

基于生物视觉特性的背景减除算法

钱生¹, 张陈斌¹, 陈宗海¹, 王智灵²

(1. 中国科学技术大学自动化系, 安徽合肥 230027; 2. 中国科学院合肥物质科学研究院先进制造技术研究所, 安徽合肥 230031)

摘要:针对构建鲁棒的背景模型和背景模型更新问题, 结合 ViBe 算法提出了一种基于生物视觉特性的背景减除算法。首先, 借鉴蛙眼视觉外部特性中的近视特性, 阐述了“区域模糊化理解”预处理操作的含义及其实现; 其次, 从人类对颜色的心理认知特性出发, 给出了一种 LUV 颜色空间中基于韦伯定律的颜色差异度量准则; 最后, 从背景建模、前景检测以及背景模型更新三个主要方面, 介绍了算法的具体实现。实验结果表明, 该算法能够提高运动目标检测的精度。

关键词:生物视觉; 区域模糊化理解; 韦伯定律; 背景减除; 运动目标检测

中图分类号: TP391 **文献标识码:** A doi:10.3969/j.issn.0253-2778.2014.04.003

引用格式: Qian Sheng, Zhang Chenbin, Chen Zonghai, et al. A background subtraction algorithm based on biological vision characteristics[J]. Journal of University of Science and Technology of China, 2014, 44(4):270-277.

钱生, 张陈斌, 陈宗海, 等. 基于生物视觉特性的背景减除算法[J]. 中国科学技术大学学报, 2014, 44(4):270-277.

A background subtraction algorithm based on biological vision characteristics

QIAN Sheng¹, ZHANG Chenbin¹, CHEN Zonghai¹, WANG Zhiling²

(1. Department of Automation, University of Science and Technology of China, Hefei 230027, China;

2. Institute of Advanced Manufacturing Technology, Hefei Institute of Physical Science, Chinese Academy of Science, Hefei 230031)

Abstract: For the problem of how to build a robust background model and update the background model, a background subtraction algorithm based on biological vision characteristics combined with the ViBe algorithm was proposed. Firstly, utilizing the extrinsic nearsightedness characteristic of the frog's visual system, the meaning of a pretreatment method called "region fuzzy" and its implementation were described. Then, considering the characteristics of color cognition by humans, a measurement criterion for color difference based on Weber's law in the LUV color space was given. Finally, specific implementations of the algorithm were introduced from three main aspects: Background modeling, foreground detection and background model updating. Experimental results show that this algorithm can improve the accuracy of moving object detection.

Key words: biological vision; region fuzzily understanding; Weber's law; background subtraction; moving object detection

收稿日期: 2013-08-06; 修回日期: 2013-09-04

基金项目: 国家自然科学基金(61005091)资助。

作者简介: 钱生, 男, 1990年生, 硕士生。研究方向: 计算机视觉与模式识别。E-mail: qsheng@mail.ustc.edu.cn

通讯作者: 陈宗海, 教授。E-mail: chenzh@ustc.edu.cn

0 引言

运动目标检测是视觉分析的核心技术,是后续各种高级处理,如物体分类、行为识别、事件检测、行为分析等视频处理和应用理解的基础.精确的运动目标检测能极大地提高其后续处理的性能.运动目标检测技术可以分为帧差法、光流法以及背景减除法.其中,背景减除法应用最为普遍,其基本思想是通过图像序列训练一个背景模型,然后通过比较当前图像和背景模型间的差异来检测运动目标.其关键之处在于构建鲁棒的背景模型和实时更新背景模型以适应复杂动态背景.背景建模方法基本上可分为参数化方法和非参数化方法两类.

参数化的背景建模通常假设像素点的潜在概率密度函数是高斯或者高斯混合函数.Wren等^[1]提出了在YUV彩色空间中使用单个高斯分布对背景中的像素进行建模,来提高抵抗背景干扰的能力;然而,这种方法不能处理多模态的复杂动态背景.Stauffer等^[2]提出的混合高斯模型是处理多模态的动态背景比较实用的方法,优点是可以在固定数量的模式中选择当前最稳定的模式作为背景,并实时在线更新各个模式的分布参数.Jepson等^[3]在此基础上提出使用增量EM算法来自适应地更新高斯分布的参数,从而提高背景模型的精度.区别于Stauffer等提出的方法中需要事先固定模式的数量,Zivkovic^[4]提出了一种利用递归方程自动选择模式数量的方法.一般来说,基于高斯模型的背景建模方法具有便于计算的特点,然而其基本出发点——像素时域分布符合高斯分布的假设,在实际应用中却不是总能得到满足.

与参数化方法不同,非参数化的背景建模不对背景模型的参数作出任何假设,而是通过保存像素一段时间内的统计样本建立背景模型,非常适用于概率密度分布未知的情况.Elgammal等^[5]使用颜色特征,利用核密度估计的方法进行背景建模.Sheikh等^[6]使用由颜色和位置信息构成的多维特征,利用核密度估计的方法同时对背景和前景进行建模.基于核密度估计的算法计算复杂度高,不适合应用于实时运动检测系统,并且在动态背景和光线突变条件下的鲁棒性不高.Kim等^[7]提出了基于码本模型的背景减除算法.该算法根据颜色失真程度及其亮度范围,将像素点的连续采样值用多个码本表示,能够适应像素剧烈变化或者移动物体,甚至更为复杂

的背景,由于只使用亮度特征,对阴影、光线变化等处理效果不好.Barnich等^[8]提出的ViBe算法使用每个像素的多个样本来描述该像素的背景模型,能够适应复杂的背景.该算法提出了背景模型的随机更新和背景信息的空间传播,但是没有区分前景区域和背景区域中背景信息空间传播的差异.Maddalena等^[9-10]提出的SOBS算法在一个神经网络框架下学习背景模型,对每个像素的背景模型更新受到其邻域像素分类标记的影响.该算法对光照有一定的鲁棒性,但背景模型比较复杂,计算复杂度高.

有些背景减除方法着重于背景建模所采用的视觉特征,而不是背景建模所采用的方法.在所采用的特征中,由于颜色和强度特征的显著性,它们的应用最为普遍.Li等^[11]提出使用颜色共生特征来描述运动的背景像素.由于纹理对局部微弱扰动和光照变化的相对不变性,Heikkilä等^[12]和Liao等^[13]提出了局部多值模式来描述像素的局部纹理特征.为了表征复杂动态的场景,多特征融合的方法在文献[14-17]中也有研究,其中文献[14]还提出了一种通过Real Boost学习算法实现显著性特征的选择方法.

为了提高运动目标检测的精度,本文从构建鲁棒的背景模型和背景模型更新两个方面出发,结合ViBe算法提出了一种基于生物视觉特性的背景减除算法.

1 区域模糊化理解

蛙眼视觉系统因其对运动目标的特殊敏感性而闻名——青蛙看不见(至少是不关注)周围世界中的静态部分细节.Lettvian等^[18]和Ingle^[19]在其论文中研究了蛙眼视网膜的神经结构和生理特性,并解释了蛙眼的这种特殊的视觉行为.王智灵等^[20]介绍了蛙眼视觉两种突出的外部特性:近视和分辨目标大小及其运动类型.目前大多数的研究集中于运用计算机信息处理技术对蛙眼视觉行为机制的模拟.

对蛙眼而言,近视的一个优势就是使得背景模糊而前景相对清晰.本文借鉴该特性提出一种“区域模糊化理解”的预处理操作,来处理场景中的背景干扰运动.与蛙眼对场景的理解方式一样,该方法的基本思路是:通过滤波、保持、分割、合并的方式,使图像场景中的干扰因素和运动目标的内部纹理逐渐模糊,甚至消失,从而保留期望的目标特征和目标外轮廓.

本文应用Mean Shift^[21]滤波过程实现该预处理操作.Mean Shift算法是一种基于核密度估计的自适应梯度算法,它从任意一点出发,沿着核密度的梯度上升方

向,以自适应的步长进行搜索,最终收敛于核密度估计函数的局部极大值处. 设输入图像为 $I(x, y)$, 以图像空间和值空间构成联合特征空间 $Z(x, y) = ((x, y), I(x, y)) \in \mathbb{R}^3$. 对于图像中的每一个像素, 记联合特征矢量为 $\mathbf{z} = (\mathbf{z}^s, \mathbf{z}^c)$. 其中, $\mathbf{z}^s = (x, y)$, $\mathbf{z}^c = I(x, y)$, 其概率密度函数的核密度估计为

$$\hat{f}(\mathbf{z}) = \frac{1}{n} |\mathbf{H}|^{-1/2} \sum_{i=1}^n k \left(\left\| \frac{\mathbf{z} - \mathbf{z}_i}{|\mathbf{H}|^{-1/2}} \right\|^p \right) \quad (1)$$

式中, $k(\cdot)$ 是高斯核函数, \mathbf{H} 为带宽矩阵, n 为迭代区域的样本数目. 采用空间域带宽与值域带宽的解耦方式^[21], 概率密度函数的估计可表示为

$$\hat{f}(\mathbf{z}) = \frac{C}{h_s^2 h_c^p} \sum_{i=1}^n k \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right) \cdot k \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right) \quad (2)$$

式中, h_s, h_c 分别为解耦后的空间域带宽和值域带宽; $p=1$ 表示输入的是单通道灰度图像, $p=3$ 表示输入的是三通道彩色图像, C 为归一化常数. 在联合特征空间内, h_s, h_c 不仅决定 Mean Shift 迭代的样本区域, 而且影响区域模糊化理解的效果. 进一步导出概率密度函数的梯度值估计:

$$\nabla \hat{f}(\mathbf{z}) = \begin{pmatrix} \nabla_{\mathbf{z}^s} \hat{f}(\mathbf{z}) \\ \nabla_{\mathbf{z}^c} \hat{f}(\mathbf{z}) \end{pmatrix} = \frac{2C}{h_s^2 h_c^p} \cdot \begin{pmatrix} \frac{1}{h_s^2} \sum_{i=1}^n (\mathbf{z}^s - \mathbf{z}_i^s) \frac{\partial}{\partial \mathbf{z}^s} k \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right) k \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right) \\ \frac{1}{h_c^{2p}} \sum_{i=1}^n (\mathbf{z}^c - \mathbf{z}_i^c) k \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right) \frac{\partial}{\partial \mathbf{z}^c} k \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right) \end{pmatrix} \quad (3)$$

令

$$g(\mathbf{z}) = \begin{pmatrix} -\frac{\partial}{\partial \mathbf{z}^s} k(\mathbf{z}^s) \cdot k(\mathbf{z}^c) \\ -k(\mathbf{z}^s) \cdot \frac{\partial}{\partial \mathbf{z}^c} k(\mathbf{z}^c) \end{pmatrix} \quad (4)$$

由式(3), (4)可得

$$\nabla \hat{f}(\mathbf{z}) = \frac{2C}{h_s^2 h_c^p} \cdot \begin{pmatrix} \sum_{i=1}^n \frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right)}{h_s^2} \left(\frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right) \cdot \mathbf{z}_i^s}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^s - \mathbf{z}_i^s}{h_s} \right\|^p \right)} - \mathbf{z}^s \right) \\ \sum_{i=1}^n \frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right)}{h_c^{2p}} \left(\frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right) \cdot \mathbf{z}_i^c}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z}^c - \mathbf{z}_i^c}{h_c} \right\|^p \right)} - \mathbf{z}^c \right) \end{pmatrix} \quad (5)$$

在上式中, 令

$$\nabla m(\mathbf{z}) = \frac{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z} - \mathbf{z}_i}{h} \right\|^p \right) \cdot \mathbf{z}_i}{\sum_{i=1}^n g \left(\left\| \frac{\mathbf{z} - \mathbf{z}_i}{h} \right\|^p \right)} - \mathbf{z} \quad (6)$$

$\nabla m(\mathbf{z})$ 即为核函数在该像素处的 Mean Shift 增量. 通过迭代计算 $\nabla m(\mathbf{z})$, 特征矢量 \mathbf{z} 将收敛到其模式中心 \mathbf{z}_c .

2 颜色差异度量准则

ViBe 算法使用像素灰度值为特征, 仅以欧氏距离和一个固定的阈值来衡量当前图像和背景模型的匹配情况. 该方法虽然在很多视频序列中取得了比较好的效果, 但是仍可以得到进一步的改进. 考虑到运动目标与背景不仅有亮度的差异, 而且也有色度的差异, 可以根据人眼对亮度和色度不同的敏感特性, 利用 LUV 颜色空间中基于韦伯定律的颜色差异度量准则, 实现当前图像和背景模型间的差异比较.

德国生理学家韦伯发现这样一个现象, 想要引起差别感觉必须使刺激差别量和原始刺激量的比值达到一个阈值, 公式表示为

$$\Delta I / I = C \quad (7)$$

式中, ΔI 是刺激差别量, I 是原始刺激量, C 为常数, 这就是韦伯定律^[22]. 它主要是用来描述物理量和心理量两者的关系. 当一个刺激差别量小于原始刺激量和这个恒定比例的乘积时, 一般就被当作背景噪声而不是一个有效的信号.

不同于 RGB 颜色空间, LUV 颜色空间中的均匀变化对应于感知颜色中的均匀变化. 其中, L 代表亮度信息, U 和 V 代表色度信息^[23]. 为了模拟人眼对亮度和色度的不同敏感特性, 可以对亮度差异和色度差异采用不同的度量准则, 并基于韦伯定律给出如下式所示的颜色差异度量准则:

$$\left. \begin{aligned} \Delta E_L &= \Delta L / L \\ \Delta E_{UV} &= \sqrt{(\Delta U)^2 + (\Delta V)^2} \end{aligned} \right\} \quad (8)$$

式中, L 为亮度的原始刺激量, $\Delta L, \Delta U$ 和 ΔV 分别为各个分量的刺激差别量, ΔE_L 和 ΔE_{UV} 分别为相对亮度差异值和色度差异值.

3 背景减除算法设计

3.1 背景建模

背景减除算法的目的在于产生一个区分前景和

背景的二进制分类掩码. 这个分类过程需要将输入图像以某种方式与一个背景模型进行比较. 每个像素的背景模型 $B(x, y)$ 包含了 N 个采样值, 可表示为

$$B(x, y) = \{b_1(x, y), \dots, b_k(x, y) \dots, b_N(x, y)\} \quad (9)$$

式中, $b_k(x, y)$ 为图像中位于 (x, y) 处的像素 $p(x, y)$ 的第 k 个背景采样值, 并记 $s(x, y)$ 为像素 $p(x, y)$ 的当前采样值.

3.2 前景检测

根据颜色差异度量准则, 如果当前像素 $p(x, y)$ 的采样值 $s(x, y)$ 与背景模型 $B(x, y)$ 中采样值的匹配数目低于一定的匹配数目阈值时, 则该像素被分类为前景, 否则为背景, 即

$$F(x, y) = \begin{cases} 1, & \#\{b_k \in B(x, y) \mid m(s, b_k) = 1\} < \#_{\min} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$$m(s, b_k) = \begin{cases} 1, & \Delta E_L < Th_L, \Delta E_{UV} < Th_{UV} \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

式中, $\Delta E_L, \Delta E_{UV}$ 分别为 $s(x, y)$ 与 $b_k(x, y)$ 的相对亮度差异值和色度差异值; Th_L, Th_{UV} 分别为预设的相对亮度差异阈值和色度差异阈值; $m(s, b_k) = 1$ 表示 $s(x, y)$ 与 $b_k(x, y)$ 匹配, 反之则不匹配; $\#\{\cdot\}$ 表示 $s(x, y)$ 与背景模型 $B(x, y)$ 中样本匹配的数目, $\#_{\min}$ 为预设的匹配数目阈值; $F(x, y)$ 为初始的分类掩码, $F(x, y) = 1$ 表示像素被分类为前景, 反之则为背景.

在获得初始分类掩码后, 就可以清除一些零星分布的由背景中运动物体的干扰与随机噪声形成的伪前景, 并填充前景目标内的细小空洞, 从而通过这一系列的操作来提高前景和背景分类的准确性. 本文通过对 $F(x, y)$ 进行形态学开闭操作得到最终的分类掩码 $Class(x, y)$, 具体步骤为: 在 $F(x, y)$ 上, 消除面积不大于 A_s 的前景区域, 并填补面积大于 A_s 且不大于 A_m 的前景区域, 产生 $Class(x, y)$. 其中, A_s, A_m 是预设的固定面积阈值, 算法中取 $A_s = 9, A_m = 25$.

3.3 背景模型更新

(I) 分类掩码与更新掩码的区别

在背景减除算法中, 背景模型更新对提高模型的准确性来说至关重要. 大多数背景减除算法直接

使用分类掩码来决定像素当前采样值是否被用来更新背景模型, 即将分类掩码直接作为背景模型更新掩码. 由于分类掩码中存在前景和背景误分类的情况, 这就会影响背景模型更新的一致性和准确性. 在获得最终分类掩码后, 就可以通过进一步抑制内嵌在前景目标内的伪背景像素的背景模型更新, 并填充前景目标内较大的空洞等操作, 提高背景模型更新的一致性和准确性. 这里通过对 $Class(x, y)$ 进行形态学开闭操作得到更新掩码 $Update(x, y)$, 具体步骤为: 在 $Class(x, y)$ 上, 填补面积大于 A_m 且不大于 A_l 的前景区域, 产生 $Update(x, y)$. 其中, A_l 是预设的固定面积阈值, 算法中取 $A_l = 50$.

(II) 模型更新方法

通过上述分析可知, 需要进行背景模型更新的像素由更新掩码 $Update(x, y)$ 确定. 本文借鉴 ViBe 算法的背景模型更新方法, 按照如下步骤对背景模型进行更新:

Step 1 自身背景模型的更新. 算法以 $\frac{1}{N}$ 的概率从 $p(x, y)$ 的背景模型中随机选取一个样本, 使用该像素的当前采样值替换选取的样本.

Step 2 背景信息的空间传播. 算法以 $1/8$ 的概率从 $p(x, y)$ 对应的 8 邻域中随机选取一个像素, 以 Step 1 中的方法更新选取像素的背景模型.

(III) 目标区域背景更新的抑制

随着图像的不断更新, ViBe 算法提出的背景模型更新方法有助于抑制目标由静止状态转变为运动状态时所产生的鬼影; 并且在目标由运动状态转变为静止状态时, 能够快速地将目标成为新背景. 在运动目标检测中, 前一种情况下鬼影应该快速被融入背景中, 而后一情况下目标应该要维持一段时间. 然而 ViBe 算法未能区分这两种不同的情况, 实验表明, 它对抑制鬼影十分有效, 但对运动目标的维持没有效果. 针对这一问题, 本文借助基于区域分类一致性的目标维持方法, 来抑制目标区域内像素的背景模型更新:

在 $Class(x, y)$ 的基础上获取目标轮廓, 对于目标轮廓外的背景像素, 抑制其对目标内部像素的背景信息的空间传播. 考虑到区域分类的一致性, 通过统计目标轮廓内部每个前景像素邻域内前景像素的数目, 以此决定该像素维持为前景的时间, 即

$$Keep(x, y) = K \cdot \sum_{-R < i < R} \sum_{-R < j < R} Class(x+i, y+j) \quad (12)$$

式中, R 表示像素 $p(x, y)$ 的邻域半径; K 为维持时间的常量比例系数; $\text{Keep}(x, y)$ 为 $p(x, y)$ 维持为前景的时间, 与其邻域内前景像素的数目成正比.

3.4 具体步骤

本文算法主要由模型初始化、前景检测和模型更新等步骤组成:

(I) 模型初始化

选取初始的连续 N 帧输入图像, 在 RGB 颜色空间中对这些图像进行“区域模糊化理解”预处理, 并将预处理的结果图像转为 LUV 颜色空间图像, 以此建立每个像素的背景模型.

(II) 前景检测

从第 $N+1$ 帧开始进行前景检测, 在 RGB 颜色空间中对输入图像进行“区域模糊化理解”预处理, 并将预处理的结果图像转为 LUV 颜色空间图像, 根据颜色差异度量准则实现当前图像与背景模型的差异比较, 实现前景检测.

(III) 模型更新

由分类掩码产生更新掩码, 根据基于区域分类一致性的目标维持方法抑制目标区域背景更新, 同时按照背景模型更新方法对非目标区域的背景进行更新.

4 实验

4.1 评测数据集和评价标准

为了评价算法的性能, 本文采用文献[24]中介绍的评测数据集和评价标准. 该数据集覆盖了运动目标检测的各种挑战, 包含如下 6 种类别的视频且每种类别有 4~6 个视频序列: ① Baseline 类别的视频包含了简单静止的背景; ② Dynamic background 类别的视频包含了复杂动态的背景; ③ Camera jitter 类别的视频是由严重抖动的相机拍摄的; ④ Intermittent object motion 类别的视频包含了间歇运动的目标; ⑤ Shadow 类别的视频包含了伴随运动目标的弱阴影和强阴影; ⑥ Thermal 类别的视频是由热成像仪拍摄的. 评价标准包括如下指标: ① 召回率: Recall (Re); ② 精度: Precision (Pr); ③ 特异性: Specificity (Sp); ④ 假阳性率: False positive rate (FPR); ⑤ 假阴性率: False negative rate (FNR); ⑥ 误分类率: Percentage of wrong classifications (PWC); ⑦ F 测度: F-measure (F-m), 可表示为

$$\left. \begin{aligned} \text{Re} &= \text{TP}/(\text{TP} + \text{FN}) \\ \text{Pr} &= \text{TP}/(\text{TP} + \text{FP}) \\ \text{Sp} &= \text{TN}/(\text{TN} + \text{FP}) \\ \text{FPR} &= \text{FP}/(\text{FP} + \text{TN}) \\ \text{FNR} &= \text{FN}/(\text{TP} + \text{FN}) \\ \text{PWC} &= \frac{100 * (\text{FP} + \text{FN})}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \\ \text{F-m} &= 2 * \text{Re} * \text{Pr}/(\text{Re} + \text{Pr}) \end{aligned} \right\} \quad (13)$$

式中, TP 为被正确检测为前景的像素数量; FP 为被错误检测为前景的像素数量; TN 为被正确检测为背景的像素数量; FN 为被错误检测为背景的像素数量.

4.2 实验结果及其分析

本文算法包含如下参数, 基于大量的实验测试结果, 给出以下参数设置:

① $N, \#_{\min}$

当 $N=40, \#_{\min}=3$ 时, 算法的性能趋于饱和. 在算法的性能趋于饱和的情况下, 再增大 N 和 $\#_{\min}$ 只会增加空间复杂度和时间复杂度. N 和 $\#_{\min}$ 对整个数据集都是固定的.

② R, K

R 和 K 共同决定了运动目标维持为前景的时间. 运动目标维持为前景的最大帧数可以达到 $K \cdot R^2$ 帧; 超过 $K \cdot R^2$ 帧后, 运动目标就逐渐融入背景中. 当 $R=5, K=25$, 即 $K \cdot R^2=625$ 时, 运动目标维持为前景的帧数能够满足数据集集中的实际情况. R 和 K 对整个数据集都是固定的.

③ h_s, h_c

随着 h_s 和 h_c 的不断增大, 区域模糊化理解的效果不断增强. h_s 和 h_c 过大则会导致前景与周围背景的差异变小, 前景像素被误检测为背景的概率增加; 过小则不能达到模糊背景并凸显前景的效果, 使抵抗背景干扰运动的能力减弱. 针对不同的视频序列, 设置不同的 h_s 和 h_c 以取得最佳的效果. 当 $h_s=5, h_c=8$ 时, 算法在整个数据集有最佳的性能体现.

④ $\text{Th}_L, \text{Th}_{UV}$

Th_L 和 Th_{UV} 共同决定了采样值 $s(x, y)$ 与背景模型 $B(x, y)$ 的匹配裕度. Th_L 和 Th_{UV} 过大则会导致前景像素被误检测为背景的概率增加, 过小则导致背景像素被误检测为前景的概率增加. 针对不同的视频序列, 设置不同的 Th_L 和 Th_{UV} 以取得最佳的效果. 当 $\text{Th}_L=0.08, \text{Th}_{UV}=5$ 时, 算法在整个数据集有最佳的性能体现.

以 Dynamic Background 类别的“canoe”视频为例,该场景的背景中含有显著的水面波动和轻微的树木摆动.本文的“区域模糊化理解”预处理操作能够将水面和树木表面的局部差异模糊化.图 1 是本文算法与其他算法的对比实验结果.其中,图 1(a)是输入视频帧;图 1(b)是运动目标检测结果的真实数据;图 1(c)是对图 1(a)进行“区域模糊化理解”的预处理结果.比较图 1(d)~图 1(i)与图 1(b)可知,相对 GMM-Ziv^[4], KDE^[5] 及 SOBS 算法,ViBe, PBAS^[25] 和本文算法没有将大量位于水面波动处和树木摆动处的像素误检测为前景.相对图 1(d)~图 1(h),图 1(i)与图 1(b)的差异最小,这表明本文方法在检测精度上要优于 ViBe 和 PBAS.

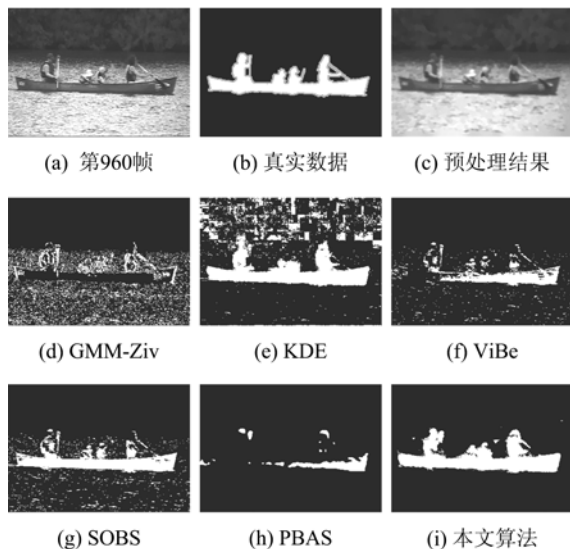


图 1 在“canoe”视频上的对比实验结果图

Fig. 1 Results of contrast experiments on the “canoe” video

以 Baseline 类别的“office”视频为例,该场景的运动目标存在了较长的一段时间且基本维持位置不变.图 2 是本文算法与其他算法的对比实验结果.其

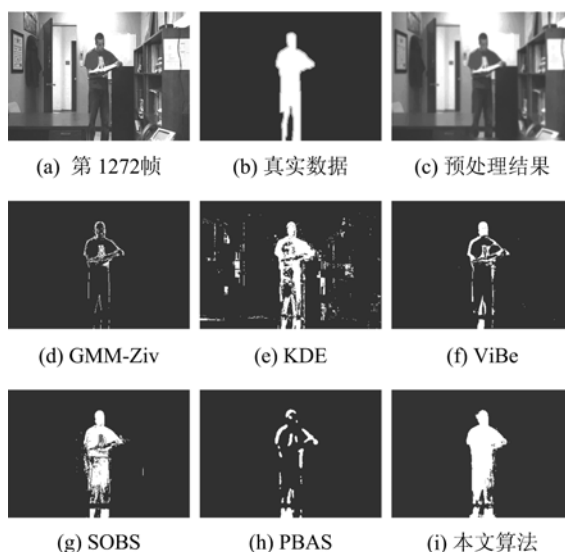


图 2 在“office”视频上的对比实验结果图

Fig. 2 Results of contrast experiments on the “office” video

中,图 2(a)是输入视频帧;图 2(b)是运动目标检测结果的真实数据;图 2(c)是对输入图像进行“区域模糊化理解”的预处理结果.由图 2(d)、图 2(f)、图 2(h)与图 2(b)的对比可知,GMM-Ziv, ViBe 及 PBAS 算法逐渐将运动目标的绝大部分区域理解为新背景,仅将其运动的边缘区域检测为前景.由图 2(e)、图 2(g)、图 2(i)与图 2(b)的对比可知,基于核密度估计的 KDE 方法,基于自组织神经网络结构的 SOBS 方法,以及本文提出的基于区域分类一致性的目标维持方法都在一定程度上抑制了运动目标被理解为新背景的过程.相对图 2(e)和图 2(g),图 2(i)与图 2(b)的差异最小,这表明本文方法在维持目标完整性和检测精度上要优于 KDE 算法和 SOBS 算法.

表 1 给出了算法在每种类别视频上的平均性能以及在数据集上的平均性能.为了与其他算法

表 1 本文算法在所有类别视频上的平均评价结果

Tab. 1 Average results of the proposed algorithm on all categories of videos

视频类别	Re	Pr	Sp	FPR	FNR	PWC	F-m
Baseline	0.932 7	0.934 1	0.998 0	0.002 0	0.017 3	0.374 7	0.933 3
Camera jitter	0.748 5	0.816 5	0.983 7	0.014 3	0.021 1	1.820 5	0.711 0
Dynamic background	0.893 5	0.852 6	0.988 3	0.001 8	0.019 5	0.323 6	0.742 0
Intermittent object motion	0.875 7	0.843 5	0.989 0	0.002 6	0.010 0	5.134 5	0.723 4
Shadow	0.842 5	0.775 8	0.962 0	0.016 7	0.132 7	1.835 6	0.823 5
Thermal	0.782 5	0.892 2	0.985 1	0.002 1	0.354 6	1.586 1	0.806 9
整个数据集	0.845 9	0.852 5	0.984 4	0.004 5	0.092 5	1.845 8	0.790 0

表 2 本文算法和其他基准算法在整个数据集的平均评价结果

Tab. 2 Average results of the proposed algorithm with other benchmark algorithms on the dataset

算法	Re	Pr	Sp	FPR	FNR	PWC	F-m
GMM-Sta ^[2]	0.710 8	0.701 2	0.986 0	0.014 0	0.020 2	3.104 6	0.662 3
GMM-Ziv ^[4]	0.696 4	0.707 9	0.984 5	0.015 5	0.019 3	3.150 4	0.659 6
KDE ^[5]	0.744 2	0.684 3	0.975 7	0.024 3	0.013 8	3.460 2	0.671 9
ViBe ^[8]	0.682 1	0.735 7	0.983 0	0.017 0	0.017 6	3.117 8	0.668 3
SOBS ^[9]	0.788 2	0.717 9	0.981 8	0.018 2	0.009 4	2.564 2	0.715 9
PBAS ^[25]	0.784 0	0.816 0	0.989 8	0.010 2	0.216 0	1.769 3	0.753 2
ViBe+ ^[26]	0.690 7	0.831 8	0.992 8	0.007 2	0.309 3	2.182 4	0.722 4
本文算法	0.845 9	0.852 5	0.984 4	0.004 5	0.092 5	1.845 8	0.790 0

进行对比,本文选择文献[24]中提到的 7 种基准算法,包括:① GMM-Sta^[2];② GMM-Ziv;③ KDE;④ ViBe;⑤ SOBS;⑥ PBAS;⑦ ViBe+^[26].表 2 给出了本文算法和以上基准算法在整个数据集上的平均评价结果,其中粗体部分显示了在特定性能指标下所有对比算法能够取得的最优值.与其他算法性能的对比可知,本文算法性能在 Re, Pr, FPR, F-measure 指标上分别为 0.845 9, 0.852 5, 0.004 5, 0.790 0,取得了最优值;在 PWC 指标上为 1.845 8,取得了次优值,仅与 PBAS 算法取得的最优值 1.769 3 相差 0.076 5.

5 结论

为了提高运动目标检测的精度,针对构建鲁棒的背景模型和背景模型更新问题,本文结合 ViBe 算法提出了一种基于生物视觉特性的背景减除算法.该算法具有以下特点:①借鉴蛙眼视觉外部特性中的近视特性,通过“区域模糊化理解”的预处理操作,提高了背景模型抵抗背景干扰运动的能力;②基于人类对颜色的心理认知特性,从区分颜色亮度属性与色度属性的度量差异角度出发,给出了一种 LUV 颜色空间中基于韦伯定律的颜色差异度量准则,提高了前景检测的鲁棒性和准确性;③通过区分分类掩码和更新掩码的差异,给出了一种基于区域分类一致性的目标维持方法,提高了背景模型更新的一致性.实验结果表明本文算法能够提高运动目标检测的精度.

参考文献(References)

- [1] Wren C R, Azarbayejani A, Darrell T, et al. Pfunder: Real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [2] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Fort Collins, USA: IEEE Press, 1999, 2: 252-258.
- [3] Jepson A D, Fleet D J, El-Maraghi T F. Robust online appearance models for visual tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1 296-1 311.
- [4] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction[C]// Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK: IEEE Press, 2004, 2: 28-31.
- [5] Elgammal A, Harwood D, Davis L. Nonparametric model for background subtraction[C]// Proceedings of the 6th European Conference on Computer Vision. Dublin, Ireland, 2000: 751-767.
- [6] Sheikh Y, Shah M. Bayesian modeling of dynamic scenes for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(11): 1 778-1 792.
- [7] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground-background segmentation using codebook model[J]. Real-time imaging, 2005, 11(3): 172-185.
- [8] Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences [J]. IEEE Transactions on Image Processing, 2011, 20(6): 1 709-1 724.
- [9] Maddalena L, Petrosino A. A self-organizing approach to background subtraction for visual surveillance applications [J]. IEEE Transactions on Image Processing, 2008, 17(7): 1 168-1 177.
- [10] Maddalena L, Petrosino A. The SOBS algorithm: What are the limits? [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, USA: IEEE Press, 2012: 21-26.

- [11] Li L Y, Huang W M, Gu I Y H, et al. Statistical modeling of complex backgrounds for foreground object detection[J]. *IEEE Transactions on Image Processing*, 2004, 13(11): 1 459-1 472.
- [12] Heikkilä M, Pietikäinen M. A texture-based method for modeling the background and detecting moving objects[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 657-662.
- [13] Liao S C, Zhao G Y, Kellokumpu V, et al. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, USA: IEEE Press, 2010: 1 301-1 306.
- [14] Parag T, Elgammal A, Mittal A. A framework for feature selection for background subtraction [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. New York, USA: IEEE Press, 2006, 2: 1 916-1 923.
- [15] Han B, Davis L S. Density-based multifeature background subtraction with support vector machine [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(5): 1 017-1 023.
- [16] Guo J M, Hsu C S. Cascaded background subtraction using block-based and pixel-based codebooks [C]// *Proceedings of the 20th IEEE International Conference on Pattern Recognition*. Istanbul, Turkey: IEEE Press, 2010: 1 373-1 376.
- [17] Zaharescu A, Jamieson M. Multi-scale multi-feature codebook-based background subtraction [C]// *Proceedings of the IEEE International Conference on Computer Vision Workshops*. Barcelona, Spain: IEEE Press, 2011: 1 753-1 760.
- [18] Lettvin J Y, Maturana H R, McCulloch W S, et al. What the frog's eye tells the frog's brain [J]. *Proceedings of the IRE*, 1959, 47(11): 1 940-1 951.
- [19] Ingle D. Disinhibition of tectal neurons by pretectal lesions in the frog[J]. *Science*, 1973, 180(4084): 422-424.
- [20] Wang Zhiling, Chen Zonghai, Xu Xiaoxiao, et al. A fuzzy region understanding tactic for object tracking based on frog's vision characteristic [J]. *Acta Automatica Sinica*, 2009, 35(8): 1 048-1 054.
王智灵, 陈宗海, 徐萧萧, 等. 基于蛙眼视觉特性的运动目标模糊化区域理解跟踪方法[J]. *自动化学报*, 2009, 35(8): 1 048-1 054.
- [21] 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.
- [22] Jain A K. *Fundamentals of Digital Image Processing* [M]. Englewood Cliffs, USA: Prentice Hall, 1989.
- [23] Sharma G, Trussell H. J. Digital color imaging [J]. *IEEE Transactions on Image Processing*, 1997, 6(7): 901-932.
- [24] Goyette N, Jodoin P, Porikli F, et al. Changedetection. net: A new change detection benchmark dataset [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Providence, USA: IEEE Press, 2012: 1-8.
- [25] Hofmann M, Tiefenbacher P, Rigoll G. Background segmentation with feedback: The pixel-based adaptive segmenter [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Providence, USA: IEEE Press, 2012: 38-43.
- [26] Van Droogenbroeck M, Paquot O. Background subtraction: Experiments and improvements for ViBe [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Providence, USA: IEEE Press, 2012: 32-37.