

# 基于改进混合高斯模型与阴影去除的目标检测

王林, 和萌

(西安理工大学 自动化与信息工程学院, 西安 710048)

**摘要:** 随着计算机视觉和摄像设备的日益普及, 目标检测技术已经成为一个重要的研究领域; 虽然提出了几种目标检测方法, 但由于其适用性与局限性, 并不能解决实际复杂场景中的各种挑战; 针对传统混合高斯模型对动态背景、光照变化和阴影敏感等问题, 提出一种混合高斯模型的改进算法, 用于视频中目标检测; 该方法首先通过传统混合高斯模型获取当前帧目标的粗略区域; 通过将双级学习率和组合权重引入混合高斯模型, 从而区分出运动区域与包含动态背景的背景区域; 然后进一步利用基于颜色特性与空间连续性的方法去除阴影; 最后通过形态学处理提取出准确的运动目标区域; 对比实验表明, 所提方法不仅能够有效去除动态背景, 而且能够有效抑制阴影和光照变化的影响。

**关键词:** 目标检测; 混合高斯模型; 动态背景; 阴影去除

## Object Detection Based on Improved Gaussian Mixture Model and Shadow Removal

Wang Lin, He Meng

(College of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** Object detection technologies have emerged as an important research area with increasing popularity of computer vision and camera devices. Even though several object detection approaches have been proposed, they cannot address various challenges in actual complex scenes owing to their applicability and restrictions. For the traditional Gaussian mixture model is sensitive to dynamic background, light change and shadow, an improved algorithm of Gaussian mixture model is proposed, which is used for object detection in video. Firstly, we extract rough region of the current frame by the traditional Gaussian mixed model. By introducing the two-level learning rate and combined weight into the Gaussian mixture model, the moving region and the background region containing the dynamic background are distinguished. Then shadow is removed based on color characteristic and spatial continuity. Finally, the complete and accurate moving object area is detected out by morphological closing operation. Comparative experiments indicate that the proposed method not only can effectively restrain the influence of dynamic background, but also suppress the influence of shadow and light changes.

**Keywords:** object detection; Gaussian mixture model; dynamic background; shadow removal

## 0 引言

运动目标检测是许多基于图像和视频的监测应用中的一个重要步骤, 如智能监控系统、智能交通系统、人体检测与跟踪系统等。针对各种应用中存在的问题, 提出几种前景检测方法。最具代表性的方案包括背景减除法<sup>[1]</sup>、光流法和帧差法。背景减除法是视频监控系统中最常用的方法。背景减除法通常假设三个条件<sup>[2]</sup>: 固定摄像机, 恒定光照条件和静态背景(即背景中没有动静或没有噪音)。这些条件的违反带来了一些挑战, 其中包括: 投影和光照变化、动态背景、噪声视频、相机抖动等。这些挑战通常会产生大量假阳性或假阴性。

基于混合高斯模型<sup>[3]</sup>的背景减除法应用广泛, 它能处理逐渐变化的光照变化和具有微小重复运动的背景, 然而缓慢的物体往往会被背景迅速吸收。为了应对复杂的动态背景或

光照变化等挑战, 研究人员提出了一些改进方法。文献 [4] 提出了一种简单的非参数自适应密度估计方法, 通过引入 Dirichlet 函数先验估计每个像素的高斯模型的适当个数, 是一种混合高斯模型的改进算法, 但是该算法只完成了对高斯模型个数的自适应, 高斯模型权重、学习率的更新方式并未完成。文献 [5] 提出一种使用在线 K-均值聚类来实现模型初始化的方法, 同时对模型更新部分做了优化, 提高了算法的收敛速度, 节约了存储空间。文献 [6] 提出一种将时域和空间信息相结合的方法, 利用广义高斯分布与共同分析前景背景的方法通过时域信息建模, 利用多尺度帧间关联分析和直方图匹配相结合的方法通过空间信息进行建模, 融合时域和空间信息, 可以应对复杂动态背景。而文献 [5-6] 虽然是基于混合高斯模型的改进, 但是并没有考虑混合高斯模型对于光照变化敏感这一问题。

在实际应用中, 区分运动目标和阴影是很有挑战的, 因为运动目标和它们的阴影往往呈现相似的运动模式。为解决这一问题, 研究人员提出了一些方法。文献 [7] 提出一种基于色度的方法, 通过计算视频帧与背景参考帧的 HSV 分量的变化率来检测阴影。该方法假设视频帧中的阴影像素和

收稿日期:2019-01-10; 修回日期:2019-02-11。

基金项目:陕西省科学技术厅(2017ZDCXL-GY-05-03)。

作者简介:王林(1963-),男,江苏省东台人,博士,教授,主要从事深度学习、社交网络和数据挖掘方向的研究。

背景参考帧中的阴影像素的色调分量相差不大, 而阴影像素的饱和度和明度分量要低很多。然而, 仅仅是颜色信息可能无法正确区分与阴影相似的深色前景物体 (例如深色的车子, 或者深色的衣服)。文献 [8] 提出的方法使用色度信息来识别阴影区域, 然后使用梯度信息来细化阴影或非阴影像素的初始估计。该方法提供了良好的阴影检测结果, 并结合色度和梯度信息对 [7] 中色度法进行了改进。但是为了获得高质量的阴影检测, 需要对前景物体和阴影进行显著的色差处理。文献 [9] 提出一种结合 HSV 与纹理特征的视频阴影消除方法, 解决了由于亮度比值不稳定而导致的前景误检为阴影的问题。

针对上述问题, 提出一种基于改进混合高斯模型与阴影去除的目标检测方法。首先使用传统混合高斯模型将图像帧区分为背景区域与粗略运动目标区域。然后将双级学习率和组合权重引入混合高斯模型, 从而区分出包含动态背景的背景区域与运动区域, 使背景模型更好地适应动态背景。同时, 为了消除阴影对目标检测精度的影响, 利用基于颜色特性与空间连续性的方法去除阴影。实验结果表明所提方法不仅能在动态背景下提供可靠的背景模型, 而且能够精确的去除掉阴影和光照变化的影响。

## 1 传统混合高斯模型

传统混合高斯模型算法考虑像素点  $(x_0, y_0)$  的值随着时间的变化过程, 该过程形成一个像素值的时间序列。  $\{X_1, X_2, \dots, X_t\}$  代表每个像素一段时间内的像素值, 其中  $X_t$  是  $t$  时刻像素点  $(x_0, y_0)$  的值。像素的概率密度函数由  $K$  个高斯分布函数的概率密度函数加权求和来表示:

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

其中:  $\mu_{i,t}$  是  $t$  时刻第  $i$  个高斯分布的均值,  $\omega_{i,t}$  是  $t$  时刻第  $i$  个高斯分布的权重,  $M_{i,t}$  是  $t$  时刻第  $i$  个高斯分布的协方差矩阵,  $\eta$  是第  $i$  个高斯分布密度函数:

$$\eta(X_t, \mu_{i,t}, M_{i,t}) = \frac{1}{(2\pi)^n |M_{i,t}|^{n/2}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T M_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (2)$$

其中:  $n$  是  $X_t$  的维数; 协方差  $M_{i,t} = \sigma_{i,t}^2 I$ , 其中  $I$  是单位矩阵;  $\sigma_{i,t}^2$  是  $t$  时刻第  $i$  个高斯分布的方差。

对于每一个新的像素  $X_t$ , 都会根据现有的  $K$  个高斯分布进行检查, 直到找到匹配。当像素点在高斯分布的 2.5 个标准差范围内, 则为匹配, 像素点被判定为背景点; 否则为不匹配, 像素点被判定为前景点。若匹配, 则按式 (3) 更新当前高斯分布的参数; 若不匹配, 则只更新权重  $\omega_{i,t}$ 。若  $K$  个高斯分布中没有一个是和当前像素值匹配, 把  $K$  个高斯分布中权重最小的分布用一个把当前值作为均值, 选取较大方差和较小权值的新的的高斯分布进行替换。

$$\begin{cases} \omega_{i,t} = (1 - \alpha)\omega_{i,t-1} + \alpha(R_{i,t}) \\ \mu_{i,t} = (1 - \beta)\mu_{i,t-1} + \beta X_t \\ \sigma_{i,t}^2 = (1 - \beta)\sigma_{i,t-1}^2 + \beta(X_t - \mu_{i,t})^T (X_t - \mu_{i,t}) \\ \beta = \alpha/\omega_{i,t} \end{cases} \quad (3)$$

其中:  $\alpha$  是学习率; 模型匹配时  $R_{i,t} = 1$ , 否则是 0;  $\beta$  是参数学习率。高斯分布按  $P = \omega_{i,t}/\sigma_{i,t}$  从大到小进行排序, 然后选择前  $B$  个高斯分布作为背景模型。其中:

$$B = \operatorname{argmin}_b \left( \sum_{i=1}^b \omega_{i,t} \geq T \right) \quad (4)$$

其中:  $T$  是阈值, 可由背景复杂度决定。

## 2 改进的混合高斯模型

本文算法首先使用混合高斯模型法, 利用混合高斯模型法能够处理逐渐变化的光照变化和具有微小重复运动的背景的特性, 有效地将图像像素区分为背景区域和粗略目标区域; 设置双级学习率并且根据像素变化频次动态调整组合权重, 适应背景图像实时动态变化, 降低动态背景的影响; 然后采用基于颜色特征和空间连续性的方法来消除光照变化和阴影的影响; 最后使用形态学处理得到准确的运动目标区域。

算法整体流程如图 1 所示。

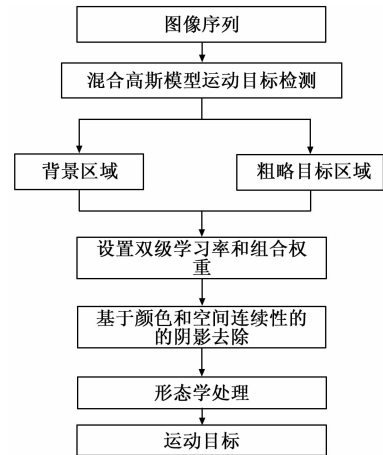


图 1 算法整体流程

### 2.1 动态背景去除

在动态背景 (摇曳的树木、喷泉、水波等) 下, 大多数目标检测方法的性能都会下降。因此, 构建一种能够忍受背景变化的方法是非常必要的。提出一种改进混合高斯模型, 通过将双级学习率和组合权重引入混合高斯模型, 使算法在动态背景下取得较好的结果。下边, 将详细介绍这种实现动态背景去除的方法。

#### 2.1.1 双级学习率

在传统的混合高斯模型中, 学习率  $\alpha$  固定, 一般由人为经验设定。学习率设置偏大, 容易导致学习震荡甚至发散, 而无法收敛; 学习率设置偏小, 容易导致学习速率慢, 收敛过于缓慢。学习速率的大小, 也会影响高斯建模的准确性<sup>[10]</sup>。为了优化学习率, 提出设置双级学习率。

实际上, 运动区域的变化比较大, 该区域应赋一个较高的学习率, 以加速模型收敛, 同时有助于消除前景中静止目标突然运动所留下的鬼影; 背景区域的变化比较小, 该区域应该赋一个较低的学习率, 以加强环境的抗干扰性, 并有效抑制噪声。

而本文中首先对图像序列进行混合高斯建模, 获取当前帧运动目标的粗略区域, 分为包含阴影的运动区域以及背景区域。其中包含阴影的运动区域认为是前景可信度高的区域, 该区域给一个较高的学习率, 定义为  $\alpha_{high}$ ; 背景区域认为是背景可信度高的区域, 该区域给一个较低的学习率, 定义为  $\alpha_{low}$ 。

### 2.1.2 组合权重

在传统的混合高斯模型中, 每个像素的高斯项首先按  $\omega_{i,t}/\sigma_{i,t}$  的降序排列, 然后背景模型由前 B 个最高权重值的高斯分布组成 (如式 (4))。当背景中有动态背景时, 前景/背景部分快速间断切换, 使得背景无法迅速收敛到两部分, 模型的性能会下降。

本文使用组合权重  $\pi_{i,t}$  来代替  $\omega_{i,t}$ 。关于  $\pi_{i,t}$  的计算, 要引入一个计数值  $d_{i,t}$ 。在给定的像素  $p$ , 对于每一个高斯分量, 我们都设置一个  $d_{i,t}$  (上限是  $K/2$ ), 其中  $i = 1, \dots, K$ 。  $d_{i,t}$  的初始值是 1。  $d_{i,t}$  反映了随着时间的变化每个高斯分量连续出现的次数。换句话说, 如果  $d_{i,t}$  在时间  $t$  和时间  $t+1$  的两个连续帧中匹配, 则  $d_{i,t}$  将增加 1; 否则, 将重置为 1。它遵循稳定的高斯分量将具有更高的持久性值。

组合权重  $\pi_{i,t}$  的计算遵循式 (5):

$$\pi_{i,t} = v_{i,t} \times \omega_{i,t} + (1 - v_{i,t}) \times v_{i,t} \times \omega_{i,t} \quad (5)$$

其中,  $v_{i,t}$  是持久性因子, 是衡量每个高斯分量稳定性的一个标准,  $v_{i,t} \in (0, 1)$ 。  $v_{i,t}$  越大, 则表示该高斯分量越稳定, 则它被判定为背景的可能性更大; 越小, 则表示该高斯分量越不稳定, 则它被判定为前景的可能性更大。动态背景  $v_{i,t}$  计算如式 (6):

$$v_{i,t} = \frac{2d_{i,t}}{K} \quad (6)$$

改进后的混合高斯模型中, 选取一个稍大的高斯分布数  $K$  ( $K = 7$ ), 以此来使持久性因子和组合权重的设置能发挥更大的作用。阈值  $T$  设置为 0.85, 它反映背景高斯成分在像素的整个概率分布中所占的最小比例。  $\alpha_{high}$  初始化为 0.05;  $\alpha_{low}$  初始化为  $10^{-5}$ 。

## 2.2 阴影去除

在上述步骤中, 混合高斯模型把图像序列粗略分为包含阴影的运动区域以及背景区域, 然后通过设置双级学习率和组合权重来降低动态背景的影响, 提取出的运动目标, 阴影还在影响目标检测的精度, 因此还需要去除阴影区域<sup>[11]</sup>。

首先, 应该对图像进行预处理, 以增强阴影和非阴影区域的对比度。然后将色彩和梯度信息相结合的方法对阴影区域进行检测。

在 RGB 空间, 亮度和色度特性没有分离。各分量之间具有高度的相关性, 并且对光照变化和噪声具有高度的敏感性。而 HSV 空间, 根据色调 (H)、饱和度 (S)、明度 (V) 三个分量来定义颜色, 它是一个近似均匀的感知色彩空间<sup>[12]</sup>, 与人类的色彩感知有很好的相关性。在均匀空间中, 测量到的色差 (欧几里得距离) 与人类对这种色差的感知成正比。相比之下, RGB 并不是统一的。因此采用 HSV 颜色

空间进行阴影消除。

### 2.2.1 图像预处理

第一步是在 HSV 颜色空间中对图像  $I$  的 V 分量进行处理, 以增强阴影和非阴影区域的对比度。设  $I = \{H^I, S^I, V^I\}$  是图像  $I$  的 HSV 分量, 处理后的 V 分量是:

$$V_m^I = \frac{V^I - \min(V^I)}{\max(V^I) - \min(V^I)} \quad (7)$$

其中:  $V_m^I$  是处理后的 V 分量,  $V^I$  是图像  $I$  的 V 分量的原始值。把原始的 V 分量用处理后的 V 分量  $V_m^I$  替代, 得到预处理后的图像  $I_m = \{H^I, S^I, V_m^I\}$ 。

### 2.2.2 基于颜色和空间连续性的阴影检测

提出的阴影检测方案的第二步是根据颜色和空间连续性对每个像素进行分类。最初, 我们通过比较 HSV 颜色空间中的当前帧和背景帧来检测阴影像素 (参见式 (8))。选择 HSV 颜色空间是因为它提供了色调和明度之间的充分分离。然后我们使用梯度信息<sup>[13]</sup>细化这些阴影像素估计, 将每个像素分类为阴影像素或非阴影像素。

一个像素  $p(i, j)$  如果满足式 (8), 则可以被认为是一个阴影像素。

$$\begin{cases} \beta_{low} \leq (F_p^V / B_p^V) \leq \beta_{high}, \text{ 且} \\ (F_p^S - B_p^S) \leq \tau_S, \text{ 且} \\ |F_p^H - B_p^H| \leq \tau_H. \end{cases} \quad (8)$$

其中:  $\{F_p^H, F_p^S, F_p^V\}$  和  $\{B_p^H, B_p^S, B_p^V\}$  分别为视频帧 F 和背景参考 B 中像素  $p(i, j)$  的 HSV 分量的值; 阈值  $\beta_{low}, \beta_{high}, \tau_S, \tau_H$  的设置将增加这个阶段的阴影像素检测的数量;  $\beta_{low}$  是当前帧和背景帧的 V 通道值的比率的最低阈值,  $\beta_{high}$  是当前帧和背景帧的 V 通道值的比率的最高阈值,  $\beta_{low}, \beta_{high} \in [0, 1]$ ;  $\tau_S$  是当前帧和背景帧的 S 通道的差值的阈值,  $\tau_H$  是当前帧和背景帧的 H 通道的差值的阈值,  $\tau_S, \tau_H \in [0, 255]$ 。更优的  $\beta_{low}, \beta_{high}, \tau_S, \tau_H$  参数值的选取, 可以达到更优的阴影检测效果。满足式 (8) 的所有像素组成一个区域, 该区域就是初步获取的阴影区域。

然后在根据梯度信息进一步得到更准确的阴影区域。计算初步获取的阴影区域中每个像素点的梯度幅值  $|\nabla_p|$  和梯度方向  $\theta_p$ 。设定梯度幅值阈值  $\tau_m$ ,  $\tau_m \in [0, 10]$ , 只有重要的像素点的梯度幅值  $|\nabla_p| \geq \tau_m$ , 更进一步考虑, 噪音的影响也降低了<sup>[14]</sup>。

对于初步获取的阴影区域中每一个像素  $p(i, j)$ , 其在当前帧 F 和背景参考帧 B 中梯度方向的差值计算如下:

$$\Delta\theta_p = \arccos \left[ \frac{\nabla_x^F \nabla_x^B + \nabla_y^F \nabla_y^B}{\{(\nabla_x^F + \nabla_x^B)(\nabla_x^F + \nabla_x^B)\}^{\frac{1}{2}} \{(\nabla_y^F + \nabla_y^B)(\nabla_y^F + \nabla_y^B)\}^{\frac{1}{2}}} \right] \quad (9)$$

在初步获取的阴影区域的每一个像素点  $p(i, j)$  处, 当前帧 F 和背景参考帧 B 的梯度方向相关性估计为:

$$c = \left\{ \sum_{p=1}^N H(\tau_a - \Delta\theta_p) \right\} / N \quad (10)$$

其中:  $N$  是初步获取的阴影区域中像素的个数;  $\tau_a$  是梯度方向阈值,  $\tau_a \in [0, \pi]$ ; 如果  $\Delta\theta_p \leq \tau_a$ ,  $H(\cdot) = 1$ , 否则是 0。  $c$  是在初步获取的阴影区域中当前帧 F 和背景参考帧 B 中梯度方

向相似的像素与总像素个数的比。验证了初步获取的阴影区域中满足  $c \geq \tau_c$  的像素, 这些像素组成的区域即为最终获取到的阴影区域。更优的  $\tau_m$ ,  $\tau_a$  和  $\tau_c$  的设置会使阴影检测率<sup>[15]</sup>更大化。如表 1 所示为阴影检测模块涉及的 7 个参数的设置。

表 1 采用的阴影检测方法使用参数的范围及建议值

参数	取值范围	参数取值
$\beta_{low}$	[0,1]	0.21
$\beta_{high}$	[0,1]	0.99
$\tau_S$	[0,255]	76
$\tau_H$	[0,255]	93
$\tau_m$	[0,10]	6
$\tau_a$	[0, $\pi$ ]	$\pi/10$
$\tau_c$	[0,1]	0.2

以上获取到的去除动态背景和阴影的目标图像的二值图像可能含有许多缺陷, 如噪音等。因此需要关注连续帧中发生的重大变化, 并消除较小的噪音。为了实现这一目标, 可以使用形态学图像处理。形态学处理是对像素进行对排序, 而不是对像素获取的数值进行排序。该算法有 2 种运算方式: 膨胀和腐蚀。

### 3 实验结果与分析

改进后的算法测试使用 Qt5.8.0 平台作为编译环境, 基于 OpenCV, 选择配置为 Inter 双核 CPU, 主频为 2.5GHz, 内存为 4GB 的 PC 机。

为了证明改进后的算法在处理动态背景、光照变化与阴影方面的优势, 使用数据集 CDW-2014<sup>[16]</sup> 进行对比实验。通过三组实验来进行对比分析: (1) 对“baseline”类别中的视频序列采用混合高斯模型法与改进后的算法进行实验对比; (2) 对“dynamic backgrounds”类别中的视频序列采用混合高斯模型法与改进后的算法进行实验对比; (3) 对“shadows”类别中的视频序列采用基于 HSV 颜色空间的阴影去除方法与改进后的算法进行实验对比。每帧图片大小 432 \* 288。

图 2 为第一组实验结果: 对“baseline”类别中的视频序列采用混合高斯模型法与改进后的算法进行实验对比。图 2 中 (a1) 为“baseline”类别中的“highway”序列中的第 445 帧原始帧, 室外光照强烈且树叶在风中抖动情况下, 车子在前进; (a2) 为传统混合高斯模型检测结果; (a3) 为改进后算法检测结果。由图 2 可知, 使用传统混合高斯模型, 对树叶抖动和投射阴影处理的效果差, 而改进后的方法中树叶抖动产生的虚影和车辆的阴影都消除了, 检测结果更准确。

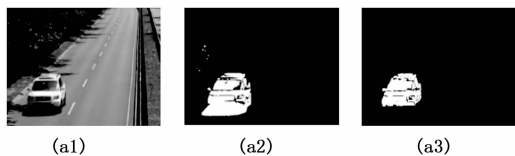


图 2 改进前后算法检测结果对比

图 3 为第二组实验结果: 对“dynamic backgrounds”类别采用混合高斯模型法与改进后的算法进行实验对比。图 3 中 (b1) 为“dynamic backgrounds”类别中的“fountain01”序列中的第 235 帧原始帧, 室外喷泉涌动; (b2) 为传统混合高斯模型检测结果; (b3) 为改进后的算法检测结果。由图 3 可知, 使用传统混合高斯模型, 对喷泉涌动等动态背景处理的效果差, 而改进后的方法消除了大多数假阳性像素点, 检测结果更准确。

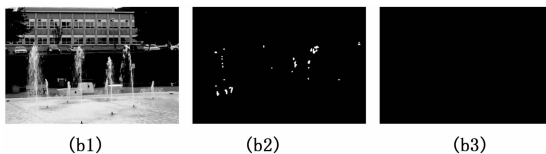


图 3 动态背景实例

图 4 为第三组实验结果: 对“shadows”类别采用混合高斯模型法与改进后的算法进行实验对比。图 4 中 (c1) 为“shadows”类别中的“peopleInShade”序列中的第 453 帧原始帧, 室外光照强烈情况下, 人在行走; (c2) 为基于 HSV 颜色空间的阴影去除方法检测结果; (c3) 为改进后的算法提取的目标图像。由图 4 可知, 使用基于 HSV 颜色空间的阴影去除方法, 对投射阴影处理的效果差, 而改进后的方法检测结果更准确。

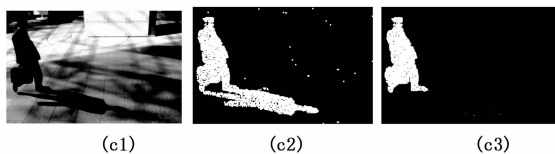


图 4 阴影去除实例

### 4 结束语

提出一种基于改进混合高斯模型与阴影去除的目标检测方法。该算法通过将双级学习率和组合权重引入混合高斯模型, 然后进一步利用基于颜色特性与空间连续性方法去除阴影; 最后通过形态学处理提取出准确的运动目标区域。对比实验表明: 文中所提检测方法能有效克服传统混合高斯模型在动态背景、光照变化和阴影的场景下检测精度低的问题。在动态场景下能够准确的检测运动目标, 消除树叶抖动、喷泉涌动或水波荡漾等虚假目标。在光照变化和阴影场景下, 改善了基于 HSV 颜色空间阴影去除方法的不足, 能够很好地对运动目标阴影进行消除。在未来的工作中, 将努力优化算法, 使其在高清视频监控实时环境中具有更好适应性。

#### 参考文献:

[1] Brutzer S, Hoferlin, Benjamin, Heidemann G. Evaluation of background subtraction techniques for video surveillance [A]. IEEE Conference on Computer Vision & Pattern Recognition [C]. IEEE Computer Society, 2011.