

一种快速鲁棒的视频序列 运动目标检测方法

秦晓燕¹,袁广林¹,李从利²,张旭¹

(1. 解放军陆军军官学院十一系,安徽合肥 230031; 2. 解放军陆军军官学院三系,安徽合肥 230031)

摘 要: 稀疏表示已经成为运动目标检测的有效方法之一,但其还没有很好地解决目标检测的快速性和鲁棒性. 本文基于最大后验概率提出了一种快速鲁棒的运动目标检测模型,并设计了该模型的求解算法. 该算法包括两个阶段:在第一阶段利用编码迁移实现稀疏系数的快速求解;在第二阶段基于运动目标的空间连续性结构,利用图切实现目标检测. 在多个具有挑战性的图像序列上的实验结果表明,与其他经典运动目标检测算法相比,本文方法在快速性和鲁棒性方面具有较优的性能.

关键词: 运动目标检测; 稀疏表示; 编码迁移; 图切

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112 (2017)10-2355-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2017.10.007

An Approach to Fast and Robust Detecting of Moving Target in Video Sequences

QIN Xiao-yan¹, YUAN Guang-lin¹, LI Cong-li², ZHANG Xu¹

(1. Eleventh Department, Army Officer Academy of PLA, Hefei, Anhui 230031, China;

2. Third Department, Army Officer Academy of PLA, Hefei Anhui 230031, China)

Abstract: Sparse representation is one of effective methods in dealing with the moving object detection. However, the quickness and robustness of object detection are far from being solved in the existing methods. In this paper, a fast and robust moving object detection model based on the maximum posteriori probability is proposed, and a two-stage detection algorithms is designed. At the first stage, sparse coefficient is quickly solved by using coding transfer; At the second stage, based on spatial continuity structure, moving object detection is achieved by using graph cut. The experimental results on several challenging image sequences show that the proposed method has better performance than the existing classical moving object detection algorithms in rapidity and robustness.

Key words: moving object detection; sparse representation; coding transfer; graph cut

1 引言

运动目标检测是计算机视觉研究的一个热门课题,在智能安防、交通监控等方面应用广泛. 近几十年,运动目标检测技术取得了快速发展,但仍然面临着光照变化、背景干扰、遮挡等难题,传统的目标检测方法并不能达到理想的检测效果. 由于稀疏表示方法符合人类视觉系统的描述方式,实现图像的有效表示,为运动目标检测提供了新思路,近年来,基于稀疏表示的运动

目标检测受到了广泛关注^[1-6].

2009年, Ma^[7]提出了基于稀疏表示的人脸识别,该方法提高了人脸识别的准确率,特别是对于有遮挡或损毁情况下的人脸具有优越的识别能力. 受稀疏表示在人脸识别中的启发,人们开始关注稀疏表示在目标检测中的应用. 基于稀疏表示的运动目标检测通过计算测试图像在超完备字典下的稀疏编码,然后利用稀疏编码重构出背景模型,实现运动目标检测. Huang^[8]利用前景的空间连续性这一先验知识,提出组稀疏理

收稿日期:2016-02-24;修回日期:2016-08-12;责任编辑:孙瑶

基金项目:安徽省自然科学基金(No. 1508085QF114, No. 1608085QF144);国家自然科学基金(No. 61379105);中国博士后科学基金(No. 2014M562535)

论,根据前景的组稀疏结构和背景在字典上的稀疏表示,实现基于稀疏表示的运动目标检测.在稀疏表示的运动目标检测方法中,字典的鲁棒性对检测性能至关重要.Zhao^[9]和 Lu^[10]通过在线非负字典学习,即采用在线重构的背景来更新背景模型,在一定程度上避免场景的动态变化对目标检测结果的影响.字典学习和稀疏编码均是迭代求解最优解的过程,运算时间较长,另外,在实际应用中,图像分辨率较高,而构建的超完备字典维数是图像分辨率的平方,导致算法的空间复杂性较大,影响了该类运动目标检测算法的快速性.

与此同时,矩阵的稀疏性也被用来解决视频序列的运动目标检测问题.2010年,Candes^[11]提出了主成分追踪方法(Principal Component Pursuit, PCP)用于视频序列的背景建模,PCP方法将待检测视频排列为一个矩阵,利用凸优化将该矩阵分解为低秩矩阵和稀疏矩阵,其中低秩矩阵用于重建背景,稀疏矩阵对应前景.考虑到观测信号中包含测量噪声,Zhou^[12]提出了稳定的主成分追踪算法(Stable Principal Component Pursuit, SPCP),PCP和SPCP算法对于挑战性的视频(如背景中包含树叶的摆动、水面波纹、扶手电梯的运动等)检测效果不理想.Zhou^[13]提出DECOLOR运动目标检测方法,该方法考虑运动目标的空间连续性约束,保持目标区域的完整性,损失函数采用 L_0 范数,有效地解决了被检测视频序列中的背景干扰问题,提高了算法鲁棒性.基于矩阵稀疏的运动目标检测方法采用批处理方式,无法进行实时检测.

基于稀疏表示的运动目标检测算法能有效提高检测性能,但同时需要付出巨大的计算时间代价,导致算法的实时性差,这也是该类运动目标检测算法在工程应用中的瓶颈.为此,本文利用编码迁移的思想快速求解稀疏表示系数,计算背景的重构误差;为保证算法的鲁棒性,考虑运动目标的空间连续性特点,利用图切算法对重构误差进行运动目标分割,实现运动目标检测.实验结果表明:本文提出的运动目标检测方法在快速性和鲁棒性两方面均达到了较好的效果.

2 运动目标检测模型

基于图像的稀疏表示理论和观测噪声模型,一幅待检测图像 $\mathbf{y} \in \mathbf{R}^m$ 在超完备字典下可表示为^[7]:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} \quad (1)$$

其中, $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \in \mathbf{R}^{m \times n}$ ($m \ll n$),是包含了 n 个样本的训练集,在运动目标检测中, \mathbf{A} 是超完备的背景字典, $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ 是 n 维稀疏系数向量, $\mathbf{e} \in \mathbf{R}^m$ 为噪声向量.

为了获得更加准确的检测结果,本文利用运动目标具有空间连续性这一先验知识,采用图切模型进行

目标分割^[14-15].为此,定义的二值向量 \mathbf{s} 作为运动目标的掩膜, $\mathbf{s} = [s_1, s_2, \dots, s_m]^T$,当 $s_i = 1$ 时, y_i 表示背景;当 $s_i = 0$ 时, y_i 表示前景.

如果 y_i 是背景($s_i = 1$),它可以表示为背景字典的线性组合,另外考虑数据采集过程中引入的高斯噪声,有 $y_i = (\mathbf{A}\mathbf{x})_i + \mathbf{e}_i$,其中 $\mathbf{e}_i \in N(0, \sigma^2)$.它的条件概率密度函数如下:

$$p(y_i | s_{i=1}, \mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{[y_i - (\mathbf{A}\mathbf{x})_i]^2}{2\sigma^2}\right\} \quad (2)$$

如果 y_i 为前景($s_i = 0$),它的条件概率是均匀分布的,且与系数向量 \mathbf{x} 无关.它的条件概率密度函数如下^[16]:

$$p(y_i | s_{i=0}, \mathbf{x}) = \begin{cases} \frac{1}{b-a}, & y_i \in [a, b] \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中, $[a, b]$ 表示图像的像素取值区间,在该区间内任意 y_i 值的概率密度相同.

定义图模型 $G = (\mathbf{V}, \mathbf{E})$,其中, $\mathbf{V} = \{y_1, \dots, y_m\}$ 表示图像的 m 个像素, \mathbf{E} 表示相邻像素之间的边集(本文采用的是四邻域边集).用图模型表示 \mathbf{s} ,由于 \mathbf{s} 是具有空间连续性的二值向量,可以利用马尔科夫随机场进行建模,基于Ising模型^[15], \mathbf{s} 的概率分布函数为:

$$p(\mathbf{s}) = \frac{1}{Z_1} \exp\left(-\sum_{(i,j) \in \mathbf{E}} \beta_{i,j} |s_i - s_j|\right) \quad (4)$$

其中, $\beta_{i,j} > 0$,是用于控制 s_i 和 s_j 之间依赖度的常量, Z_1 是归一化常量.

另外,为保证解的稳定,需要对系数 \mathbf{x} 进行稀疏约束,其概率分布函数^[17]为:

$$p(\mathbf{x}) = \frac{1}{Z_2} \exp\left(-\beta \sum_{k=1}^n |x_k|\right) \quad (5)$$

由于 \mathbf{x} 和 \mathbf{s} 满足独立同分布,因此后验概率 $p(\mathbf{s}, \mathbf{x} | \mathbf{y})$ 表示如下:

$$\begin{aligned} p(\mathbf{s}, \mathbf{x} | \mathbf{y}) &\propto p(\mathbf{y} | \mathbf{s}, \mathbf{x}) p(\mathbf{x}) p(\mathbf{s}) \\ &= \left[\prod_{i=1}^m p(y_i | s_i, \mathbf{x}) \right] p(\mathbf{x}) p(\mathbf{s}) \\ &= \left[\prod_{i=1}^m p(y_i | s_{i=1}, \mathbf{x})^{s_i} p(y_i | s_{i=0}, \mathbf{x})^{1-s_i} \right] p(\mathbf{x}) p(\mathbf{s}) \\ &= \left\{ \prod_{i=1}^m \left[\frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(y_i - (\mathbf{A}\mathbf{x})_i)^2}{2\sigma^2}\right) \right]^{s_i} \left(\frac{1}{b-a}\right)^{1-s_i} \right\} \\ &\quad \times \frac{1}{Z_1} \exp\left(-\beta \sum_{k=1}^n |x_k|\right) \times \frac{1}{Z_2} \exp\left(-\sum_{(i,j) \in \mathbf{E}} \beta_{i,j} |s_i - s_j|\right) \end{aligned} \quad (6)$$

由于 \mathbf{x} 和 \mathbf{s} 的求解可以通过最大化后验概率 $p(\mathbf{s}, \mathbf{x} | \mathbf{y})$ 获得,也等同于最小化 $-\log p(\mathbf{s}, \mathbf{x} | \mathbf{y})$.由式(6)推导得到, $-\log p(\mathbf{s}, \mathbf{x} | \mathbf{y}) = C + \frac{1}{\sigma^2} J(\mathbf{s}, \mathbf{x})$,其中, C 为常

数, $J(\mathbf{s}, \mathbf{x})$ 定义如下:

$$J(\mathbf{s}, \mathbf{x}) = \sum_{i=1}^m \left\{ s_i \frac{[y_i - (\mathbf{A}\mathbf{x})_i]^2}{2} + (1 - s_i) \frac{\lambda^2}{2} \right\} + \gamma \sum_{k=1}^n |x_k| + \sum_{(i,j) \in E} \lambda_{i,j} |s_i - s_j| \quad (7)$$

其中, $\lambda = (2\sigma^2 \log \frac{b-a}{\sqrt{2\pi\sigma}})^{\frac{1}{2}}$, $\gamma = \sigma^2 \beta$, $\lambda_{i,j} = \sigma^2 \beta_{ij}$.

3 运动目标检测算法

为了保证系数 \mathbf{x} 的稀疏约束, 并利用目标的空间连续性结构, 本文分步迭代求解式(7), 得到 \mathbf{x} 和 \mathbf{s} 的最优解.

引理 1 给定 $\hat{\mathbf{s}}$, 则可用基追踪算法 (BP) 求解 \mathbf{x} .

证明: 当 \mathbf{s} 已知, 式(7)中的目标函数只需要考虑前两项, 最后一项看作常量, 因此目标函数的最小化的问题等同于如下目标函数的最小化:

$$F(\mathbf{x}) = \sum_{s_i \neq 0} \frac{[y_i - (\mathbf{A}\mathbf{x})_i]^2}{2} + \gamma \sum_{k=1}^n |x_k| \quad (8)$$

该问题即为 L_1 范数最小化问题:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \gamma \|\mathbf{x}\|_1 \quad (9)$$

式(9)的优化问题可用 BP 算法求解^[18].

在实际应用中, 稀疏编码求解的训练集 \mathbf{A} 的维数通常在 10^8 以上, 这显然会给稀疏编码的求解带来巨大的内存开销和计算负担. 针对这一问题, 本文利用编码迁移实现稀疏表示系数的快速求解.

Yang 等^[19]指出高、低分辨率图像块相对于各自的过完备字典能够由同一稀疏系数表示出来. 因此, 对于给定的低分辨率图像, 首先求出与之对应的低分辨率字典的稀疏表示系数, 然后利用高分辨率字典乘以该稀疏表示系数, 即可恢复出对应的高分辨率背景图像, 这就是稀疏编码迁移的思想.

利用相同的采样率, 分别对背景字典 \mathbf{A} 和待检测图像 \mathbf{y} 进行采样, 得到一个维数较低的 \mathbf{A}_L 和 \mathbf{y}_L . 求解 \mathbf{y}_L 在 \mathbf{A}_L 上的稀疏编码系数 $\hat{\mathbf{x}}$, 该算法的具体描述见算法 1.

算法 1 稀疏编码迁移算法 SCT (Spare Coding Transfer)

输入: 观测图像 $\mathbf{y} \in \mathbf{R}^m$, 背景字典 $\mathbf{A} \in \mathbf{R}^{m \times n}$,

下采样因子 γ ;

输出: 稀疏编码系数 $\hat{\mathbf{x}} \in \mathbf{R}^{n \times 1}$

Step1 对检测图像和背景字典进行降维:

$\mathbf{y}_L = \text{sample}(\mathbf{y}, \gamma)$; $\mathbf{A}_L = \text{sample}(\mathbf{A}, \gamma)$;

Step2 求解 $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y}_L - \mathbf{A}_L \mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1$;

引理 2 给定 $\hat{\mathbf{x}}$, \mathbf{s} 可以用最大流/最小切算法求解.

证明: 假如 \mathbf{x} 已知, 式(7)目标函数的最小化的问题等同于如下目标函数的最小化:

$$G(\mathbf{s}) = \sum_{i=1}^m \left\{ |0 - s_i| \frac{e_i^2}{2} + |1 - s_i| \frac{\lambda^2}{2} \right\} + \sum_{(i,j) \in E} \lambda_{i,j} |s_i - s_j| \quad (10)$$

其中 $e_i = y_i - (\mathbf{A}\hat{\mathbf{x}})_i$, 式(10)的目标函数可以被视为一阶二元马尔科夫随机场的能量函数^[15], 因此该目标函数的最小化问题就是 \mathbf{s} 的能量最小化问题, 利用图切算法即可求解.

由引理 1 和引理 2, 式(7)可分两个步骤求解: 首先固定目标掩模向量 \mathbf{s} , 求解背景的稀疏系数 \mathbf{x} ; 然后固定 \mathbf{x} , 求解 \mathbf{s} . 循环迭代这两个步骤直至算法满足终止条件, 图 1 给出了该算法流程, 算法 2 给出了算法的具体实现过程.

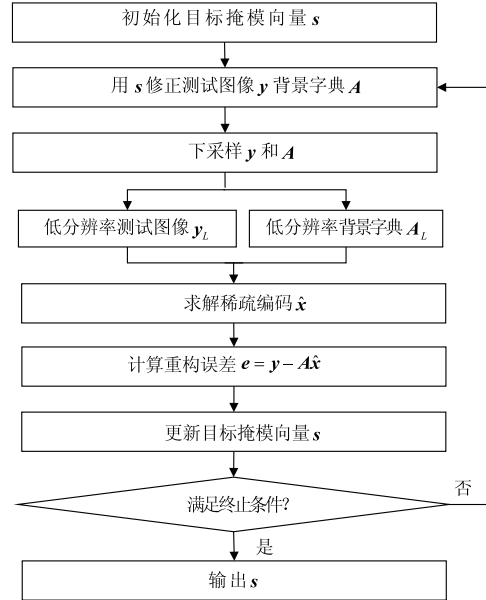


图1 本文算法流程图

算法 2 快速鲁棒运动目标检测算法

输入: 观测图像 $\mathbf{y} \in \mathbf{R}^m$, 背景字典 $\mathbf{A} \in \mathbf{R}^{m \times n}$, 下采样因子 γ ;

输出: 运动目标二值化图像 $\mathbf{s} \in \mathbf{R}^m$

步骤 1. 初始化 $\mathbf{s}^{(0)} = \mathbf{1}_m$, $t = 0$, $\max = 20$;

repeat;

步骤 2. $t = t + 1$;

步骤 3. 用 \mathbf{s} 的值修正 \mathbf{A} 和 \mathbf{y} :

$$\mathbf{y}^* = \mathbf{y}[\mathbf{s}^{(t-1)} = 1], \mathbf{A}^* = \mathbf{A}[\mathbf{s}^{(t-1)} = 1];$$

步骤 4. 用算法 1 求解稀疏编码系数: $\mathbf{x} = \text{SCT}(\mathbf{y}^*, \mathbf{A}^*, \gamma)$;

步骤 5. 计算重构误差 $\mathbf{e} = \mathbf{y} - \mathbf{A}\mathbf{x}$;

步骤 6. 利用图切算法更新 \mathbf{s} :

$$\mathbf{s}^{(t)} = \arg \min_{\mathbf{s} \in \{1,0\}^m} \sum_{i=1}^m \left(|0 - s_i| \frac{e_i^2}{2} + |1 - s_i| \frac{\lambda^2}{2} \right) + \sum_{i,j \in E} \lambda_{i,j} |s_i - s_j|;$$

步骤 7. 计算 $\mathbf{s}^{(t)}$ 的能量值 $E^{(t)}$;

$$E^{(t)} = \sum_{i=1}^m \left\{ |0 - s_i^{(t)}| \frac{e_i^2}{2} + |1 - s_i^{(t)}| \frac{\lambda_i^2}{2} \right\} + \sum_{(i,j) \in E} \lambda_{i,j} |s_i^{(t)} - s_j^{(t)}|;$$

until $\text{abs}(E^{(t)} - E^{(t-1)}) < 1e-3$ 或者 $t > \max$

步骤 8. 令 $s = s^{(t)}$

为简单起见,在检测之前,假定待检图像中没有目标,根据目标掩模向量 s 的定义,该值初始化为全 1,当检测到目标时,将 s 中与之对应的值置为 0,算法结束时, s 即为运动目标的二值图像.

算法 2 中的终止条件是相邻两次迭代得到的能量差值小于设定的阈值,另外,如果迭代次数超过设定的最大迭代次数,则强行终止.对多组视频序列进行实验,发现算法 2 在迭代 2 至 3 次即可满足终止条件,说明算法具有较好的收敛性,算法收敛性实验详见 4.2.3 节.

4 实验分析

4.1 实验说明

以 Matlab R2009a 为开发工具,在 Inter(R) Forth-Core 2.50GHz CPU,4G 内存的台式机上,实现了本文提出的运动目标检测方法.用 3 个代表性的视频序列(Airport、Hall 和 Watersurface),以及一个炮弹炸点视频(Bomb)对提出的运动目标检测方法进行了实验验证(表 1 给出了 4 个实验数据详细情况),并从检测准确性和检测时间两方面分别与文献[9]中的 BSRDL 算法,文献[11]中的 PCP 算法和文献[13]中的 DECOLOR 算法进行比较.

表 1 实验数据说明

实验视频	大小 × 帧数	描述
Hall	[144 × 176] × 48	场景复杂:运动目标较多; 背景动态变化:运行中的手扶电梯.
Airport	[144 × 176] × 48	场景复杂:运动目标较多.
Watersurface	[128 × 160] × 24	背景动态变化:水面波纹的运动.
Bomb	[288 × 352] × 30	背景动态变化:视频采集 设备抖动带来的背景变化.

4.2 实验结果与分析

本文研究的主要目的是提高运动目标检测的实时性,同时算法在处理光照变化、背景干扰等复杂环境下的视频序列时,具有较高的鲁棒性.下面主要从算法检测的时间和鲁棒性两方面进行实验分析,另外,在本节最后给出算法收敛性实验.

4.2.1 算法实时性实验

表 2 给出了 4 种算法在 4 组数据上的检测时间,为方便比较,该表中给出的时间是每一帧图像的检测时间.由于 DECOLOR 和 PCP 算法采用是批处理方式,处

理的数据是一段视频序列,因此其取值为每一帧的平均检测时间.而本文算法和 BSRDL 算法采用在线处理方式,每次处理一帧图像,其时间为算法实际运行时间.通过表 2 中的数据对比可知,本文算法在实时性上优于其他算法,比 BSRDL 算法的检测时间大约快了近 10 倍,这是由于本文算法利用编码迁移的思想,降低了求解问题的规模,从而提高了稀疏编码求解速度.另外需要说明的是,利用矩阵分解的算法(DECOLOR 和 PCP)采用批处理方式,实际的检测时间为每帧的平均检测时间 × 视频帧数.

表 2 算法运行时间对比

实验数据	运行时间(秒/帧)			
	BSRDL	PCP	DECOLOR	本文算法
Hall	3.1	1.01	2.1	0.28
Airport	2.87	0.53	0.86	0.23
Watersurface	2.80	0.49	0.73	0.23
Bomb	5.59	2.68	3.6	0.8

4.2.2 算法鲁棒性实验

图 2 给出了 4 种检测方法对 4 个视频序列中运动目标检测的结果.从左至右,第一列是原始序列及其对应的 ground truth,第二列是 BSRDL 的检测结果,第三列是 PCP 检测结果,第四列是 DECOLOR 的检测结果,最后一列是本文算法的检测结果.其中,BSRDL 和 PCP 显示的是对前景进行阈值分割后的结果,在实验中发现,BSRDL 和 PCP 算法的检测结果受阈值的影响很大,图 2 中 BSRDL 和 PCP 的检测结果是基于大津法^[20]全局阈值分割后的结果. DECOLOR 和本文算法的检测结果没有经过任何处理.

从实验结果看,BSRDL 和 PCP 重建的背景与真实的背景存在差距,特别是 PCP 算法重建的背景出现明显的鬼影,以上两种算法虽然能检测到目标,但是检测到的目标区域不完整,存在空洞现象.对于复杂场景下的图像序列,例如 Airport 视频中存在手扶电梯的运动,BSRDL 和 PCP 算法均将手扶电梯误检为运动的目标,虽然经过形态学处理会消除部分误检的目标,但形态学的处理也会导致真实目标信息的损失,影响算法性能.从四组实验数据的检测结果来看,本文算法与 DECOLOR 算法检测结果相似,均较好地检测出运动目标,特别是在检测存在背景干扰的视频序列时(例如 Airport 和 Watersurface),运动的手扶电梯和水面波纹并没有被误检,表现出较好的鲁棒性.

需要说明的是,由于本文算法在实时性和鲁棒性方面的改进,在一定程度上影响了目标检测的精度,例如当视频序列中存在位置较为集中的小目标时,出现了小目标合并现象(如 Airport 和 Hall 的检测结果).

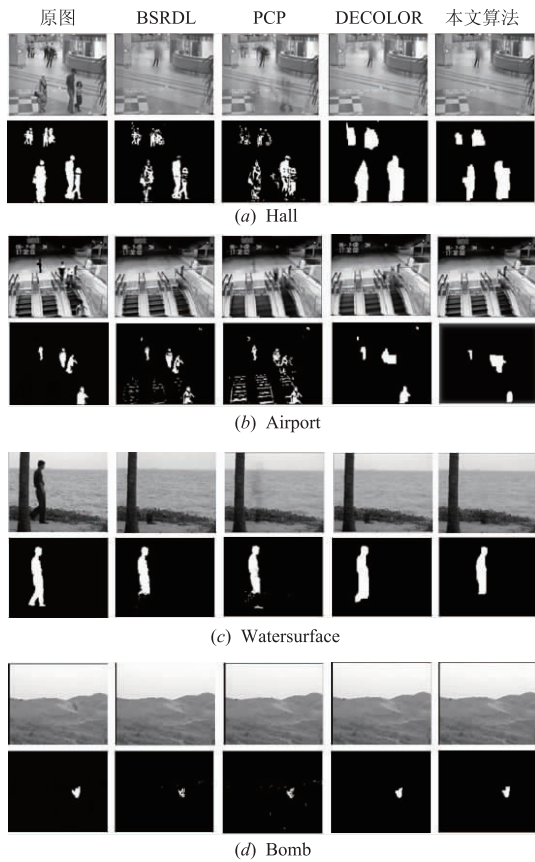


图2 运动目标检测结果比较

为进一步验证算法性能,本文将运动目标检测作为二分类问题来看待,其中目标对应二分类中的正样本,背景对应负样本,使用准确率 (precision) 和召回率 (recall) 对算法的检测性能进行评估比较,准确率和召回率定义如下:

$$\text{precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{recall} = \frac{TP}{TP + FN}$$

其中 TP 为正确检测到的目标个数;FP 为漏检的目标个数;FN 为误检的目标个数.为简单起见,本文使用平均检测性能 (F-measure) 进行算法评价,F-measure 是和检测准确率和召回率均相关的评价指标,定义如下:

$$\text{F-measure} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

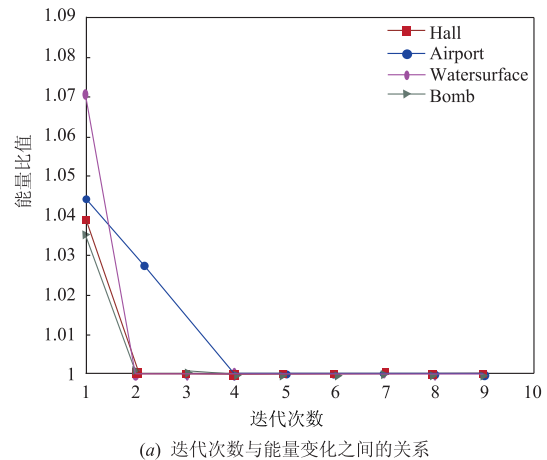
F-measure 的值越高表示检测准确率越高.表 3 列出四种检测算法在 4 组数据上的 F-measure 值,对比发现,DECOLOR 算法与本文算法的 F-measure 值接近,并远远高于 BSRDL 和 PCP 算法的 F-measure 值,特别是对于背景动态变化的图像检测效果的差距尤为明显.

表 3 不同方法在 4 个数据上的 F-measure 值比较

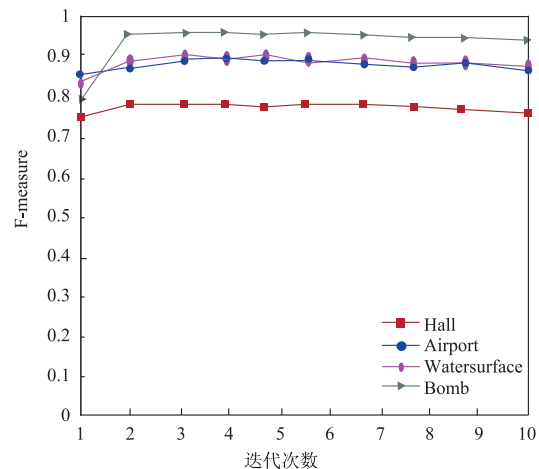
	BSRDL	PCP	DECOLOR	本文算法
Hall	0.72	0.70	0.92	0.91
Airport	0.55	0.49	0.82	0.80
Watersurface	0.90	0.83	0.91	0.91
Bomb	0.80	0.76	0.95	0.95

4.2.3 算法收敛性实验

在第 3 节提到算法 2 具有较好的收敛性,下面给出算法 2 在 4 组数据集上的收敛性实验,为进一步说明情况,实验中设置算法的迭代次数为 10 次,并记录每一次迭代的能量值 $E^{(t)}$,图 3(a) 给出了迭代次数与能量变化之间的关系,其中横坐标是迭代次数,纵坐标是相邻两次能量比值 ($E^{(t)}/E^{(t+1)}$).从实验结果看,算法迭代 2~3 次后,能量比值接近 1,这说明此时能量值趋于稳定,几乎不再发生变化,算法满足终止条件.图 3(b) 给出随着迭代次数的增加,算法检测性能发生的变化.从 4 组不同数据集的实验结果看,检测性能在迭代 2~3 次后均趋于稳定,这与能量值稳定



(a) 迭代次数与能量变化之间的关系



(b) 迭代次数对检测性能的影响

图3 算法收敛性分析

时的迭代次数相吻合,说明算法收敛时,检测性能也达到最优.

5 结论

本文在稀疏表示的框架下,提出了一种快速鲁棒的视频序列运动目标检测算法,首先从降低问题规模的角度出发,引入了编码迁移的思想,快速求解测试图像在字典上的表示系数,进而提高了算法的实时性;其次,利用目标的空间连续性特点,采用图切算法进行图像分割,提高了算法鲁棒性.实验结果表明,本文算法适用于处理对实时性要求较高,或背景存在干扰的视频序列的运动目标检测问题.本文为提高检测时间的快速性,没有考虑背景字典的在线更新,这导致对全局背景变化的视频序列的检测效果不够理想.

参考文献

- [1] KULCHANDANI J S, DANGARWALA K J. Moving object detection: Review of recent research trends [A]. Proceedings of International Conference on Pervasive Computing [C]. Pune: IEEE Press, 2015: 1 - 5.
- [2] LUO Yuhua, ZHANG Hong. Sparse learning for robust background subtraction of video sequences [A]. Proceedings of Intelligent Computing Theories and Methodologies [C]. Fuzhou, China: Springer International Publishing, 2015: 400 - 411.
- [3] CUI X, HUANG J, ZHANG S, et al. Background subtraction using low rank and group sparsity constraints [A]. Proceedings of IEEE Conference on Computer Vision [C]. Florence, Italy: IEEE Computer Society Press, 2012: 612 - 625.
- [4] BO Xin-yuan, TIAN Yizhou, WANG Wen-gao. Background subtraction via generalized fused lasso foreground modeling [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Boston, MA: IEEE Computer Society Press, 2015: 4676 - 4684.
- [5] 高仕博,程咏梅,肖利平,等.面向目标检测的稀疏表示方法研究进展[J].电子学报,2015,43(2):320-332.
GAO Shi-bo, CHENG Yong-mei, XIAO Li-ping, et al. Recent advances of sparse representation for object detection [J]. Acta Electronica Sinica, 2015, 43(2): 320 - 332. (in Chinese)
- [6] 蒋建国,金玉龙,齐美彬,等.基于稀疏表达残差的自然场景运动目标检测[J].电子学报,2015,43(9):1738-1744.
JIANG Jian-guo, JIN Yu-long, QI Mei-bin, et al. Moving target detection in natural scene based on sparse representation of residuals [J]. Acta Electronica Sinica, 2015, 43(9): 1738 - 1744. (in Chinese)
- [7] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210 - 227.
- [8] HUANG Jun-zhou, HUANG Xiao-lei, METAXAS Dimitris. Learning with dynamic group sparsity [A]. Proceedings of IEEE International Conference on Computer Vision [C]. Los Alamitos: IEEE Computer Society Press, 2009: 64 - 71.
- [9] ZHAO Cong, WANG Xiao-gang, CHAM Wai-Kuen. Background subtraction via robust dictionary learning [J]. EURASIP Journal on Image and Video Processing, 2011, 2011(2): 1 - 12.
- [10] LU Cewu, SHI Jian-ping, JIA Jia-ya. Online robust dictionary learning [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Portland, OR: IEEE Press, 2013: 415 - 422.
- [11] CANDÈS E J, LI X D, MA Y, et al. Robust principal component analysis? [J]. Journal of the ACM, 2011, 58(3): 1 - 37.
- [12] ZHOU Z H, LI X D, WRIGHT J, et al. Stable principal component pursuit [A]. Proceedings of the 2010 IEEE International Symposium on Information Theory [C]. Austin, TX: IEEE Press, 2010: 1518 - 1522.
- [13] ZHOU Xiao-wei, YANG Can, YU Wei-chuan. Moving object detection by detecting contiguous outliers in the low-rank representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(3): 597 - 610.
- [14] SCHICK A, BAUML M, STIEFELHAGEN R. Improving foreground segmentations with probabilistic superpixel markov random fields [A]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops [C]. Providence, RI: IEEE Press, 2012: 27 - 31.
- [15] LI S Z. Markov Random Field Modeling in Image Analysis [M]. London: Springer, 2009.
- [16] 平岡和幸,堀玄著,陈筱烟译.程序员的数学2 概率统计[M].北京:人民邮电出版社,2015.
- [17] ELAD Michael. Sparse and Redundant Representations From Theory to Application in Signal and Image Processing [M]. New York: Springer, 2010.
- [18] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit [J]. SIAM Journal on Scientific Computing, 2001, 20(1): 33 - 61.
- [19] YANG J, WRIGHT J, HUANG T, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861 - 2873.
- [20] 何志勇,孙立宁,陈立国. Otsu 准则下分割阈值的快速

计算[J]. 电子学报, 2013, 41(2): 267 - 272.
HE Zhi-yong, SU Li-ning, CHEN Li-guo. Fast computa-

tion of threahold based on otsu criterion [J]. Acta Elec-
tronica Sinica, 2013, 41(2): 267 - 272. (in Chinese).

作者简介



秦晓燕 女, 1980 年 2 月出生, 安徽淮北人. 2004 年毕业于解放军炮兵学院获工学硕士学位、讲师. 主要研究方向为图像处理、目标检测、机器学习及应用.
E-mail: 70853559@ qq. com



袁广林 男, 1973 年 8 月出生, 河南周口人. 2013 年毕业于合肥工业大学获博士学位, 副教授. 研究方向为图像处理、计算机视觉、机器学习及其应用.
E-mail: yuanguanglin1008@ sina. com