

文章编号: 2095-2163(2022)08-0001-07

中图分类号: TP391

文献标志码: B

# 基于卫星视频影像的车辆提取算法评价

谢晓欢<sup>1</sup>, 陈仁喜<sup>2</sup>, 刘世杰<sup>1</sup>

(1 同济大学 测绘与地理信息学院, 上海 200092; 2 河海大学 地球科学与工程学院, 南京 210098)

**摘要:** 视频卫星是一种新型的对地观测卫星, 其以视频的形式获取拥有高时间分辨率的动态信息, 为运动目标的检测和分析提供数据。但卫星视频影像幅宽大, 运动车辆目标在影像上非常小, 颜色和纹理信息较弱, 并且视频动态场景复杂, 背景呈现像素级、甚至亚像素级的不均匀移动, 给运动目标检测算法提出了挑战。由于具有与传统地面监控视频不同的特点, 常用的动态目标检测方法在卫星视频上会表现出不同的性能。本文基于 SkySat 和吉林一号视频影像, 对 6 种常用的运动目标检测算法结果的精准率、召回率和  $F_1$  分数进行了对比评价, 分析了各算法的特点和优劣, 为卫星视频影像动态目标提取算法的选择和改进提供了有益参考。

**关键词:** 视频卫星; 运动目标检测; SkySat; 吉林一号

## Evaluation of vehicle extraction algorithms based on satellite videos

XIE Xiaohuan<sup>1</sup>, CHEN Renxi<sup>2</sup>, LIU Shijie<sup>1</sup>

(1 College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China;

2 School of Earth Science and Engineering, Hohai University, Nanjing 210098, China)

**[Abstract]** Video satellite is a new type of earth observation satellite, which obtains dynamic information with high time resolution in the form of videos and provides data for moving object detection and analysis. However, because of the wide view of satellite videos, the moving vehicles are very tiny, showing weak distinctive color and texture information. Under complex and dynamic scenarios, the background presents uneven movement at pixel level or even the sub-pixel-level, which poses a challenge to the moving target detection algorithm. Due to the different characteristics from traditional ground surveillance videos, the commonly used moving object detection algorithms will perform differently on satellite videos. Based on SkySat and Jilin-1 satellite videos, this paper compares and evaluates the precision, recall and  $F_1$  scores of six commonly used moving target detection algorithms. The algorithm evaluation in this paper also analyzes the characteristics, as well as pros and cons of each algorithm, providing a useful reference for algorithm selection and improvement in moving object extraction from satellite videos.

**[Key words]** satellite videos; moving objects detection; SkySat; Jilin-1

## 0 引言

随着遥感技术的迅速发展, 对地观测卫星的时间分辨率得到有效提升, 数据获取能力也在不断提高<sup>[1]</sup>。视频卫星是一种新型的对地观测卫星, 以视频的形式获取拥有高时间分辨率的动态信息, 用于运动目标检测和分析<sup>[2]</sup>。

运动目标检测是图像处理与计算机视觉领域的一个重要研究课题。近几十年来, 研究者们提出了许多基于不同理论的运动目标检测方法。

20 世纪 70 年代末, Jain 等人<sup>[3]</sup>提出使用帧间差分法 (Frame Difference) 来提取运动目标。1997 年, Wren 等人<sup>[4]</sup>提出利用单高斯模型 (Single

Gaussian Model) 进行背景建模, 进而检测运动目标。1999 年, Stauffer 等人<sup>[5]</sup>提出了经典的自适应混合高斯模型 (Gaussian Mixture Models, GMM), 使用加权的混合高斯分布, 通过参数来调整背景模型。此后, 许多研究者在高斯混合模型的基础上进行改进, 例如: 2004 年, Zivkovic 等人<sup>[6]</sup>提出了混合高斯模型个数自适应的算法。2006 年, Zivkovic 等人<sup>[7]</sup>从贝叶斯的角度进行改进使其完全自动地适应场景, 检测效果得到改善。

然而, 现实条件下的图像背景变化快速且复杂, 有时候不符合高斯分布, 故而利用高斯分布的背景建模方法具有许多局限性。

2000 年, Chen 等人<sup>[8]</sup>提出了同步分割和类参

**基金项目:** 中国科学院太空应用重点实验室开放基金 (LSU-KFJJ-2018-10); 国家自然科学基金 (41471276)。

**作者简介:** 谢晓欢 (1998-), 女, 硕士研究生, 主要研究方向: 摄影测量与遥感图像处理; 陈仁喜 (1976-), 男, 博士, 副教授, 硕士生导师, 主要研究方向: 遥感与图像处理、机器学习; 刘世杰 (1982-), 男, 博士, 教授, 博士生导师, 主要研究方向: 航天测绘遥感与深空探测。

**通讯作者:** 陈仁喜 Email: chenrenxi@hhu.edu.cn

收稿日期: 2022-02-18

数估计 (Simultaneous Partition and Class Parameter Estimation, SPCPE) 算法, 使用统计学的方法将分割参数和分类参数变量进行联合估计, 用于识别目标及其空间关系。为了得到更精确的提取结果, 2003年, Zhang 等人<sup>[9]</sup>提出了基于 SPCPE 算法的自适应背景学习 (Adaptive Background Learning, ABL) 算法。在此基础上, 2011年, Ng 等人<sup>[10]</sup>提出了一种在动态场景中更新背景模型的自适应选择性背景学习 (Adaptive Selective Background Learning, ASBL) 算法, 增加了动态场景下算法的鲁棒性且降低了计算量。

2009年, Barnich 等人<sup>[11]</sup>提出了一种名为视觉背景提取 (Vision Background Extractor, ViBe) 的新背景建模算法。2011年, 通过将一些经典的运动目标检测算法与视觉背景提取算法进行比较, 实验证明了此算法的高效率<sup>[12]</sup>。

近年来, 基于统计与分析的运动目标检测方法快速发展, 由 Cover 等人<sup>[13]</sup>提出的经典的机器学习算法 K-近邻 (K-Nearest Neighbor, KNN) 算法也被用于运动目标检测。

尽管运动目标检测方法一直都在不断地改进和发展, 但目前没有任何一种方法能够适应所有场景。

卫星视频可以方便地监测城市规模的动态场景, 实现许多潜在的应用, 研究中的关键任务则在于提取和跟踪卫星视频中的运动车辆。虽然目前的运动目标检测算法在传统地面监控视频中获得了较好的效果, 但卫星视频具有不同的特点<sup>[14]</sup>:

(1) 卫星视频视野广阔, 运动目标非常小, 通常只占据几个像素, 没有任何明显的颜色或纹理, 加上噪声干扰, 原有的技术无法有效地检测这种微型车辆。

(2) 卫星视频画面覆盖的范围大, 提供的动态场景复杂, 包括道路、建筑、植被以及各种各样的交通条件等。

(3) 由于卫星的运动, 卫星视频帧的背景呈现

亚像素级的不均匀移动。卫星视频帧本质上是卫星平台的复杂三维运动在二维平面上的投影, 因此卫星视频的相对运动非常复杂。另外, 由于视频卫星距离地球十分遥远, 在连续的视频帧之间观察到的移动非常缓慢, 总是亚像素级的。

目前还没有在卫星视频中检测微小移动车辆的技术, 适用于传统场景的运动检测算法不一定适用于卫星视频。因此, 本文以卫星视频影像为对象, 对现有经典运动目标检测算法进行定性与定量的评价。

## 1 运动目标检测算法

运动目标检测是指从图像或视频序列中检测出发生动态变化的区域, 并将运动目标 (前景) 与背景分离。常用的方法主要有帧间差分法 (Frame Difference)、背景建模法 (Background Subtraction) 和光流法 (Optical Flow) 等<sup>[15]</sup>。

### 1.1 帧间差分法

帧间差分法是将相邻 2 帧或多帧图像对应的像素值相减, 得到差分图像, 如果对应像素值的变化大于预先设置的阈值, 则认为此处是前景, 反之则是背景<sup>[16]</sup>。主要分为两帧差分法和三帧差分法。对此拟做探讨分述如下。

两帧差分法的算法示意如图 1 所示。两帧差分法是设相邻的 2 帧图像第  $t-1$  帧图像位置  $(i, j)$  处的像素值为  $I_{t-1}(i, j)$ , 第  $t$  帧图像位置  $(i, j)$  处像素值为  $I_t(i, j)$ , 两者之差小于设定阈值时, 判定为背景像素, 否则判定为前景像素<sup>[17]</sup>, 用到的数学公式可写为:

$$F_t(i, j) = \begin{cases} 1 & |I_t(i, j) - I_{t-1}(i, j)| \geq T \\ 0 & |I_t(i, j) - I_{t-1}(i, j)| < T \end{cases} \quad (1)$$

其中,  $F_t(i, j)$  表示获得的二值差分图像;  $(i, j)$  表示像素所在位置;  $I_t(i, j)$  和  $I_{t-1}(i, j)$  表示像素值; 二值图像中 1 (白色) 表示前景, 0 (黑色) 表示背景;  $T$  为设定阈值。

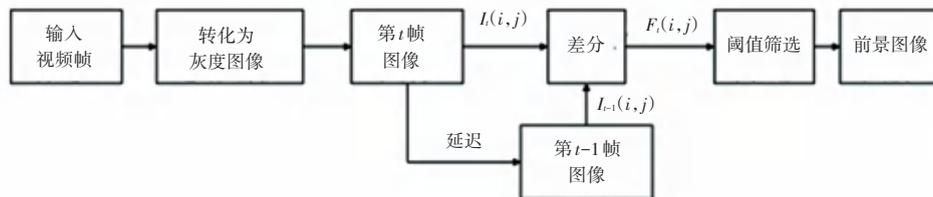


图 1 两帧差分算法示意图

Fig. 1 Schematic diagram of two-frame difference algorithm

三帧差分法的算法示意如图 2 所示。三帧差分法是对连续的 3 帧图像, 用第  $t$  帧图像的像素值减去第  $t-1$  帧图像的像素值, 得到二值图像  $G_1(i, j)$ , 再用第  $t+1$  帧图像的像素值减去第  $t$  帧图像的像素

值, 得到二值图像  $G_2(i, j)$ , 将二者进行“与”运算, 得到三帧差分图像  $F(i, j)$ <sup>[18]</sup>, 研究推导得到的数学公式分别如下:

$$G_1(i, j) = \begin{cases} 1 & |I_t(i, j) - I_{t-1}(i, j)| \geq T \\ 0 & |I_t(i, j) - I_{t-1}(i, j)| < T \end{cases} \quad (2)$$

$$G_2(i, j) = \begin{cases} 1 & |I_t(i, j) - I_{t+1}(i, j)| \geq T \\ 0 & |I_t(i, j) - I_{t+1}(i, j)| < T \end{cases} \quad (3)$$

$$F(i, j) = G_1(i, j) \otimes G_2(i, j) = \begin{cases} 1 & G_1(i, j) \cap G_2(i, j) \neq 0 \\ 0 & G_1(i, j) \cap G_2(i, j) = 0 \end{cases} \quad (4)$$

其中,  $F(i, j)$  为得到的三帧差分二值图像;  $G_1(i, j)$  和  $G_2(i, j)$  为两帧差分二值图像;  $(i, j)$  表示像素所在位置;  $I_{t-1}(i, j)$ 、 $I_t(i, j)$  和  $I_{t+1}(i, j)$  表示像素值; 二值图像中, 1(白色) 表示前景, 0(黑色) 表示背景;  $T$  为设定阈值。

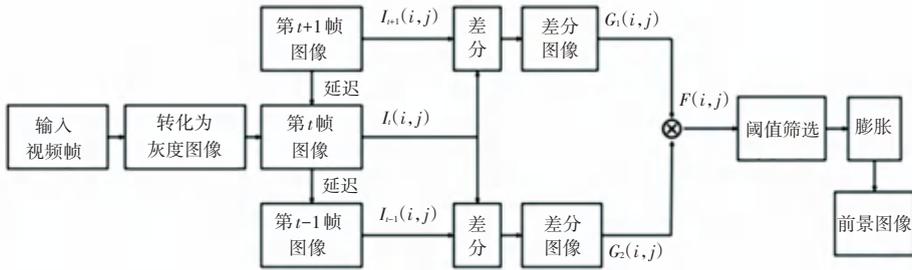


图 2 三帧差分算法示意图

Fig. 2 Schematic diagram of three-frame difference algorithm

1.2 高斯混合模型

高斯混合模型(GMM)算法示意如图 3 所示。GMM 算法为每一个像素都建立加权的混合高斯分布, 基于高斯模型的期望和方差判断哪一个高斯模型更可能对应当前的像素, 不符合各个高斯模型的像素被判断为前景, 符合的被当作背景<sup>[7]</sup>。大致分为 4 步: 模型初始化、模型学习、增加或替换高斯分量以及判断背景<sup>[19]</sup>。

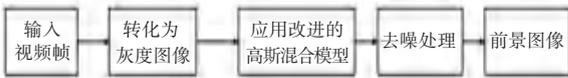


图 3 高斯混合模型算法示意图

Fig. 3 Schematic diagram of Gaussian mixture models algorithm

首先, 提取在视频序列  $N$  帧含有运动目标的图像, 逐个像素求平均值, 并将结果作为背景; 视频图像像素值序列  $\{X_1, X_2, \dots, X_t\}$  用  $K$  个高斯分布描述, 每个高斯分布都有权重  $\omega_{i,t}$ , 将每个像素的新像素  $X_t$  与  $K$  个高斯分布逐个匹配, 若某像素的  $K$  个高斯分布与当前像素值  $X_t$  都不匹配, 则剔除权重最低的高斯分布, 引入一个新的高斯分布, 同时更新  $K$  个高斯分布的权重。将  $K$  个高斯分布按照  $\omega_{i,t} / \sigma_{i,t}$  从大到小排列, 用前  $b$  个高斯分布表示背景模型<sup>[6]</sup>, 对此可表示为:

$$B = \operatorname{argmin}_b \left( \sum_{i=1}^b \omega_{i,t} > T \right) \quad (5)$$

其中,  $\operatorname{argmin}_b$  表示加权求和大于  $T$  的最小值;  $T$  是判定像素是否属于背景的最低阈值, 即背景阈值。

1.3 K-近邻算法

K-近邻(KNN)算法中, 选取训练样本集, 样本

集中每一个数据都有其所属分类, 通过测量待分类样本到邻近的训练样本集中样本之间的距离, 找到特征空间中距离最近的几个样本, 根据各类样本数目多少来决策分类, 即待分类样本所属的类别就是邻近样本中出现次数最多的那个分类<sup>[13]</sup>。设有  $N$  个样本分布到  $c$  个类中, 即为:  $\omega_1, \omega_2, \dots, \omega_c$ , 每个类有  $N_i$  个样本, 其中  $i = 1, 2, \dots, c$ 。该算法主要分为 4 步:

- (1) 计算待分类样本与各个邻近样本之间的距离。
- (2) 选取距离最近的近邻。
- (3) 确定近邻出现的频率。
- (4) 预测分类。

K-近邻算法的分类过程示意如图 4 所示。图 4 中,  $X_0$  为待分类样本,  $\omega_1, \omega_2, \omega_3$  为 3 个不同的分类, 箭头表示待分类样本到各近邻之间的距离。

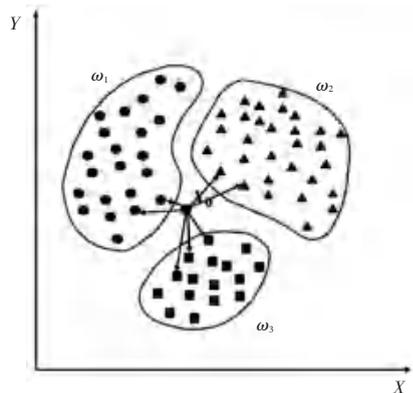


图 4 K-近邻算法分类过程示意图

Fig. 4 Schematic diagram of K-Nearest Neighbor classification

## 1.4 自适应选择性背景学习算法

自适应选择性背景学习算法(ASBL)根据每一帧图像不同部分有不同的动态,为每个像素分配不同的学习率,逐像素更新背景模型,计算当前帧与背景模型的差异,与自适应阈值进行比较,分割前景和背景<sup>[10]</sup>。

设像素的学习率为  $\alpha_{ad,t}(x,y)$ ,由2个加权参数  $\alpha_1, \alpha_2$  决定,计算公式见式(6):

$$\alpha_{ad,t}(x,y) = \omega_1 \alpha_1 + \omega_2 \alpha_2 \quad (6)$$

其中,  $\omega_1$  和  $\omega_2$  分别为  $\alpha_1$  和  $\alpha_2$  的权重,且  $\omega_1 + \omega_2 \leq 1$ 。

$\alpha_1$  取决于背景模型像素和当前帧之间的差异,差异越大,则  $\alpha_1$  越小;  $\alpha_2$  取决于像素被分类为背景像素所持续的时间,持续时间越长,则  $\alpha_2$  越大。

## 1.5 视觉背景提取算法

视觉背景提取算法(ViBe)的流程:对于每个像素点,为其建立一个样本集,选取该像素点历史像素值和其邻近点的像素值作为样本集的采样值;将每一个新的像素值与样本集进行比较,预测新的像素值分属于前景点、还是背景点<sup>[11]</sup>。

ViBe 算法模型像素样本集如图5所示。图5中,用  $v(x)$  表示在位置  $x$  处的像素值,随着时间的推移,该处的背景像素值构成一个大小为  $N$  的背景样本集  $M(x)$ , 数学表达式可写为:

$$M(x) = \{v_1, v_2, \dots, v_N\} \quad (7)$$

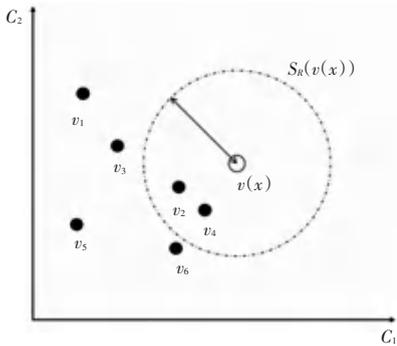


图5 ViBe 算法模型像素样本集

Fig. 5 Schematic diagram of vision background extractor

以当前像素值  $v(x)$  为圆心,  $R$  为半径,在空间坐标系下构造一个圆球  $S_R(v(x))$ , 这个球体与背景样本集  $M(x)$  的交集的元素个数记为式(8):

$$\{S_R(v(x)) \cap \{v_1, v_2, \dots, v_N\}\} \quad (8)$$

设定一个阈值  $\min$ ,如果上式大于  $\min$ ,那么当前值  $v(x)$  就更新为背景,存入背景样本集  $M(x)$  中。

对模型初始化,计算新像素点到背景样本集  $M(x)$  中样本的距离,如果距离小于  $R$ , 则该点为近似样本点,如果近似样本点的个数多于  $\min$ ,则新的

像素点被判断为背景。遵循时间衰减原则和空间一致性原则,对模型进行更新。

## 2 算法评价

### 2.1 评价指标

将运动目标的检测看作是对每个像素的二分类,结果有4种,分别是:真正类(True Positives, TP)、真负类(True Negatives, TN)、假负类(False Negatives, FN)、假正类(False Positives, FP)。

评价算法性能常用3个参数,即:精准率、召回率和  $F_1$  分数。文中拟展开剖析论述如下。

(1)精准率(Precision)。主要反映检测的准确性,即被预测准确的正类占有所有预测正类的比例,又称为查准率,定义公式具体如下:

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

(2)召回率(Recall)。主要反映检测的全面程度,即被预测准确的正类占有所有实际正类的比例,又称为查全率,定义公式具体如下:

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

(3)  $F_1$  分数。精准率和召回率在不同的应用场景下面的关注点是不同的,因而总会出现矛盾,而  $F_1$  分数( $F_1$ -Score)采用了调和平均数的方式来综合考虑,因此  $F_1$  分数能更好地衡量正样本的预测效果,  $F_1$  分数越大,预测效果越好。定义公式具体如下:

$$F_1 - Score = \frac{2TP}{2TP + FP + FN} \quad (11)$$

### 2.2 Ground-Truth 图

由于视频分帧后得到的帧数甚多,并且视频帧中的运动车辆所占像素较小,相邻帧之间目标的运动多是亚像素级,很难逐帧地人工标注视频范围内的所有车辆,权衡标注工作量与标注精度,随机选取样本帧进行标注。

首先,将运动车辆检测结果的二值图像与对应的单帧图像叠加,同时将真实运动目标在单帧图像上用红色来做标记;其次,人工删除错误的标记结果、添加未标记的目标并补齐“空洞”,制作 Ground-Truth 图用于实验结果评价。

### 2.3 噪声处理

由于噪声的存在对评价结果会产生较大的影响,因此在进行定量评价之前需要进行处理。首先,设置一个面积阈值,将小于面积阈值的“目标”删

除。为了防止误删,面积阈值的设置需要根据实际运动目标在图像上连通域的大小来进行设置。

由于空间分辨率的大小会影响卫星视频中的运动车辆所占像素的多少。假设视频卫星的分辨率为  $a$  m, 运动目标的最小尺寸为  $w$  m  $\times$   $l$  m, 那么检测得到的区域大小  $d$  至少应该占据  $\frac{aw}{\xi a \varnothing} \times \frac{\ell l}{\xi a \varnothing}$  个像素点。

在卫星视频中最小的运动车辆约为 2.5 m  $\times$  3 m, 假定分辨率为 1.1 m 左右 (SkySat-1 视频卫星的分辨率为 1.1 m), 则普通车辆的连通域面积大约为 4~80 个像素点。如果检测到的目标所占像素小于 4, 则判断为噪声, 使用形态学处理的方法进行消除, 提高运动车辆提取的准确度<sup>[20]</sup>。

为了得到较为客观的评价结果, 本文实验的面积阈值根据算法不同酌情选择, 尽可能使每一种算法都达到较好的评价结果。

### 3 实验与分析

随机选取视频帧检测结果作为评价样本, 设置面积阈值  $min-size$ , 将小于面积阈值的“目标”认为是噪声进行删除, 与相应的 Ground-Truth 图进行对比, 正确的运动目标标记为红色, 错误的标记为黄色, 未检测到的标记为蓝色, 输出图像与评价指标数值。

本文基于 6 种算法的结果进行评价, 包括自适应选择性背景学习法 (ASBL)、两帧差分法 (Diff2)、三帧差分法 (Diff3)、K-近邻算法 (KNN)、高斯混合模型 (GMM) 和视觉背景提取法 (ViBe), 从定性和定量两个角度评价算法的性能。

#### 3.1 SkySat-1 卫星视频实验与结果分析

2013 年 11 月, 美国 Skybox Imaging 公司发射了 SkySat-1 卫星, 是世界上首颗亚米级视频卫星, 可拍摄空间分辨率为 1.1 m 的黑白卫星视频。本实验使用的卫星视频拍摄于 2014 年 3 月 25 日, 地点为美国拉斯维加斯 (Las Vegas) 地区局部。选取部分卫星视频帧作为检测底图, 6 种车辆检测结果如图 6 所示。

根据图 6 车辆检测结果图定性对比得出, 不同的方法都存在一定程度上的漏检和误检。误检位置一般集中在建筑物的顶部, 主要是由于光线变化、建筑物阴影和高层建筑物的顶端位移, 产生了大量的虚假目标。漏检目标多是小型车辆, 在图像上所占连通域较小, 一方面运动检测算法对微小的运动目标不敏感, 另一方面, 形态学处理过程中可能会将其误判为噪声进行删除。



图 6 SkySat-1 卫星视频 6 种算法车辆检测结果  
Fig. 6 Vehicles detection results of six algorithms in SkySat-1 satellite videos

误检方面, 自适应选择性背景学习法和视觉背景提取算法得到的结果中存在较大量的虚假目标, 尤其是自适应选择性背景学习法中存在连片的虚假目标区域, 误检率高。漏检方面, 两帧差分法和视觉背景提取算法漏检目标相对较多, 而自适应选择性背景学习法和 K-近邻算法漏检的目标极少, 不过各方法之间的差异并不大, 漏检比例整体可以控制在一个较低的水平。

6种算法车辆检测定量对比见表1,算法的精准率整体较低,只有K-近邻算法的精准率可以达到80%,而其余多数算法精准率都低于70%,这说明各种算法都存在了较多的误检目标。除了两帧差分法

和视觉背景提取算法,其余算法的召回率都可以达到70%以上,漏检目标较少。总体来看,K-近邻算法的 $F_1$ 分数可以达到0.8以上,检测精确度较高,两帧差分算法检测效果最差,整体精度不高。

表1 SkySat-1卫星视频6种算法车辆检测定量对比

Tab. 1 Quantitative comparison of six algorithms for vehicles detection in SkySat-1 satellite videos

评价指标	选择性背景学习法	两帧差分法	三帧差分法	K-近邻算法	高斯混合模型	视觉背景提取法
Precision	0.585 1	<b>0.363 9</b>	0.698 2	<b>0.810 9</b>	0.663 1	0.618 7
Recall	<b>0.884 8</b>	<b>0.514 5</b>	0.714 7	0.883 5	0.755 7	0.671 9
$F_1 - Score$	0.704 4	<b>0.426 3</b>	0.706 4	<b>0.845 6</b>	0.706 4	0.644 2

### 3.2 吉林一号卫星视频实验与结果分析

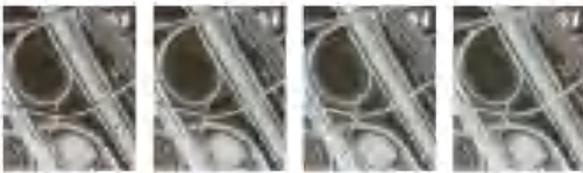
2015年10月,中国长光卫星公司发射了吉林一号视频卫星组,包含2颗视频卫星灵巧01和灵巧02,可拍摄空间分辨率为1.13 m的彩色卫星视频。本实验使用的卫星视频拍摄于2017年,地点是美国亚特兰大(Atlanta)地区局部。选取部分卫星视频帧作为检测底图,6种车辆检测结果如图7所示。



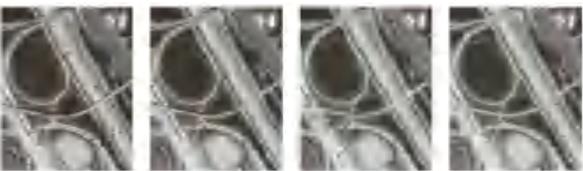
(a) 选择性背景学习法



(b) 两帧差分法



(c) 三帧差分法



(d) K-近邻算法



(e) 高斯混合模型



(f) 视觉背景提取法

图7 吉林一号卫星视频6种算法车辆检测结果

Fig. 7 Vehicles detection results of six algorithms in Jilin-1 satellite videos

根据图7车辆检测结果图定性对比得出,不同的方法都存在一定程度上的漏检和较少部分的误检。

误检方面,自适应选择性背景学习法、两帧差分法和视觉背景提取算法得到的结果中存在少量的虚假目标。漏检方面,由于截取的视频图像并不清晰,图像序列中的运动目标面积很小,6种算法都出现了比较多的漏检目标,其中高斯混合模型算法和视觉背景提取算法漏检目标相对较多。

吉林一号卫星视频6种算法车辆检测结果见表2。定量分析可知,K-近邻算法和混合高斯模型算法精准率可到达90%以上,误检较少。自适应选择性背景学习法和K-近邻算法的召回率较高,在80%以上,漏检比例不高。K-近邻算法的 $F_1$ 分数可以达到0.8以上,检测精确度较高,两帧差分算法的 $F_1$ 分数较低,检测效果相对较差。

表2 吉林一号卫星视频6种算法车辆检测定量对比

Tab. 2 Quantitative comparison of six algorithms for vehicles detection in Jilin-1 satellite videos

评价指标	选择性背景学习法	两帧差分法	三帧差分法	K-近邻算法	高斯混合模型	视觉背景提取法
Precision	0.784 9	<b>0.634 2</b>	0.712 2	<b>0.916 4</b>	0.909 7	0.781 3
Recall	0.805 8	<b>0.542 9</b>	0.552 3	<b>0.837 1</b>	0.688 6	0.648 5
$F_1 - Score$	0.795 2	<b>0.585 0</b>	0.622 2	<b>0.875 0</b>	0.783 9	0.708 7

## 4 结束语

本文应用了多种经典的运动目标检测算法,利用卫星视频进行了运动车辆检测实验,并通过定性对比和定量分析,将不同算法之间的性能差异直观地显示出来。

本文实验一定程度上证明了现有经典算法的有效性。根据运动车辆提取结果,可以看出几种算法对于运动车辆的捕捉、提取具有不同的效果。其中,K-近邻算法表现优秀,检测结果相对准确,两帧差分法整体检测效果较差,对噪声极其敏感,误检率较高。

目前,利用监控视频对运动车辆进行检测的应用已相对成熟,但监控视频无法宏观地展示大视野场景车流的动态变化,而卫星视频结合遥感技术可以为运动车辆的提取和分析提供更加宏观、多源的数据,还可以提供更加细致的地物信息用于辅助,得到全面的分析结果<sup>[21-22]</sup>。

本文对几种运动检测算法在卫星视频中的应用做了综合的对比和分析,发现以下几点问题:

(1)由于空间分辨率的限制,卫星视频中的运动车辆都非常小,通常只占据几个像素,缺乏纹理信息,对运动车辆的提取造成了一定的困难。

(2)复杂的环境条件(例如:阴影、光照等)和场景变化加大了运动车辆的提取难度,影响检测的精确度,造成误检和漏检的情况。

(3)传统的运动目标检测算法受噪声和应用场景限制大,部分场景下无法得到优良的效果,需要进一步研发新的算法。

许多研究者在传统算法的理论基础上,不断地进行改进优化,结合新兴的技术改进传统的算法,使得检测效果不断加强,但是对于现实中复杂多变的场景的处理仍存在许多不足。只有不断地优化算法或者结合现有新兴技术提出新的算法,才能为卫星视频影像中运动车辆的检测提供更多的可能。

## 参考文献

[1] 袁益琴,何国金,王桂周,等.背景差分与帧间差分相融合的遥感卫星视频运动车辆检测方法[J].中国科学院大学学报,2018,35(01):50-58.

[2] 李德仁,王密,沈欣,等.从对地观测卫星到对地观测脑[J].国际太空,2014,9(02):50-56.

[3] JAIN R, NAGEL H H. On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1979, 1(2):206-214.

[4] WREN C R, AZARBAYEJANI A, DARRELL T, et al. Pfindex;

real-time tracking of the human body[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7):780-785.

[5] STAUFFER C, GRIMSON W. Adaptive background mixture models for real-time tracking [C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Fort Collins: IEEE, 1999:246-252.

[6] ZIVKOVIC Z. Improved adaptive Gaussian mixture model for background subtraction[C]// Proceedings of the 17<sup>th</sup> International Conference on Pattern Recognition. Cambridge: IEEE, 2004:28-31.

[7] ZIVKOVIC Z, HEIJDEN F. Efficient adaptive density estimation per image pixel for the task of background subtraction[J]. Pattern recognition letters, 2006, 27(7):773-780.

[8] CHEN S C, SRINIVAS SISTA, MEI-LING SHYU, et al. An indexing and searching structure for multimedia database systems [C]// Storage and Retrieval for Media Databases. San Jose, CA, USA: SPIE, 2000:262-270.

[9] ZHANG Chengcui, CHEN S C, SHYU M L, et al. Adaptive background learning for vehicle detection and spatio-temporal tracking [C]//4<sup>th</sup> International Conference on Information, Communications & Signal Processing, and 4<sup>th</sup> Pacific-Rim Conference on Multimedia (ICICS-PCM 2003). Singapore: IEEE, 2003:797-801.

[10] NG K K, DELP E J. Background subtraction using a pixel-wise adaptive learning rate for object tracking initialization[C]// Visual information processing and communication II. San Francisco, CA, USA: SPIE, 2011:78820I:1-78820I:9.

[11] BARNICH O, VANOGENBROECK M. ViBe: A powerful random technique to estimate the background in video sequences [C]// IEEE International Conference on Acoustics, Speech & Signal Processing. Taipei: IEEE, 2009:945-948.

[12] 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.

[13] COVER T, HART P. Nearest neighbor pattern classification[J]. IEEE Transactions on Information Theory, 1967, 13(1):21-27.

[14] AO Wei, FU Yanwei, XU Feng. Detecting tiny moving vehicles in satellite videos[J]. arXiv preprint arXiv:1807.01864, 2018.

[15] QI Yunguang, AN Gang. Infrared moving targets detection based on optical flow estimation[C]// Proceedings of 2011 International Conference on Computer Science and Network Technology. Harbin, China: IEEE, 2011:2452-2455.

[16] 范超男,李士心,张海,等.基于改进的混合高斯模型运动目标检测算法研究[J].天津职业技术师范大学学报,2019,29(4):25-31.

[17] 丁业兵.运动目标检测算法综述[J].科技资讯.2019,17(30):21-23.

[18] 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(06):82-85,91.

[19] 王健,王晓东,郭磊.基于混合高斯模型的运动目标检测[J].宁波大学学报:理工版,2019,32(03):35-39.

[20] 于澜博,张涛,郭立红,等.卫星视频运动目标检测算法[J].液晶与显示,2017,32(02):138-143.

[21] 李贝贝,韩冰,田甜,等.吉林一号视频卫星应用现状与未来发展[J].卫星应用,2018(03):23-27.

[22] 刘珠妹,刘亚岚,谭衢霖,等.高分辨率卫星影像车辆检测研究进展[J].遥感技术与应用,2012,27(01):8-14.