(i,j)|^2 \quad (4)$$

其中: 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。

PSNR 的计算如式(5)所示。PSNR 值越大,就代表失真越少;反之 PSNR 值越小,就代表失真越多。

$$PSNR = 10 \times \log_{10}^{255^2/MSE} \quad (5)$$

4 实验结果和分析

通过对多组视频序列进行 MATLAB 实验仿真,验证在环境变化、目标迂回移动、多目标等情况下本文算法的有效性,并同文献[13]算法的实验效果进行统计比较,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60帧作为训练序列进行背景重构。重构背景如图5(c)所示,为了与重构背景进行效果对比,必须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,本文用常用的图像处理软件 Photoshop 7.0 手动提取了背景^[15],即本文以下所简称的“理想背景”,如图5(a)所示。

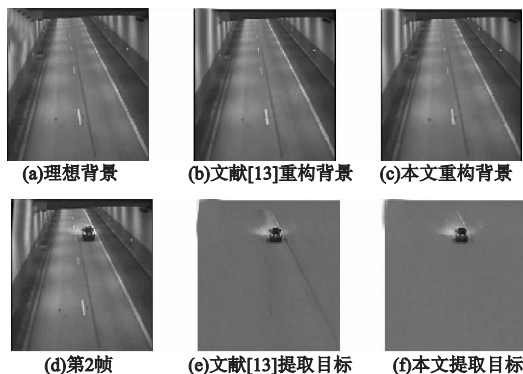


图5 测试序列tunnel video

表1是用理想背景分别与本文重构背景、文献[13]重构背景进行相似性测试后的数据。

视频1中随着汽车的驶入,虽然汽车的前照灯会对路面的

重构成造成影响,但是本文算法还是能提取出准确的目标,文献[13]由于阈值选取不当,提取后的图像中存在道路分道线被处理后留下的黑线。

表1 MSE & PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	10.704 6	37.835 1
理想背景	文献[13]重构背景	13.440 4	36.846 7

视频2为PETS2000标准测试视频,使用视频的100帧作为训练序列进行实验仿真。视频中运动目标从停车处移动到目的地再返回。如图6(b)所示,文献[13]重构的背景中明显存在拖影,而本文提取的背景较为理想。



图6 PETS2000标准测试视频

表2数据是用理想背景分别与本文重构背景、文献[13]重构背景进行相似性测试的结果。

表2 MSE & PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	19.305 2	35.274 1
理想背景	文献[13]重构背景	31.845 1	33.100 4

图7中视频3为PETS2000标准测试视频,使用视频的100帧作为训练序列进行背景重构,进行多目标提取实验。

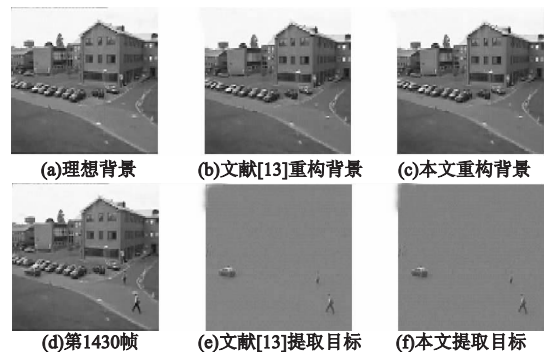


图7 PETS2000标准测试视频

表3是用理想背景分别与本文重构背景、文献[13]重构背景进行相似性测试后的数据。

表3 MSE & PSNR 对比

理想背景	重构背景	MSE	PSNR
理想背景	本文重构背景	22.007 7	34.705 1
理想背景	文献[13]重构背景	23.379 7	34.442 4

以上实验中考虑了环境变化、目标迂回移动、多目标运动的情况,分别用本文提出的算法与文献[13]的算法进行仿真,通过提取的目标可以看出,本文重构的背景均比较理想, MSE 和 PSNR 统计的数据也充分说明了本文算法的有效性。

5 本文算法适用性讨论

在“背景像素总是以较高的频率在图像序列中出现”的假设下,本文能重构到理想的背景。但是当不满足此假设条件时,虽然能重构完整的背景,但是存在混合现象。图8是标准

Hall (CIF) 测试序列分别用本文算法和文献[13]算法进行仿真实验的效果图。在此序列中第148帧之前箱子在手上随运动目标移动,背景中不含有箱子,从第148帧开始,箱子被放下来。如图8(b)(c)所示,文献[13]和本文算法均把箱子处理成背景。此外,从提取的效果来看,提取出来的左侧目标腿部含有背景,存在混合现象。

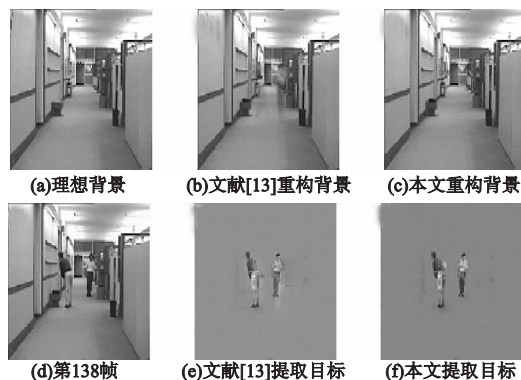


图8 Hall标准测试视频

6 结束语

本文从降低视频处理复杂度方面阐明了 DWT 变换的多分辨率优点,基于 DWT 变换提出一种自适应在线聚类的运动目标提取方法。该算法在像素级上结合改进的自适应聚类算法和相似类合并,无须建立复杂的背景模型,即可实现背景重构。通过对环境变化、目标迂回移动、多目标运动情况的讨论,借鉴图像匹配的参数对重构的背景进行评价,验证了本文方法的有效性,说明了本文方法能完成对运动目标的完整提取。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,这是笔者下一步要努力解决的问题。

参考文献:

- [1] 薛丽霞,罗艳丽,王佐成.基于帧间差分的自适应运动目标检测[J].计算机应用研究,2011,28(4):151-159.
- [2] SILVA L S, SCHARCANSKI J. Video segmentation based on motion coherence of particles in a video sequence[J]. IEEE Trans on Image Processing,2010,19(4):1036-1049.
- [3] SMITH S M, BRADY J M. ASSET-2: real-time motion segmentation and shape tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence,1995,17(8):814-820.
- [4] CEVHER V, SANKARANARAYANAN A, DUARTE M, et al. Compressive sensing for background subtraction[C]//Proc of European Conference on Computer Vision. Berlin:Springer-Verlag,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] BARNICH O, Van DROOGENBROECK M. ViBe: a universal background subtraction algorithm for video sequences[J]. IEEE Trans on Image Processing,2011,20(6):1709-1724.
- [7] HASSANPOUR H, SEDIGHI M, MANASHTY A R. Video frame's background modeling: reviewing the techniques[J]. Journal of Signal and Information Processing,2011,2(2):72-78.
- [8] ELGAMMAL A, HARWOOD D, DAVIS L. Non-parametric model for background subtraction[C]//Proc of European Conference on Computer Vision. Berlin:Springer-Verlag,2000:751-767.
- [9] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE,1999:246-252.
- [10] 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.
- [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] 夏洁,吴健,陈建明,等.基于自适应在线聚类的背景提取[J].计算机工程,2011,37(3):169-171.
- [14] CHENG F H, CHEN Y L. Real time multiple objects tracking and identification based on discrete wavelet transform[J]. Pattern recognition,2006,39(6):1126-1139.
- [15] 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.