

一种基于证据推理的视频运动目标融合检测算法

韩德强^{a,b}, 程博^{a,b}, 杨艺^{c,d}

(西安交通大学 a. 智能网络与网络安全教育部重点实验室, b. 电子与信息工程学院,
c. 机械振动与强度国家重点实验室, d. 航天航空学院, 西安 710049)

摘要: 为了解决复杂场景下传统的运动目标检测问题, 利用证据推理——谨慎有序加权平均方法(COWA-ER), 提出一种综合使用混合高斯、均值滤波和码本的多方法融合的检测方法. 该融合检测算法以上述3种检测方法为准则建立一个多准则决策框架, 通过双阈值检测法来表征检测过程中的不确定性, 最终利用COWA-ER方法进行决策级融合, 实现多种方法的优势互补. 实验表明, 所提出的融合检测算法具有更理想的目标检测效果, 能有效应对诸如阴影及光照突变等问题对检测性能的影响.

关键词: 视频目标检测; 不确定性; COWA-ER; 融合检测; 多准则决策

中图分类号: TP273

文献标志码: A

A fusion approach for video moving object detection based on evidential reasoning

HAN De-qiang^{a,b}, CHENG Bo^{a,b}, YANG Yi^{c,d}

(a. Ministry of Education Key Lab for Intelligent Networks and Network Security, b. School of Electronic and Information Engineering, c. State Key Laboratory of Mechanical Vibration and Strength, d. School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China. Correspondent: HAN De-qiang, E-mail: deqhan@mail.xjtu.edu.cn)

Abstract: To handle the problems encountered in traditional moving object detecting algorithms under complex scenarios, by comprehensively using the Gaussian mixture model(GMM), mean filter and codebook(CB), a fusion-based detection approach is proposed based on cautious ordered weighted averaging with evidential reasoning(COWA-ER). In the proposed approach, a multi-criteria decision-making framework is established, where the three detection algorithms are used as multiple criteria. The double threshold method is used to model the uncertainty in the detection. The decision-level fusion for detection is finally accomplished by using COWA-ER, where complementary advantages of the three algorithms can be fully used. Experimental results show that the proposed approach can achieve better performance for the detection, and effectively deal with the influence of detection performance caused by shadow and light change.

Keywords: video moving object detection; uncertainty; cautious ordered weighted averaging with evidential reasoning; fusion-based detection; multi-criteria decision making(MCDM)

0 引言

视频运动目标检测^[1]旨在从图像序列(视频)中将运动目标从背景图像中尽可能完整地分割提取, 是视频及图像后期处理过程(如目标分类、跟踪和行为理解)的前提和基础, 引发了越来越多的关注.

常见的视频运动目标检测方法有光流法、帧间差分法和背景差分法等^[1-3]. 背景差分技术是最为

常用的运动目标检测方法之一, 其基本思想是构建背景图像或背景模型, 利用当前帧与背景模型的差异检测出运动目标区域. 代表性方法包括混合高斯检测法^[4-6]、均值滤波检测法^[7]、码本检测法^[8-9]. 其中Stauffer等^[4]提出的混合高斯法是将每个像素按照多个高斯分布混合建模. 均值滤波法是先建立一个视频流滑窗, 用来缓存 L 张视频帧, 然后将缓存中

收稿日期: 2015-07-23; 修回日期: 2015-10-23.

基金项目: 国家973计划项目(2013CB329405); 国家自然科学基金项目(61573257, 61203222); 陕西省科技计划项目(2013KJXX-46); 教育部博士点基金项目(20120201120036); 中央高校基本科研业务费专项资金项目(xjj2012104, xjj2014122).

作者简介: 韩德强(1980—), 男, 副教授, 博士生导师, 从事信息融合、模式分类等研究; 程博(1991—), 男, 硕士生, 从事视频运动目标检测与跟踪的研究.

所有视频帧同位置像素的平均值作为背景图像。Kim等^[10-11]提出的码本法是通过长时间观察彩色视频序列,利用量化和聚类技术构建码本背景模型。

传统的背景建模算法应对复杂场景时都存在各自的缺点:混合高斯模型不能适应快速的背景变化,对噪声变化比较敏感,易产生“虚影”现象^[6];均值滤波算法背景训练时间较长,检测时阈值难以精确确定,自适应能力差;码本模型难以适应如光照变化、背景突变等复杂的动态背景情况。各算法检测阈值都为主观选取,使得检测结果均具有一定的不确定性。针对上述算法的缺陷,本文提出一种基于证据推理的融合检测算法,期望通过对多种检测方法的综合利用来应对复杂场景下目标检测中的阴影、光线突变等问题,以提高算法在复杂环境下的鲁棒性和准确性。实验结果表明,本文所提出的方法合理有效。

1 基于背景差分的目标检测方法概述

背景差分原理相对简单,当背景较稳定时检测效果较好,故成为目标检测中最常用的一种方法。其实现流程如图1所示。

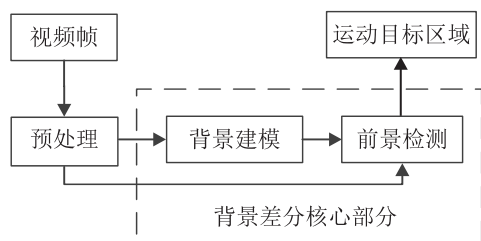


图1 背景差分法目标检测流程

背景模型的构建和更新方法直接影响背景模型对场景变化的适应性和前景目标检测的效果。下面分别介绍3种有代表性的背景建模算法。

1.1 混合高斯背景建模

混合高斯模型的基本思想是:对每个像素点,定义 K 个状态分布(K 的建议取值为3~5)。每个状态分布由高斯函数模型表示,通过加权拟合 K 个分布得到每个像素点的联合概率密度函数

$$P(X_t) = \sum_{k=1}^K \omega_{k,t} \cdot \eta(X_t, \mu_{k,t}, \sigma_{k,t}^2 | I). \quad (1)$$

其中: X_t 为像素点灰度值; K 为单高斯模型总数, $\omega_{k,t}$ 为第 k 个高斯模型在 t 时刻的权值; $\eta(X_t, \mu_{k,t}, \sigma_{k,t}^2 | I)$ 为第 k 个高斯模型在 t 时刻的概率密度函数,其均值为 $\mu_{k,t}$,协方差矩阵为 $\sigma_{k,t}^2 | I$ 。

基于混合高斯背景建模的目标检测算法^[4-5]具体步骤如下。

Step 1: 混合高斯模型初始化。选取视频第1帧图像对混合高斯模型中的第1个高斯模型的均值进行初始化,其他高斯模型的均值取值为零。所有高斯模

型的方差取一个相等且较大的初始值,权重相等,取为 $1/K$ 。

Step 2: 模型参数更新。假设通道相互独立,对输入新的像素灰度值 I_t ,按下式判断是否与高斯模型匹配(c 为匹配系数):

$$\left| \frac{I_t - \mu_{k,t-1}}{\sigma_{k,t-1}} \right| \leq c. \quad (2)$$

若 I_t 与该模型匹配,则更新该高斯模型;若 I_t 与所有的高斯模型均不匹配,则将权值最小的模型删除,加入一个新的高斯模型,然后将高斯模型的权值 $\omega_{k,t}$ 归一化。

Step 3: 背景模型的生成。将高斯模型按照 $\omega_{k,t}/\sigma_{k,t}$ 的值由大到小排序,选择前 B 个模型作为背景模型,其中

$$B = \arg \min_b \left(\sum_{k=1}^b \omega_{k,t} > T \right), \quad (3)$$

T 为权重阈值。

Step 4: 前景背景分割。将当前点像素值与背景模型中 B 个高斯模型依次进行匹配。若存在匹配,则判断为背景点;若无匹配,则判断为前景点。

1.2 均值滤波背景建模

传统均值滤波^[7]的基本思想是:首先选取多帧视频图像的平均作为背景图像,然后将后续新输入图像和背景图像相减进行背景消除。故其需要存储之前 N 帧的图像,这样不仅需要较大的存储空间而且运算速度也较慢。有学者提出了改进的均值滤波方法^[8],引入背景更新系数 λ 来更新背景图像,基本步骤如下。

Step 1: 背景图像初始化。对视频中前 N 帧采用均值滤波法提取背景图像,作为初始背景图像。

Step 2: 背景更新。通过下式:

$$BG_{k+1}(x, y) = (1 - \lambda)BG_k(x, y) + \lambda f_k(x, y) \quad (4)$$

进行背景模型的更新,其中 (x, y) 为像素点的位置, BG_k 为第 k 帧图像对应的背景图像, f_k 为第 k 帧图像, λ 为背景更新系数。

Step 3: 目标提取。采用阈值分割法将前景从背景中分离出来,即

$$B_k(x, y) = \begin{cases} 0, & D_k(x, y) \leq T; \\ 1, & D_k(x, y) > T. \end{cases} \quad (5)$$

其中: D_k 为第 k 帧图像与背景的差分图像; B_k 为二值图像; T 为分割阈值,可采用大津法^[9](OTSU)求取阈值。

1.3 改进码本背景建模

Kim等^[10-11]提出的码本检测算法基本思想是:根据像素点的颜色失真程度及亮度范围将背景像素用一个或者多个码本表示,然后将新输入像素值与其

对应码本进行比较判断,从而提取出前景运动目标.码本检测可以在 RGB 色彩空间实现,亦可在 YUV 色彩空间^[12]进行.

YUV 色彩空间下的码本背景模型简介如下.

对每一个像素点建立一个码本模型,这个码本模型中包含若干个码字,记为 $C = \{c_1, c_2, \dots, c_N\}$,每个码字包含一个 YUV 向量 v_k 和一个 10 维向量 u_k ,即

$$\begin{cases} v_k = (\overline{Y}_k, \overline{U}_k, \overline{V}_k), \\ u_k = (Y_k^{\min}, Y_k^{\max}, U_k^{\min}, U_k^{\max}, \\ V_k^{\min}, V_k^{\max}, f_k, \lambda_k, p_k, q_k). \end{cases} \quad (6)$$

其中: k 为码本中码字的索引; $Y_k^{\min}, Y_k^{\max}, U_k^{\min}, U_k^{\max}, V_k^{\min}, V_k^{\max}$ 为像素的 YUV 分量的最小、最大值; f_k 为该码字被匹配上的次数; λ_k 为该码字最长不更新的时间; p_k 为码字第 1 次被匹配上的时间; q_k 为码字最后 1 次被匹配上的时间.

YUV 色彩空间码本检测算法步骤如下.

Step 1: 码本模型建立过程.

Step 1.1: 选取视频的前 L 帧作为训练序列,将每个像素点的码本置空,即 $C = \emptyset$.

Step 1.2: 使用前 L 帧视频序列依次进行码本模型训练.如果像素点的码本为空,则为像素点创建一个新的码字;如果码本不为空,则判断码本中是否有码字与当前像素值匹配.判断条件如下式所示:

$$\begin{cases} Y_{\text{Low}} \leq Y_t \leq Y_{\text{High}}, \\ U_{\text{Low}} \leq U_t \leq U_{\text{High}}, \\ V_{\text{Low}} \leq V_t \leq V_{\text{High}}, \end{cases} \quad (7)$$

其中 $[Y_{\text{Low}}, Y_{\text{High}}], [U_{\text{Low}}, U_{\text{High}}], [V_{\text{Low}}, V_{\text{High}}]$ 为对应的匹配区间.若无匹配码字,则为该像素点创建一个新的码字;若有匹配的码字 c_k ,则对码字进行更新.

Step 1.3: 训练结束后,利用码字中最长的不更新时间来消除冗余码本,得到精炼后的初始背景码本

$$C = \{c_k | c_k \in C \cap \lambda_k \leq T_M\}, \quad (8)$$

其中阈值 T_M 一般取训练帧数 L 的一半.

Step 2: 运动目标提取.根据式 (7) 的条件判断像素点新输入像素值与码本是否匹配,若存在码字匹配,则判断为背景点,否则为前景点.

Step 3: 码本模型更新.为了适应环境的变化和目标运动变化引起的背景改变,在运动目标提取过程的同时,可定期对码本背景模型进行更新,更新方法同 Step 1.在码本模型定期更新过程中,会将当前帧运动目标区域的像素点值引入码本的背景模型中,可能导致后续一段时间内检测结果出现闪烁或目标空洞的情况.因此,每个像素点除了背景码本,还需要建立一个缓冲码本,用于存储未匹配的像素点值.

若缓冲码本中的码字更新频繁,表明是新出现的背景,则将其移到背景码本中,同时将其从缓冲码本中剔除;若码字长时间不更新,则将其从缓冲码本中剔除.为了进一步精炼背景码本,将长时间不更新的冗余码字从背景码本中删除.

1.4 3 种检测方法不确定性分析

混合高斯检测法通过式 (2) 判断当前像素点是否与背景模型中的高斯模型匹配来实现前景背景分割,而均值滤波检测法是通过式 (5) 来实现前景背景分割,这两种方法均采用了阈值判断法.然而,阈值选取的不同,将会给检测的结果带来不确定性.码本检测法通过式 (7) 判断像素点当前值是否与背景码本中的码字匹配来实现前景背景分割.每个通道的匹配区间的设定也具有不确定性,并且当前值与区间匹配还存在匹配程度的问题.

综上,上述 3 种检测方法都存在一定的不确定性,会对目标检测性能造成影响.针对不确定性这个问题,本文提出双阈值法对单一检测方法中的不确定性进行建模与描述,同时采用一种基于证据推理的融合方法 (COWA-ER) 实现多算法的融合检测,以期实现优势互补,提升检测性能.

2 基于 COWA-ER 的融合检测算法

本文在对单一检测算法的不确定性分析与建模基础上,提出采用 COWA-ER 融合规则将上述 3 种检测算法进行融合以实现运动目标融合检测的算法,基本流程如图 2 所示.

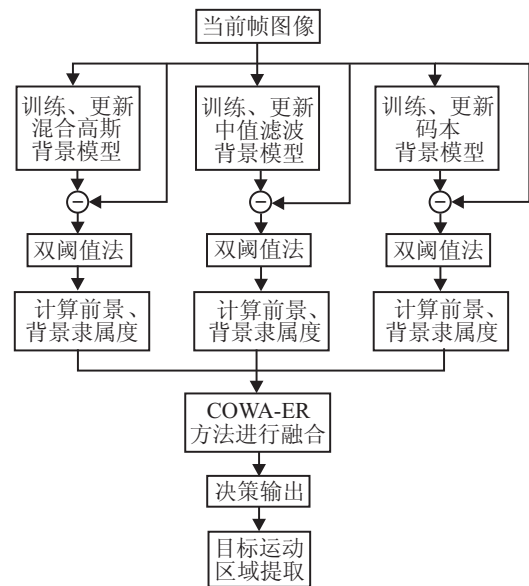


图 2 基于 COWA-ER 的视频运动目标融合检测算法

按照图 2 中的流程顺序,这里首先介绍不确定性建模方法.

2.1 背景差分法不确定性的双阈值建模

上文分析了 3 种基于背景差分的检测方法均具

有一定的不确定性, 都受检测阈值 T 的影响. 为了兼顾前景和背景的检测效果, 本文采用双阈值方法应对此问题, 即选择下判阈值 T_1 和上判阈值 T_2 进行背景检测 ($T_1 < T_2$).

基于背景差分技术的目标检测算法, 在对当前像素点进行分类时, 若当前像素值与背景像素值的差值小于检测阈值, 则判断为背景. 故当判断为背景点时, 检测阈值越小则此判断结果越可靠, 背景检测越精确; 检测阈值越大则此判断结果越不可靠, 背景检测越不精确, 因此前景检测越精确.

由上述分析可知, 下判阈值 T_1 较小, 进行背景判别时检测较为严格, 故背景检测效果较好. 若像素点被判断为背景, 则其可信度较高 (P_{BT1} 较大), 导致检测结果中背景区域较小、噪声点较多. 而上判阈值 T_2 较大, 则进行背景判别时检测较为宽松, 故背景检测效果较差, 从而前景检测效果较好. 若像素点被判断为前景, 则其可信度较低 (P_{BT2} 较小), 反之则被判断为前景的可信度较高 ($P_{FT2} = 1 - P_{BT2}$ 较大), 导致检测结果中前景区域较小、噪声点较少. 因此, 最终像素在当前决策准则下的背景隶属度、前景隶属度可由下式计算得到:

$$P_B = \max(P_{BT1}^1, P_{BT1}^2, \dots, P_{BT1}^n), \quad (9)$$

$$P_F = \max(P_{FT2}^1, P_{FT2}^2, \dots, P_{FT2}^n). \quad (10)$$

其中: P_{BT1}^i 为由下判阈值 T_1 所得的属于不同模型的背景概率, $P_{FT2}^i = 1 - P_{BT2}^i$ 为由上判阈值 T_2 所得的属于不同模型的前景概率.

2.2 检测算法的隶属度输出建模

使用双阈值方法时, 应首先分别计算不同检测算法在不同阈值条件下的背景隶属度.

2.2.1 基于分段线性变换的可能性计算

上述 3 种目标检测算法都是基于背景建模的思想, 即在建立好背景模型后, 根据当前像素是否在其背景模型的一定波动范围内来判断像素点是否与背景模型匹配.

对于差分图像像素值, 若等于检测阈值, 则认为

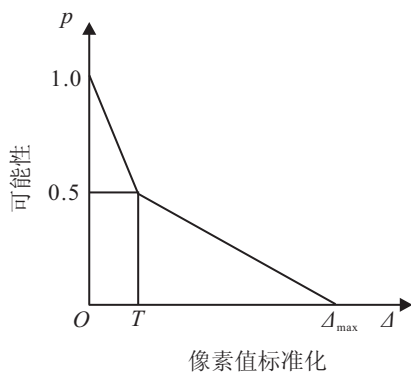


图 3 像素值属于背景的可能性计算

其属于背景和前景的可能性一样, 即 $p_F = p_B = 0.5$; 若小于检测阈值, 则认为属于背景可能性较大, 即 $0.5 < p_B < 1$; 若大于检测阈值, 则认为属于背景可能性较小, 即 $0 < p_B < 0.5$. 因此, 为了计算当前像素值属于背景的可能值大小, 采用分段线性变换方法^[13], 其中分段点为检测阈值 T . 根据当前像素值、背景模型和检测阈值, 采用图 3 的分段线性变换映射计算当前像素点属于背景及前景的可能性 p .

2.2.2 不同检测算法的隶属度计算

计算各检测算法下像素点属于前景点或背景点的可能值, 是后续融合检测算法最关键的一步.

1) 准则 S_1 : 混合高斯模型方法.

本文基于混合高斯模型的目标检测方法是基于灰度图像进行计算的, 该方法通过式 (2) 判断当前像素点是否与背景模型中 n 个高斯模型匹配来判断当前像素点属于背景点还是前景点.

通过下式计算当前像素点在背景模型中第 k 个高斯模型下属于前景点和背景点的可能值:

$$\begin{cases} \Delta_{\max}^1 = \left| \frac{255 - \mu_k}{\delta_k} \right|, \\ \Delta_{\max}^2 = \left| \frac{0 - \mu_k}{\delta_k} \right|; \end{cases} \quad (11)$$

$$\Delta_{\max k} = \begin{cases} \Delta_{\max}^1, & \Delta_{\max}^1 \geq \Delta_{\max}^2; \\ \Delta_{\max}^2, & \Delta_{\max}^1 < \Delta_{\max}^2; \end{cases} \quad (12)$$

$$\Delta_k = \left| \frac{I - \mu_k}{\delta_k} \right|. \quad (13)$$

通过分段线性化函数计算属于背景的可能性, 这里选取不同的阈值 T_1 、 T_2 来计算 ($T_1 < T_2$), 有

$$P_{BT_i}^k = \begin{cases} 1 - \frac{1}{2} \cdot \frac{\Delta_k}{T_i}, & \Delta_k \leq T_i; \\ \frac{1}{2} \cdot \frac{\Delta_{\max k} - \Delta_k}{\Delta_{\max k} - T_i}, & \Delta_k > T_i; \end{cases} \quad (14)$$

其中 $i = 1, 2$.

计算准则 S_1 下前景 c_{11} 、背景 c_{21} 决策的隶属度

$$\begin{cases} c_{21} = P_{BT1} = \max(P_{BT1}^1, \dots, P_{BT1}^n), \\ c_{11} = P_{FT2} = 1 - P_{BT2} = \\ 1 - \max(P_{BT2}^1, \dots, P_{BT2}^n). \end{cases} \quad (15)$$

2) 准则 S_2 : 均值滤波方法.

本文中基于均值滤波目标的检测算法也是基于灰度图像的, 由式 (5) 可知, 前景点和背景点的判断是通过差分图像的阈值分割实现的.

设当前差分图像像素点 (x, y) 的值为 $\Delta = D_k(x, y)$, 范围为 $[0, 255]$. 根据大津法计算每一帧的最优阈值 T , 然后在此基础上选择合适的下判阈值 T_1 和上判阈值 T_2 ($T_1 < T_2$).

通过分段线性化函数计算当前像素值属于前景点和背景点的可能值, 即

$$P_{FT_i} = \begin{cases} 1 - \frac{1}{2} \cdot \frac{\Delta}{T_i}, & \Delta \leq T_i; \\ \frac{1}{2} \cdot \frac{\Delta_{\max} - \Delta}{\Delta_{\max} - T_i}, & \Delta > T_i. \end{cases} \quad (16)$$

其中: $i = 1, 2, \Delta_{\max} = 255$.

计算准则 S_2 下前景 c_{12} 、背景 c_{22} 决策的隶属度

$$\begin{cases} c_{12} = P_B = 1 - P_{FT_2}, \\ c_{22} = P_F = P_{FT_1}. \end{cases} \quad (17)$$

3) 准则 S_3 : 改进码本方法.

本文方法通过式 (7) 判断像素点新输入值是否与背景码本中的 n 个码字匹配来判定该像素点属于前景还是背景. 通过如下方式计算得到属于背景点和前景点的可能值.

在 YUV 空间下, 为了计算 3 个通道当前值与背景码本中第 $k(k = 1, 2, \dots, n)$ 个码字匹配区间距离之和 d_{Bk} 、非匹配区间距离之和 d_{Fk} , 引入区间数及区间数距离的概念^[14], 以区间的形式描述取值的不确定性.

若 $\tilde{b} = [b_1, b_2] = \{x | b_1 \leq x \leq b_2\}, b_1, b_2 \in R$, 则称 \tilde{b} 为一个区间数. 特别地, 若 $b_1 = b_2$, 则 \tilde{b} 退化为一个实数.

设 $A = [a_1, a_2]$ 和 $B = [b_1, b_2]$ 是两个区间数, 则区间数 A 和 B 之间距离的二次方为

$$d^2 = \left[\frac{a_1 + a_2}{2} - \frac{b_1 + b_2}{2} \right]^2 + \frac{[(a_2 - a_1) + (b_2 - b_1)]^2}{12}. \quad (18)$$

在计算实数点到不同区间数的距离时, 为消除区间长度差异的影响, 本文采用相对距离, 反应了实数点隶属于区间的程度, 有

$$d^2 = \frac{(x - (b_1 + b_2)/2)^2 + (b_2 - b_1)^2/12}{(b_2 - b_1)^2}. \quad (19)$$

对于 Y 通道, 匹配区间为 $[Y_{\text{Low}}, Y_{\text{High}}]$, 非匹配区间为 $[0, Y_{\text{Low}}]$ 和 $[Y_{\text{High}}, 255]$, 用相对距离公式依次计算得到 $d_{Bk}^Y, d_{Fk1}^Y, d_{Fk2}^Y$. 当前像素在 Y 通道与匹配区间的相对距离为 d_{Bk}^Y , 与非匹配区间的相对距离为

$$d_{Fk}^Y = \begin{cases} d_{Fk1}^Y, & d_{Fk1}^Y \leq d_{Fk2}^Y; \\ d_{Fk2}^Y, & d_{Fk1}^Y > d_{Fk2}^Y. \end{cases} \quad (20)$$

同理可得, U 通道 d_{Bk}^U, d_{Fk}^U 和 V 通道 d_{Bk}^V, d_{Fk}^V . 当前像素值在第 k 码字下属于背景点和前景点的可能值为

$$\begin{cases} d_{Bk} = d_{Bk}^Y + d_{Bk}^U + d_{Bk}^V, \\ d_{Fk} = d_{Fk}^Y + d_{Fk}^U + d_{Fk}^V. \end{cases} \quad (21)$$

找到 $\{d_{Bk} | k = 1, 2, \dots, n\}$ 中最小值对应的索引

$$N = \arg \min_k \{d_{Bk} | k = 1, 2, \dots, n\}, \quad (22)$$

得到当前像素点值与码本匹配区间的距离和非匹配区间的距离分别为

$$\begin{cases} d_B = d_{BN}, \\ d_F = d_{FN}. \end{cases} \quad (23)$$

对距离求倒数并归一化, 计算当前准则 S_3 下前景 c_{13} 、背景 c_{23} 决策的隶属度

$$\begin{cases} c_{13} = \frac{d_B}{d_F + d_B}, \\ c_{23} = \frac{d_F}{d_F + d_B}. \end{cases} \quad (24)$$

完成了对上述 3 种单一检测方法中的不确定性建模之后, 将利用信息融合技术来实现单一方法的优势互补及改善检测效果.

2.3 基于 COWA-ER 的融合检测算法

2.3.1 COWA-ER

COWA-ER^[16-17]是一种多准则决策框架下的不确定推理方法^[18]. 该方法以证据理论^[19]为基础, 在具有不确定性的多准则决策框架下, 首先通过乐观和悲观两种态度下的有序加权平均^[15](OWA)方法, 构造不同决策备选方案的期望代价区间; 然后通过简单自然的转换方式分别得到不同决策的信度函数; 最后基于 Dempster 组合规则和判决规则作出决策.

COWA-ER 方法综合利用最乐观和最悲观两种态度的结果来进行后面的融合, 最终结论是在综合意义下, 而非片面强调某一方面(悲观或乐观), 这也是 Cautious(谨慎)一词的由来.

COWA-ER 算法步骤如下.

定义多准则决策框架 C 为

$$A_i \begin{bmatrix} S_1 & \dots & S_j & \dots & S_n \\ c_{i1} & \dots & c_{ij} & \dots & c_{in} \\ \vdots & & \vdots & & \vdots \\ c_{q1} & \dots & c_{qj} & \dots & c_{qn} \end{bmatrix} = C. \quad (25)$$

其中: $A_i(i = 1, 2, \dots, q)$ 为可能出现的决策结果, $S_j(j = 1, 2, \dots, n)$ 为不同的准则, c_{ij} 为准则 S_j 下判断为决策结果 A_i 的可能值.

1) 从 C 中每种决策结果对应的一行中选取最大值和最小值, 得到

$$E[C] = \begin{bmatrix} c_1^{\min} & c_1^{\max} \\ c_2^{\min} & c_2^{\max} \\ \vdots & \vdots \\ c_q^{\min} & c_q^{\max} \end{bmatrix}. \quad (26)$$

对矩阵 $E[C]$ 的所有元素进行归一化计算, 得到

$$E^{\text{Imp}}[C] = \begin{bmatrix} N_1^{\min} & N_1^{\max} \\ N_2^{\min} & N_2^{\max} \\ \vdots & \vdots \\ N_q^{\min} & N_q^{\max} \end{bmatrix}, \quad (27)$$

其中左列和右列分别表示最悲观和最乐观的期望代价向量.

2) 生成基本信度分配. 假设辨识框架 $\Theta = \{A_1, A_2, \dots, A_q\}$, 根据 $E^{Imp}[C]$ 及下式计算各决策 A_i 的基本信度分配:

$$\begin{cases} m_i(A_i) = a, \\ m_i(\bar{A}_i) = 1 - b, \\ m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a, \end{cases} \quad (28)$$

其中 \bar{A}_i 是在 Θ 中 A_i 的补集.

3) 利用 Dempster 组合规则融合上述各 $mass_{A_i}(\cdot)$ 函数, 得到新的融合 $mass_{Dempster}(\cdot)$ 函数.

4) 将融合 $mass$ 函数转换为 Pignistic 概率, 选择概率最大的决策备选作为最终的决策.

2.3.2 融合检测算法

这里结合 COWA-ER 实现融合检测, 其中乐观和悲观两种态度分别对应着同一决策结果中不同检测算法计算出的最大和最小隶属度.

1) 多准则决策框架. 如果将上文中 3 种不同的检测方法{混合高斯, 均值滤波, 改进码本}当作 3 个决策准则 $\{S_1, S_2, S_3\}$, 将{前景, 背景}当作两种决策结果 $\{A_1, A_2\}$, 则可以得到多准则决策框架 C 如下:

$$\begin{matrix} & S_1 & S_2 & S_3 \\ \begin{matrix} A_1 \\ A_2 \end{matrix} & \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \end{matrix} = C, \quad (29)$$

其中 c_{ij} 为准则 S_j 下判断为决策结果 A_i 的可能值 ($i = 1, 2, j = 1, 2, 3$).

2) 不同决策的隶属度计算. 基于双阈值检测算法, 根据分段线性映射变换计算各检测算法下像素点属于前景点及背景点的隶属度, 即利用式 (15)、(17)、(24) 计算得到多准则框架下不同决策的隶属度.

3) 决策级融合. 基于求得的多准则决策框架, 利用 COWA-ER 方法完成当前帧图像每个像素点属于前景点还是背景点的决策, 实现融合目标检测.

基于上述步骤实现了像素点的融合决策. 由于融合方法需要同时计算 3 种单一检测算法的前景、背景隶属度, 且需要进行融合计算, 故增加了每个像素点决策的运算复杂度, 但这并不影响本文方法的实用性, 原因如下所述.

首先, 融合检测方法中涵盖的 3 种单一检测算法可以并行实现, 从而在一定程度上降低了运算复杂度; 其次, 针对融合部分, 在算法中增加了一个降低融合频度的机制, 在有必要融合检测时, 再启动融合决策算法. 具体实现描述如下.

在多准则决策框架 C 下, 定义决策 $A_i (i = 1, 2)$ 的置信度为: 决策 A_i 的悲观隶属度 (3 种准则下决策

A_i 的最小隶属度) 与决策 $A_j (j \neq i)$ 的乐观隶属度 (3 种准则下决策 A_j 的最大隶属度) 之差. 若决策 A_i 的置信度大于零, 则最终的融合结果为 A_i . 因此, 若 3 种检测方法对某一像素点的检测结果均明显 (明显程度用上述置信度定义来描述) 趋向于同一决策, 则无需进行后续的融合计算, 即可由此结论直接得出最终的决策结果.

而对于每一帧图像大部分“无争议”像素点, 即 3 种检测方法的检测结果均以较大置信度趋向于同一决策 (满足上述结论) 时, 只需进行第 1 步的隶属度计算后, 即可基于逻辑判断直接得出最终决策结果. 只有在复杂场景下, 即传统单一算法无法有效应对的场合, 如运动目标区域边缘、内部空洞、变化的背景、噪声等较小部分“有争议”的像素点处, 3 种检测方法的检测结果发生不一致, 才启动后续的融合决策计算, 使得最终的检测结果更加理想. 故对于每一帧图像可通过上述计算优化节省大量时间.

由上述分析可知, 基于 COWA-ER 的融合检测算法较单一检测算法的计算复杂度有所提升, 但是通过并行计算以及降低融合频度机制的设计, 仍然能够有效降低运算复杂度, 满足实时要求.

2.3.3 基于 COWA-ER 的融合算例

对于图像中的某个像素点而言, 假设通过计算得到的多准则决策框架为

$$\begin{matrix} & S_1 & S_2 & S_3 \\ \begin{matrix} A_1 \\ A_2 \end{matrix} & \begin{bmatrix} 0.5 & 0.3 & 0.4 \\ 0.2 & 0.4 & 0.6 \end{bmatrix} \end{matrix} = C.$$

利用 COWA-ER 进行决策级融合判断.

Step 1: 计算得到

$$E[C] = \begin{bmatrix} 0.3 & 0.5 \\ 0.2 & 0.6 \end{bmatrix},$$

$$E^{Imp}[C] = \begin{bmatrix} 0.500 & 0.833 \\ 0.333 & 1.000 \end{bmatrix}.$$

Step 2: 利用上文所述方法生成基本信度分配, 见表 1.

表 1 各决策的基本信度分配 (BBA)

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.5	0.167	0.333
A_2	0.333	0	0.667

Step 3: 利用 Dempster 规则得到 $mass_{Dempster}(\cdot)$, 见表 2.

表 2 BBA 的 Dempster 规则融合结果

Alternatives	$mass_{Dempster}(\cdot)$
A_1	0.400
A_2	0.333
$A_1 \cup A_2$	0.267

Step4: 生成 Pignistic 概率, 见表 3. A_1 (前景) 为最终的决策结果.

表 3 生成 Pignistic 概率

Alternatives	BetP(·)
A_1	0.533
A_2	0.467

3 实验结果与分析

目前对于运动目标检测效果好坏的评价, 尚未形成统一的评价标准, 主要的评价方式有基于客观真值的评价以及主观目视检测效果的评价. 除检测效果外, 时间消耗以及计算资源消耗程度也是较为常用的评价标准.

基于客观真值的检测效果评价, 顾名思义, 需要用到理想参考结果图(前景真值图、背景真值图), 在此基础上可以对各种方法得出的检测结果进行定量评价. 然而, 对视频序列的每一帧建立理想参考结果图非常困难, 在已有的实验条件下对检测结果与参考结果进行匹配对比难度较大.

基于主观的目视效果, 是通过人为主观分析对比, 实际上是利用人脑的智能形成主观的真值, 这也

是目前实际应用中采用较多的评估方法.

本文对算法的性能评估, 主要是对存在阴影、空洞、尾影、噪声时的多种典型场景, 通过主观目视效果以及运算复杂度(内存、CPU 消耗情况)来评价本文方法较传统方法的优势. 为了更好地、直接地比较各检测方法, 本文所有算法都是在原始彩色视频上进行实验, 并且每一帧的检测结果都没有进行任何的形态学滤波和连通性分析处理.

3.1 室外校园视频

本文的实验选取 UCSD 大学的标准视频, 是室外校园视频, 分别做了基于混合高斯、均值滤波、改进码本和基于 3 种方法融合的目标检测实验, 并进行比较. 校园视频中包含了车辆(大目标)和行人(较小目标), 运动目标都存在阴影, 并且背景中有物体在一段时间内移动和停止.

从室外校园视频这组实验效果可以看出, 融合检测算法的效果总体上都要优于 3 个单一的检测算法. 图 4 为第 68 帧视频检测结果, 可以看出融合检测算法背景最纯净, 运动的汽车轮廓检测比较完整. 图 5 为第 162 帧检测结果, 由于背景中警戒栏杆突然出现运

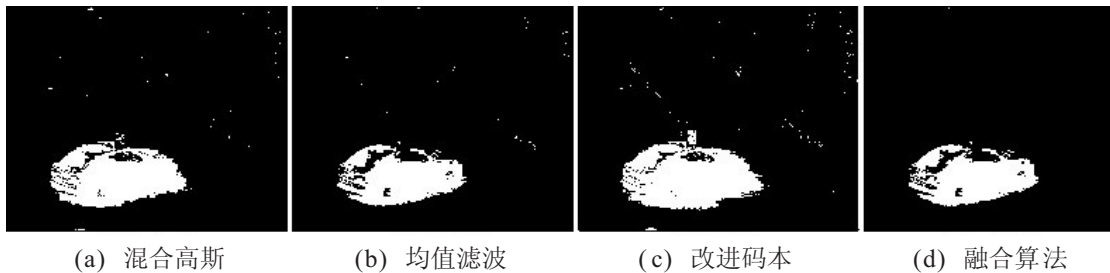


图 4 室外校园视频第 68 帧检测效果

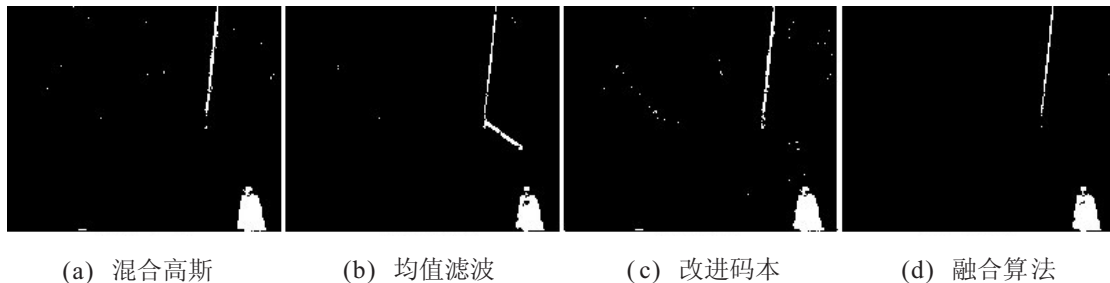


图 5 室外校园视频第 162 帧检测效果

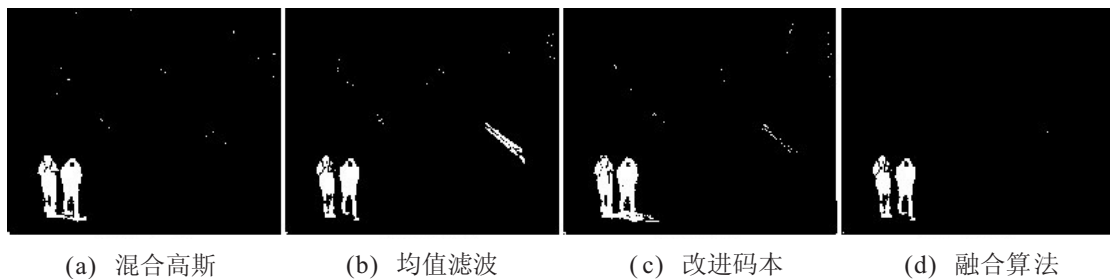


图 6 室外校园视频第 413 帧检测效果

动的情况,使得栏杆位置变化后,栏杆原位置区域像素值发生突变而出现误检测.如均值滤波法在该视频实验中背景更新系数取值非常小,相当于几乎不存在背景更新,所以在栏杆初始位置一直存在误检测,而混合高斯和改进码本两种方法背景有更新,一段时间后误检测区域会消失.融合检测算法的背景更新速度与背景更新速度最快的混合高斯相当,并且背景中误检测的杂点最少,行人检测效果也较好.图6为第413帧检测结果,混合高斯检测结果阴影区域明显,背景杂点也较多,改进码本检测结果也出现一定程度的阴影区域,而融合检测结果中行人轮廓完整,内部充实,几乎没有阴影区域,背景杂点也较少.

3.2 室内光照变化视频

本文还做了另一组实验,实验采用自行拍摄的具有室内光照变化效果的视频,以增加视频目标检测的难度和复杂度,用以进一步验证基于COWA-ER融合检测算法较单一检测算法的优越性.

在本视频中,环境变化主要是光线的突然变化.在图7所示的第215帧中,光线突然开始变暗,3种单一检测方法出现了不同程度背景区域突变为前景,改进码本方法误检范围最大,而融合检测方法检测效果稍好.在图8所示的第230帧中,整个场景中光线变暗一段时间,每种检测方法都进行了背景更新,融合检测算法效果最好,背景较为纯净.

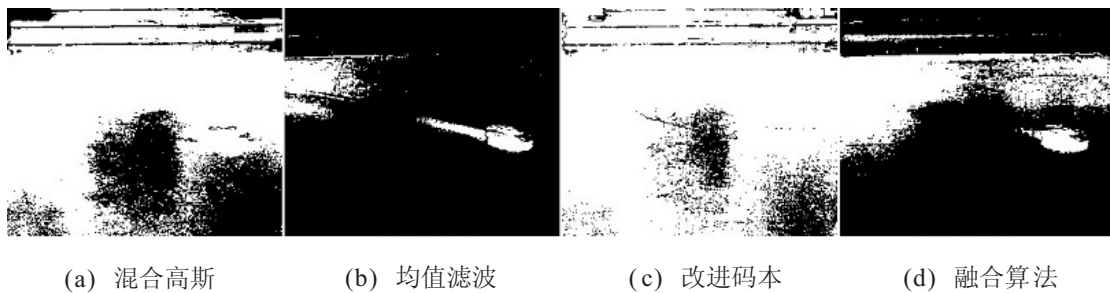


图7 室内视频第215帧检测效果

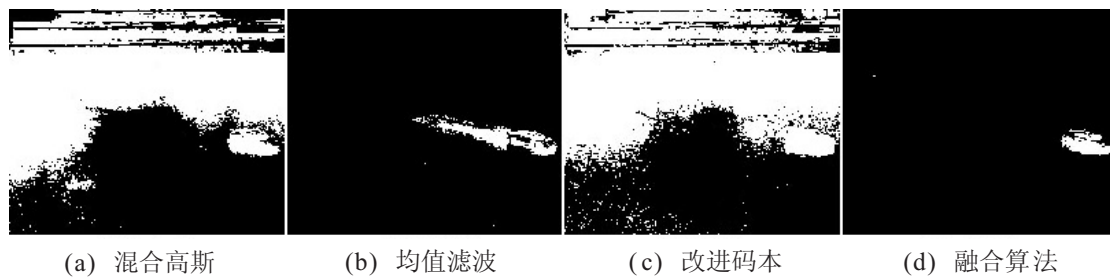


图8 室内视频第230帧检测效果

上述两段视频序列包含了传统目标检测方法无法有效应对的几个代表性的复杂场景,即光照变化、阴影、噪声干扰、对象由动到静或由静到动、目标较小或运动较慢5种场景.本文融合检测算法相对于传统检测算法,均取得了较理想的检测结果.限于篇幅,罗列的只是一些典型帧、典型场景下(特别是传统方法无法有效应对的情形)的检测结果,用以验证本文提出的融合检测算法相对于传统检测算法在鲁棒性及有效性方面的优势.

4 结 论

本文基于信息融合的思想,提出了一种基于COWA-ER的多方法融合的运动目标检测算法.由于混合高斯、均值滤波和码本3种检测方法均具有一定的不确定性,本文构建了具有2种可能决策和3种准则的多准则决策框架,通过双阈值法描述不确定性,利用COWA-ER实现3种检测方法的融合检测.实验

表明,本文的融合检测算法是合理有效的,并且在目标完整性、背景杂点及背景突变适应性等方面,本文的融合检测算法的检测效果均优于上述3种单一的检测方法.

尽管本文的融合检测算法取得了较好的检测效果,但还有进一步提高的空间.考虑到信息源之间的互补性和差异性是实现有效融合的重要前提和基础,在今后的研究工作中,将重点研究各单一检测方法之间的互补性和差异性,以获得更加理想的融合检测效果.

参考文献(References)

- [1] 万纛,韩毅,卢汉清.运动目标检测算法的探讨[J].计算机仿真,2006,23(10):221-226.
(Wan Y, Han Y, Lu H Q. The methods for moving object detection[J]. Computer Simulation, 2006, 23(10): 221-226.)

- [2] Piccardi M. Background subtraction techniques: A review[C]. 2004 IEEE Int Conf on Systems, Man, and Cybernetics. Hague: IEEE, 2004: 3099-3104.
- [3] Dai K X, Li G H, Tu D, et al. Prospects and current studies on background subtraction for moving objects detection from surveillance video[J]. J of Image and Graphics, 2006, 11(7): 919-927.
- [4] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]. 2007 IEEE Int Conf on Computer Vision and Pattern Recognition(CVPR). Minneapolis, 2007: 2246.
- [5] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction[C]. Proc of the 17th Int Conf on Pattern Recognition ICPR 2004. Cambridge: IEEE, 2004: 28-31.
- [6] Chen Z J, Chen X J, He H. Moving object detection based on improved mixture Gaussian models[J]. J of Image and Graphics, 2007, 12(9): 1585-1589.
- [7] 钟珞, 刘剑. 基于混合高斯和均值滤波法的运动检测方法[J]. 武汉理工大学学报: 信息与管理工程版, 2010, 32(5): 691-693.
(Zhong L, Liu J. Moving object detection based on mixture Gaussian model and median filter method[J]. J of Wuhan University of Technology: Information & Management Engineering, 2010, 32(5): 691-693.)
- [8] Zhang X, Xiong Y. Impulse noise removal using directional difference based noise detector and adaptive weighted mean filter[J]. IEEE Signal Processing Letters, 2009, 16(4): 295-298.
- [9] 李了了, 邓善熙, 丁兴号. 基于大津法的图像分块二值化算法[J]. 微计算机信息, 2005, 21(8): 76-77.
(Li L L, Deng S X, Ding X H. Binarization algorithm based on image partition derived from Da-Jing method[J]. Microcomputer Information, 2005, 21(8): 76-77.)
- [10] Kim K, Chalidabhongse T H, Harwood D, et al. Background modeling and subtraction by codebook construction[C]. Int Conf on Image Processing. Singapore, 2004: 3061-3064.
- [11] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground-background segmentation using codebook model[J]. Real-Time Imaging, 2005, 11(3): 172-185.
- [12] Qi M B, Yang A L, Jiang J G, et al. A vehicle detection and tracking algorithm based on improved codebook[J]. J of Image and Graphics, 2011, 16(3): 406-412.
- [13] 邸慧, 于起峰, 张小虎. 一种基于灰度变换的红外图像增强算法[J]. 应用光学, 2006, 27(1): 12-14.
(Di H, Yu Q F, Zhang X H. An algorithm for infrared image enhancement based on gray scale transform[J]. J of Applied Optics, 2006, 27(1): 12-14.)
- [14] 康兵义, 李娅, 邓勇, 等. 基于区间数的基本概率指派生成方法及应用[J]. 电子学报, 2012, 40(6): 1092-1096.
(Kang B Y, Li Y, Deng Y, et al. Determination of basic probability assignment based on interval numbers and its application[J]. Acta Electronica Sinica, 2012, 40(6): 1092-1096.)
- [15] 韩崇昭, 朱红艳, 段战胜, 等. 多源信息融合[M]. 第2版. 北京: 清华大学出版社, 2010: 86-87.
(Han C Z, Zhu H Y, Duan Z S, et al. Multi-source information fusion[M]. 2nd ed. Beijing: Tsinghua University Press, 2010: 86-87.)
- [16] 韩德强, 杨艺, 韩崇昭. DS证据理论研究进展及相关问题探讨[J]. 控制与决策, 2014, 29(1): 1-11.
(Han D Q, Yang Y, Han C Z. Advances in DS evidence theory and related discussions[J]. Control and Decision, 2014, 29(1): 1-11.)
- [17] Yager R R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking[J]. IEEE Trans on Systems, Man, and Cybernetics, 1988, 18(1): 183-190.
- [18] Tacnet J M, Dezert J. Cautious OWA and evidential reasoning for decision making under uncertainty[C]. Proc of the 14th Int conf on Information Fusion. Chicago, 2011: 1-8.
- [19] Han D, Dezert J, Tacnet J M, et al. A fuzzy-cautious OWA approach with evidential reasoning[C]. Proc of the 15th Int Conf on Information Fusion. Singapore, 2012: 278-285.

(责任编辑: 孙艺红)