

基于改进 Census 变换的多特征背景建模算法

郭治成^{1,2}, 党建武^{1,2*}, 王阳萍^{1,2}, 金静^{1,2}

¹兰州交通大学电子与信息工程学院, 甘肃 兰州 730070;

²甘肃省人工智能与图形图像处理工程研究中心, 甘肃 兰州 730070

摘要 针对视频图像易受噪声干扰和背景变化复杂的特点,改进传统 Census 变换特征值对中心像素的依赖问题,建立 Census 模板以保持 Census 变换对光线变化的稳健性。将改进后的 Census 变换特征值、图像像素值、更新频数、最近更新时间和动态指数等多种特征融合,建立了一种新的背景建模算法。利用帧间亮度差,自适应选择融合多种特征更新背景模型,依据动态指数衡量背景变化复杂程度,建立不同的更新规则,提升模型对光线突变和复杂场景处理的稳定性。经测试多组标准视频序列,本算法检测精度优于其他算法,有效改善了光线突变对前景目标提取的影响,提高了对光线突变和复杂场景的稳健性,减少了运动目标的孔洞和像素漂移产生的假前景。

关键词 机器视觉; 视频处理; Census 变换; 运动目标检测; 背景建模

中图分类号 TP391.41

文献标识码 A

doi: 10.3788/AOS201939.0815003

Multi-Feature Background Modeling Algorithm Based on Improved Census Transform

Guo Zhicheng^{1,2}, Dang Jianwu^{1,2*}, Wang Yangping^{1,2}, Jin Jing^{1,2}

¹School of Electronic and Information Engineering, Lanzhou Jiaotong University,

Lanzhou, Gansu 730070, China;

²Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing,

Lanzhou, Gansu 730070, China

Abstract In view of the noise interference to video images and the complexity of background change, the traditional Census transform eigenvalue dependence on the central pixel is improved, and the Census template is established to maintain the robustness of the Census transform to light changes. A new background modeling method is established by combining the improved Census transform eigenvalue, image pixel value, update frequency, latest update time and dynamic index. The background texture difference is adaptively selected and fused with multiple features to update the background model. According to the dynamic index, the background change complexity is established, and different update rules are established to improve the stability of the model for light mutation and complex scene processing. After testing multiple sets of standard video sequences, the detection accuracy of this algorithm is better than that of other algorithms, which effectively improves the influence of light mutation on foreground target extraction, increases the robustness to light mutations and complex scenes, and reduces the false foreground caused by holes and pixel shift of the moving target.

Key words machine vision; video processing; Census transform; motion detection; background modeling

OCIS codes 150.0155; 110.4155; 330.4150

1 引 言

背景建模技术是机器视觉和视频处理的一个重要研究方向,目标是将感兴趣区域从视频序列中标识出来,为目标检测^[1]、目标跟踪^[2-3]和行为分析理

解等算法提供重要依据。稳健背景模型的建立和更新算法直接决定前景目标的检测效果。国内外很多学者对背景建模方法进行了深入的研究^[4-6],这些模型主要分为两类:一类是参数化方法,为每个像素点建立带参数的模型;另一类是非参数方法,用已获取

收稿日期: 2019-03-06; 修回日期: 2019-03-24; 录用日期: 2019-04-08

基金项目: 国家自然科学基金(61661026,61841303)、甘肃省科技计划项目(18JR3RA104)、甘肃省教育厅(2017D-08)、兰州交通大学青年基金(2016005)

* E-mail: lzjdgr@163.com

样本像素的关系建立模型。

Wren 等^[7]提出为每个像素值时域信息建立带参单高斯背景模型,该方法适用于变化缓慢及单一的背景,当背景像素发生扰动变化且较为复杂时,会产生大量假前景,检测效果较差。对于单高斯背景模型处理复杂背景效果差的问题,Stauffer 等^[8]提出为每个像素值建立多个高斯分布的混合高斯背景模型,模型对复杂背景处理效果较好,但由于多个高斯分布参数调整困难,易造成前景目标丢失,计算复杂度,处理速度慢。CodeBook(码本模型)为视频帧中每个像素建立一个码本结构^[9],每个码本结构由多个码字组成,计算量小,处理时间起伏较好,但背景动态变化程度较高时易造成码字更新错误,适应能力差,内存消耗大。ViBe(Visual Background Extractor)用单帧视频建立初始背景模型^[10-11],依据邻域像素点空间分布特性更新背景模型,采用随机替换策略使模型能够快速适应场景的简单变化,但在光线突变的场景下,会产生大量的假前景。PBAS(Pixel-Based Adaptive Segmenter)引入控制论思想,结合 SACON(Sample Consensus)背景建模方法和 ViBe 随机替换策略^[12],在检测性能上优于 SACON 和 ViBe,具有建模速度快、准确率较高的优点,但更新率计算和判断阈值复杂度较高,实时性较差。

目前大多数背景建模算法对光线亮度变化较为敏感,尤其在光线突变场景中目标获取效果较差,Census 变换^[13-14]能够减少光线差异引起的误匹配,对亮度偏差较为稳健,同时计算量小,满足视频处理实时性。本文提出一种基于改进 Census 变换的多特征背景建模算法,改进传统 Census 变换特征值对中心像素的依赖问题,引入 Census 模板,提高了

Census 变换对视频图像处理的稳健性,将改进后的 Census 变换特征值、图像像素值、更新频数、最近更新时间 and 动态指数等多种特征融合,建立了一种新的背景建模方法。该方法能够依据光线对视频序列的影响自适应地选择有效的特征值,改善了光线突变对目标提取的影响,在复杂场景中提取运动目标具有较强的适应性和稳健性。

2 改进 Census 变换

Census 变换是一种基于图像局部信息的非参数变换方法,其基本原理是在图像中选取一个像素点 p ,以 p 为中心建立一个矩形窗口,将窗口内中心像素的邻域像素 p' 与 p 进行比较,并将结果用 0 和 1 来表示。Census 变换的实质是通过邻域像素与中心像素的大小关系,将图像像素值编码成二进制码流,公式为

$$C_{\text{Cen}}(p) = \bigotimes_{p' \in W(p)} \zeta[I(p), I(p')], \quad (1)$$

式中: $W(p)$ 为中心像素 p 对应的 Census 变换窗口; $I(p)$ 为 p 的像素值; $I(p')$ 为 p 周边像素的像素值; \bigotimes 符号表示按位连接 $\zeta[I(p), I(p')]$,满足

$$\zeta[I(p), I(p')] = \begin{cases} 1, & I(p) < I(p') \\ 0, & I(p) \geq I(p') \end{cases} \quad (2)$$

中心像素灰度值经过 Census 变换后被替换为 $C_{\text{Cen}}(p)$,利用 Hamming 距计算 $C_{\text{Cen}}(p)$ 之间的差异程度,测量两点之间的距离。 $C_{\text{Cen}}(p)$ 与变换窗口中心像素和其他各个像素灰度大小有关,因此,Census 变换对于幅度失真与光线突然变换有着较强的稳健性。

传统的 Census 变换对中心像素依赖性较强,如图 1(a)和(b)所示,窗口内邻域像素值不变,中心像素值变化,Census 变换波动较大,容易引起误配。

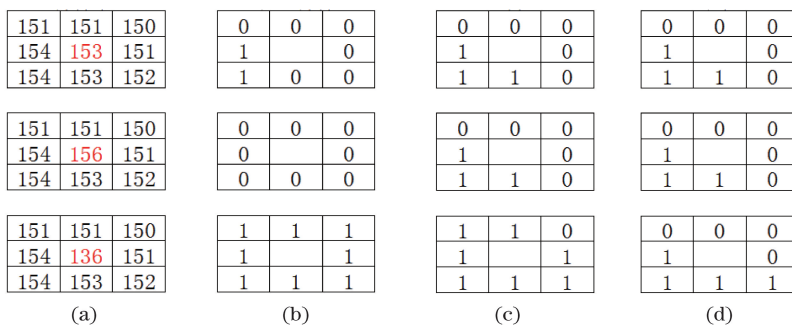


图 1 3×3 窗口像素值。(a)窗口内邻域像素相同,中心像素不同;(b)中心像素为原始灰度值;(c)中心像素为窗口各像素值的平均灰度值;(d)中心像素为窗口内各像素值的中位数

Fig. 1 Pixel value in 3×3 window. (a) Fields in the window with the same pixels and the different center pixels; (b) center pixel is the original gray value; (c) center pixel is the average gray value of each pixel value of the window; (d) center pixel is the median value of each pixel value in the window

为改变视频序列图像受噪声干扰和背景变化复杂带来 Census 变换误配的问题,本文提出一种改进 Census 变换算法,其核心思想是取窗口内各像素值的中位数替代中心像素灰度值 $I(p)$:

$$\zeta[M(p), I(p')] = \begin{cases} 1, & M(p) < I(p') \\ 0, & M(p) \geq I(p') \end{cases}, \quad (3)$$

$$M(p) = m_{\text{median}}[I(p), I(p')], \quad (4)$$

式中: $M(p)$ 为窗口内所有像素值的中位数; $m_{\text{median}}(\cdot)$ 为求中位数的函数。

将窗口中邻域各像素的灰度值与参考值进行比较,如图 1(b)~(d)所示,分别取窗口内中心像

素值、所有像素平均值、所有像素中位数为中心像素参考值,以中位数取代中心像素的 Census 变换明显降低了对中心像素的依赖,提高了算法的稳健性。

采集视频信息设备硬件缺陷和外界电磁波干扰,易引起视频图像在同一坐标位置不同帧之间像素值的抖动。图 2(a)矩形区域中心像素 p (坐标为(185,122)),连续 140 帧内像素值在 146~152 之间的抖动如图 2(b)所示,Census 变换结果如图 2(c)~(e)所示,中心像素 $M(p)$ 不变,邻域像素值抖动,Census 变换状态稳定性较差。

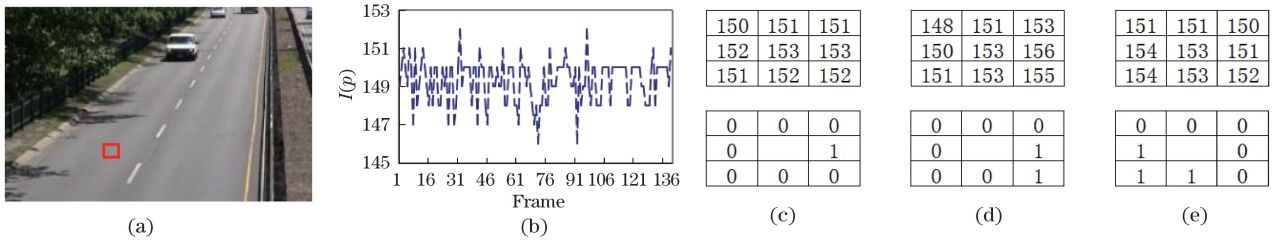


图 2 邻域像素值抖动。(a)视频序列;(b)矩形区域像素值变化;(c)中心像素为 153 第 t 帧 Census 变换;(d)中心像素为 153 第 $t+1$ 帧 Census 变换;(e)中心像素为 153 第 $t+2$ 帧 Census 变换

Fig. 2 Neighborhood pixel value change. (a) Video sequence; (b) rectangular area pixel value change; (c) center pixel is 153, the t -th frame Census transform; (d) center pixel is 153, the $(t+1)$ th frame Census transform; (e) center pixel is 153, the $(t+2)$ th frame Census transform

为改进 Census 变换的稳定性,本文提出 Census 模板,定义如下:

$$C_{\text{cen}}(p) = \begin{cases} C_{\text{cen}}(p_t), & \max[|M(p) - M(p_t)|, |I(p') - I(p'_t)|] < \epsilon \\ C_{\text{cen}}(p), & \max[|M(p) - M(p_t)|, |I(p') - I(p'_t)|] \geq \epsilon \end{cases}, \quad (5)$$

式中: $C_{\text{cen}}(p_t)$ 为第 t 帧像素 p 的 Census 变换值; $M(p_t)$ 为 t 帧像素窗口内所有像素值的中位数; $I(p'_t)$ 为 t 帧 p 周边像素的像素值; ϵ 为常数,取 5~8; $\max(\cdot)$ 为求最大值函数。

Census 模板更新规则为

$$M(p) = \frac{M(p_t) + M(p)}{2}, \quad (6)$$

$$I(p') = \frac{I(p'_t) + I(p')}{2}. \quad (7)$$

处理结果如图 3 所示,因为差异小于 ϵ ,所以图 3(a)~(c)公用一个 census 模板,减少了像素抖动对前景目标提取的影响,增强了 Census 变换的稳定性。

3 基于改进 Census 变换的多特征背景建模方法

3.1 背景模型原理

背景模型常用像素值、梯度、目标表象等信息为特征建立模型^[15],如 Liao 等^[16]以纹理为特征建立

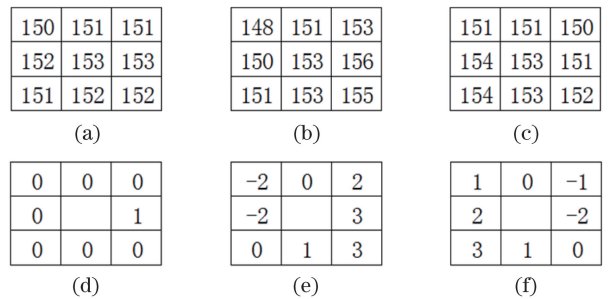


图 3 Census 模板。(a) Census 模板像素值;(b) t_1 帧像素值;(c) t_2 帧像素值;(d) Census 模板变换值;(e) $I(p') - I(p'_{t_1})$;(f) $I(p') - I(p'_{t_2})$

Fig. 3 Census template. (a) Census template pixel value; (b) t_1 frame pixel value; (c) t_2 frame pixel value; (d) Census template transform value; (e) $I(p') - I(p'_{t_1})$; (f) $I(p') - I(p'_{t_2})$

背景模型,Barnich 等^[10]利用像素邻域空间关系更新背景模型等。这些算法用短时间内场景变化的统计特性来增强背景模型的适应性,却忽略了长时间内场景变化呈现周期性和受光线突变引起的场景变

化。这些复杂场景变化导致背景模型退化,误检率上升,算法性能下降。为解决以上问题,本文将改进后的 Census 变换特征和视频多种特征融合,提出了一种新的背景建模算法。

视频是由多幅图像组成的序列,这些图像称为帧,连续帧在同一位置上的像素值在时间轴上构成了一个集合 $V(x_t) = \{x_1, x_2, \dots, x_t \mid t = 1, \dots, A\}$, 其中 $V(x_t)$ 为像素 x_t 的值, x_t 为图像中的一个像素, t 为视频序列中第 t 帧, A 为视频序列帧数。模型为视频图像中的每一个 x_i 建立一个样本集 $M(x_i) = \{C(x_i), P(x_i), F(x_i), T(x_i), S(x_i)\}$, 其中:

$C(x_i) = \{c_1, c_2, \dots, c_N\}$ 为 x_i 处背景样本集中每个样本的 Census 变换特征值, N 为样本集大小;

$P(x_i) = \{p_1, p_2, \dots, p_N\}$ 为 x_i 处背景样本集中每个样本的像素值;

$F(x_i) = \{f_1, f_2, \dots, f_N\}$ 为 x_i 处背景样本集中每个样本的频数;

$T(x_i) = \{t_1, t_2, \dots, t_N\}$ 为 x_i 处背景样本集中每个样本最近更新时间;

$S(x_i) = \{s\}$ 为 x_i 像素的动态指数。

模型假设像素值在 $M(x_i)$ 外分布的为前景像素, 在 $M(x_i)$ 内分布的为背景像素, 前景、背景像素的判定边界由 $M(x_i)$ 构成, t 帧输入的像素 $V(x_t)$ 通过下式来判断是否为背景像素,

$$V(x_t) = \begin{cases} G_F, \{V(x_t) \cap M(x_i)\} < S(x_i) \\ G_B, \{V(x_t) \cap M(x_i)\} \geq S(x_i) \end{cases} \quad (8)$$

式中: G_F 为前景像素; G_B 为背景像素。

3.2 模型的初始化

初始化背景模型用视频序列的第 1 帧图像, 取以像素 x_i 为中心的 3×3 窗口内所有像素初始化 x_i , Census 变换特征值 $C(x_i) = \{C_{cen}(x_i), C_{cen}(x_j) \mid j = 1, \dots, 8\}$, 其中 i 为 1 帧中第 i 个像素, 其中 j 为 x_i 第 j 个邻域像素, $C_{cen}(x_i)$ 为 x_i 的 Census 变换特征值。取像素 x_i 为中心的 5×5 窗口中满足

$$P(x_i) = \{V(x_i), V'_j(x_i) \mid j = 1, \dots, N\} \quad (9)$$

的像素值初始化 $P(x_i)$, $F(x_i) = \{f_i = 1 \mid i = 1, \dots, N\}$, $T(x_i) = \{t_i = 1 \mid i = 1, \dots, N\}$, N 为样本集大小, 满足

$$V'_j(x_i) \in \{|V'_k(x_i) - V(x_i)| < \epsilon, k = 1, \dots, 24\} \quad (10)$$

的 x_i 元素个数, 动态指数满足

$$S(s_i) = \frac{\sum |V(x_i) - V'_j(x_i)|}{N}, j = 1, \dots, N, \quad (11)$$

式中: $V'_k(x_i)$ 为像素 x_i 的 5×5 邻域像素; $S(s_i)$ 为 x_i 像素的动态指数; j 为第 j 个满足(10)式的邻域像素; k 为第 k 个邻域像素, $V(x_i)$ 为窗口中心像素的像素值; $V'_j(x_i)$ 为满足(10)式的邻域像素; ϵ 为常数, 取值同(5)式。

为减少背景中复杂动态变化, 如: 水波、光线和树叶晃动等因素产生的大量错误前景目标, 建立衡量背景动态变化程度的标准 $D(x_i)$, 满足

$$D(x_i) = \frac{\sum_{k=1}^N |V(x_i) - p_k|}{N}, \quad (12)$$

(12)式为当前图像像素值与背景模型的样本值先差分求和, 得到欧氏距离和后的均值。式中 p_k 为背景模型 $P(x_i)$ 的样本值。

动态指数 $S(x_i)$ 利用不同复杂度区域中像素反馈的信息来衡量背景动态变化程度。动态变化程度低, 如图 4(a) 中区域 A (左下方框标记, 中心坐标为 (166, 36)), 动态指数 $S(x_i)$ 应较小, 降低误识别背景点; 动态变化程度高, 如图 4(a) 中区域 B (左上方框标记, 中心坐标为 (187, 95)), 动态指数 $S(x_i)$ 应较大, 避免产生错误的前景点, $S(x_i)$ 满足

$$S(x_i) = \begin{cases} S(x_i) - D(x_i), & S(x_i) \leq a \cdot D(x_i) \\ S(x_i) - D(x_i), & S(x_i) > a \cdot D(x_i) \end{cases} \quad (13)$$

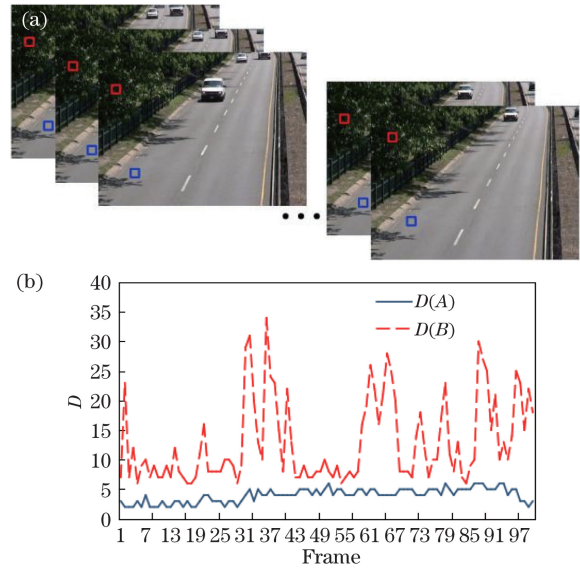


图 4 视频不同动态区域。(a)连续 100 帧;

(b)区域 A、B 的复杂度

Fig. 4 Different dynamic areas in the video. (a) Continuous 100 frames; (b) complexity of areas A and B

式中: a 为固定系数。

3.3 模型更新策略

前景目标的提取质量主要取决于背景模型的更新策略,模型更新慢会引起停滞物体和运动目标 ghost 现象,模型更新快会出现大量假前景和运动目标内孔洞现象。算法利用亮度差来衡量视频序列图像中是否发生光线突变,调整建模特征更新模型,有效去除了光线突变造成的大量假前景,提升了前景目标识别率;利用动态指数 $S(x_t)$ 来衡量不同区域动态程度,结合样本频数 $F(x_t)$ 、最近更新时间 $T(x_t)$ 控制背景模型在不同区域的更新速度,有效地减少了 ghost 和运动目标内孔洞现象。

模型更新策略如下:

步骤 1: 设 $d_t(h, w)$ 为第 t 帧和第 $t-1$ 帧第 h 行第 w 列像素差异, x_t 为第 t 帧第 h 行第 w 列像素的像素值,有

$$d_t(h, w) = \begin{cases} 1, & |x_t - x_{t-1}| \neq 0 \\ 0, & |x_t - x_{t-1}| = 0 \end{cases}, \quad (14)$$

$$1 \leq h \leq H, 1 \leq w \leq W,$$

式中: W 为图像宽度; H 为图像高度。通过

$$L(x_t) = \frac{\sum d_t(h, w)}{W \times H} \times 100\% \quad (15)$$

计算输入图像与上一帧的亮度差 $L(x_t)$, 衡量是否有光线突变发生, 如 $L(x_t)$ 满足

$$L(x_t) < \eta, \quad (16)$$

执行步骤 2, 否则执行步骤 4. η 为固定常数, 一般设为 0.4~0.6 之间。

步骤 2: x_t 像素点未被分类。

像素 $V(x_t)$ 满足

$$\sum_{j=1}^8 (|V(x_t) - V'_j(x_t)| < \delta) = 0, \quad (17)$$

x_t 为噪声会干扰像素分类, $V(x_t)$ 表示为

$$V(x_t) = m_{\text{median}}[V'_j(x_t)], j = 1, \dots, 8, \quad (18)$$

式中: $V'_j(x_t)$ 为像素 x_t 3×3 窗口内的邻域像素值; δ 为固定常数。

步骤 3: x_t 在 $M(x_t)$ 中找到满足

$$|V(x_t) - p_i(x_t)| < S(x_t), i = 1, \dots, N \quad (19)$$

的值, 则 x_t 为背景像素, 样本集中第 i 个样本频数自增 1, 更新时间清零, 其他样本的频数自减 1, 更新时间自增 1, 式中 $p_i(x_t)$ 为 $P(x_t)$ 中第 i 个样本的像素值。

如果 x_t 在 $M(x_t)$ 中未找到满足(19)式的值, 则 x_t 为前景像素, x_t 的 3×3 邻域元素 x_j 满足

$$C_{\text{cen}}(x_t) \in C(x_t) \quad (20)$$

的值设置为前景像素, $F(x_t)$ 所有元素自减 1, $T(x_t)$ 所有元素自增 1, 执行步骤 5。

步骤 4: x_t 在 $M(x_t)$ 中找到满足(20)式的值, x_t 为背景, 否则为前景, 初始化 $P(x_t)$ 、 $F(x_t)$ 、 $T(x_t)$ 、 $S(x_t)$ 。

步骤 5: 更新 $M(x_t)$ 。

4 实验与结果分析

为验证本文所提背景建模算法的性能, 将该算法在多组数据上进行实验, LightSwitch、Bootstrap、TimeOfDay、WavingTrees 来源于 Microsoft Wallflower paper 数据集, Highway、Canoe、Fountain02 来源于 CDNet2014^[17] 数据集。LightSwitch、Bootstrap、TimeOfDay 为室内场景, 分辨率为 160 pixel \times 120 pixel, WavingTrees、Highway 为室外场景, 分辨率为 160 pixel \times 120 pixel 和 320 pixel \times 240 pixel, Canoe、Fountain02 为复杂场景, 分辨率为 320 pixel \times 240 pixel 和 432 pixel \times 288 pixel。

实验在酷睿 i3 3.2 GHz 6.0 GB 的计算机上, VS2010、C++ 和 OpenCV3.3.0 环境实现算法, 与文献[9]中的 CodeBook、文献[8]的混合高斯模型 MOG、文献[12]的 PBAS 和文献[10]的 ViBe 4 种算法作对比实验, CodeBook 和 MOG 建模帧数为 15 帧, PBAS 和 ViBe 维持原文献中的参数。

4.1 光线突变场景实验及分析

光线突变场景实验中本文选取 LightSwitch 的闪烁显示器(755 帧)、开关灯(801 帧)和摄像头遮挡(827 帧)3 个场景, 结果如图 5 所示。

定量分析采用 3 个信息检索邻域内反映检索效果的重要指标对 5 种背景建模进行评估, 包括: 准确率(P), 表示检测出前景目标的正确率; 召回率(R), 表示所有前景目标有多少被检测出来; 误检百分比(P_{PWC}), 表示前景和背景点的错误率。

$$P = \frac{t_{\text{tp}}}{t_{\text{tp}} + f_{\text{fp}}}, \quad (21)$$

$$R = \frac{t_{\text{tp}}}{t_{\text{tp}} + f_{\text{fn}}}, \quad (22)$$

$$P_{\text{PWC}} = 100 \times \frac{f_{\text{in}} + f_{\text{fp}}}{t_{\text{tp}} + t_{\text{tn}} + f_{\text{fp}} + f_{\text{fn}}}, \quad (23)$$

式中: t_{tp} 为检测正确前景点数; t_{tn} 为检测错误前景点数; f_{fp} 为检测正确背景点数; f_{fn} 为检测错误背景点数。

CodeBook、MOG、PBAS 和 ViBe 4 种背景建模

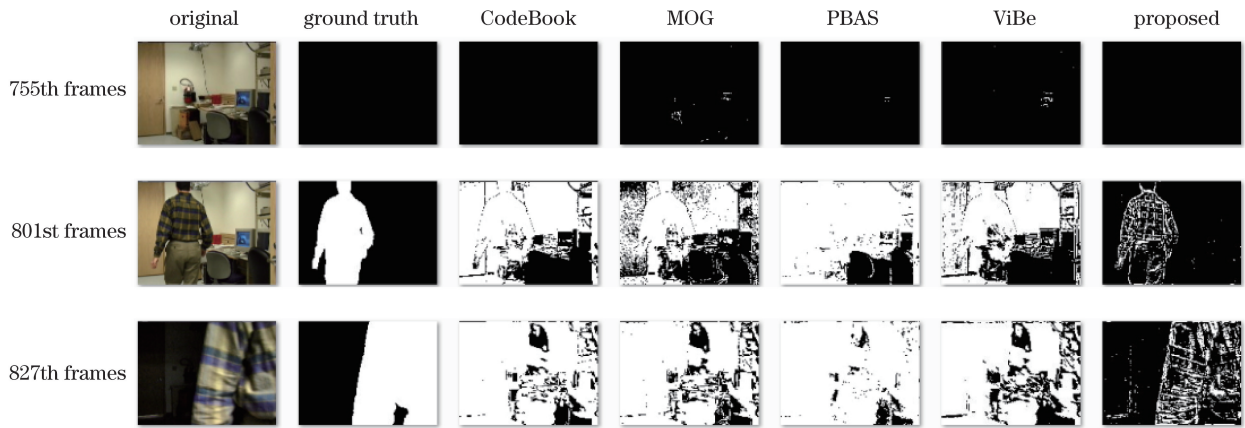


图 5 光线突变场景

Fig. 5 Light mutation scenes

算法以像素值为主要特征识别前景目标。从图 5 可以看出,在光线突变场景下,由光线突变引起当前背景图像内大量像素的像素值大幅度变化,导致应为背景的像素无法在模型中正确匹配,产生大量假前景,无法正确识别前景目标,召回率较低,误检百分

比较高。本文算法利用 Census 变换对光线变化不敏感的特点融合多种特征,能够在光线突变场景中正确识别前景目标,从图 5 和表 1 可以看出,本文算法召回率较高、误检百分比较低,明显优于其他 4 种算法,对光线突变场景具有较强的稳健性。

表 1 光线突变场景准确性验证

Table 1 Light mutation scene accuracy verification

Method	755th frames			801st frames			827th frames		
	<i>R</i>	<i>P</i>	<i>P_{PWC}</i>	<i>R</i>	<i>P</i>	<i>P_{PWC}</i>	<i>R</i>	<i>P</i>	<i>P_{PWC}</i>
CodeBook	0	0	0	0.89	0.34	51	0.78	0.5	55
MOG	0	0	0.35	0.82	0.38	41	0.71	0.47	59
PBAS	0	0	0.042	0.97	0.33	55	0.92	0.53	48
ViBe	0	0	0.22	0.79	0.33	48	0.68	0.47	59
Proposed	0	0	0	0.25	0.9	21	0.32	0.95	38

4.2 普通背景和复杂背景实验及分析

普通和复杂背景实验数据选取 Bootstrap 共 3054 张, TimeOfDay 共 5889 张, WavingTrees 共 286 张, Highway 共 1700 张, Canoe 共 1189 张和 Fountain02 共 1499 张, 结果如图 6 所示。

将初始化模型耗费时间(M_{MT})和模型对单帧的处理时间(H_{HT})进行对比, 本文算法初始化模型耗费时间较短, 处理单帧速度较快。结果如表 2 所示。

Bootstrap、TimeOfDay、WavingTrees 和 Highway 4 组普通场景的视频序列实验结果如表 3 所示。Canoe、Fountain02 出现不平静的水面、喷泉等复杂背景的实验结果如表 4 所示。由图 6 可以看出, 本文算法前景目标提取较为完整, 背景识别错误率较低, 假前景较少, 在复杂背景环境下前景目标识别程度也较高。对表 3 和表 4 的定量分析也显示本文算法的准确率、召回率和误检百分比优于其

他算法。

5 结 论

提出一种基于改进 Census 变换的多特征背景建模算法, 用 3×3 窗口中位数取代中心像素值降低了 Census 变换特征值对中心像素的依赖, 针对处理对象为连续多幅图像序列, 建立 Census 模板和模板更新规则, 提高了 Census 变换对视频图像处理的稳健性, 将改进后的 Census 变换特征值、图像的像素值、频数、更新时间 and 动态指数等多种特征融合, 建立了一种新的背景建模方法。该方法能够依据光线变化对视频序列的影响程度自适应地选择融合有效的特征值, 更新背景模型并识别前景目标, 改善了光线突变对目标提取的影响, 利用动态指数来衡量视频不同区域的复杂度, 为不同复杂度的区域设置不同的更新规则, 实验结果表明, 该算法提高了在光线

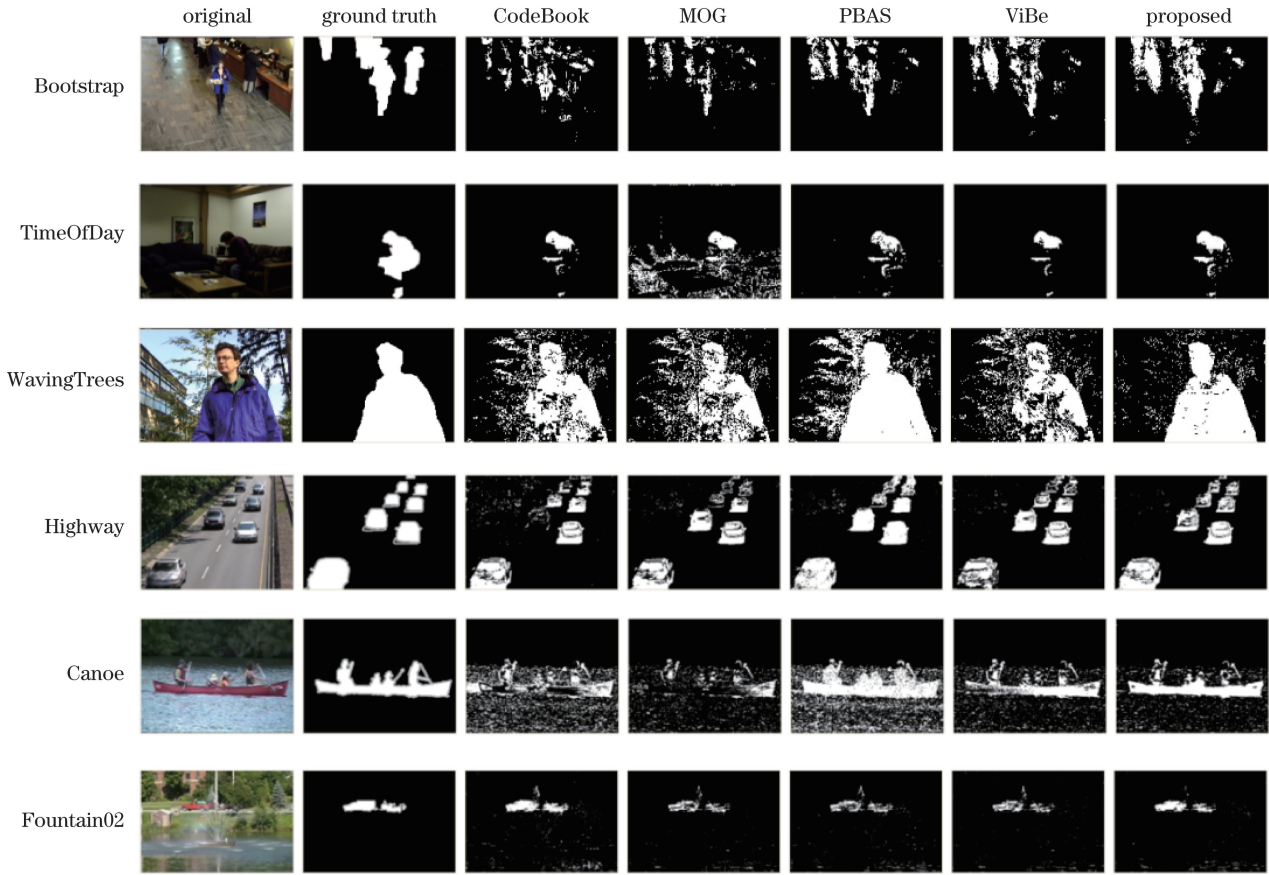


图 6 5 种算法处理结果对比

Fig. 6 Processing results comparison of five algorithms

表 2 5 种算法处理速度对比

Table 2 Processing speed comparison of five algorithms

Method	WavingTrees		Bootstrap		Highway		Fountain02	
	M_{MT}	H_{HT}	M_{MT}	H_{HT}	M_{MT}	H_{HT}	M_{MT}	H_{HT}
CodeBook	0.59	0.05	0.62	0.05	1.25	0.08	1.39	0.10
MOG	2.31	0.29	2.30	0.26	8.50	1.30	8.98	1.63
ViBe	0.09	0.03	0.11	0.03	0.45	0.14	0.49	0.18
PBAS	0.08	0.37	0.09	0.36	0.31	0.58	0.46	0.61
Proposed	0.07	0.08	0.06	0.08	0.36	0.28	0.35	0.39

表 3 普通场景对比实验

Table 3 Common scene comparison experiment

Method	Bootstrap			TimeOfDay			WavingTrees			Highway		
	R	P	P_{PWC}	R	P	P_{PWC}	R	P	P_{PWC}	R	P	P_{PWC}
CodeBook	0.374	0.626	11.3	0.366	0.998	4.32	0.766	0.727	16.00	0.703	0.891	4.81
MOG	0.290	0.626	11.7	0.494	0.310	11.00	0.695	0.693	18.70	0.748	0.962	3.55
PBAS	0.386	0.568	12.0	0.404	0.981	4.11	0.997	0.677	14.60	0.807	0.98	2.63
ViBe	0.397	0.525	12.8	0.315	0.998	4.67	0.654	0.688	19.70	0.706	0.937	4.29
Proposed	0.588	0.576	11.2	0.435	0.981	3.90	0.962	0.878	5.25	0.845	0.972	2.25

表 4 复杂背景对比实验

Table 4 Complex scene comparison experiment

Method	Canoe			Fountain02		
	R	P	P_{PWC}	R	P	P_{PWC}
CodeBook	0.530	0.341	15.90	0.926	0.577	1.570
MOG	0.295	0.495	10.70	0.609	0.707	1.340
PBAS	0.949	0.410	15.10	0.582	0.589	1.720
ViBe	0.711	0.663	6.91	0.608	0.684	1.400
Proposed	0.833	0.729	5.070	0.809	0.755	0.944

突变和动态复杂场景中提取运动目标的适应性和稳健性。因此,本文算法可为视频检测系统的背景建模提供有力的技术支持和参考。

由于本文主要针对较为理想的场景,在雾霾、沙尘或雨雪等极端气候下,需要考虑恶劣的室外环境对前景目标提取的影响,探索适用程度更高的背景建模算法。

参 考 文 献

- [1] Wang H Y, Wang L, Yin W R, *et al.* Multi-scale correlation filtering visual tracking algorithm combined with target detection[J]. *Acta Optica Sinica*, 2019, 39(1): 0115004.
王红雨, 汪梁, 尹午荣, 等. 结合目标检测的多尺度相关滤波视觉跟踪算法[J]. *光学学报*, 2019, 39(1): 0115004.
- [2] Ge B Y, Zuo X Z, Hu Y J. Long-term object tracking based on feature fusion[J]. *Acta Optica Sinica*, 2018, 38(11): 1115002.
葛宝义, 左宪章, 胡永江. 基于特征融合的长时目标跟踪算法[J]. *光学学报*, 2018, 38(11): 1115002.
- [3] Yang C Y, Li C, Liang Y C, *et al.* Blurred object detection based on improved particle filter in coal mine underground surveillance[J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2017, 47(6): 1976-1985.
杨超宇, 李策, 梁胤程, 等. 基于改进粒子滤波的煤矿视频监控模糊目标检测[J]. *吉林大学学报(工学版)*, 2017, 47(6): 1976-1985.
- [4] Sun X W, Xu Q S, Cai Y, *et al.* Sea sky line detection based on edge phase encoding in complicated background[J]. *Acta Optica Sinica*, 2017, 37(11): 1110002.
孙熊伟, 徐青山, 蔡熠, 等. 基于边缘相位编码的复杂背景下海天检测[J]. *光学学报*, 2017, 37(11): 1110002.
- [5] Sui Z S, Li J S, Zhang J, *et al.* Video foreground detection of tensor low-rank representation and

spatial-temporal sparsity decomposition[J]. *Optics and Precision Engineering*, 2017, 25(2): 529-536.

隋中山, 李俊山, 张姣, 等. 张量低秩表示和时空稀疏分解的视频前景检测[J]. *光学精密工程*, 2017, 25(2): 529-536.

- [6] Ma J J, Pan Q, Liang Y, *et al.* Object detection based on improved Grassberger entropy random forest classifier[J]. *Chinese Journal of Lasers*, 2019, 46(7): 0704011.
马娟娟, 潘泉, 梁彦, 等. 基于改进 Grassberger 熵随机森林分类器的目标检测[J]. *中国激光*, 2019, 46(7): 0704011.
- [7] Wren C R, Azarbayejani A, Darrell T, *et al.* Pfinder: real-time tracking of the human body[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780-785.
- [8] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]//1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 23-25, 1999, Fort Collins, CO, USA. New York: IEEE, 1999: 246-252.
- [9] Huo D H, Yang D, Zhang X H, *et al.* Principal component analysis based codebook background modeling algorithm[J]. *Acta Automatica Sinica*, 2012, 38(4): 591-600.
霍东海, 杨丹, 张小洪, 等. 一种基于主成分分析的 Codebook 背景建模算法[J]. *自动化学报*, 2012, 38(4): 591-600.
- [10] Barnich O, van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences [J]. *IEEE Transactions on Image Processing*, 2011, 20(6): 1709-1724.
- [11] Bi G L, Xu Z J, Chen T, *et al.* Complex background model and foreground detection based on random aggregation[J]. *Acta Physica Sinica*, 2015, 64(15): 150701.
毕国玲, 续志军, 陈涛, 等. 基于随机聚类的复杂背景建模与前景检测算法[J]. *物理学报*, 2015, 64(15): 150701.
- [12] Zhang Z B, Yuan X B. An improved PBAS algorithm for dynamic background[J]. *Electronic Design Engineering*, 2017, 25(3): 35-40.
张泽斌, 袁晓兵. 一种改进反馈机制的 PBAS 运动目标检测算法[J]. *电子设计工程*, 2017, 25(3): 35-40.
- [13] Stein F. Efficient computation of optical flow using the census transform[M]//Rasmussen C E, Bühlhoff H H, Schölkopf B, *et al.* *Pattern recognition. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2004, 3175: 79-86.
- [14] Yan L, Wang R, Liu H, *et al.* Stereo matching

- method based on improved cost computation and adaptive guided filter[J]. *Acta Optica Sinica*, 2018, 38(11): 1115007.
- 闫利, 王芮, 刘华, 等. 基于改进代价计算和自适应引导滤波的立体匹配[J]. *光学学报*, 2018, 38(11): 1115007.
- [15] Guo Z C, Dang J W, Wang Y P, *et al.* Background modeling method based on multi-feature fusion[J]. *Opto-Electronic Engineering*, 2018, 45(12): 180206.
- 郭治成, 党建武, 王阳萍, 等. 基于多特征融合的背景建模方法[J]. *光电工程*, 2018, 45(12): 180206.
- [16] 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]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 1301-1306.
- [17] Wang Y, Jodoin P M, Porikli F, *et al.* CDnet 2014: an expanded change detection benchmark dataset[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 393-400.