采用边界评价的红外视频运动目标时空域分割方法

闵超波,张俊举,常本康,孙 斌,李英杰

(南京理工大学 电子工程与光电技术学院,江苏 南京 210094)

摘 要:为了在红外视频中准确分割运动目标,提出了一种基于边界评价的红外运动目标时空域分割的新方法。首先,利用运动目标在时域差分图像中的“空洞”效应,提取出最有意义运动目标种子点。重点是运动目标的空间分割,利用种子区域整体与局部的关系,在提取出的种子上进行区域生长,可以得到不同生长阈值下的运动目标分割掩膜。为确定最佳生长阈值,提出了一种无需先验知识的红外目标分割掩膜边界评价准则,并采用“分割-评价-再分割-再评价”的循环迭代模式,利用“由粗到精”的搜索方法,找出最佳的生长阈值,同时得到最佳的运动目标分割掩膜。实验证明,所提出的方法能在红外视频中准确分割出运动目标区域,效果良好,性能鲁棒。

关键词:时空域分割; 图像分割; 性能评价; 红外视频

中图分类号:TP391 **文献标志码**:A **文章编号**:1007-2276(2013)10-2636-06

Spatiotemporal segmentation method of moving-object using boundary evaluation in infrared video

Min Chaobo, Zhang Junju, Chang Benkang, Sun Bin, Li Yingjie

(School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: In this paper, a new method was presented for spatiotemporal segmentation of moving-object using boundary evaluation in infrared video. At first, the ideal seeds of every moving object were extracted based on the "holes" effect of temporal difference, respectively. The wok focus was spatial segmentation. On the basis of the relationship between the global and local standard deviation of seeds, the segmented masks could be grown form the ideal seeds by using different growing thresholds. For determination of the best growing threshold, a criterion was constructed for evaluating the boundary of the segmented mask of infrared moving-object without prior knowledge. According to the proposed criterion, an iterative model which was "segmentation-evaluation-segmentation-evaluation" and the search method called as "coarse to fine" were applied to find the best growing threshold. Meanwhile the best segmented mask was obtained too. The experiment results show that the proposed method is superior and effective on segmentation of moving object in infrared video.

Key words: spatiotemporal segmentation; image segmentation; performance evaluation; infrared video

收稿日期:2013-02-05; **修订日期**:2013-03-03

基金项目:国家自然科学基金(61101195);江苏省自然科学基金(SBK201343283)

作者简介:闵超波(1987-),男,博士生,主要从事图像处理、光电系统设计等领域方面的研究。Email:chaobomin@gmail.com

导师简介:常本康(1950-),男,教授,博士生导师,主要从事光电发射、光电系统等领域方面的研究。Email:bkchang@njust.edu.cn

0 引 言

运动目标检测分割是计算机视觉的一个重要环节[1]。由于红外热像仪对热目标敏感,能穿透烟雾识别隐藏的目标 [2], 因此针对红外图像的运动目标分割应用广泛。最常用的方法有帧差法、背景差法和光流法[3]。帧差法虽然实时性好,但检测出来的运动目标掩膜不完整;背景差法易受噪声干扰;光流法效果不错,但计算量大,且需满足运动物体表面亮度不变的假设[4]。

混合高斯(MOG)背景差法[5]是如今公认比较好的算法之一。但是对于具有复杂背景和低信噪比的红外图像,MOG 法不能获得令人满意的效果。参考文献[6]提出了一种基于马尔卡夫随机场(MRF)的时空域运动目标分割方法, 利用先前帧的 MRF 模型对当前帧进行重新估计, 然后利用估计帧进行时域分割,并利用先前帧对分割结果进行补偿和修正。但当目标运动幅度过大时,该方法无法分割出完整的运动目标掩膜,且在低信噪比情况下受噪声干扰较为严重。

为了能降低红外图像噪声对运动目标分割的干扰,提取出完整的运动目标轮廓,文中提出了一种基于边界评价的运动目标分割方法, 构建基于对比度的红外目标分割掩膜边界评价准则,通过“分割-评价-再分割-再评价”的循环迭代模式,得到最佳的分割掩膜。最后,实验结果证明了利用该方法可以在红外视频中提取精确的运动目标轮廓, 并且能很好地抑制噪声。

1 基于对比度的红外目标分割掩膜边界评价准则

对于红外目标分割掩膜来说, 其边界的准确程度,可以用来衡量分割掩膜的性能。因为边界反映了分割掩膜轮廓与实际运动目标轮廓的契合程度。将红外原图用 $\mathbf{I}_{original}$ 来表示,分割掩膜定义为 $\mathbf{I}_{mask}$,其中 1 代表目标区域,0 代表背景区域。首先,利用 3×3 的 Sobel 算子对红外目标分割掩膜进行边界检测,得到分割掩膜的边界二值化图像。然后,在边界图中,隔 $\mathbf{n}_r$ 行隔 $\mathbf{n}_c$ 列取边界点为特征点, 令 $\mathbf{L}=\{\mathbf{z}_k:\mathbf{z}_k\in\mathbf{R}^2\}_{k=1}^{\mathbf{N}}$ 表示分割掩膜边界特征点集合, 其中 $\mathbf{N}$ 表示特征点个数,如图 1(a)、(b)所示。

单独的边界特征点无法准确表征边界在某一点的性能。因此,对于每个边界特征点,以特征点 $\mathbf{z}_k$ 为圆心,$\mathbf{r}$ 为半径,取圆形区域作为特征点 $\mathbf{z}_k$ 的特征区域,记为 $\mathbf{V}_k$。同时,以分割掩膜为参照对象,将特征区域 $\mathbf{V}_k$ 分为内部特征区域 $\mathbf{V}_k^{in}$ 和外部特征区域 $\mathbf{V}_k^{out}$。然后, 分别定义内部和外部特征区域的局部图像灰度特征:

$$\mathbf{G}(\mathbf{V}_k^{in})=\frac{1}{\mathbf{N}_k^{in}}\sum_{q\in\mathbf{V}_k}\xi_k^{in}(q)\cdot\mathbf{I}_{original}(\mathbf{z}_k+q) \tag{1}$$

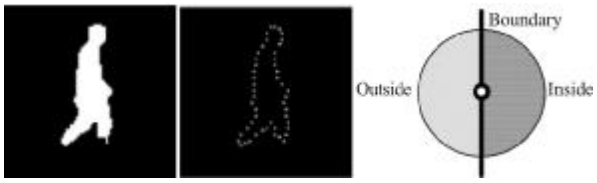
$$\mathbf{G}(\mathbf{V}_k^{out})=\frac{1}{\mathbf{N}_k^{out}}\sum_{q\in\mathbf{V}_k}\xi_k^{out}(q)\cdot\mathbf{I}_{original}(\mathbf{z}_k+q) \tag{2}$$

式中:q 表示在特征区域内 $\mathbf{V}_k$ 径向向量;$\xi_k^{in}(q)$和 $\xi_k^{out}(q)$ 分别表示内部和外部特征区域中每个点的权重函数。

$$\xi_k^{in}(q)=1[\mathbf{z}_k+q\in\mathbf{V}_k^{in}]\cdot\exp\left\{-\frac{\|q\|^2}{2\sigma^2}\right\} \tag{3}$$

$$\xi_k^{out}(q)=1[\mathbf{z}_k+q\in\mathbf{V}_k^{out}]\cdot\exp\left\{-\frac{\|q\|^2}{2\sigma^2}\right\} \tag{4}$$

式中:$1[\mathbf{z}_k+q\in\mathbf{V}_k^{x}]$表示若 $\mathbf{z}_k+q$ 在 $\mathbf{V}_k^{in}/\mathbf{V}_k^{out}$ 中,则该值取 1,否则取 0。该函数表示区域中每个点对相应区域灰度特征的贡献度。图 1(c)为特征区域的例子。



(a) 分割掩膜　(b) 边界特征点　(c) 特征区域

(a) Segmented mask　(b) Feature points of boundary　(c) Feature region

图 1 边界特征点例子

Fig.1 Example of feature points of boundary

接下来, 可以定义分割掩膜边界的平均对比度 $\mathbf{C}(\mathbf{I}_{mask})$:

$$\mathbf{C}(\mathbf{I}_{mask})=\frac{1}{\mathbf{N}}\sum_{k=1}^{\mathbf{N}}\left|\frac{\mathbf{G}(\mathbf{V}_k^{in})}{\mathbf{G}(\mathbf{V}_k^{out})}-1\right| \tag{5}$$

在红外图像中,目标的灰度一致性较好。因此,若分割掩膜与实际目标区域完全一致, 则分割掩膜边界处内外灰度特征的对比度应该最大; 若出现过分割或不完全分割情况时, 分割掩膜边界处的内外灰度特征的相似度会增加。因此,公式(5)可以变为:

$$\hat{\mathbf{I}}_{mask}=\arg\max_{\mathbf{I}_{mask}}\mathbf{C}(\mathbf{I}_{mask}) \tag{6}$$

式中:$\hat{\mathbf{I}}_{mask}$ 为具有最佳边界的分割掩膜。

文中关注的是分割掩膜自身的性能，而不是分割算法的好坏。因此,利用手工选择种子点,并以不同生长阈值进行区域生长[7]来得到不同性能的分割掩膜如图 2 和图 3 所示，并将所提出准则与错分率(ME)进行对比,以测试边界评价准则的有效性。ME 定义如下所示。

$$ME=1-\frac{|B_o\cap B_r|+|F_o\cap F_r|}{|B_o|+|F_o|} \tag{7}$$

式中:B_o 和 F_o 分别为实际背景与实际目标;B_r 和 F_r 分别为分割背景和分割目标。ME 为一种需要先验知识的分割评价准则，表示的是错误分割的像素的比例,显然 ME 越小,分割效果越好。

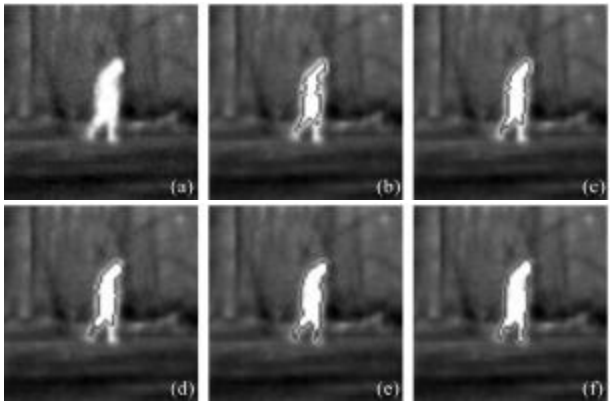


图 2 实验数据 1

Fig.2 Dataset #1

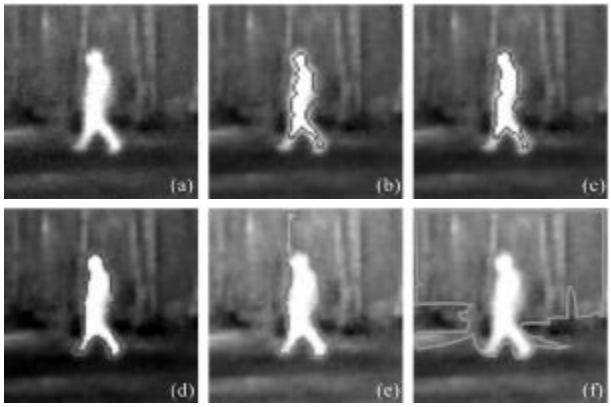


图 3 实验数据 2

Fig.3 Dataset #2

ME 衡量的是分割掩膜客观实际性能。从表 1 和表 2 可以看出，所提出的边界评价准则的评价结果与主观感知和客观实际吻合得很好，证明了边界评价准则的有效性。而且,边界评价准则的最大优越性在于计算过程中不需要以先验知识为参考，因此可以用于分割掩膜在线实时评价。

表 1 实验数据 1 的评价结果

Tab.1 Evaluation results of dataset #1

	Fig.2(b)	Fig.2(c)	Fig.2(d)	Fig.2(e)	Fig.2(f)
Mask	0.878 9	1.209 2	1.297 4	1.540 9	1.697 7
ME	1.813 7%	1.310 8%	1.112 6%	0.335 3%	0.106 7%

表 2 实验数据2 的评价结果

Tab.2 Evaluation results of dataset #2

	Fig.3(b)	Fig.3(c)	Fig.3(d)	Fig.3(e)	Fig.3(f)
Mask	0.885 1	1.121 4	1.603 4	1.371 1	0.568 9
ME	3.856 1%	2.362 4%	1.082 2%	1.646 1%	48.437 7%

2 基于边界评价的空间分割

2.1 基于“空洞”效应的种子提取

将红外视频中第 k 帧图像定义为 f_k,则时域分割的差分图 d_k 可以定义为:

$$d_k=|f_k-f_{k-1}| \tag{8}$$

利用最大类间差法(Otsu)[8]从 d_k 中检测出有效变化像素,形成变化检测掩膜(CDM),其中 1 代表变化像素点,0 代表不变像素点。由于差分图像的性质,CDM 中所检测到的运动目标中间会出现“空洞”效应。这些“空洞”是时域分割检测运动目标的缺点,但对于区域生长的分割方法来说,“空洞” 中包含的像素恰恰是与运动目标最相关的“种子”。因此,将这些像素点提取出来，就可以作为区域生长最有意义的种子点。具体方法分为以下几个步骤:

(1) 利用$[(2\cdot r_a+1)^2-1]$的邻域模型,对 CDM 进行扩大的连通区域检测，一个连通区域对应一个运动目标的变化像素区域 CDM_i,$i=1,2,\cdots,M$,其中 M 为图像中运动目标个数;

(2) 在 CDM_i 中检测每一行或每一列中是否有两个不连续的变化像素,若有,则将这两点用直线相连;

(3) 对区域内所有行列检测完毕之后，在 CDM_i 中标记为 0 且为连接线相交的点，即是 CDM_i 对应运动目标的种子点。每个运动目标的种子点掩膜定义为 S_i,1 代表种子点,0 代表非种子点;

(4) 重复步骤(2)、(3),直到对所有运动目标提取种子点完毕。图 4 为提取种子点的例子。

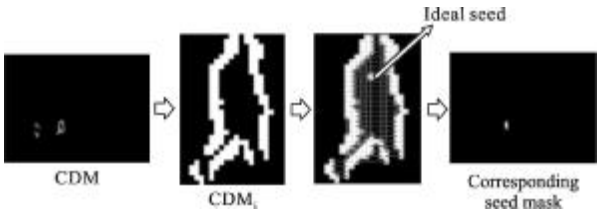


图 4 提取种子点流程

Fig.4 Processing of seed extraction

2.2 区域生长

为了避免多余的计算和误差，在每个运动目标的 $\mathbf{S}_i$ 中，以标记为 1 的像素区域的质心为中心,开 $(2\cdot l_r+1)\times(2\cdot l_c+1)$ 大小窗口,其中 l_r 和 l_c 分别表示 $\mathbf{S}_i$ 中种子点坐标行方向和列方向的最大距离，窗口种子掩膜用 $\mathbf{Sw}_i$ 表示。同时,在相同的位置,以相同的大小在 f_k 上也开一个窗口,用 f_k^i 表示。对一个运动目标的种子掩膜 $\mathbf{S}_i$ 进行生长的具体实现方法如算法 1 所示。令 $\mathbf{T}$ 表示生长阈值。

算法 1

do

　　令 $\mathbf{Q}=\mathbf{Sw}_i$;

　　for $x=2:2\cdot l_r$

　　for $y=2:2\cdot l_c$

　　if $\mathbf{Sw}_i(x,y)\neq 1$

令 A 为 $\mathbf{Sw}_i$ 中以 $\mathbf{Sw}_i(x,y)$ 为中心的 8 邻域矩阵;

令 B 为 f_k^i 中以 $f_k^i(x,y)$ 为中心的 8 邻域矩阵;

令 $\mathbf{N}_A$ 表示 A 中标记为 1 的元素的个数;

　　if $1\leqslant \mathbf{N}_A\leqslant 7$

　　for ii=1:3

　　for jj=1:3

$\mathbf{C}=A(ii,jj)\cdot B(ii,jj)$;

$\mathbf{D}(ii,jj)=|f_k^i(x,y)-\mathbf{C}|$;

　　end

end

令 $\mathbf{D}_m$ 表示 $\{\mathbf{D}(ii,jj)\}$ 中最小值;

计算 f_k^i 平均灰度值,记为 $\mathrm{mean}(f_k^i)$;

令 $\mathbf{D}_z=|f_k^i(x,y)-\mathrm{mean}(f_k^i)|$;

计算 B 的标准差,记为 std(B);

计算 f_k^i 的标准差,记为 $\mathrm{std}(f_k^i)$;

　　if $\mathbf{D}_m\leqslant \mathbf{T}\cdot \mathrm{std}(B)$ and $\mathbf{D}_z\leqslant \mathbf{T}\cdot \mathrm{std}(f_k^i)$

　　$\mathbf{Q}(x,y)=1$;

end

end

end

end

end

　　$\mathbf{Sw}_i=\mathbf{Sw}_i\cup \mathbf{Q}$;

While $\mathbf{Sw}_i$ 中标记为 1 的元素个数不断增加。

当循环结束时所得到的 $\mathbf{Sw}_i$ 则为相应运动目标在阈值 $\mathbf{T}$ 下的分割掩膜,记为 $\mathbf{M}_i$。

2.3 “由粗到精”的迭代搜索方法

生长阈值 $\mathbf{T}$ 对于区域生长来说是一个非常重要的参数,它决定着最后分割掩膜的精确度。由于所提出的分割掩膜边界评价准则无需先验知识作为参考,因此,可以用于确定阈值 $\mathbf{T}$。利用“由粗到精”的迭代搜索方法来确定最佳的阈值。对于一个运动目标的种子掩膜窗口 $\mathbf{Sw}_i$,具体步骤如下所示。令 t 表示迭代步数,t=0,1,2,…,step 代表搜索步长。

(1) 粗调迭代过程:

(a) 设定 $\mathbf{T}$ 的初始值 $\mathbf{T}(0)=0$,令 step=α;

(b) 用 $\mathbf{T}(t)$ 对种子掩膜窗口 $\mathbf{Sw}_i$ 进行区域生长(如 2.2 节所示),得到分割掩膜 $\mathbf{M}_i(t)$;

(c) 计算 $\mathbf{M}_i(t)$ 的边界评价准则 $\mathbf{C}(\mathbf{M}_i(t))$;

(d) 当 t<1 时,则 $\mathbf{T}(t+1)=\mathbf{T}(t)+\mathrm{step}$,并重复步骤(b)、(c)和(d);当 t≥1 时,计算 $\Delta\mathbf{C}=\mathbf{C}(\mathbf{M}_i(t))-\mathbf{C}(\mathbf{M}_i(t-1))$;

(e) 若 $\Delta\mathbf{C}\geqslant 0$,则 $\mathbf{T}(t+1)=\mathbf{T}(t)+\mathrm{step}$ 并重复步骤(b)、(c)和(d);若 $\Delta\mathbf{C}<0$,则停止粗调迭代过程,得到精调过程阈值的初始值 $t^f=\mathbf{T}(t-1)$,并令 t=0;

(2) 精调迭代过程:

(f) 设定 $\mathbf{T}$ 的初始值 $\mathbf{T}(0)=\mathbf{T}^f$,令 step=$\beta$;

(g) 重复步骤(b)、(c)和(d);

(h) 若 $\Delta\mathbf{C}\geqslant 0$,则 $\mathbf{T}(t+1)=\mathbf{T}(t)+1$,并继续重复步骤(b)、(c)和(d);若 $\Delta\mathbf{C}<0$,则停止精调迭代过程,得到最佳生长阈值 $\mathbf{T}^{best}=\mathbf{T}(t)-1$,同时得到运动目标最佳分割掩膜 $\mathbf{M}_i(t-1)$;

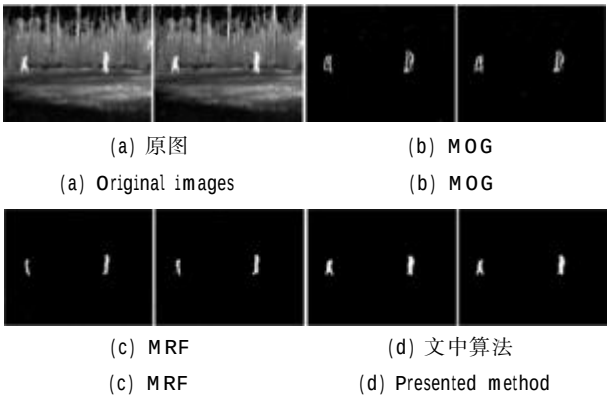
其中,α 和 β 分别表示粗调过程和精调过程的迭代步长。这种迭代方法的好处是可以控制阈值选择的走向,避免无效阈值的计算,以提高效率。

需要注意的是，得到的运动目标分割掩膜是窗口图像,因此,还必须根据种子掩膜质心的位置,将所有运动目标最佳分割掩膜贴到相应的位置，合并成一幅完整的运动目标分割掩膜。

3 实验结果与讨论

3.1 实验结果

利用三种红外视频序列对所提出的方法进行测试。为了验证所提出算法的有效性,将其与混合高斯背景差法(MOG)[3]和一种基于 MRF 的运动目标分割算法[4]进行了对比。图 5~7 为各种算法的实验结果。



(a) 原图 (b) MOG

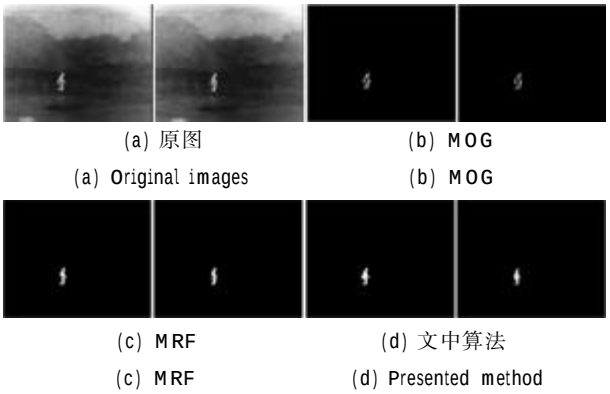
(a) Original images (b) MOG

(c) MRF (d) 文中算法

(c) MRF (d) Presented method

图 5 实验结果与对比 1

Fig.5 Comparison #1 of experiment results



(a) 原图 (b) MOG

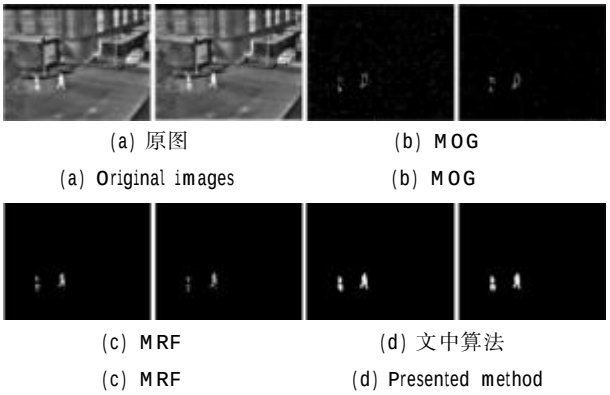
(a) Original images (b) MOG

(c) MRF (d) 文中算法

(c) MRF (d) Presented method

图 6 实验结果与对比 2

Fig. 6 Comparison #2 of experiment results



(a) 原图 (b) MOG

(a) Original images (b) MOG

(c) MRF (d) 文中算法

(c) MRF (d) Presented method

图 7 实验结果与对比 3

Fig.7 Comparison #3 of experiment results

利用 ME 来客观定量地衡量各种算法的性能。表 3 分别示出了三种算法对不同视频处理之后所得分割掩膜的平均 ME。从图 5~7 的分割结果的实际对比和表 3 中平均 ME 的比较可以看出, 无论是单目标图像还是多目标图像, 所提出的算法在分割掩膜的完整性和运动目标的准确性上都要优于其他两种算法。

表 3 各算法的平均 ME 对比

Tab.3 Comparison of average ME

	MOG	MRF	Presented method
Experiment 1	3.150 9%	0.887 2%	0.032 8%
Experiment 2	2.716 3%	0.207 6%	0.050 1%
Experiment 3	3.637 5%	0.219 1%	0.072 6%

3.2 分析与讨论

图 5(b)~7(b)为利用混合高斯背景差法处理之后的结果,检测结果比较准确,但由于其易受背景干扰,因此,分割掩膜不完整,且多有噪声干扰。图 5(c)~7(c)为 MRF 方法处理的结果。它先将图像利用MRF重新标记,再进行帧差,因此其分割结果较为准确,但分割掩膜仍旧不完整,且有少量的噪声干扰。从图5(d)~7(d)可以看出,文中方法所得分割掩膜轮廓最为完整,且几乎没有噪声干扰,这是由于采用的是边界评价和由内而外的生长方式,因此,可以得到准确的边界轮廓,并能很好地抑制噪声。

特别指出的是,在整个实验过程中,只是用了一组设定参数: 特征区域半径 $r=5$, 邻域半径 $r_a=5$,粗调步长 $\alpha=1$, 精调步长 $\beta=1$。没有做任何的参数调整。在 Core i3-4.01 GHz 处理器、2 GB 内存的计算平台上,对于 240×320 图像大小的视频,文中算法速度达到 23 帧/s;对 400×512 图像大小的视频,计算速度达 19 帧/s。

4 结 论

基于边界评价的红外运动目标时空域分割方法,利用所提出的基于对比度的分割掩膜边界评价准则,可以很好地克服背景干扰,排除噪声,准确提取出运动目标轮廓边界,有效减少虚警率。实验证明了该方法的优越性和鲁棒性。今后要对算法速度进行改进,

并使其可以用于可见光视频的运动目标检测。

参考文献:

[1] Tian Rui, Hao Zhicheng, Wu Chuan. Target area segmenting using dynamic chained list in binary images[J]. Infrared and Laser Engineering, 2011, 40(2): 344-349. (in Chinese)
田睿, 郝志成, 吴川. 利用动态链接表的二值图像目标区域分割[J]. 红外与激光工程, 2011, 40(2): 344-349.

[2] Zhang Junju, Chang Benkang, Zhang Baohui, et al. Long-distance image fusion system for infrared and LLL/visible bands [J]. Infrared and Laser Engineering, 2012, 41(1): 20-24. (in Chinese)
张俊举, 常本康, 张宝辉, 等. 远距离红外与微光/可见光融合成像系统[J]. 红外与激光工程, 2012, 41(1): 20-24.

[3] Lu Zhimao, Liu Minghua, Liu Chen. Method for moving object detection based on construction of feature frames[J]. Infrared and Laser Engineering, 2012, 41 (7): 1959-1963. (in Chinese)
卢志茂, 刘明华, 刘晨. 基于特征帧构建的运动目标检测方法[J]. 红外与激光工程, 2012, 41(7): 1959-1963.

[4] Barron J J L, Fleet D J, Beauchemin S S. Performance of optical flow techniques[J]. Int J Comput Vis, 1994, 12: 43-77.

[5] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C]//Proc Internat Conf on Computer Vision and Pattern Recognition, 1999: 2246-2252.

[6] Badri Narayan Subudhi, Pradipta Kumar Nanda, Ashish Ghosh. A change information based fast algorithm for video object detection and tracking [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(7): 993-1004.

[7] Rafael C. Gonzalez, Richard E Woods. Digital Image Processing [M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2002: 66-71. (in Chinese)

[8] Otsu N. A threshold selection method from gray-level histograms [J]. IEEE Trans Syst Man Cybern, 1979, 9(1): 62-66.