

基于滑动窗非负矩阵分解的运动目标检测方法

祝加祥¹, 胡鹏程¹, 何璇¹, 王莹冠²

(1. 安徽大学 计算机科学与技术学院, 安徽 合肥 230601;

2. 中国科学院上海微系统与信息技术研究所 中国科学院无线传感网与通信重点实验室, 上海 200050)

摘要:主要研究了智能视频监控系统中的运动目标检测算法, 试图将非负矩阵分解算法引入运动目标检测算法中, 通过非负矩阵分解算法对视频序列的背景进行建模, 使用背景差分法将当前视频帧图像与建立的背景模型比较获得运动目标。针对运动目标检测中基本非负矩阵分解批处理算法的不足, 提出一种基于滑动窗非负矩阵分解的运动目标检测算法。通过滑动窗处理控制非负矩阵分解模型中被分解矩阵的规模, 降低了算法的计算复杂度和空间复杂度, 并在一定程度上增加了模型的非记忆性。实验结果表明, 该算法能够更好地自适应背景模型的动态改变, 并且在视频场景中存在光照突变和较小运动目标时具有较好的检测效果。

关键词:非负矩阵分解; 运动目标检测; 背景差分; 滑动窗

中图分类号: TP391.9

文献标识码: A

文章编号: 1673-629X(2017)01-0020-05

doi: 10.3969/j.issn.1673-629X.2017.01.005

Moving Target Detection Method Based on Non-negative Matrix Factorization of Sliding Window

ZHU Jia-xiang¹, HU Peng-cheng¹, HE Xuan¹, WANG Ying-guan²

(1. School of Computer Science and Technology, Anhui University, Hefei 230601, China;

2. Key Lab of Wireless Sensor Network and Communication, Shanghai Institute of

Microsystem and Information Technology, Chinese Academy of Science, Shanghai 200050, China)

Abstract: The algorithms of detecting the moving objects in the intelligent video surveillance system are mainly studied. The Non-negative Matrix Factorization (NMF) algorithm is introduced into the moving target detection algorithm for modeling the background of a video sequences, and the background subtraction method is used to obtain the moving target by comparing the differences of the current video frame and the background model. In order to solve the problem of the NMF algorithm with a batch process, a moving target detection algorithm of NMF based on sliding window is put forward, which controls the size of the decomposed matrix in NMF matrix decomposition model by adjusting the length of the sliding window. The proposed algorithm can reduce the computation and space complexity, and to some extent, it can increase non-memory characteristic of the model. The experiments show that the proposed method can adaptively change the background model and has better detection effect when there is light change and small moving target in video scene.

Key words: non-negative matrix factorization; moving object detection; background difference; sliding window

0 引言

智能视频监控系统^[1]是智能视频分析中的一个重要应用, 同时也是计算机视觉领域中一直活跃的研究课题。智能视频监控对摄像机等视频采集设备获取的视频图像信息进行分析, 试图在一系列视频序列中检测出运动目标, 对目标进行识别, 跟踪并对其行为进行理解并做出相应反应, 最终取代传统视频监控中由值

班人员监控摄像机的现状。其中, 运动目标检测^[2]是智能视频监控系统的基础, 只有实时、准确、可靠地检测出运动目标, 才能保证后续目标定位、识别和行为理解等步骤的可靠性。

目前常用于智能视频系统中的运动目标检测方法有光流法^[3]、帧差法^[4]和背景差分法^[5]。帧差法利用相邻两帧或者多帧之间的差异来提取运动目标, 是一

收稿日期: 2015-06-03

修回日期: 2015-09-23

网络出版时间: 2017-01-04

基金项目: 安徽省科技攻关强警专项(1101b0403030); 中国科学院上海微系统与信息技术横向研发基金课题

作者简介: 祝加祥(1991-), 男, 硕士, 研究方向为图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20170104.1017.008.html>

种基于时间域的运动目标检测算法。帧差法比较简单,计算复杂度较低,易于实现,该算法对光照的变化不敏感并且对于环境的动态改变具有很强的鲁棒性。但是,帧差法不能提取出运动目标的所有像素点,只能检测出运动目标的部分轮廓,容易出现“空洞”现象。光流法是在适当的平滑性约束条件下,根据图像序列时空梯度估算运动场^[6],通过分析运动场的变化对运动目标和场景进行检测和分割。光流法的优点在于光流不仅携带了运动物体的运动信息,而且还携带了有关场景三维结构的丰富信息,它能够在不知道场景任何信息的情况下检测出运动目标。但是采用光流法计算耗时,实时性和实用性都较差。

使用背景差分法进行运动目标检测。首先对视频序列中的场景进行建模得到背景模型,然后通过计算当前帧与背景模型之间的差异来获得运动目标区域。当前帧图像与背景模型显著变化的区域认为有运动目标出现,变化区域的像素被标记为前景像素点以供下一步处理。背景差分法的关键在于背景模型的建立和更新,背景模型建立的好坏直接影响运动目标检测率的高低。目前,常用于背景模型的建立和更新算法有单高斯背景模型^[7](SGM)、混合高斯背景模型^[8](GMM)等。其中,SGM在多模态场景下存在背景模型不能更好地自适应、检测率不完整的问题;GMM能够在一定程度上解决视频场景中的多模态问题,但是随着高斯分布个数的增加,计算复杂程度也随之增加,并且在光照突变时效果也较差。

文中在非负矩阵分解批处理算法基础上,提出一种基于滑动窗非负矩阵分解的运动目标检测算法。通过滑动窗非负矩阵分解算法进行视频场景背景模型建立及更新,最终使用背景差分法获得运动目标。

1 非负矩阵分解算法原理与步骤

1.1 算法描述

非负矩阵分解算法^[9-11](NMF)可以看作是解决以下问题的算法:在给定非负矩阵 V 的情况下,寻找两个非负矩阵 W 和 H 使得:

$$V \approx WH \quad (1)$$

其中, $V = [v_1, v_2, \dots, v_m]$ 是 $n \times m$ 维的非负矩阵; $W = [w_1, w_2, \dots, w_r]$ 是 $n \times r$ 维的非负矩阵; $H = [h_1, h_2, \dots, h_m]$ 是 $r \times m$ 维的非负矩阵。

多元数据统计分析中,NMF可以被理解为:给定一系列的 n 维数据向量,这些向量作为列向量组成矩阵 $V \in R^{n \times m}$,这里的 m 是数据样本集的大小。矩阵 V 被分解成两个非负矩阵 $W \in R^{n \times r}$ 和 $H \in R^{r \times m}$ 的乘积。通常 r 的选择远远小于 n 和 m ,因此分解后的 W 和 H 矩阵的规模都小于原始矩阵 V 。该过程实际上是对原

始数据矩阵 V 的压缩过程。

更重要的是式(1)可以重写成两个列向量乘积的形式: $v \approx Wh$ 。其中, v 和 h 是矩阵 V 和 H 的对列。换句话说,每个数据向量 v 可以近似认为是 W 矩阵中列向量的线性组合,并且相应的权重是 h 。因此 W 可以看作对 V 进行线性组合近似的一组基向量, H 可以认为是对 V 进行线性组合近似的系数。所以在非负矩阵分解中, W 称作是基矩阵, H 称作是系数矩阵。由于原始大规模的数据矩阵 V 由相对规模较小的基矩阵 W 来表示,因此只有当基向量能够表示原始数据的隐含结构信息时才能得到有效的近似分解。

1.2 代价函数

为了能够得到式(1)中对于 V 的近似分解,首先需要定义一个代价函数来衡量分解的近似程度。可以使用两个非负矩阵的距离程度来构建一个代价函数用来衡量两个矩阵的近似程度。文中使用的度量公式是测量矩阵 V 和 WH 的欧氏距离的平方,即:

$$F(V \| WH) = \sum_{i,j} (V_{ij} - WH_{ij})^2 \quad (2)$$

矩阵 V 和 WH 的欧氏距离的平方大于等于0,并且仅当 $V = WH$ 时为0。

因此将NMF问题可以转化成在 W 和 H 非负约束下求解代价函数最小值的优化问题,问题描述如下:

在 $W, H \geq 0$ 非负约束下,寻找 W 和 H 使得 $F(V \| WH)$ 达到最小值,即

$$\begin{cases} \arg \min_{W, H} F(V \| WH) \\ \text{s. t. } W, H \geq 0 \end{cases} \quad (3)$$

1.3 乘性迭代规则

尽管代价函数 $F(V \| WH)$ 和 $D(V \| WH)$ 在单一求解 W 和 H 的情况下是凸优化^[12]问题,但是同时求解 W 和 H 的情况下并不是一个凸优化问题。所以通过设计一个算法来寻找式(3)的全局最小值是不切实际的。但是有很多数值优化算法可以用来求解上述问题的局部最小值。

梯度下降算法是用来解决上述优化问题最易于实现的方法,但是梯度下降算法收敛速度较慢。共轭梯度算法的收敛速度相比于梯度下降算法较快,至少在局部最小值区域很显著,但是^[13]实现起来更加复杂。另外基于梯度的优化算法的收敛性对步长的选择比较敏感,这在大型应用场景中很不方便。因此解决上述优化问题时,在速度和实现复杂度上采取了上述两种方法的折中:乘性迭代算法^[14],采用迭代规则分别更新 W 和 H ,在求取 W 的时候,固定矩阵 H ,然后利用 W 计算下一步的 H ,保证算法收敛快并且算法复杂度低。

以欧氏距离的平方作为代价函数时,将

$F(\mathbf{V} \parallel \mathbf{WH})$ 两边分别对 \mathbf{W} 和 \mathbf{H} 求偏导,得到:

$$\begin{cases} \frac{\partial F}{\partial \mathbf{H}_{au}} = -2((\mathbf{W}^T \mathbf{V})_{au} - (\mathbf{W}^T \mathbf{WH})_{au}) \\ \frac{\partial F}{\partial \mathbf{W}_{ia}} = -2((\mathbf{V} \mathbf{H}^T)_{ia} - (\mathbf{WHH}^T)_{ia}) \end{cases} \quad (4)$$

采用梯度下降数值优化方法,以梯度方向的负方向为下降方向,分别使用 η_{au} 和 δ_{ia} 作为梯度法的步长,也称为学习率,得到:

$$\begin{cases} \mathbf{H}_{au} \leftarrow \mathbf{H}_{au} + \eta_{au} [(\mathbf{W}^T \mathbf{V})_{au} - (\mathbf{W}^T \mathbf{WH})_{au}] \\ \mathbf{W}_{ia} \leftarrow \mathbf{W}_{ia} + \delta_{ia} [(\mathbf{V} \mathbf{H}^T)_{ia} - (\mathbf{WHH}^T)_{ia}] \end{cases} \quad (5)$$

梯度法中步长的选择对算法的收敛速度影响大,当式(5)中步长选择较大时,可能成之字形收敛甚至发散不能达到局部最小值;根据优化理论,式(5)中的学习率选取充分小的正数时,代价函数的值会逐渐减小,但是并不能保证每部迭代后的结果非负,乘性迭代规则正好能克服这个缺点。在乘性迭代规则中,只要 \mathbf{W} 和 \mathbf{H} 的初始值非负,在后续迭代过程中得到的 \mathbf{W} 和 \mathbf{H} 都为非负。为了得到乘性迭代公式,式(5)中学习率 η_{au} 和 δ_{ia} 分别选取为:

$$\begin{cases} \eta_{au} = \frac{\mathbf{W}_{au}}{(\mathbf{W}^T \mathbf{WH})_{au}} \\ \delta_{ia} = \frac{\mathbf{W}_{ia}}{(\mathbf{WHH}^T)_{ia}} \end{cases} \quad (6)$$

将式(6)代入式(5)中,得到使用欧氏距离的平方作为代价函数的 NMF 算法乘性迭代规则为:

$$\begin{cases} \mathbf{H}_{au} \leftarrow \mathbf{H}_{au} \frac{(\mathbf{W}^T \mathbf{V})_{au}}{(\mathbf{W}^T \mathbf{WH})_{au}} \\ \mathbf{W}_{ia} \leftarrow \mathbf{W}_{ia} \frac{(\mathbf{V} \mathbf{H}^T)_{ia}}{(\mathbf{WHH}^T)_{ia}} \end{cases} \quad (7)$$

2 基于滑动窗非负矩阵分解运动目标检测算法

在运动目标检测中,通常认为一个视频场景由背景和包含运动目标的前景组成,在摄像头固定的情况下,一段连续视频序列的背景基本变化不大,并且背景在视频序列中所占的比例较大,而前景所占比例较小。NMF 方法可以对数据的整体进行感知,基于此理论,可以通过 NMF 方法对视频序列的背景建模,然后通过背景差法比较当前帧与背景模型之间的差异提取出运动目标。

在实际的运动目标检测应用中,随着视频序列的增加,非负矩阵分解模型中被分解矩阵的规模增大,基本非负矩阵分解算法所需要的内存空间和计算时间都线性增加,这在实际应用中是不切实际的。针对基本非负矩阵分解的批处理算法的不足,以及视频序列在时间上的连续性,之前的视频内容对后续背景模型的

影响逐渐减小等特点,提出一种基于滑动窗非负矩阵分解的运动目标检测算法。算法步骤如下(其中矩阵运算采用 Matlab 语言中矩阵运算表示方式):

(1) 读取视频信息,视频的总帧数为 NumberOfVideo,初始化用于分解的矩阵 $\mathbf{V} = []$;

(2) 读取当前帧视频,判断当前帧序号 NumberOfFrame 是否大于预定的训练样本数量 NumberOfTrain,大于,跳转至步骤(4);

(3) 读取第 NumberOfFrame 帧中的灰度图像 frame,对 frame 进行预处理,并转换成列向量的形式 ColFrame,将当前 ColFrame 加入到被分解矩阵中,即 $\mathbf{V} = [\mathbf{V}, \text{ColFrame}]$;

(4) 对 \mathbf{V} 进行非负矩阵分解得到分解结果的基矩阵 \mathbf{W} 和系数矩阵 \mathbf{H} ,并进行重构得到矩阵 ReconstructedData = $\mathbf{W} * \mathbf{H}$;

(5) 获取重构后的矩阵 ReconstructedData 的最后一列作为背景模型信息,重构成图像形式 BackgroundFrame;

(6) 当前帧图像 frame 与 BackgroundFrame 进行比较提取出运动目标,并进行后续处理;

(7) 读取下一帧图像,如果 NumberOfFrame 大于 NumberOfVideo,算法结束;

(8) 读取第 NumberOfFrame 帧中的灰度图像 frame,对 frame 进行预处理,并转换成列向量的形式 ColFrame,将当前 ColFrame 以滑动窗的形式加入到被分解矩阵中,即 $\mathbf{V} = [\mathbf{V}(2:\text{end}, :), \text{ColFrame}]$,跳转至步骤(4)。

3 实验结果

为了验证基于滑动窗非负矩阵分解的运动目标检测方法的性能,使用 2 段视频序列进行实验测试,同时将其与帧差法、SGM 算法和 GMM 算法进行对比。实验平台为 Matlab 2012b,处理器为 Intel(R) Core(TM) i5-2450M CPU@2.5 GHz。采用的基本非负矩阵分解方法中,基矩阵 \mathbf{W} 和系数矩阵 \mathbf{H} 随机初始化为非负正数,迭代终止条件选取为不超过预定的迭代次数,在该实验中预定迭代次数选取为 10;滑动窗大小为 20;帧差法中的阈值设置为 25。SGM 算法中参数选取为: $K = 15$, $\alpha = 0.02$;GMM 算法中的参数选取为: $K = 3$, $\beta = 0.2$, $\tau = 0.8$, $c = 2.9$, $T = 15$, $\alpha = 0.02$ 。

实验 1 中使用的视频数据为 Matlab 2012b 中自带的道路交通监控视频,视频中的运动目标主要为行驶的车辆和光照的变化。对比测试结果如图 1 所示。其中,图(a)~(c)为视频序列中第 30、46、77 帧原始视频图像;图(d)~(f)为采用帧差法得到的前景图像;图(g)~(i)为使用 SGM 算法得到的前景图;图(j)~

(1)为使用 GMM 算法得到的前景图;图(m)~(o)为

使用基于滑动窗非负矩阵分解得到的前景图。

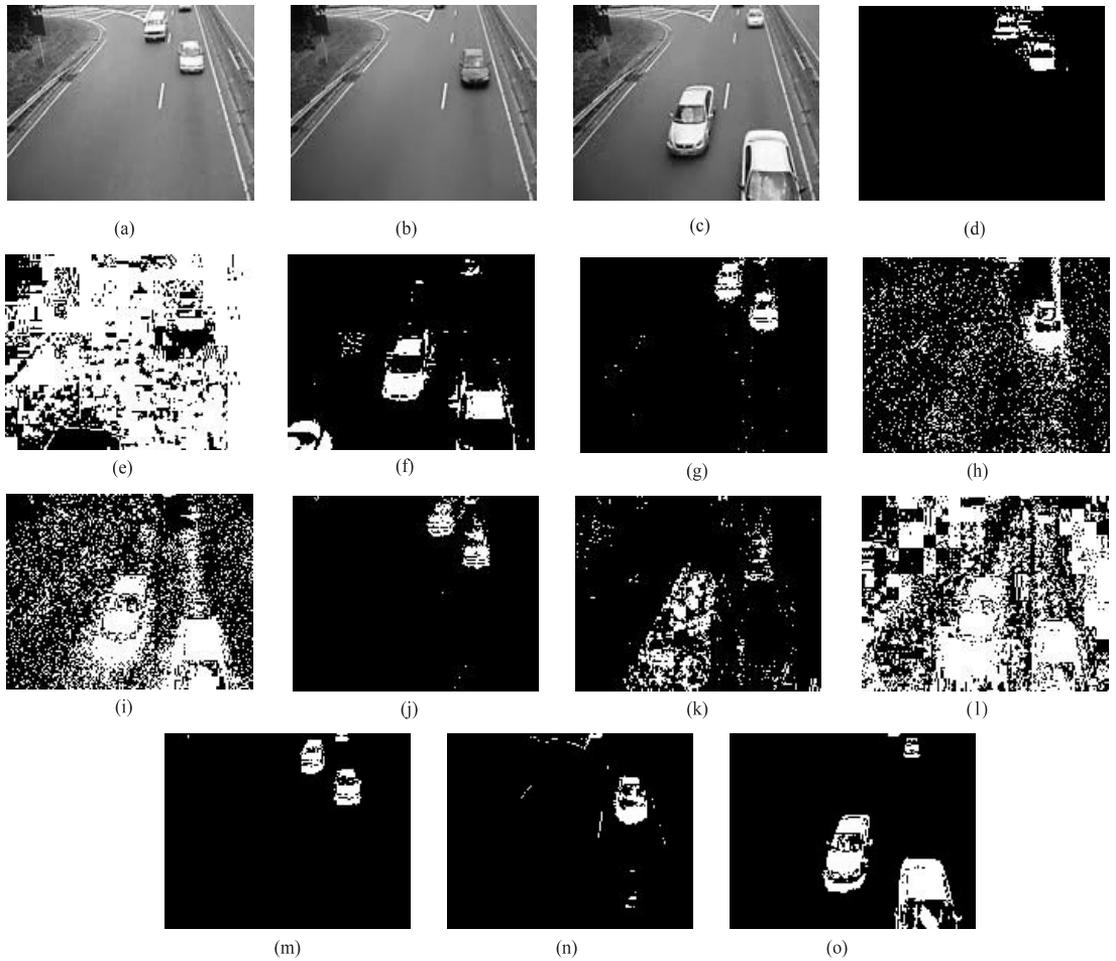


图1 实验1结果对比图

通过实验结果可知,在视频场景相对简单、摄像头固定的情况下,基于滑动窗非负矩阵分解的运动目标检测算法能够在一定程度上有效地检测出运动目标。例如在第34帧中,当运动目标较小时也能比较完整地检测出运动车辆。相较于帧差法,滑动窗 NMF 算法能够得到更完整的运动目标,并且能够有效地抑制光照变化产生的噪声。在第46帧和第77帧时,视频序列中光照发生整体性突变,SGM 和 GMM 均无法很好地抑制噪声,导致大量的背景信息被误判为前景目标。NMF 算法作为一种盲源分离算法,能够提取出观测数据中的主要成分,单帧图像的突变对模型的影响不大,基于滑动窗 NMF 的运动目标检测算法在整体光照发生突变时仍能正确地检测出运动目标。

为了更好地验证基于滑动窗非负矩阵分解的运动目标检测方法的性能,选取了一段视频场景比较复杂的视频序列对滑动窗 NMF 算法进行实验测试,并且将其与帧差法、SGM 算法和 GMM 进行对比。实验所选取的视频为一段合成的仿真视频,视频中运动目标包括行驶的车辆、跳跃的行人以及爬行的猫,并且视频场景中出现了树枝摇曳和光照闪烁变化等多模态现象。

图2为对比实验结果图。其中,图(a)~(c)为视频序列中第245、285、473帧原始视频图像;图(d)~(f)为采用帧差法得到的前景图像;图(g)~(i)为使用 SGM 算法得到的前景图;图(j)~(l)为使用 GMM 算法得到的前景图;图(m)~(o)为使用基于滑动窗非负矩阵分解得到的前景图。

从实验结果对比图来看,滑动窗 NMF 算法得到的运动目标比帧差法检测结果更加完整。在视频场景中摇曳的树枝和闪烁的灯光造成的运动目标噪声,基于滑动窗 NMF 的检测结果都能很好地处理,检测效果与 GMM 算法接近。因此,在复杂的环境下,基于滑动窗非负矩阵分解的运动目标检测方法能够在一定程度上抑制光照闪烁变化等多模态现象带来的噪声,具有较好的检测效果。

4 结束语

在非负矩阵分解批处理算法的基础上进行改进,使用滑动窗非负矩阵分解算法对视频场景中背景进行建模和更新,然后使用背景差分法获得运动目标。该算法有效降低了基本非负矩阵分解批处理算法的计算

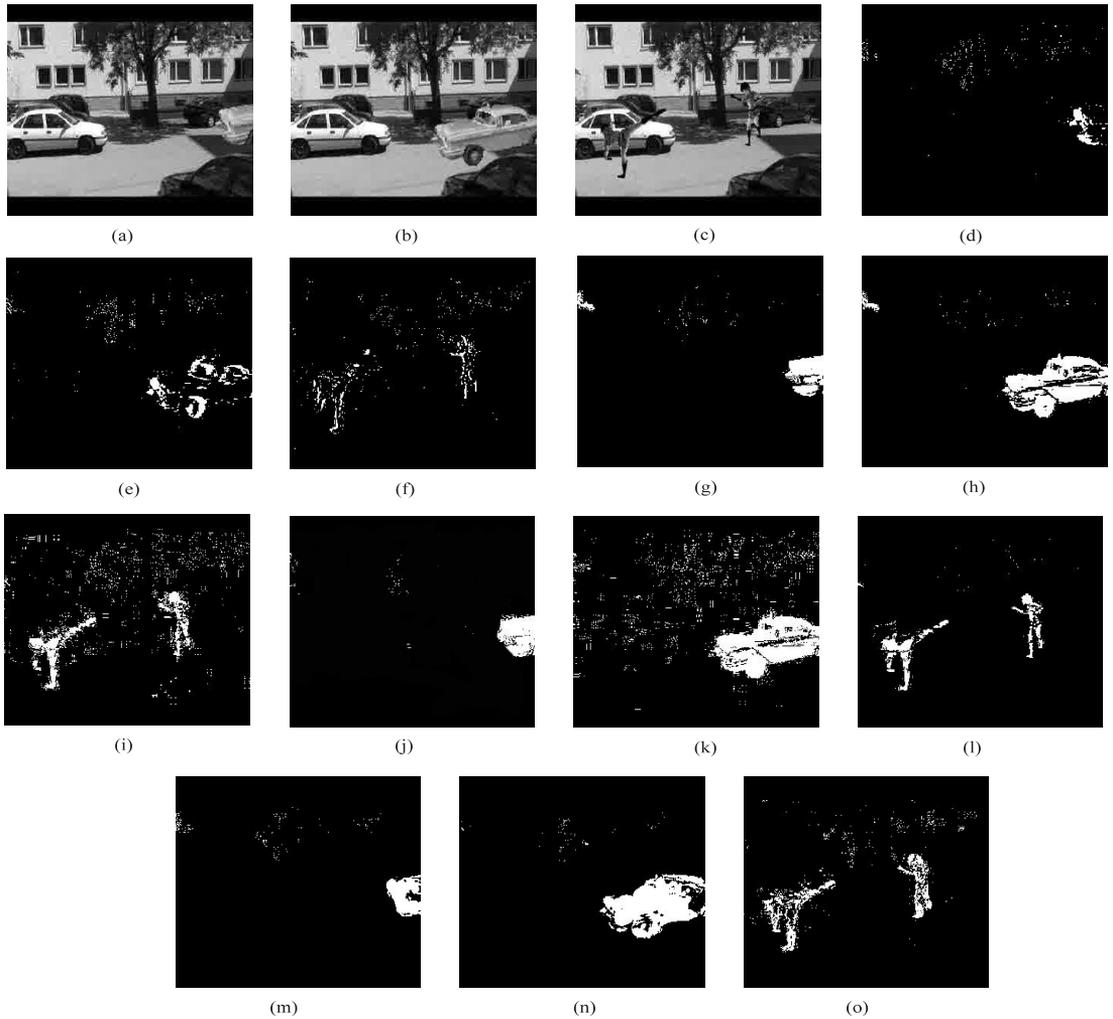


图 2 实验 2 结果对比图

复杂度和空间复杂度,通过对比实验可知,该算法能更好地自适应背景模型的动态改变,且在视频场景存在光照突变和较小运动目标时都具有较好的检测效果。

参考文献:

[1] 黄斯茜,李光. 智能视频监控系统运动目标检测技术研究综述[J]. 信息通信,2012(4):57-58.

[2] 侯宏录,李宁鸟,刘迪迪,等. 智能视频监控中运动目标检测的研究[J]. 计算机技术与发展,2012,22(2):49-52.

[3] 杨叶梅. 基于改进光流法的运动目标检测[J]. 计算机与数字工程,2011,39(9):108-110.

[4] 王孝艳,张艳珠,董慧颖,等. 运动目标检测的三帧差分算法研究[J]. 沈阳理工大学学报,2011,30(6):82-85.

[5] 汪冲,席志红,肖春丽. 基于背景差分的运动目标检测方法[J]. 应用科技,2009,36(10):16-18.

[6] Kaehler A, Bradski G. Learning OpenCV[M]. [s. l.]: O'Reilly Media, Inc., 2014.

[7] 王小平,张丽杰,常 佶. 基于单高斯背景模型运动目标检测方法的改进[J]. 计算机工程与应用,2009,45(21):118-120.

[8] 马义德,朱望飞,安世霞,等. 改进的基于高斯混合模型的运动目标检测方法[J]. 计算机应用,2007,27(10):2544-2546.

[9] 李乐,章毓晋. 非负矩阵分解算法综述[J]. 电子学报,2008,36(4):737-743.

[10] Albright R, Cox J, Duling D, et al. Algorithms, initializations, and convergence for the nonnegative matrix factorization[R]. USA: North Carolina State University, 2006.

[11] Yang S, Yi Z, Ye M, et al. Convergence analysis of graph regularized non-negative matrix factorization[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(9): 2151-2165.

[12] Boyd S, Vandenberghe L. 凸优化[M]. 王书宁,译. 北京:清华大学出版社,2013.

[13] 伏思华,姜广文,龙学军,等. 随机并行梯度下降图像匹配方法[J]. 实验力学,2013,28(6):663-668.

[14] Mao Yijun, Zhao Zhijun, Shang Junna. Blind source separation algorithms based on nonnegative matrix factorization using different learning rates[J]. Hans Journal of Wireless Communications, 2015, 5(5): 91-97.