



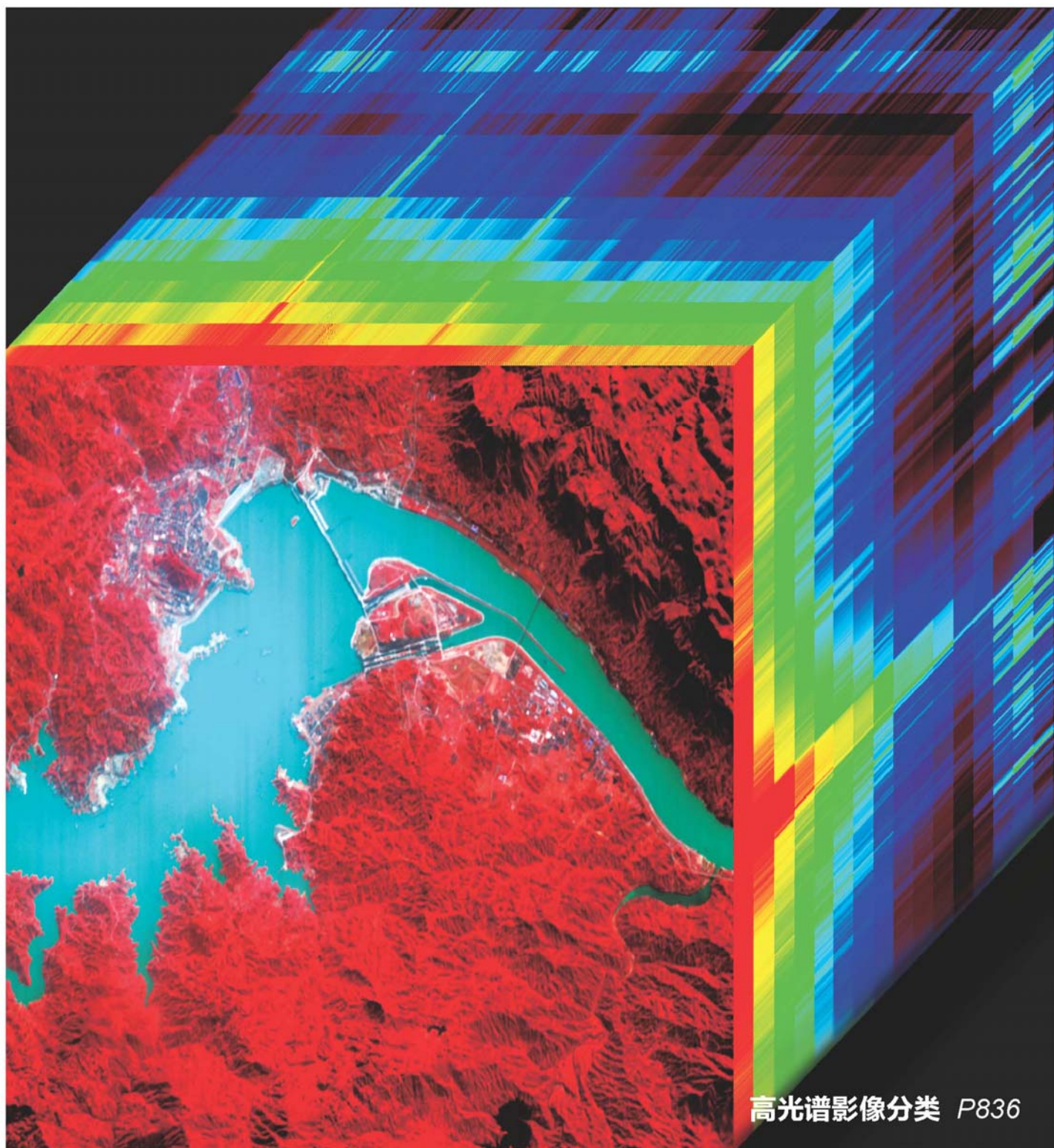
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

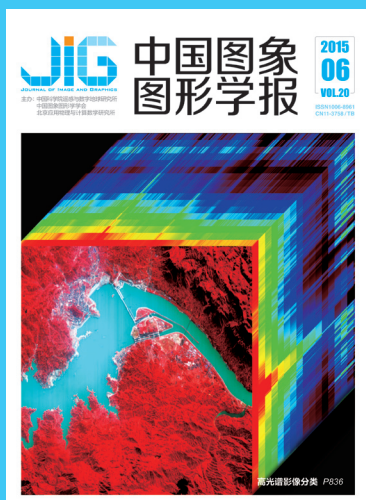
中国图象图形学报

2015
06
VOL.20

ISSN1006-8961
CN11-3758/TB



高光谱影像分类 P836



第20卷第6期（总第230期）

2015年6月16日

中国精品科技期刊
中国影响国际优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院遥感与数字地球研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

编辑出版《中国图象图形学报》编辑出版委员会

邮政信箱 北京9718信箱

邮 编 100101

电子信箱 jig@radi.ac.cn

电 话 010-64807995

网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

（邮政信箱：北京399信箱 邮编：100048）

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing and Digital Earth, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen

Editor, Publisher Editorial and Publishing Board of Journal of

Image and Graphics

P.O.Box 9718, Beijing, P.R.China

Zip code 100101

E-mail jig@radi.ac.cn

Telephone 010-64807995

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

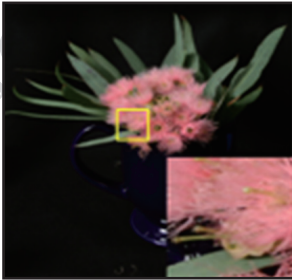
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

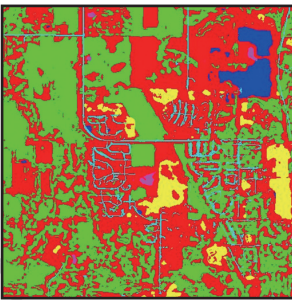
国内定价 50.00元



边缘结构保持的加权BDTV
全光场图像超分辨率重建
(第0733页)



基于超像素的多主体图像交
互分割(第0764页)



谐波分析光谱角制图高光谱
影像分类(第0836页)

图像处理和编码

边缘结构保持的加权BDTV全光场图像超分辨率重建

张旭东, 李梦娜, 张骏, 胡良梅, 王一 0733

基于Arnold映射的分块双层自适应扩散图像加密算法

徐亚, 张绍武 0740

结合梯度信息的特征相似性图像质量评估

苗莹, 易三莉, 贺建峰, 邵党国 0749

图像分析和识别

复杂场景下自适应背景减除算法

邵奇可, 周宇, 李路, 陈庆章 0756

基于超像素的多主体图像交互分割

伯彭波, 袁野, 王宽全 0764

采用金字塔纹理和边缘特征的图像烟雾检测

李红娣, 袁非牛 0772

结合手指检测和HOG特征的分层静态手势识别

刘淑萍, 刘羽, 於俊, 汪增福 0781

形状角计算改进及其在矩形目标识别中的应用

孟伟荣, 李辉, 方黎勇, 白金平 0789

融合图像感知哈希技术的运动目标跟踪

李子印, 朱明凌, 陈柱 0795

利用快速傅里叶变换的目标尺度自适应回归跟踪

张浪, 侯志强, 余旺盛, 许婉君 0805

基于点云属性信息的平面标靶特征提取

刘燕萍, 贾东峰 0815

融合局部思想和协作表达的鲁棒分类

楼宋江, 赵小明, 于海涛, 张石清 0822

计算机图形学

草绘3维人体建模的模板形变方法

李毅, 刘兴川, 孙亭 0828

遥感图像处理

谐波分析光谱角制图高光谱影像分类

杨可明, 刘飞, 孙阳阳, 魏华锋, 史钢强 0836

倾斜航空影像的城区DSM生成

吴军, 程门门, 姚泽鑫, 彭智勇, 李俊, 马峻 0845

封面图片由中国科学院遥感与数字地球研究所高光谱遥感应用技术研究室张霞研究员提供

CONTENTS

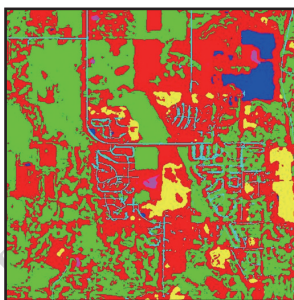
JOURNAL OF IMAGE AND GRAPHICS



Edge preserved light field image super-resolution based on weighted BDTV model (P0733)



Interactive multiphase image segmentation based on superpixels (P0764)



Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping (P0836)

Image Processing and Coding

Edge preserved light field image super-resolution method based on weighted BDTV model

Zhang Xudong, Li Mengna, Zhang Jun, Hu Liangmei, Wang Yi 0733

Encryption algorithm of image blocking and double adaptive diffusion with Arnold mapping

Xu Ya, Zhang Shaowu 0740

Image quality assessment of feature similarity combined with gradient information

Miao Ying, Yi Sanli, He Jianfeng, Shao Dangguo 0749

Image Analysis and Recognition

Adaptive background subtraction approach of Gaussian mixture model

Shao Qike, Zhou Yu, Li Lu, Chen Qingzhang 0756

Interactive multiphase image segmentation based on superpixels

Bo Pengbo, Yuan Ye, Wang Kuanquan 0764

Image based smoke detection using pyramid texture and edge features

Li Hongdi, Yuan Feiniu 0772

Hierarchical static hand gesture recognition by combining finger detection and HOG features

Liu Shuping, Liu Yu, Yu Jun, Wang Zengfu 0781

Improved shape angle computation and its application in detection of rectangular

Meng Weirong, Li Hui, Fang Liyong, Bai Jinping 0789

Object tracking algorithm based on perception hash technology

Li Ziyin, Zhu Mingling, Chen Zhu 0795

Scale adaptive regression tracking method based on fast Fourier transform

Zhang Lang, Hou Zhiqiang, Yu Wangsheng, Xu Wanjuan 0805

Feature extraction of flat target based on point cloud attribute information

Liu Yanping, Jia Dongfeng 0815

Locality constrained collaborative representation classification for robust classification

Lou Songjiang, Zhao Xiaoming, Yu Haitao, Zhang Shiqing 0822

Computer Graphics

Sketch-based method for 3D human modeling using template deformation

Li Yi, Liu Xingchuan, Sun Ting 0828

Remote Sensing Image Processing

Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping

Yang Keming, Liu Fei, Sun Yangyang, Wei Huafeng, Shi Gangqiang 0836

Automatic generation of high-quality urban DSM with airborne oblique images

Wu Jun, Cheng Menmen, Yao Zexin, Peng Zhiyong, Li Jun, Ma Jun 0845

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2015)06-0756-08

论文引用格式: Shao Q K, Zhou Y, Li L, Chen Q Z. Adaptive background subtraction approach of Gaussian mixture model[J]. Journal of Image and Graphics, 2015, 20(6): 0756-0763. [邵奇可, 周宇, 李路, 陈庆章. 复杂场景下自适应背景减除算法[J]. 中国图象图形学报, 2015, 20(6): 0756-0763.] [DOI:10.11834/jig.20150604]

复杂场景下自适应背景减除算法

邵奇可, 周宇, 李路, 陈庆章

浙江工业大学计算机学院, 杭州 310023

摘要: 目的 复杂场景下的背景减除是智能视频监控研究领域研究重点和热点之一。针对混合高斯模型中高斯分布个数固定和参数初始化粗糙问题, 提出一种应用于复杂场景中的基于混合高斯模型的自适应背景减除算法 (AMGBS)。方法 通过灰度值归类算法自适应调整模型的高斯分布个数, 使得背景模型能够适应场景的变化, 并且结合在线 K 均值 (online K-means) 算法和在线期望最大化 (online EM) 算法初始化混合高斯模型参数。结果 针对灰度值统计结果调整高斯分布数, 以及采用优化参数初始化过程, 实验表明, 本文方法的平均查准率和平均查全率比传统的混合高斯算法高出 10% 左右, 比其他改进的混合高斯算法高出 2% 左右。结论 提出一种新的自适应背景减除算法, 针对灰度值统计结果调整高斯分布数, 以及采用优化参数初始化过程。实验结果表明, 该方法对复杂场景有较强的适应能力, 能够有效快速地完成背景减除, 进而实现运动目标的提取。

关键词: 背景减除; 混合高斯模型; online K-means; online EM; 灰度值

Adaptive background subtraction approach of Gaussian mixture model

Shao Qike, Zhou Yu, Li Lu, Chen Qingzhang

College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

Abstract: **Objective** Background subtraction is an important step in object detection for many computer vision applications, including intelligent surveillance and human detection. The purpose of this process is to segment moving objects from complex scenes. Performance mainly depends on the background modeling algorithm; however, the background is a complex environment that usually includes distracting motions. Thus, background subtraction is complicated, and an adaptive method is proposed to address this problem. **Method** The method is based on the Gaussian mixture model. In their approach, each pixel is modeled by a mixture of K Gaussian distributions. An online learning technique is employed to update background models. In their approach, online K-means is applied to initialize the parameters of the Gaussian mixture model. The number of Gaussian distributions cannot be changed in the process of detection. The initialization of the model parameters significantly influences foreground detection, and the fixed Gaussian distribution cannot accommodate the changing background. In this study, we initialize the parameters of the Gaussian mixture model by combining the online K-means and the online expectation-maximization (EM) algorithms. The outcome of online K-means algorithm is the input of the online EM algorithm. The former rapidly generates the parameter values that are close to the reasonable value, whereas the online EM rapidly and accurately astringes the result that is obtained through online K-means to a reasonable range. In addition, this paper also presents a gray-value classification algorithm to adjust the number of Gaussians to adapt to the dynamic envi-

收稿日期: 2014-11-28; 修回日期: 2015-03-05

基金项目: 国家自然科学基金项目 (61104095)

第一作者简介: 邵奇可 (1977—), 男, 副教授, 2009 年于浙江工业大学获控制理论和控制工程博士学位, 主要研究方向为网络控制、无线传感器网络、时滞系统、智能交通。E-mail: sqk@zjut.edu.cn

Supported by: National Natural Science Foundation of China (61104095)

ronment. Recent statistics regarding gray value are obtained for each pixel. Then, this algorithm classifies these gray values into different categories. Finally, this method updates the number of Gaussian distributions on the basis of the number of categories. In this paper, we conduct several experiments with four video datasets to evaluate the proposed background subtraction algorithm. Three of these datasets are standard test videos that are widely used in the video monitoring field. The videos entitled “Waving Trees” and “Bootstrapping” are derived from the Wallflower dataset. The pedestrian video is selected from the Change detection dataset. Another video is captured by a local bus station downtown. Moreover, a quantitative analysis is conducted with the general evaluation criteria of Precision and Recall. **Result** In the conventional Gaussian mixture model, the number of Gaussian distributions is fixed for each pixel and the system cannot adjust to the changing background. Some static regions apply only one or two Gaussian distributions, whereas moving regions need more Gaussian distributions to maintain the model. However, maintaining additional distributions consumes resources, whereas a lack of distributions may result in false detections. **Conclusion** Test results show that the proposed method performs better than the conventional Gaussian mixture model. Experimental findings also suggest that the proposed approach adapts more effectively to complex scenes than those presented in reviewed literature. The background can be segmented effectively and rapidly, and the results for Precision-to-Recall ratio demonstrate the superiority of the method to analogous algorithms. The proposed approach provides a new direction for research on background subtraction. In future work, we will attempt to control learning rate with effective strategies and improve Precision-to-Recall ratio.

Key words: background subtraction; Gaussian mixture model; online K-means; online EM; gray value

0 引言

背景减除(BS)是智能视频监控获取信息的重要步骤,其目的是减除实际场景中因各种原因产生的噪声和非目标背景元素,从而提取出监测区域的运动目标,以作为高一级任务的输入,如分类、跟踪、或行为识别等高级视频分析任务^[1]。BS方法的复杂性主要表现在实际场景通常存在动态变化的影响因素^[2],例如光照强度的变化、树叶的摇晃等,背景减除的关键在于设计一个稳定的、有效的背景模型。

目前,已有许多背景模型算法用于完成背景减除^[3]。Stauffer等人提出了混合高斯模型(GMM)^[4],该模型充分考虑了像素点在时间上的连续性,并且假设像素点的时间序列可以用多个高斯分布来描述。混合高斯背景模型能够有效地适应光照的变化,在复杂场景里具有一定的适应性和稳定性,因此成为应用最为广泛的背景建模算法之一。但在实际的应用中,该方法有一定的局限性,主要体现在简单的模型参数初始化方法精确度不高^[5]和固定的高斯分布个数往往不能适应场景的变化^[6]。对于模型参数初始化,文献[7-8]在GMM的初始化阶段采用了期望最大化(EM)算法对参数进行了估计,该方法可以将高斯分布的均值和方差收敛到合理范围内,提高了该模型对复杂场景的适应性,但其

计算过程比较复杂,计算速度较慢。另外文献[9]以 Online K-means^[10]算法和 EM^[11]算法相结合的形式分阶段初始化参数值,解决了 EM 算法计算速度慢的缺点,但是需要消耗一定的内存空间。针对高斯分布个数问题, Lee 等人提出了一种根据像素值变化的程度,动态地调整高斯分布个数的方法,有利于减少运算时间^[12];文献[13]提出的自适应混合高斯模型(AGMM)根据场景变化,通过极大似然估算出优化的高斯分布个数,但文献[12-13]未对高斯模型的均值、方差和权值初始化进行改进。

除了混合高斯背景模型,Elgammal等人提出了一种非参数背景建模算法——核密度估计(KDE)^[14],该方法利用一段视频帧中某一点的像素值序列,通过高斯内核非参数化地对该位置像素的概率密度进行估计,以此来判断当前像素是否属于背景像素。KDE 算法有利于新训练样本的加入,但其计算复杂度过大,且实时性较差,不适合应用于实际场景中。

针对文献[9]中提出的 GMM 算法的局限性,提出一种应用于复杂场景中的基于混合高斯模型的自适应背景减除算法。算法主要包括以下两个主要部分:1)模型参数初始化。结合 online K-means 和 online EM 算法初始化模型中高斯分布的均值、方差和权值,使得模型参数值能够较快地收敛到合理值范围内,保证可以较为快速、准确地建立背景模型;

2) 高斯分布个数自适应更新。根据一段时间内像素点灰度值的归类结果, 自适应调整高斯分布的个数, 从而使算法可以适应场景中背景的变化。

1 基于 GMM 的自适应背景减除算法

本文算法主要分为 3 个阶段。第 1 阶段为建立背景模型和模型初始化阶段, 首先用 K 个混合高斯分布来描述视频帧中每个点在一段时间内的像素值序列, 其次结合 online K-means 算法和 online EM 算法初始化模型参数; 第 2 阶段是前景检测, 根据与各个高斯分布匹配的情况判别输入的像素为背景还是前景, 同时调整模型参数值; 第 3 阶段为动态调整参数 K , 通过统计归类一段时间内像素点的灰度值, 类别数作为参数 K 的新值。

1.1 建立背景模型和模型初始化

对于视频帧中的一点 $p(x, y)$ 在一段时间内的像素值序列 $\{X_1, \dots, X_t\}$ 可以用 K 个高斯分布来描述。其中 X_t 表示 t 时刻点 p 的像素值。像素值可以是 1 维的、2 维的或者 3 维的, 当前像素值的概率为

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \times \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

式中, K 是该模型中高斯分布的个数, $\omega_{i,t}$ 是 t 时刻第 i 个高斯分布的权重, $\mu_{i,t}$ 是 t 时刻第 i 个高斯分布的均值, $\Sigma_{i,t}$ 是 t 时刻第 i 个高斯分布的协方差矩阵, η 是一个高斯概率密度函数, 其公式为

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu)^T \Sigma^{-1}(X_t - \mu)} \quad (2)$$

K 由摄像机的存储空间和计算能力来决定, 目前普遍为 3~5。为了简便计算, 假定协方差矩阵

$$\Sigma_{i,t} = \sigma_k^2 I \quad (3)$$

这里假定 X_t 各维数据之间相互独立, 并且拥有相同的方差。

Stauffer 等人在文献[4]中使用 online K-means 算法初始化参数 μ 、 ω 和 Σ , 但是文献[8]验证了该方法在复杂的场景中误差较大。EM 算法精确度较高, 但是如同文献[6]中提到的, 对于实时系统其计算量太大。另外, EM 算法对初始化非常敏感, 初始值的设定会对结果产生很大影响^[9]。为了避免不

精确的初始化以及得到一个可靠的结果, 将初始化阶段分为两步: 第 1 步, 在初始的若干帧里使用 online K-means, 得到初步的模型参数值, 避免初始值的设定影响 EM 的运算结果; 第 2 步, 将第 1 步中得到的结果作为 EM 算法的输入, 然后根据接下去的视频帧迭代 EM 算法, 让估计的参数值快速逼近精确值, 直到初始化阶段结束或者收敛到合理范围之内。

1) 使用 online K-means 算法初步初始化参数, 分为以下 4 个步骤:

(1) 猜测每一类别的均值, 标记为 $\{M_1, M_2, \dots, M_k\}$;

(2) 为每一个类别设置一个计数变量, 标记为 $n_1, n_2, \dots, n_k$, 并且初始化为 0;

(3) 每新进一个样本 x , 依次比较 x 与 $\{M_1, M_2, \dots, M_k\}$ 之间的距离, 如果 M_i 更加靠近 x , 则 n_i 的值增加 1, 并且将 M_i 的值替换为 $M_i + (1/n_i) \times (x - M_i)$, 直到样本结束, 否则循环执行此步骤;

(4) 用各类的均值作为 μ , 并计算协方差 Σ , 样本的比例作为 ω 。

2) online EM 算法进行细致迭代, 优化参数值。设待定参数集为 θ , 由此可以将式(1)写为

$$p(x) = \sum_{i=1}^K \omega_i \eta(x | \theta_i) \quad (4)$$

常用的参数估算方法有最大似然估计和矩估计方法, 似然方法比较常用。给定样本 $X = \{x_1, x_2, \dots, x_N\}$, 则似然函数为

$$L(\theta | X) = \prod_{i=1}^N p(x_i | \theta) \quad (5)$$

式中, $\theta = \{\theta_1, \theta_2, \dots, \theta_k\}$, 对式(5)两边取对数, 得

$$\begin{aligned} \log(L(\theta | X)) &= \\ \sum_{i=1}^N \log(p(x_i | \theta)) &= \\ \sum_{i=1}^N \log \left\{ \sum_{k=1}^K \omega_k \eta(x_i | \theta_k) \right\} & \end{aligned} \quad (6)$$

由于在对数函数里面有加和, 没法直接用求导解方程的办法直接求得最大值。为了解决这个问题, 引入一个隐含变量 z , 这是一个 K 维向量, 如果第 k 个分布被选中了, 那么将其第 k 个元素置为 1, 其余的全为 0。同时考虑 x 和 z , 最大似然函数也将变为

$$\prod_{i=1}^N p(x_i, z_i | \theta) \quad (7)$$

注意到 x 和 z 的联合分布 $p(x, z)$ 可以表示为关于 z

的边缘分布和关于 $\mathbf{x}$ 的条件分布, 即

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{z})p(\mathbf{x} | \mathbf{z}) \quad (8)$$

式中 $\mathbf{z}$ 的边缘分布为

$$p(\mathbf{z}_k = 1) = \omega_k \quad (9)$$

这里的 $\mathbf{z}_k$ 指的是某个 $\mathbf{z}_i$ 向量中值为 1 的元素, 由于 $\mathbf{z}$ 是一个 K 维的向量, 因此将 $\mathbf{z}$ 的边缘分布表示成

$$p(\mathbf{z}) = \prod_{k=1}^K \omega_k^{z_k} \quad (10)$$

于是可以写出 $\mathbf{x}$ 的条件分布

$$p(\mathbf{x} | \mathbf{z}) = \prod_{k=1}^K \eta(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k} \quad (11)$$

这样就得到似然函数为

$$\sum_{i=1}^N \sum_{k=1}^K \mathbf{z}_{i,k} \{ \log \omega_k + \log \eta(\mathbf{x}_i | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \} \quad (12)$$

由于 $\mathbf{z}$ 是一个隐藏变量, 无法直接得到。于是 EM 算法假设 $\mathbf{z}$ 的分布由步骤 1) (online K-means 所得结果) 的参数确定, 并求 $\mathbf{z}$ 的期望 (E 步), 然后将似然函数最大化以获得新的参数 (M 步), 这样通过不断迭代执行, 直到达到收敛。

E 步: 通过后验概率 $p(\mathbf{z} | \mathbf{x})$ 来求期望, 即

$$E(\mathbf{z}) = \gamma(\mathbf{z}_{n,k}) = \frac{\omega_k \eta(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \omega_j \eta(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \quad (13)$$

M 步: 通过似然函数求参数, 即

$$\boldsymbol{\mu}_k^{\text{new}} = \frac{1}{N_k} \sum_{n=1}^N \gamma(\mathbf{z}_{n,k}) \mathbf{x}_n \quad (14)$$

$$\boldsymbol{\Sigma}_k^{\text{new}} = \frac{1}{N_k} \times$$

$$\sum_{n=1}^N \gamma(\mathbf{z}_{n,k}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}})^T \quad (15)$$

$$\omega_k^{\text{new}} = N_k / N \quad (16)$$

$$N_k = \sum_{n=1}^N \gamma(\mathbf{z}_{n,k}) \quad (17)$$

式(14)~(16)用于处理成批的数据, 在实时监控系统中需要将式(14)~(16)改写为

$$\boldsymbol{\mu}_k^{\text{new}} = \boldsymbol{\mu}_k + \frac{\gamma(\mathbf{z}_{n,k})}{N_k} \cdot (\mathbf{x}_n - \boldsymbol{\mu}_k) \quad (18)$$

$$\boldsymbol{\Sigma}_k^{\text{new}} = \boldsymbol{\Sigma}_k + \frac{\gamma(\mathbf{z}_{n,k})}{N_k} \cdot$$

$$((\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{\text{new}})^T - \boldsymbol{\Sigma}_k) \quad (19)$$

$$\omega_k^{\text{new}} = \omega_k + \frac{1}{N} \cdot (\gamma(\mathbf{z}_{n,k}) - \omega_k) \quad (20)$$

1.2 前景检测

在完成模型初始化之后, 开始进行运动目标的

检测与提取。首先, 需要对 K 个高斯分布以 $r_j = \omega_j / \sigma_j$ 值的大小进行一个排序, 目的是确定哪几个高斯分布属于背景模型。考虑到在一段时间内, 某一位置的像素点展现为背景的时间比展现为运动目标的时间要长, 并且作为背景时该点的像素值将会更加稳定、连续, 因此背景像素具有更大权重和更小的方差。选取前 B 个分布构成背景模型, B 的取值为

$$B = \arg \min_b \left(\sum_{i=1}^b \omega_{i,t} > T \right) \quad (21)$$

式中, B 等于满足括号内不等式的最小的 b 值。 T 的选取会影响背景模型高斯分布的个数, 若 T 值较小, 则有可能变成单峰模型, 即单高斯背景模型; 如果 T 较大, 则会增加背景模型分布数, 可以适应复杂变化的场景。

每一个新到达的像素值 $\mathbf{X}_{t+1}$, 都会与 K 个高斯分布一一比较, 直到找到匹配的分布为止。匹配的条件是

$$\sqrt{(\mathbf{X}_{t+1} - \boldsymbol{\mu}_{i,t})^T \cdot \boldsymbol{\Sigma}_{i,t}^{-1} \cdot (\mathbf{X}_{t+1} - \boldsymbol{\mu}_{i,t})} < k \sigma_{i,t} \quad (22)$$

式中, k 普遍认为取值为 2.5 时效果最佳。当匹配过程结束, 将会出现两种情况: 1) 匹配到某个分布; 2) 没有匹配到任何分布。针对第一种情况, 如果匹配到了某个分布, 那么要更新该分布的参数, 即

$$\omega_{i,t+1} = (1 - \alpha) \omega_{i,t} + \alpha \quad (23)$$

式中, α ($0 < \alpha < 1$) 是权重学习速率, 是为了根据匹配情况动态调整各个高斯分布的权重值。

$$\boldsymbol{\mu}_{i,t+1} = (1 - \rho) \boldsymbol{\mu}_{i,t} + \rho \cdot \mathbf{X}_{t+1} \quad (24)$$

$$\sigma_{i,t+1}^2 = (1 - \rho) \sigma_{i,t}^2 +$$

$$\rho (\mathbf{X}_{t+1} - \boldsymbol{\mu}_{i,t+1}) \cdot (\mathbf{X}_{t+1} - \boldsymbol{\mu}_{i,t+1})^T \quad (25)$$

式中, $\rho = \alpha \cdot \eta(\mathbf{X}_{t+1} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ 表示参数更新速率。同样在第一种情况中, 其他未匹配的分布, 其 $\boldsymbol{\mu}$ 和 σ 不变, 其权值更新为

$$\omega_{j,t+1} = (1 - \alpha) \omega_{j,t} \quad (26)$$

针对第 2 种情况, 没有匹配到任何分布, 则用一个新分布来替代概率最小的那个分布, 新分布的参数设置为

$$\omega_{k,t+1} = w_{\text{low}}$$

$$\boldsymbol{\mu}_{k,t+1} = \mathbf{X}_{t+1} \quad (27)$$

$$\sigma_{k,t+1}^2 = \sigma_{\text{high}}^2$$

式中, w_{low} 表示一个较低的权重值, σ_{high} 表示一个较大的方差。

1.3 动态调整参数 K

固定的高斯分布数 K 一般不能适应不断变换的背景。场景中不存在前景目标的区域,其内像素值在一段时间就保持稳定,这些像素点只需 $1 \sim 2$ 个的高斯分布就可以描述该位置像素值的序列变化,维持多余的高斯分布既不会提高检测效果且消耗额外的计算资源。反之,场景中那些存在目标的区域,则需要更多的高斯分布数来适应场景变换。因此引入灰度值归类方法来动态调整参数 K ,该方法可根据一定时间内场景的实际变化,有差别地对每个像素点设定高斯分布数。视频帧中的某个点,在一段时间内,它会呈现出不同的模态。在同一模态里,像素的灰度值会保持在一定的浮动空间里。这里对若干帧的灰度值进行统计^[15],把这些灰度值进行分类。算法可以分为 3 个步骤:

1) 划分灰度平稳区间。从视频序列中等间隔的抽取 $N+1$ 帧图像,标记为 $\{G_0, G_1, \dots, G_N\}$ 。

$G_i(x, y)$ 表示第 i 帧像素点 (x, y) 处的灰度值,则有

$$a_i(x, y) = \begin{cases} 1 & |G_i(x, y) - G_{i-1}(x, y)| > \xi \\ 0 & |G_i(x, y) - G_{i-1}(x, y)| \leq \xi \end{cases} \quad i = 1, \dots, N \quad (28)$$

式中, ξ 为判断像素点 (x, y) 处的灰度值是否发生变化的阈值,是算法中唯一需要调整的参数。实验表明,当图像为 256 级灰度,则 $\xi = 10 \sim 25$ 级灰度差时,效果良好。另外如果 $a_i(x, y) = 1$,则 $G_i(x, y)$ 与 $G_{i-1}(x, y)$ 的灰度值不一致,不属于同一个灰度区间;如果 $a_i(x, y) = 0$,则 $G_i(x, y)$ 与 $G_{i-1}(x, y)$ 的灰度值一致,属于同一区间。

2) 计算各平稳区间的平均灰度值。假设总共有 p 个灰度平稳区间,各灰度区间所含像素数标记为 $\{m_1, m_2, \dots, m_p\}$,则第 s 个区间的平均灰度值为

$$\bar{l}_s(x, y) = \frac{\sum_{i=b}^{b+m_s} G_i(x, y)}{m_s} \quad (29)$$

当 $s = 1$ 时, $b = 0$, 当 $s = 2, \dots, p$ 时, $b = m_1 + \dots + m_{s-1}$, 由此得到了每个区间的平均灰度值。

3) 将平均灰度值相近的归为一类。接着根据 p 个区间的平均灰度值,合并灰度值一致的全部区间,若

$$|\bar{l}_i(x, y) - \bar{l}_j(x, y)| < \xi \quad (i = 1, \dots, p; j = 1, \dots, p; i \neq j) \quad (30)$$

则第 i 个区间和第 j 个区间为一致区间,式中 ξ 同式(28)。假设归类后还有 q 个区间, q 就是参数 K 的更新值。

复杂场景下基于混合高斯模型的自适应背景减除算法会对每一帧图像的每个像素点执行以上步骤,完成图像的背景减除,直到所有视频帧都处理完成。

2 实验结果与应用

为了检测本文算法的有效性,对 3 个标准视频和实际视频进行了测试。标准视频分别来自 Wallflower^[16]测试集中的 WT(waving trees)和 Bootstrapping,以及来自 ChangeDetection^[17]的 Pedestrians,因为这些场景和公交车站较为相似,实际视频(Station)拍摄于市区的公交车站。本实验仅对前景图像进行比较,未对图像进行预处理,也不包括阴影去除。实验在 4 个视频上测试了 5 种背景建模方法,分别是经典的混合高斯背景模型(GMM)^[4]、IABMM(improved adaptive background mixture model)^[18]、改进的混合高斯模型 AGMM^[13]、非参数内核密度估计(KDE)^[14]和本文算法(AMGBS)。其中 GMM 模型只应用了 online K-means 算法,而 IABMM 模型则只应用了 online EM 算法。运动目标前景效果如图 1。

图 1 展示了 5 类方法在 4 个测试集进行应用的结果,针对每一个视频集,给出了测试图像、标准前景图像和 5 类方法检测的前景效果,其中,最后一列为市区公交车站拍摄的视频,由于没有标准前景图像,因此图 1 中第 2 行最后一列表现为空白。从图 1 中可以看出,AMGBS 算法同 GMM 方法相比,该方法在 WT 和 Pedestrians 测试集上的检测效果要远远好于 GMM 算法,GMM 更容易产生空洞现象且抗干扰能力较差,说明本文方法的改进是合理的、有效的,Station 测试集上也体现了 AMGBS 的优点。IABMM 的检测效果虽然比 GMM 有了明显的提升,但是与 AMGBS 相比,依然存在大量的空洞。AMGBS 与 AGMM 相比,在测试集 WT、Bootstrapping 和 Pedestrians 上检测结果近似,但是在 Station 中,由于实际场景中光照、阴影等扰动因素的增加,AGMM 算法误检的情况大大增加,可见 AMGBS 方发更具适应性。类似的,KDE 与 AMGBS 在标准测试集上



图1 前景比较图

Fig. 1 Comparison of foreground in different scenes

的检测效果相差不大,但是在实际视频中,本文方法较之KDE更具稳定性。

为了定量地分析实验结果,实验结果采用通用的评估指标查准率(Pre)和查全率(Re)进行比较。其中查全率是检测的正确数与手工标注的总数的比值,查准率是检测的正确数和总的检测数的比值。两个指标的值同时到达较高的值时说明背景减除的精确度高。5种方法的查准率和查全率如表1所示。从表1中可以看到,本文算法的平均查准率和平均查全率都要高于其他4类方法。平均查准率超过AGMM、KDE和IABMM约3%,高于GMM约12%,平均查全率比KDE与AGMM高约2%,优于

GMM和IABMM约14%。在场景Waving Trees中,AMGBS算法的效果远远好于GMM和IABMM,略优于AGMM与KDE。在场景Bootstrapping中,本文算法与AGMM、KDE效果相近,略优于GMM和IABMM。而在Pedestrians中,4种方法之间并没有较大的差距。从算法的执行速度来看,AMGBS算法平均可达到2.5 ms/帧,快于GMM(3 ms/帧)算法和KDE算法(4 ms/帧),慢于AGMM算法,与IABMM方法相当。与GMM和IABMM相比,本文算法优化了参数初始化方法,提高了检测效果。另外本文算法动态调整高斯分布个数,提升了计算速度。

表 1 查准率和查全率比较
Table 1 Precision and Recall for different methods

		算法					/%
评价指标		AGMM	GMM	IABMM	KDE	AMGBS	
WT	Pre	92.60	64.69	91.17	91.44	96.64	
	Re	95.72	67.00	84.64	95.32	98.27	
Bootstrapping	Pre	92.05	86.88	92.47	87.01	91.89	
	Re	89.16	89.02	77.67	92.28	88.74	
Pedestrians	Pre	88.94	88.59	87.30	88.79	87.76	
	Re	82.96	82.79	76.77	82.76	86.46	
平均值	Pre	89.36	81.89	90.31	89.08	92.10	
	Re	89.18	79.71	79.69	90.12	91.16	

由此可见,虽然其他 4 种方法在某一特定的场景里比 AMGBS 拥有更好的检测效果,但是在其他场景中却不能达到本文算法的效果。因此,本文算法更加稳定,更加适应复杂的场景化。

3 结 论

提出了一种复杂场景下基于混合高斯模型的自适应背景减除算法。该算法结合 online K-means 和 online EM 算法初始化模型参数,提高了参数初始化精确度,并且通过像素灰度归类方法对高斯分布的个数进行动态调整,适应动态的场景变化。在公共测试集以及实际应用环境中进行测试,实验结果表明,对比其他背景减除算法,本文算法拥有较好的查准率与查全率,并且在各类复杂的场景中都有较好的适应能力。为背景减除算法提供了新的研究方向。

参考文献 (References)

[1] Apolinário L, Armesto N, Cunqueiro L. An analysis of the influence of background subtraction and quenching on jet observables in heavy-ion collisions [J]. Journal of High Energy Physics, 2013, 22: 1-33.

[2] Shen Y, Hu W, Liu J, et al. Efficient background subtraction for real-time tracking in embedded camera networks [C] // Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems. Toronto, ON, Canada: ACM, 2012: 295-308.

[3] Bouwmans T. Recent advanced statistical background modeling for foreground detection-a systematic survey [J]. Recent Patents on Computer Science, 2011, 4(3): 147-176.

[4] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE, 1999, 22(3): 747-757.

[5] Shah M, Deng J, Woodford B. Illumination invariant background model using mixture of gaussians and SURF features [C] // Computer Vision-ACCV 2012 Workshops. Berlin Heidelberg: Springer, 2013: 308-314.

[6] Liu X, Liu H, Qiang Z P, et al. Adaptive background modeling based on mixture gaussian model and frame subtraction [J]. Journal of Image and Graphics, 2008, 13(4): 729-734. [刘鑫, 刘辉, 强振平, 等. 混合高斯模型和帧间差分相融合的自适应背景模型 [J]. 中国图象图形学报, 2008, 13(4): 729-734.] [DOI:10.11834/jig.20080422]

[7] Liu Z, Huang K, Tan T. Foreground object detection using top-down information based on em framework [J]. IEEE Transactions on Image Processing, 2012, 21(9): 4204-4217.

[8] Li Y, Li L. A novel split and merge EM algorithm for gaussian mixture model [C] // Proceedings of the 5th International Conference on Natural Computation. Tianjin: IEEE, 2009, 6: 479-483.

[9] Singh A, Jaikumar P, Mitra S K, et al. Detection and tracking of objects in low contrast conditions [C] // Proceedings of IEEE National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics. Gandhinagar, India: IEEE, 2008: 98-103.

[10] Alpaydin E. Introduction to Machine Learning [M]. Boston: MIT press, 2004: 278-280.

[11] Moon T K. The expectation-maximization algorithm [J]. Signal

- processing magazine, IEEE, 1996, 13(6): 47-60.
- [12] Nonaka Y, Shimada A, Nagahara H, et al. Evaluation report of integrated background modeling based on spatio-temporal features [C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI: IEEE, 2012: 9-14.
- [13] Zivkovic Z, van der Heijden F. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. Pattern recognition letters, 2006, 27(7): 773-780.
- [14] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction [M]//Computer Vision—ECCV 2000. Dublin, Ireland:Springer, 2000: 751-767.
- [15] Hou Z Q, Han C Z. A background reconstruction algorithm based on pixel intensity classification[J]. Journal of Software, 2005, 16(9):1568-1576. [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.]
- [16] Toyama K, Krumm J, Brumitt B, et al. Wallflower: principles and practice of background maintenance [C]// Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra: IEEE, 1999, 1: 255-261.
- [17] Goyette N, Jodoin P M, Porikli F, et al. Changedetection-net: a new change detection benchmark dataset[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI: IEEE, 2012: 1-8.
- [18] KaewTraKulPong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection [M]//Video-Based Surveillance Systems. Berlin: Springer, 2002: 135-144.