

Typical Video Target Detection Method Based on Improved Motion Template and Histogram Matching

Yongmei Zhang¹, Chao Feng¹, Jianzhe Ma²

¹School of Computer Science and Technology, North China University of Technology, Beijing

²Beijing Hualongtong Science and Technology Co., Ltd, Beijing

Email: zhangym@ncut.edu.cn, 2016309040102@mail.ncut.edu.cn, tyut_mjz@hotmail.com

Received: Dec. 31st, 2018; accepted: Jan. 14th, 2019; published: Jan. 21st, 2019

Abstract

The existing video moving target detection algorithms cannot detect the complete regions of moving targets and the hole phenomena in extracting the contour of video moving targets using classical motion template method and the false detection rate is higher when slower moving speed, the paper presents a video moving target detection method based on fusion of improved motion template and HSV histogram matching, adopts the three-frame difference and improved OSTU adaptive threshold method to obtain the contours of the potential moving targets, constructs the motion history template according to the time series relationship between each pixel point and the target contour retention time, segments the potential moving target areas using the stepwise flooding method, and calculates the Bhattacharyya distance between the potential target area and the target template area so as to detect the typical targets. The experiment results show the proposed method can help to solve the problems of the hole phenomena and incomplete contours for the typical target using classical motion template, and the average rate of detection accuracy is obviously higher than that of the background subtraction method and the motion template method respectively.

Keywords

Improved Motion Template, Three-Frame Difference Method, Improved OSTU Adaptive Threshold, Stepwise Flooding Method, Bhattacharyya Distance

改进运动模板与直方图匹配的视频典型目标检测方法

张永梅¹, 冯超¹, 马健喆²

¹北方工业大学计算机学院, 北京

²北京华龙通科技有限公司, 北京

Email: zhangym@ncut.edu.cn, 2016309040102@mail.ncut.edu.cn, tyut_mjz@hotmail.com

收稿日期: 2018年12月31日; 录用日期: 2019年1月14日; 发布日期: 2019年1月21日

摘要

针对经典运动模板法提取运动目标轮廓出现的空洞和不能提取完整轮廓的问题, 以及运动速度缓慢时误检率高的问题, 本文提出了一种改进的运动模板和HSV直方图匹配相融合的视频运动目标检测方法。采用三帧差分和改进的OSTU自适应阈值法获得潜在的运动目标轮廓, 并根据每个像素点与目标轮廓停留时间的时序关系构建运动历史模板, 利用逐级洪水泛滥法分割目标潜在区域, 并计算目标潜在区域与目标模板的Bhattacharyya距离, 从而检测典型目标区域。实验结果表明, 该方法有助于解决传统运动模板法提取运动目标轮廓产生的空洞和不能提取完整轮廓的问题, 而且精确率比背景减除法、运动模板法明显提高。

关键词

改进的运动模板, 三帧差分法, 改进的OSTU自适应阈值, 逐级洪水泛滥法, Bhattacharyya距离

Copyright © 2019 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

视频目标的检测与识别是计算机视觉领域的一个重要课题[1], 引起了国内外学者的广泛关注, 国内外学者在视频目标检测方面取得了一定成果。目前常用的视频目标检测方法包括基于帧差法的轮廓检测、基于直方图匹配法的运动目标检测、基于光流法的运动目标检测以及一些组合方法, 这些方法虽然有很好的检测效果, 但在某些方面仍存在一些缺陷。

为了更好地解决复杂场景中运动目标检测准确率低的问题, Qingyu Guo, Zheng Zhang 提出一种融合高斯混合模型和帧间差分法的视频运动目标检测方法, 该方法使用帧间差分法、高斯混合模型和 K-means 聚类检测视频运动目标的轮廓, 利用 C#和 EmguCV 框架检测运动目标[2]。该方法实现简单, 检测的准确率比经典的高斯混合模型和帧差法的准确率较高, 但是该方法并不能检测视频中运动速度缓慢或静止的目标, 而且当间隔时间设置不合理时, 该方法检测的准确率较低。

为了解决动态场景中经典运动目标检测算法检测准确率低和鲁棒性差的问题, Wang Huibin、Chen Zhe、Lu Miao 等人提出一种基于运动显著概率图的运动目标检测方法。该方法在时间尺度上构建了包含运动目标信息的长短期时间序列组, 并利用 TFT (Temporal Fourier Transform)方法计算运动目标的显著性值, 根据显著性值得到条件运动的显著性概率图, 采用全概率公式计算后验概率值, 从而得到运动目标的显著性概率图, 以概率图为基础, 对像素的空间信息进行建模, 检测运动目标。该方法能够在复杂噪声环境下准确检测运动目标区域, 并且能够提高运动目标检测的鲁棒性和普适性, 但是该方法并不能检测到运动目标的完整区域[3]。

为了有效地解决运动目标检测出现的空洞现象以及无法得到完整运动目标的问题, 贾建英、董安国提出了一种基于联合直方图的运动目标检测算法, 利用联合直方图构造相邻帧的相似性, 对相邻两帧视频进行分块, 并计算每块的联合直方图, 同时结合相似性指标消除背景块, 从而检测视频中的运动目标[4]。该方法在无背景遮挡的情况下能够很好地检测到完整的运动目标, 但是当有背景遮挡时, 并不能完整地检测到运动目标, 检测效果比经典的检测算法还要差。

针对运动目标检测领域中帧差法和背景差分法的缺陷, 孙挺、齐迎春、耿国华提出了一种融合帧差法和背景差分的运动目标检测算法, 该算法采用混合高斯建立背景模型提取运动目标轮廓, 利用连续三帧差分法以及自适应差分阈值的方法提取运动目标的轮廓, 将得到的两种差分结果进行融合并进行形态学处理提取运动目标[5]。该方法能够有效地抑制算法产生的噪声, 解决产生的空洞现象, 而且具有普适性, 但是当有遮挡时, 提取的运动目标区域并不完整。

目前视频目标检测方法主要存在两个问题: 由于有遮挡或光照干扰时导致提取的运动目标轮廓出现空洞现象进而导致提取的轮廓不完整甚至无法检测到完整的运动目标轮廓; 由于运动速度缓慢导致运动目标检测精确率低。

针对上述问题, 本文提出一种基于改进的运动模板和直方图匹配的视频运动目标检测方法, 本文使用三帧差分法和自适应阈值代替传统运动模板中的帧差和二值化法来提取运动目标的轮廓, 有效地解决了提取运动目标轮廓时产生的空洞和提取运动目标轮廓不完整的问题。该方法既能克服运动模板检测不到视频中运动速度缓慢运动目标的缺陷, 又能克服直方图匹配法对光线敏感的不足, 同时又能够充分发挥运动模板和直方图匹配各自的优势。该方法不仅能够解决提取轮廓时产生的空洞和提取运动目标轮廓不完整的问题, 而且能够有效地检测到运动目标的完整区域, 从而提高了运动目标检测的精确率。

2. 基于改进运动模板和直方图匹配的视频典型目标检测方法

本文提出了一种基于改进运动模板和直方图匹配的视频典型目标检测方法, 利用改进运动模板法检测视频中运动目标的潜在区域, 采用 HSV 直方图匹配法对潜在目标区域进行二次确认, 从而检测最终的典型目标区域, 该方法有助于提高视频中典型目标检测的精确率。本文的视频典型目标主要包括坦克和装甲车。

2.1. 传统运动模板法

运动模板法由麻省理工学院多媒体实验室的 A. Bobick 和 J. Davis 于 1996 年提出[6], 为了提高该算法对运动目标检测的实时性, J. Davis 对该算法进行了改进, 并将运动目标的方向信息加入到运动模板算法, 从而形成了现在常用的运动模板检测法。运动模板是一种检测运动目标并计算其运动方向的模型, 该模型分为运动目标的轮廓图像(silhouette, silh)获取、运动历史图像(Motion History Image, mhi)的构建、计算运动目标的梯度方向和运动目标的分割四部分。

利用运动模板法检测运动目标时, 需要获取运动目标的轮廓图像。采用帧差法对相邻两帧的灰度图像进行减法处理, 从而获得差分图像, 对差分图像进行二值化处理和去噪处理, 从而获得运动目标的轮廓图像。

在获得运动目标的轮廓图像后, 根据公式(1)对运动目标的轮廓图像进行处理, 得到运动目标的运动历史图像。运动历史图像中的像素由运动目标轮廓图像中的像素和轮廓图像滞留时间确定, 表示轮廓图像随时间的变化关系, 具体的处理方法如公式(1)所示。

$$\text{mhi}(x, y) = \begin{cases} \text{timestamp} & \text{silh}(x, y) \neq 0 \\ 0 & \text{silh}(x, y) = 0 \text{ and } \text{mhi} < (\text{timestamp} - \text{duration}) \end{cases} \quad (1)$$

公式(1)中的 *timestamp* 表示当前时间滞留值, *duration* 表示滞留时间的最大值[7]。当轮廓图像在当前时间滞留值范围内出现时, *mhi* 像素值设为当前时间滞留值; 当运动目标轮廓图像未出现或在很长时间前出现过时, *mhi* 像素值设为 0。

本文利用运动模板法将 *mhi* 图像分割为独立的区域以便于运动目标检测。分割算法具体步骤如下:

(1) 扫描 *mhi* 图像直到寻找到当前时间滞留范围内的像素点, 该像素点为最新当前时间的轮廓区域的边界点;

(2) 继续搜索当前时间最新轮廓区域的边界以寻找紧邻当前边界外围的最近运动区域。当发现这样的运动区域时, 采用逐级洪水泛滥法分割感兴趣目标的当前位置, 逐级洪水泛滥分割算法步骤如下:

① 选取 *mhi* 图像中的一点标记作为种子像素点;

② 检测种子像素点的颜色, 并对该像素点进行 8 方向搜索, 检测其邻近像素点的颜色是否满足设置的阈值范围, 若满足, 则将该像素点的颜色填充为种子像素点的颜色值, 即合并当前像素点与种子像素点构成连通区域, 否则不进行处理;

③ 重复①、②直至 *mhi* 图像都被处理;

④ 采用阈值分割法从 *mhi* 图像中分割感兴趣的目标, 本文统计图像中各像素点的值, 并计算图像中各像素点值的均值与方差, 将均值与 3 倍方差之和作为分割阈值;

(3) 保存(2)分割的感兴趣目标的当前位置;

(4) 重复步骤(2)和(3), 直至找到所有区域。

采用运动模板法得到含有运动方向的运动目标的潜在区域, 但是, 这些潜在区域中可能没有含有运动目标, 因此, 为了提高运动目标的检测精确率, 本文采用 HSV 直方图匹配法进一步检测潜在区域是否含有运动目标以降低误检率。

2.2. 直方图匹配法

颜色特征最显著的特点是具有旋转不变性, 并且在大小和方向上具有鲁棒性; 另外, 其处理速度较快, 适合应用在对速度要求较高的领域[8]。由于视频帧变化较快, 帧与帧之间时间间隔较短, 在检测视频运动目标时应采用速度较快的算法, 因此, 本文利用颜色直方图匹配算法来检测视频运动目标。

常用的颜色直方图匹配算法包括 RGB 直方图匹配算法、HSV 直方图匹配算法。由于 RGB 颜色空间是由 0~255 这 256 个数值表示, 不够直观, 而且根据 RGB 数值很难判断该数值所代表的颜色, RGB 颜色空间分布不均匀, 当 2 个颜色相近时其 R、G、B 数值可能相差很大。HSV 颜色空间更符合人类感知颜色的色彩空间, 其表示方法类似于人类感知颜色的方式, 具有很强的感知度[9], 并且其计算相对简单。本文采用 HSV 颜色空间直方图匹配算法检测视频运动目标, 检测结果更符合人眼观测效果。

HSV 颜色空间将颜色分为 3 个独立的空间向量, *H* 表示色彩的基本属性; *S* 表示色彩的纯度, *S* 值越高其色彩越纯; *V* 表示色彩的明亮程度。由于色彩的明亮程度易受光照的影响, 在有光照影响的条件下, *V* 的检测值会与真实值相差很大, 因此, 本文选择使用 *H*、*S* 色彩向量作为直方图匹配算法的基础色彩向量。

由于视频帧为 RGB 彩色图像, 因此, 在使用 HSV 直方图匹配算法时, 首先将 RGB 彩色图像转化为 HSV 图像, 需要分别统计潜在区域的 *H*、*S* 色彩直方图以及目标的 *H*、*S* 色彩直方图。由于典型目标的颜色构成较单一, 因此, 本文将 *H* 的取值范围分成 $N = 10$ 个区间, *S* 的取值范围分成 $P = 5$ 个区间, 则 *H* 分量的每个区间的色度数量为 $k = 360/N$, 记作 K_1, \dots, K_k , *S* 分量的每个区间的色度数量为 $m = 1/P$, 记作 M_1, \dots, M_m , 分别统计每个区间的色度像素个数, 建立颜色直方图, 则 *H*、*S* 的色彩空间表示如公式(2)、(3)所示。

$$H = (H_1, H_2, \dots, H_{10}) \quad (2)$$

$$S = (S_1, S_2, \dots, S_5) \quad (3)$$

H 、 S 色彩直方图中每个柱的值可以表示为 $H_i (i=1, 2, \dots, k)$ ， $S_j (j=1, 2, \dots, m)$ ，每个柱值的计算方法如公式(4)、(5)所示。

$$H_i = \sum f_i(h(x, y)), i=1, 2, \dots, k \quad (4)$$

$$S_j = \sum f_j(h(x, y)), j=1, 2, \dots, m \quad (5)$$

公式(4)、(5)中的 f 函数为分段函数，当像素满足相应的区间时，其值为 1，反之则为 0。当潜在区域和目标区域的 H 、 S 色彩直方图都计算完成后，根据 Bhattacharyya 距离计算潜在区域和目标区域在 H 、 S 色彩空间上的相似度，将 H 、 S 色彩空间上的相似度分别与相应权重相乘，得到一个新的相似度值，并根据相似度的大小来判别潜在区域是否含有典型目标，从而检测潜在区域是否含有典型目标。Bhattacharyya 距离公式如公式(6)、(7)所示。

$$D_B(p, q) = -\ln(BC(p, q)) \quad (6)$$

其中：

$$BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)} \quad (7)$$

公式(6)中的 D_B 为 Bhattacharyya 距离，表示潜在运动目标区域与目标区域的相似性， $BC(p, q)$ 为 Bhattacharyya 系数，该系数可以用来表示两个样本相对接近的程度，其中 p 和 q 分别表示视频帧图像提取的潜在目标区域和模板图像在 HSV 色彩空间中 $H(S)$ 色彩分量上各像素值构成的向量。

2.3. 融合改进运动模板和直方图匹配的视频典型目标检测方法

传统运动模板法是一种常用的运动目标检测算法，利用帧差和二值化法获得运动目标的轮廓，该算法具有帧差法对光线不敏感、能适应各种动态环境的优点[10]。因此，运动模板法常被应用于姿态识别，是一种有效的识别运动目标方法。但该算法在提取运动目标轮廓时会形成空洞现象而导致提取的目标轮廓不完整，因此，本文在传统运动模板的基础上进行了改进，采用基于三帧差分 and 自适应阈值法的运动模板来检测运动目标。三帧差分公式如(8)、(9)、(10)所示。

$$P_k(x, y) = |f_k(x, y) - f_{k-1}(x, y)| \quad (8)$$

$$R_k(x, y) = \begin{cases} 0, & P_k(x, y) < T_3 \\ 1, & P_k(x, y) \geq T_3 \end{cases} \quad (9)$$

$$R_k(x, y) = \begin{cases} 1, & R_k(x, y) \cap R_{k+1}(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

公式(8)表示相邻两帧帧差分的计算公式，公式(8)中 $P_k(x, y)$ 表示帧间差分图像， $R_k(x, y)$ 表示 1 二值化图像， T_3 是分割阈值， $f_k(x, y)$ 、 $f_{k-1}(x, y)$ 分别表示第 k 帧、 $k-1$ 帧视频。

由于分割阈值的选择对运动目标轮廓的提取影响很大，传统方法阈值的选取需要根据视频、背景和运动目标的不同来相应地选择分割阈值。当三者中任何一个因素发生变化，选择的阈值可能就不再适用。为了消除固定分割阈值对提取的运动目标轮廓影响，本文采用自适应阈值方法来设定分割阈值，该方法充分考虑了视频中运动目标的类型、背景等因素，并按照一定的方式自动选取阈值。

本文采用改进的 OSTU 自适应阈值法进行分割, 在设定阈值时首先统计视频帧图像中像素值的中位数、最大像素点值和最小像素点值, 根据公式(11)求出初始阈值 Th_0 , 根据初始阈值进行分割, 分别计算分割后图像非目标区域和目标区域各像素值的均值 μ_1 、 μ_2 , 根据公式(12)计算分割后两类间的差异度, 并判断是否达到迭代数, 若达到迭代数, 则选择最大 s 值对应的 Th_0 作为分割阈值, 否则将 Th_0 增加一个步长, 迭代数加 1, 继续执行迭代, 直至达到迭代数。一般设置的迭代数为 15, 步长值为 0.1。阈值选取的具体步骤如下。

- 1) 设定初始阈值 Th_0 、迭代数 $T = 0$ 及每步阈值增加的步长 $step$, 阈值的设定如公式(11)所示;

$$Th_0 = (\text{median}[P_k(x, y)] + \max[P_k(x, y)] + \min[P_k(x, y)]) / 4 \quad (11)$$

公式(11) median 函数表示帧差图像各像素点的中位数, \max 函数表示各像素点的最大值, \min 函数表示各像素点的最小值。

- 2) 按照公式(12)计算分割后两类间的差异度;

$$S = \frac{(Th_0 - \mu_1)(\mu_2 - Th_0)}{(\mu_2 - \mu_1)^2} \quad (12)$$

公式(12)中, μ_1 、 μ_2 分别表示分割后两部分对应像素的均值。

- 3) 检查是否达到迭代数, 若达到迭代数则停止迭代, 并选取最大的 s 对应的阈值 Th_0 作为分割阈值 T_3 , 否则继续执行步骤(4);

- 4) 按照公式(13)增加阈值, 公式(14)增加迭代数, 并重复步骤(2)、(3);

$$Th_0 = Th_0 + step \quad (13)$$

$$T = T + 1 \quad (14)$$

改进后的运动模板提取视频中运动目标轮廓的实验结果如图 1 所示。由图 1 可以看到, 改进的运动模板法能够较为完整地提取视频中运动目标的轮廓, 而且提取的目标轮廓比传统运动模板法提取的目标轮廓较好。传统运动模板法虽然提取了视频中运动目标的轮廓, 但是由于视频速度较快, 相邻帧背景变换较大, 因此, 传统运动模板法检测到的目标轮廓存在较多的非运动目标区域, 会对后续运动目标潜在区域的检测造成很大影响。本文改进的运动模板法采用三帧差分法和改进的 OSTU 自适应阈值法提取目标潜在区域, 从图(c)、(f)可以看到本文方法可以很好地提取目标的轮廓区域, 尤其在复杂背景时, 效果更为明显, 提取的运动目标轮廓更加清晰, 从而有助于提高运动目标检测的精确率。

针对逐级洪水泛滥分割法需要人工选择初始像素点作为种子像素点带来的不便, 本文提出了一种直方图阈值法选择种子像素点的逐级洪水泛滥的分割方法, 从而提高逐级洪水泛滥分割法的自动化程度。本文首先通过直方图阈值法来选取种子像素点, 然后对种子像素点进行 8 方向搜索并合并像素点颜色处于设定阈值范围内的邻近像素点构成连通域, 最后采用阈值分割法从当前背景区域分割感兴趣目标。

由于运动模板法对运动速度较慢的目标不敏感, 甚至会丢失运动目标, 因此, 本文首先利用改进的运动模板法检测运动目标的潜在区域, 利用 HSV 直方图匹配法对潜在区域进行二次确认, 提高了运动目标检测的精确率。

颜色直方图描述的是色彩在图像中的分布, 具有平移、旋转不变性, 对图像中子对象所在的位置和方向的变化不敏感, 具有很强的鲁棒性[11]。直方图匹配算法利用图像局部直方图的颜色特征作为衡量标准, 对比并计算图像局部区域与运动目标之间的相似度[12], 从而得到图像局部区域与运动目标在颜色特征方面的相似度因子, 该因子作为判别图像局部区域是否为运动目标的依据。该方法计算简单, 只利用图像中的颜色特征不受运动目标运动速度的影响, 具有很高的检测率。但是, 该算法容易受到光照影响,

在有光照影响时，会极大地降低运动目标检测的准确率。本文提出了一种融合改进运动模板和 HSV 直方图匹配法的视频典型目标检测方法，该融合方法的流程图如图 2 所示。

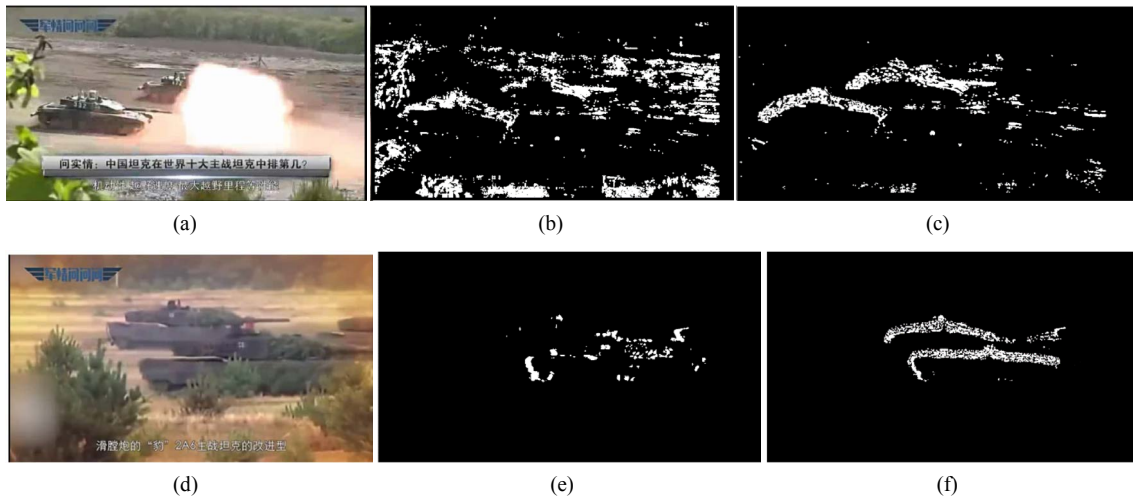


Figure 1. The contour images extracted from different backgrounds. (a) The video frame; (b) The contour images by classical motion template; (c) The contour images by improved motion template; (d) The video frame; (e) The contour images by classical motion template; (f) The contour images by improved motion template

图 1. 不同背景下提取的轮廓图像。(a) 原视频帧；(b) 传统运动模板法提取的轮廓图像；(c) 改进的运动模板提取的轮廓图像；(d) 原视频帧；(e) 传统运动模板提取的轮廓图像；(f) 改进后运动模板提取的轮廓图像

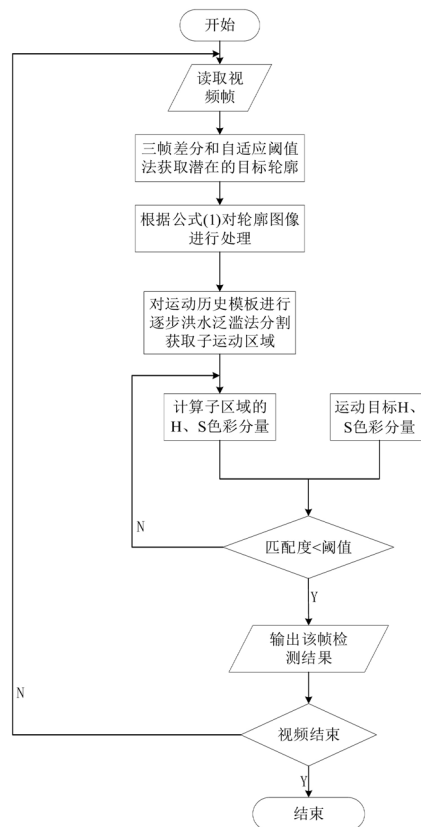


Figure 2. The flow chart of improved motion template and histogram matching method for typical target detection
图 2. 改进的运动模板和直方图匹配的典型目标检测方法流程图

首先读取视频图像，并对视频图像进行三帧差分处理和自适应阈值处理，获取典型目标的潜在轮廓图像，根据公式(1)对潜在轮廓图像进行处理，获取运动目标的历史图像，采用逐级洪水泛滥分割法将运动历史图像分割为子独立区域，从而得到可能包含有运动目标的潜在区域，提取子潜在区域的 HSV 色彩空间在 H 、 S 色彩向量的直方图，根据公式(6)计算子潜在区域与运动目标在 H 、 S 色彩向量上的 Bhattacharyya 距离，并根据公式(15)得到最终的相似度，并根据相似度值，判别子独立区域是否含有运动目标，从而有助于提高运动目标检测的精确率。HSV 直方图匹配的阈值越低，表示两幅图像中的目标区域越相似，本文通过查阅大量文献将阈值设为 0.2。

$$\begin{cases} D_B = D_{BH} * w_1 + D_{BS} * w_2 \\ w_1 + w_2 = 1 \end{cases} \quad (15)$$

3. 实验结果与分析

本文利用 Python 编程语言以及 OpenCV-Contrib 视频处理库设计实现了融合运动模板和 HSV 直方图匹配的视频运动目标检测。数据主要来自网上获取到的视频数据，主要利用两种场景下共 100 个视频来验证本文方法的有效性，每个视频大约 2 分钟时长，而且每个视频包含 400 多帧。两种场景分别为：典型目标运动缓慢的视频，以及有背景干扰的视频，并分别与运动模板法、背景减除法的检测结果进行了比较。检测结果如图 3、4 所示。

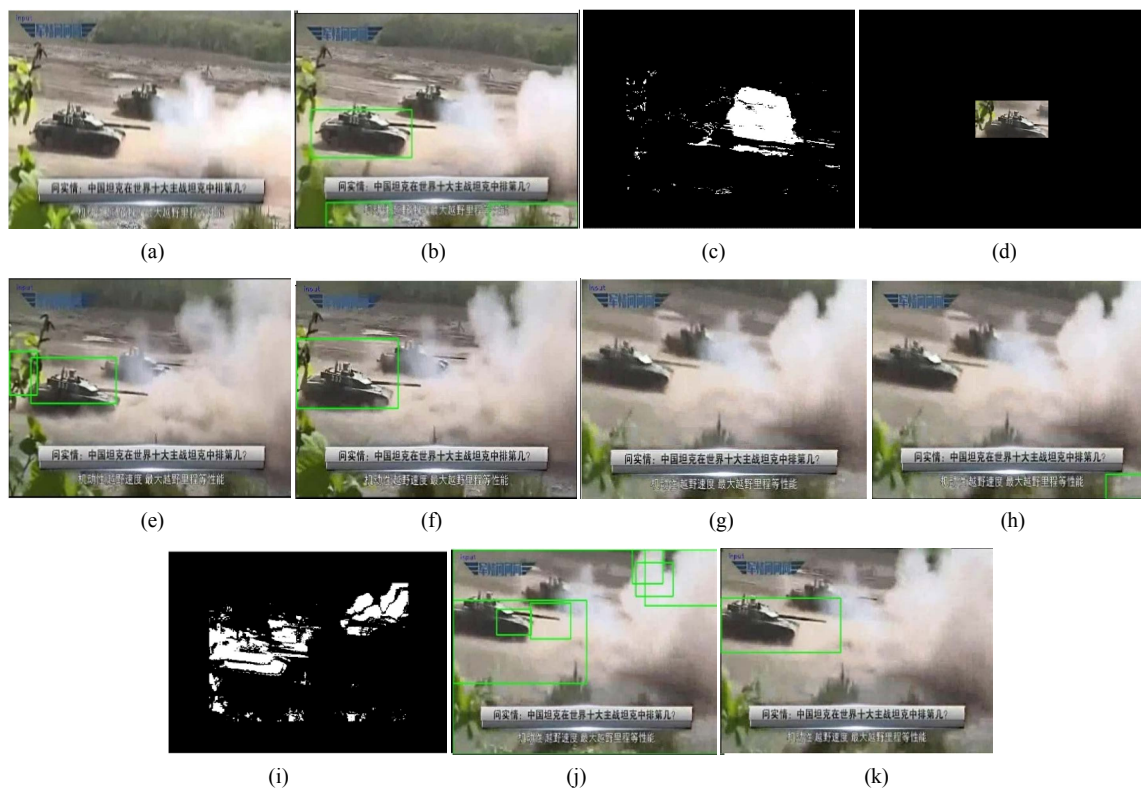


Figure 3. The experiment results of slow motion video. (a) The 90th video frame; (b) The result of background subtraction; (c) The history motion template; (d) The histogram matching template; (e) The result of motion template; (f) The result of the paper; (g) The 95th video frame; (h) The result of background subtraction; (i) The history motion template; (j) The result of motion template; (k) The result of the paper

图 3. 运动速度缓慢视频的实验结果。(a) 第 90 帧视频；(b) 背景减除法检测结果；(c) 历史运动模板；(d) 直方图匹配模板；(e) 运动模板法检测结果；(f) 本文方法检测结果；(g) 第 95 帧视频；(h) 背景减除法检测结果；(i) 历史运动模板；(j) 运动模板检测结果；(k) 本文方法检测结果

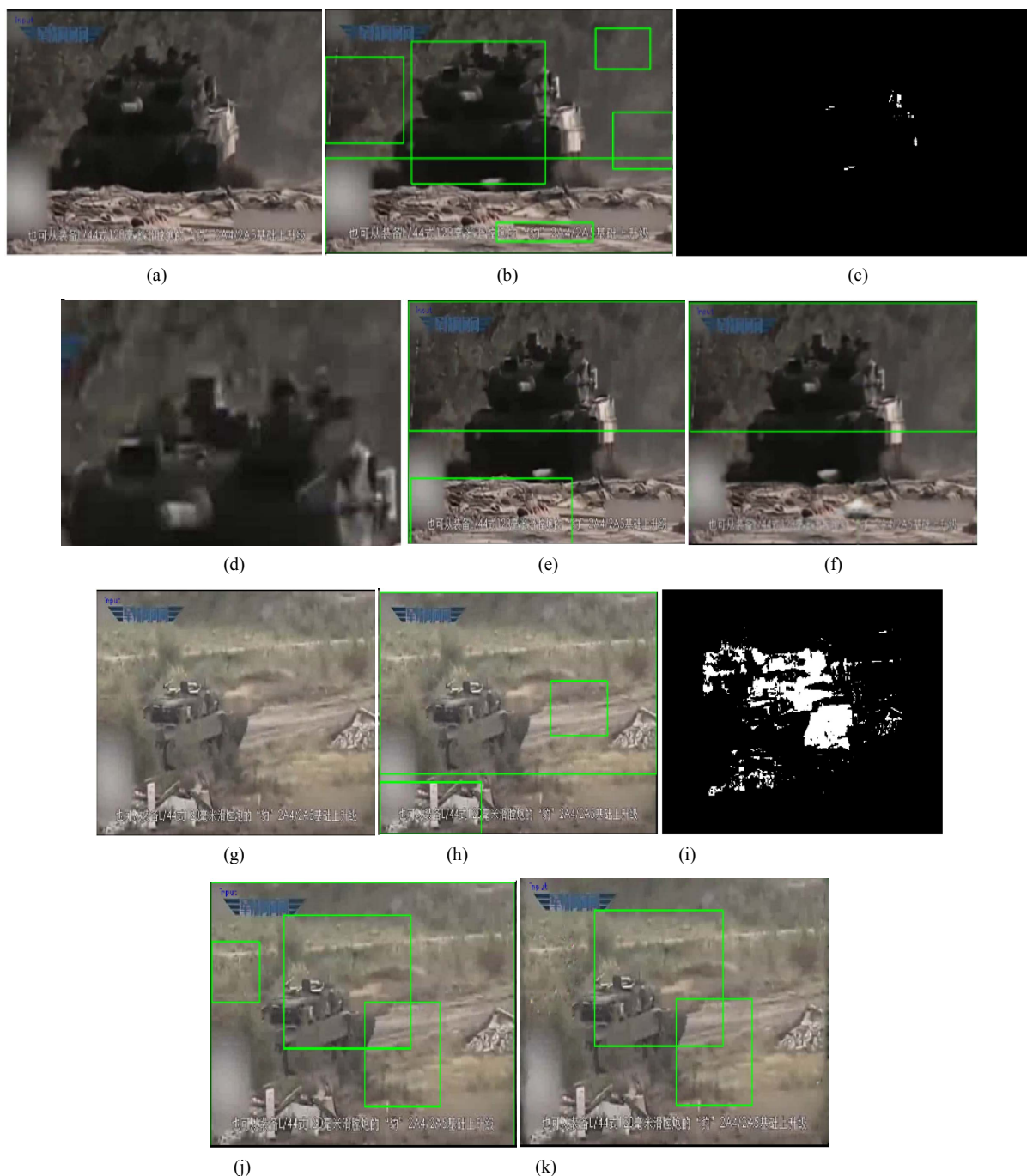


Figure 4. The experiment results of color similarity between background and moving target. (a) The 90th video frame; (b) The result of background subtraction; (c) The history motion template; (d) The histogram matching template; (e) The result of motion template; (f) The result of the paper; (g) The 205th video frame; (h) The result of background subtraction; (i) The history motion template; (j) The result of motion template; (k) The result of the paper

图 4. 背景与运动目标颜色相似视频的实验结果。(a) 第 130 帧视频; (b) 背景减除法检测结果; (c) 历史运动模板; (d) 直方图匹配模板; (e) 运动模板法检测结果; (f) 本文方法检测结果; (g) 第 205 帧视频; (h) 背景减除法检测结果; (i) 历史运动模板; (j) 运动模板法检测结果; (k) 本文方法检测结果

图 3 为同一类型的视频, 采用(d)的直方图匹配模板。(c)、(i)分别表示第 90 帧和第 95 帧视频运动模板检测法采用的运动模板, 图中白色区域表示运动模板法检测的潜在运动区域, 图中若包含非连通区域,

则该方法将相应部分的非连通区域标记出来作为检测结果。

从图 3 检测结果可以看到, 本文方法的检测结果较好, 能够较为完整地提取视频中的典型目标区域; 由于背景减除法检测的结果易受背景模板的影响, 而且检测准确率低, 尽管背景减除法在第 90 帧完整地提取了目标区域, 但也检测出了两个非目标区域, 在第 95 帧时背景减除法并没有提取出目标区域; 运动模板检测法易受运动速度影响, 运动模板法在第 90 帧和第 95 帧虽然都提取出了目标, 但是提取的非目标区域也较多, 而且提取的目标区域并不完整。

图 4 是同一类型视频, 采用(d)直方图匹配模板。(c)、(i)分别为第 130 帧和第 205 帧的历史运动模板, 白色部分当前时刻表示运动目标, 黑色部分表示非目标区域, 或者过去时刻的运动区域。由(i)可以看到, 白色部分较多, 导致运动模板法检测准确率较低。

从图 4 检测结果可以看到, 本文方法的检测结果较好, 由于图 4 中的背景与目标的颜色较为相似, 所以背景减除法检测的结果中包含了非目标区域, 而且包含的非目标区域比检测正确的目标区域还多, 所以导致其检测准确率较低; 而运动模板使用帧差法来提取目标的轮廓, 所以其检测结果也相应地受到了背景干扰, 进而导致其检测结果也存在误检测的区域。本文方法融合了改进的运动模板和 HSV 直方图匹配法, 虽然其检测结果也会受到背景的影响, 但是 HSV 直方图匹配法则可以进一步对运动模板法检测的区域进行匹配, 从而在一定程度上剔除了误检的区域, 进而有助于提高运动模板法检测的准确率。

本文方法在理论上与文献[13]方法进行了对比。为了解决混合高斯背景模型由于光照突变导致运动目标出现误检的问题, 以及由于运动目标突然运动出现“鬼影”的问题, 文献[13]提出了一种基于三帧差分的混合高斯背景模型来检测运动目标[13]。该方法使用三帧差分法对图像进行处理以获取目标的轮廓特征, 采用混合高斯背景模型来检测运动目标。在构造混合高斯背景模型时, 利用基于统计方法来初始化混合高斯模型的参数, 在初始化模型参数过程中, 该方法利用 A_1 、 A_2 两个累加器矩阵来记录灰度图像中满足相应规则的像素点的个数。 A_1 、 A_2 各个位置上的元素值的更新方式如公式(16)所示。

$$A(x) = \begin{cases} A_{1i} + 1 & |\chi - \mu| < \alpha\lambda \\ A_{2i} + 1 & |\chi - \mu'| < \alpha\lambda \end{cases} \quad (16)$$

利用公式(16)对 A_1 、 A_2 各个位置的元素值更新后, 还需对其所对应的均值进行更新。式中的 α 、 λ 分别为常数, α 取值为 4~10, λ 取值为 2.5。

初始化模型参数后, 需要挑选出混合高斯模型中的一个高斯分布, 按公式(17)进行初始化, 并计算其对应的方差 σ^2 。同时按照公式(18)初始化混合高斯模型中其他的高斯分布。

$$\mu = \begin{cases} \mu A_{1i} > A_{2i} \\ \mu' A_{1i} \leq A_{2i} \end{cases} \quad (17)$$

$$\omega_k = \omega_M$$

$$\begin{cases} \mu_k = 0 \\ \sigma_k^2 = \sigma_M \\ \omega_k = \varepsilon \end{cases} \quad (18)$$

当获取到一帧新的图像后, 且其像素点满足混合高斯模型中的某一高斯分布后, 需要按照公式(19)更新该高斯分布参数。

$$\begin{cases} \omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha \\ \mu_{k,t} = (1-\rho)\mu_{k,t-1} + \rho\chi_t \\ \sigma_{k,t}^2 = (1-\rho)\sigma_{k,t-1}^2 + \rho(\chi_t - \mu_{k,t}) \end{cases} \quad (19)$$

式中 α 代表权重学习率, $0 \leq \alpha \leq 1$, 且 α 决定权重的更新速度, 并且该值也会影响高斯分布的更新速度, ρ 表示高斯分布参数的学习率, $\rho = \alpha/\omega_{k,t}$ 。

为了消除极端值的影响, 更新混合高斯背景模型后, 需要归一化混合高斯模型中各个高斯分布的权重参数。计算各个高斯分布权重参数的和, 当求出的权重参数和大于给定阈值时, 就判定对应的高斯分布为背景, 其余的高斯分布为运动目标对应的高斯分布, 从而检测图像中的运动目标。该方法虽然解决了混合高斯背景模型由于光照突变导致运动目标出现误检的问题, 以及由于运动目标突然运动出现“鬼影”的问题, 但其精确率并不高。

本文从精确率、召回率和运行时间等评测指标与背景减除法、运动模板法进行了定性比较[14]。表中的平均检测精确率、平均检测召回率的计算方式如公式(20)、(21)所示。

$$\text{平均检测精确率: Pre} = \frac{1}{n} \sum_{i=1}^n \frac{TT_i}{TT_i + FT_i} \quad (20)$$

$$\text{平均检测召回率: Rec} = \frac{1}{n} \sum_{i=1}^n \frac{TT_i}{TT_i + FF_i} \quad (21)$$

公式(20)、(21)中的 TT_i 表示实际为目标区域检测结果也为目标区域的个数, FT_i 表示实际为非目标区域却检测为目标区域的个数, FF_i 表示实际为目标区域却未检测出的区域个数, 本文结合视觉观察效果统计上式中三个变量的值; 平均检测精确率反映了检测到的目标区域个数与所检测区域总数的比例, 该值用来评价目标检测算法检测效果的好坏, 该值越大, 表示该算法的检测效果越好; 平均检测召回率反映了检测到的目标区域个数与正确目标区域总数的比例, 该值用来评价目标检测算法的查全率, 该值越大, 表示该算法的检测效果越好; n 表示检测的次数, 本文取 n 为 10 次, 三种方法对比结果如表 1 所示。

Table 1. The comparison result of three methods
表 1. 三种方法对比统计表

视频类型	算法	平均检测精确率(%)	平均检测召回率(%)	平均运行时间(s)
运行速度缓慢	背景减除法	78.33	56.04	20
	运动模板法	84.47	62.52	15
	本文方法	93.75	73.59	18
复杂背景	背景减除法	73.07	58.42	22
	运动模板法	82.52	60.58	14
	本文方法	90.08	72.34	17

由表 1 可以看到, 在不同类型视频中, 本文方法视频典型目标检测的精确率比其他两种方法较高, 而且召回率也比其他两种方法较高, 但是运行时间比运动模板法较长。

在检测精确率方面, 由于背景减除法的检测结果与选择的背景模板有关, 当选择的背景模板不合

适时, 会极大地降低检测的精确率。运动模板法对目标的运动速度极为敏感, 当其速度运动缓慢时, 会导致其检测精确率下降。本文方法采用融合运动模板和 HSV 直方图匹配方法, 利用 HSV 直方图匹配法弥补了运动模板法对目标运动速度敏感的缺陷, 在一定程度上提高了运动模板法检测的精确率。

在召回率方面, 由于本文方法融合了运动模板和 HSV 直方图匹配, 运动模板检测法易受运动速度影响, HSV 直方图匹配法易受光照影响, 两者融合在一定程度上提高了检测的准确率和召回率, 但是召回率提高并不多。而背景减除法与其背景模板的选择有关, 当背景模板选择不合适, 导致其召回率并不高。而在运行时间上, 背景减除法的运行时间一部分花费在背景模板的构建, 一部分花费在目标检测, 所以其运行时间相对较长。本文方法融合了运动模板和 HSV 直方图匹配, 所以其运行时间比运动模板法较高。

综合上述三种方法的平均检测精确率、平均检测召回率和平均运行时间来看, 本文方法比其他两种方法较优。相对于其他两种方法, 本文方法不仅可以检测到运动目标的完整区域, 而且平均检测精确率、平均检测召回率比其他两种方法较高, 虽然平均运行时间比运动模板法长, 但比背景减除法短。因此, 综合所有评价指标, 本文方法较优。

4. 结论

本文提出了一种基于改进运动模板和 HSV 直方图匹配的视频典型目标检测算法, 融合方法能够解决传统运动模板在提取运动目标轮廓时出现的空洞现象和提取运动目标轮廓不完整的问题, 而且具有较高的检测精确率, 从而有利于后续的视频解析和视频挖掘。

基金项目

本文由国家自然科学基金(61371143), 北方工业大学 2018 年教育教学改革和课程建设研究项目(18XN009-002), 教育部高等教育司产学合作协同育人项目(201801121002), 北方工业大学优势学科项目(217051360018XN044)资助。

参考文献

- [1] Pan, F., Wang, X.J. and Wang, W.H. (2016) Research and Restoration Technology of Video Motion Target Detection Based on Kernel Method. *International Journal on Smart Sensing and Intelligent Systems*, **7**, 1516-1534.
- [2] Guo, Q.Y. and Zhang, Z. (2016) Improvement of the Gaussian Mixture Model Based on EmguCV Motion Target Detection Design. *International Conference on Management*, **10**, 561-565.
- [3] Wang, H.B., Chen, Z., Lu, M., *et al.* (2018) Object Detection Method Based on Motion Saliency Probability Map. *Journal of Image and Graphics*, **23**, 79-88.
- [4] 贾建英, 董安国. 基于联合直方图的运动目标检测算法[J]. 计算机工程与应用, 2016, 52(5): 199-203.
- [5] 孙挺, 齐迎春, 耿国华. 基于帧间差分 and 背景差分的运动目标检测算法[J]. 吉林大学学报(工学版), 2016, 46(4): 1325-1329.
- [6] Bobick, A., Davis, J., Intille, S., *et al.* (1996) Action Recognition in an Interactive Story Environment, PerCom TR 398. MIT Media Lab.
- [7] Murtaza, F., Yousaf, M.H. and Velastin, S.A. (2017) Multi_view Human Action Recognition Using 2D Motion Templates Based on MHIs and Their HOG Description. *IET Computer Vision*, **10**, 758-767. <https://doi.org/10.1049/iet-cvi.2015.0416>
- [8] 吴飞虎. 基于多摄像机的人体目标跟踪技术实现[D]: [硕士学位论文]. 杭州: 浙江理工大学, 2017.
- [9] 张伟伟, 杨正洪, 牛少彰. 基于 HSV 和 CLAHE 的复制粘贴篡改检测算法[J]. 北京邮电大学学报, 2016, 39(6): 27-32.

-
- [10] 王培玉, 李峰, 周书仁, 等. 复杂码头环境下的船舶检测与跟踪算法[J]. 计算机工程与科学, 2017, 39(5): 992-998.
- [11] 王建文, 林劼. 基于颜色直方图金字塔的图像自动标注方法[J]. 计算机工程, 2016, 42(6): 235-240.
- [12] 何颖妮, 邵党国, 刘东权. 直方图匹配算法在超声弹性成像上的应用研究[J]. 计算机应用研究, 2013, 30(4): 1266-1270.
- [13] 李晓瑜, 马大中, 英杰. 基于三帧差分混合高斯背景模型运动目标检测[J]. 吉林大学学报(信息科学版), 2018, 36(4): 414-422.
- [14] 王宝珠, 胡洋, 郭志涛, 等. 基于朗斯基函数的混合高斯模型运动目标检测[J]. 计算机应用研究, 2016, 33(12): 3880-3883.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2161-8801, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: csa@hanspub.org