

文章编号: 1001- 2486(2008) 03- 0109- 07

# 一种融入运动特性的显著性特征提取方法<sup>\*</sup>

张 焱, 张志龙, 沈振康

(国防科技大学 ATR 实验室, 湖南 长沙 410073)

**摘要:** 针对复杂背景下图像序列中运动目标的跟踪前检测问题, 提出一种融入运动特性的显著性特征提取方法, 该方法综合目标的灰度、细节和运动等形成稳健的显著性特征, 并得到一组显著性特征提取图。在这组图像中, 具有这种多特征的区域得到加强, 其他区域受到抑制, 从而能够轻松地检测出可引起人类视觉注意的运动目标。由于该算法中的显著性特征来源于目标的多种底层特性, 因此该算法具有很强的稳健性。实验证明, 这种方法具有较强的稳定性和实用性, 抗干扰能力强。从视觉效果的角度出发, 能够较大地提高运动目标在复杂背景中的信杂比。

**关键词:** 显著性; 特征提取; 目标检测; 复杂背景

中图分类号: TP391 文献标识码: A

## A Salient Feature Extraction Algorithm Fusing the Motion Characteristic of Objects

ZHANG Yan, ZHANG Zhi-long, SHEN Zhen-kang

(ATR Lab, National Univ. of Defense Technology, Changsha 410073, China)

**Abstract:** Focusing on the problem of moving object detecting before tracking in image sequences with complex background, a new salient feature extraction algorithm that fuses the motion characteristics of objects is proposed. The algorithm fuses the gray-level and the detail and motion of the objects to form a robust salient image. The regions with salient multi-features are enhanced and the others are restrained, so moving objects will easily be detected in the salient image. The salience of the objects in the algorithm is derived from the local difference of multi-bottom features of the image, so the algorithm is robust. The experimental results show that the new algorithm is practical and stable, and has the strong anti-jamming capacity. The visual effect of the experiments demonstrates that this method can improve the signal-clutter ratio.

**Key words:** salience; feature extraction; detecting objects; complex background

科学家们都梦想有朝一日人类能够制造出智能机器, 这种智能机器能对可视的世界进行观察和理解, 能够像人脑一样轻而易举地处理许多视觉问题。视觉注意模型正是基于这种思想, 模拟人类视觉系统, 采用串行的处理策略, 通过选择性注意机制, 按照底层图像特征的显著性由大到小的顺序, 从场景图像提供的大量视觉数据中选取若干注意区域, 并对注意区域进行更加精细的分析和处理。基于显著性的视觉注意模型最初是由 Koch 和 Ullman 提出的, 后来 Itti 和 Koch 又提出了相应的计算模型<sup>[1-3]</sup>。在他们所建立的计算模型中, 非常真切地展示了基于位置的视觉注意机制的多尺度和多特征融合等特点, 能够以多种不同尺度提取多种图像特征的局部对比度, 并将其融合为一幅反映场景中各个位置视觉显著性的显著图, 然后通过返回抑制和注意焦点转移获得一系列显著位置<sup>[4]</sup>。因此根据这一思想, 将场景图像中能够引起视觉注意的灰度、细节、边缘、方向、运动等特征提取出来, 经过融合处理形成目标的显著性特征, 并将其作为运动目标的检测特征将非常有力。由于灰度、细节、边缘、方向、运动等特征属于底层的图像特征, 因此其融合结果—目标的显著性特征较之单个特征具有很强的稳健性及抗干扰能力。

本文借鉴视觉注意模型中有关特征提取、显著图计算的思想, 根据灰度图像序列特点, 利用图像的

\* 收稿日期: 2007- 10- 08

基金项目: 国家部委预研基金资助项目(9140A21040306KG0195, 9140C8002010704)

作者简介: 张焱(1975—), 女, 讲师, 博士生。

静态特征—灰度、细节等,融入目标的运动特性,建立起一种运动目标的显著性特征提取方法。

## 1 融入运动特性的显著性特征提取方法基本原理及描述

视觉显著性是由多种视觉敏感特征引起的一种局部反差,反差越明显,显著性越强;反之亦然。对于没有颜色特征的红外或部分可见光图像,视觉敏感的特征包括灰度、细节、运动等。视觉显著性的产生需要两个条件:一是在某些特征维,中心区域的特征与周围区域的特征相比存在很大不同;二是在这些特征维,周围区域的特征具有较好的一致性。如果将中心区域和周围区域合称为感受野,那么人类视觉系统就具备以不同大小的感受野感知特征局部反差的能力,也即感知的多尺度特性。视觉显著图是描述场景图像中各个位置视觉显著性的一幅二维图像。它融合了多种视觉敏感特征在多个尺度上的局部差异性。计算视觉显著图,首先需要建立图像的多通道多尺度表示;然后提取局部对比度特征图;最后将不同类型的特征图融合为视觉显著图<sup>[5-8]</sup>。针对灰度图像序列,采用以下方法提取显著性特征。

### 1.1 静态显著性的计算

为了提取灰度和细节等静态特征的局部显著性,可以采用 Center-surround<sup>[1-3]</sup>算子与灰度和细节特征图卷积的方法。这是因为 Center-surround 算子能够模拟生物视觉系统对于刺激的感受方式,其形状为 DOG 函数:

$$DOG(x, y, \sigma_c, \sigma_s) = \frac{1}{2\pi\sigma_c^2} \exp\left[-\frac{r^2}{2\sigma_c^2}\right] - \frac{1}{2\pi\sigma_s^2} \exp\left[-\frac{r^2}{2\sigma_s^2}\right] \quad (1)$$

其中,  $r = \sqrt{x^2 + y^2}$ ,  $\sigma_c$  控制着中心兴奋区的大小,  $\sigma_s$  控制着周围抑制区的大小。场景图像中与周围特征存在反差的位置经过 Center-surround 算子的作用将产生较强的响应。因此, Center-surround 算子与特征图的卷积,实现了从视觉敏感特征向视觉刺激的转化。具体实现时,以原图像的多通道、多尺度表示为基础,利用粗尺度图像代表周围抑制区的图像特征,以细尺度图像代表中心兴奋区的图像特征,通过“尺度间差”运算可以得到与 Center-surround 算子相同的结果。

假设输入图像  $f_k(x, y)$  为灰度图像,首先采用 Mallat 算法<sup>[1,4]</sup>对原图像进行多级小波分解,将各级小波分解中的低频分量组成灰度金字塔  $G_k(s)$ ,将小波分解中的三个高频分量组成细节金字塔  $D_k(s, d)$ ,其中,  $s \in [0..8]$  表示尺度,  $o_d \in \{HL, LH, HH\}$  表示 3 个细节的方向。经过以上处理所求得的金字塔  $G_k(s)$  和  $D_k(s, d)$  便构成了原图像的多通道多尺度表示,其中的 4 个金字塔分别代表图像在多个尺度上的灰度和细节属性。

以图像的多通道多尺度表示为基础,可以定义任意两个尺度  $s$  和  $c$  (其中  $s$  可以表示精细尺度,  $c$  可以表示粗尺度)表示的特征图  $F(c)$  和  $F(s)$  的“尺度间差”<sup>[1,4]</sup>为

$$F(c, s) = F(c) \ominus F(s) \quad (2)$$

其中,  $\ominus$  的含义是:将  $F(s)$  插值为和  $F(c)$  具有相同大小的图像区域,然后逐点相减并取绝对值。

“尺度间差”  $F(c, s)$  体现了对于特征  $F$ , 由  $c$  和  $s$  决定的特定尺度的 Center-surround 算子所感受到的场景中各个位置的局部特征对比度。根据上述定义得到灰度特征和细节特征为:

$$G_k(c, s) = G_k(c) \ominus G_k(s) \quad (3)$$

$$D_k(c, s, o_d) = D_k(c, o_d) \ominus D_k(s, o_d) \quad (4)$$

由于在计算过程中,灰度和细节两种特征图的动态范围不一致,相互之间无法进行比较,也给各种特征之间的相互融合带来困难,故需要进行规范化处理,将其动态范围调整一致。设  $N(\cdot)$  为规范化函数,通过  $N(\cdot)$  的运算,将灰度和细节两类特征的动态范围均调整为  $[0..1]$  区间内。同时,为了能够将不同尺度的特征图进行融合处理,定义一个“尺度间和”<sup>Y</sup>运算,将粗尺度的特征图差值到与细尺度特征图相同的大小,然后逐个像素点相加。于是通过式(5)~(7)得到了一幅幅可以反映场景中各个位置灰度和细节显著性相融合的静态视觉显著性特征  $S1_k$ , 其对应的二维图像称为静态视觉显著图。

$$\tilde{G}_k = \sum_{c, s | c \neq s} N(G_k(c, s)) \quad (5)$$

$$\tilde{D}_k = \sum_{o_d \in \{HL, LH, HH\}} N\left(\frac{\dot{Y}}{c, slc \neq s} N(D_k(c, s, o_d))\right) \quad (6)$$

$$S1_k = N(N(\tilde{G}_k) + N(\tilde{D}_k)) \quad (7)$$

其中,  $N(\cdot)$  表示规范化函数,  $\tilde{G}$  表示灰度显著性特征,  $\tilde{D}$  表示细节显著性特征。

## 1.2 动态显著性的计算

本文中的动态显著性是指由目标运动与全局运动的差异引起的视觉显著性。提取目标的动态显著性, 需要对图像序列进行运动估计和运动补偿, 以消除全局运动的影响。采用灰度投影法实现。其基本思想是分别在  $x$  方向和  $y$  方向对图像像素的灰度值进行投影, 把二维图像转换为两条一维特性曲线, 然后对邻近帧的特性曲线进行相关运算, 得到两帧之间的平移量。设  $F_k(i, j)$  表示第  $k$  帧图像中第  $(i, j)$  像素的灰度值,  $dx$  和  $dy$  表示两帧图像之间的平移量, 则  $x$  和  $y$  方向的特征曲线  $X_k$  和  $Y_k$  可以表示为

$$X_k(j) = \frac{1}{nH} \sum_i F_k(i, j), \quad j \in [1, W] \quad (8)$$

$$Y_k(i) = \frac{1}{nW} \sum_j F_k(i, j), \quad i \in [1, H] \quad (9)$$

为了估计两帧图像之间的平移量  $dx$  和  $dy$ , 建立如下的匹配相似性测度:

$$S_x(dx) = 1/[1 + \sum_j |X_k(j) - X_{k-m}(j + dx)|] \quad (10)$$

$$S_y(dy) = 1/[1 + \sum_i |Y_k(i) - Y_{k-m}(i + dy)|] \quad (11)$$

显然, 如果两帧图像内容相似, 其特性曲线也应相同。由于图像序列中邻近两帧的大部分内容相同, 因此其特性曲线也基本相似。但是由于两帧图像之间的整体运动, 会引起对应特性曲线产生平移。当平移量估计正确, 对应特性曲线的匹配相似性测度应取最大值, 因此得到以下平移量估计公式:

$$\hat{dx} = \arg \max_{dx} S_x(dx) \quad (12)$$

$$\hat{dy} = \arg \max_{dy} S_y(dy) \quad (13)$$

其中,  $\hat{dx}$  和  $\hat{dy}$  是估计得到的平移量。根据平移量估计值, 可以得到第  $k$  帧的运动显著图  $S2_k(i, j)$  为

$$S2_k(i, j) = |F_k(i, j) - F_{k-m}(i - \hat{dy}, j - \hat{dx})| \quad (14)$$

## 1.3 融入运动特性的显著性特征计算

由于单一的静态显著性特征的提取方法加强的是图像中心区域与周围区域相比灰度值较高或较低、细节变化剧烈的区域, 而对目标运动特性没有考虑, 因此在静态显著图中会有很多静止区域因其灰度及细节变化剧烈而得到加强。假设我们关注的是运动目标的检测问题, 则这些被加强的静止区域将对真正的运动目标产生干扰。而对于单一的动态显著图来说, 它所加强的是目标运动与全局运动有差异的区域, 因此, 除了真正的运动目标被检测出来外, 还可能产生一些干扰区域, 这些干扰区域可能是由于相机微小抖动或是其他因素所引起的图像各帧之间的微小变化所产生的, 而这种干扰对尺度较小的运动目标将产生较强的影响。因此, 为了克服上述两种单一显著性特征提取方法在复杂背景下进行运动目标检测的局限性, 需将静态、动态显著性特征提取方法相融合, 得到融合后的显著图便是加强了具有运动、灰度和细节显著性、足以引起视觉注意的运动目标区域, 具体融合方法如下。

设  $M(i, j)$  为某一显著性特征,  $\bar{M}$  为其均值,  $M_{\max}$  为其全局极大特征值, 则定义一显著性加强因子  $w_M$  为

$$w_M(i, j) = M(i, j) - [\theta_M(M_{\max} - \bar{M}) + \bar{M}] \quad (15)$$

经过加强因子作用后的显著图为

$$M_{\text{res}}(i, j) = w_M(i, j) + M(i, j) \quad (16)$$

根据公式 (15) 及 (16) 可知, 这种修正的全局加强法将全局极大显著性与平均显著性相比较, 对于在该特征下呈现强显著性的区域进一步加强, 而显著性弱的区域则被抑制。而且, 修正因子  $\theta_M$  的取值范围为  $0 < \theta_M < 1$ , 它可根据图像的实际情况以及需要注意捕获的显著性区域  $A_0$  的大小而确定, 与  $A_0$  的

大小成反比。该修正因子的作用示意图如图 1 所示, 当  $\theta_M^1 > \theta_M^2$  时, 被进一步加强的区域  $A_{\theta^1} < A_{\theta^2}$ 。这种可根据实际问题选择注意捕获的显著性区域大小的能力, 在确定所关注的场景图像中的运动目标的数量时相当有用, 从而对利用动态显著性特征实现粒子滤波的多目标跟踪非常有用。因为, 在实现运动目标跟踪的过程中, 当场景中的运动目标的数量过多时, 必须根据实际问题对众多的目标进行取舍, 把有限的计算资源用于一定数量的目标跟踪处理中。

根据上述修正的全局加强方法, 利用公式 (17) 实现对静态显著性特征  $S1$  和运动显著性特征  $S2$  的融合, 得到动态显著性特征  $S$ :

$$(S(i, j))_k = N(S1(i, j) + ws1(i, j)) \dot{Y} N(S2(i, j) + ws2(i, j)) \quad (17)$$

其中,  $N(\cdot)$  表示规范化函数, 将不同特征归一到区间上  $[0, 1]$ ; “ $\dot{Y}$ ” 为“尺度间和”运算, 其定义如上一小节。 $(S)_k$  为本文获取的最终的第  $k$  帧场景图像动态显著性特征值, 它反映了第  $k$  帧图像中各个位置上灰度、细节和运动等局部特性相融合后的动态显著性, 因此对于灰度、细节和运动特征同时显著的位置具有很好的指示作用, 故根据  $(S)_k$  的结果可获得场景图像中各个像素点所对应的动态视觉显著图。

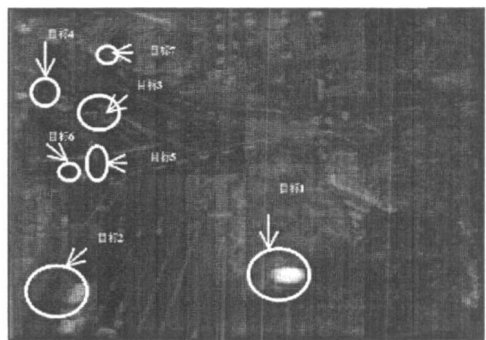
## 2 实验结果

依据以上算法机理, 通过以下实验来证明本文提出的显著性特征提取方法的可行性及有效性。

实验一 根据本文提出的算法, 首先对图像序列提取静态的灰度及细节显著性特征, 如图 2 所示, 图 2(a)、(c)、(e)、(g)、(i) 是一组交通监控序列图中的第 5、32、45、89、147 帧, 在这组图像中, 运动目标数量众多, 大小不一, 图 2(b)、(d)、(f)、(h)、(j) 是利用本文阐述的提取图像静态显著性特征方法得到的相应各帧的静态显著图, 从两组图像的对比中可以看出, 当考虑目标的灰度、细节等静态信息时, 模拟人眼的视觉注意机制, 图像有明显变化, 能引起视觉注意的区域得到加强, 其他区域受到抑制, 可以看出, 在图 2(b) 中, 目标 1、目标 2 最为突出, 目标 3 及目标 4 等几个典型目标较为突出, 但其显著性明显低于目标 1、2, 目标 5、6 及目标 7 则较之目标 1 更加不显著。在图 2 第 32、89 及 147 帧中, 分别出现了新目标——目标 8、9 及目标 10 和 11, 但由于目标与背景的灰度及细节变化不明显, 因此在静态显著图中目标 8 及 9 只是隐约可见。同时, 在第 147 帧中, 目标 3 和 4 不再运动, 但由于其灰度和细节的变化特征仍很明显, 故而在静态显著图中仍然很显著。而且, 在图 2 第 32、45、89 帧中, 目标 4 被标志牌从局部遮挡→几乎全遮挡→局部遮挡的过程中, 由于其边缘细节变化的显著性较强, 可以看出在几幅显著图中目标 4 均可见。



(a) 第 5 帧原始图像



(b) 第 5 帧静态显著图

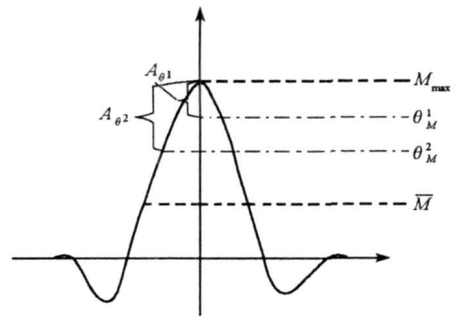
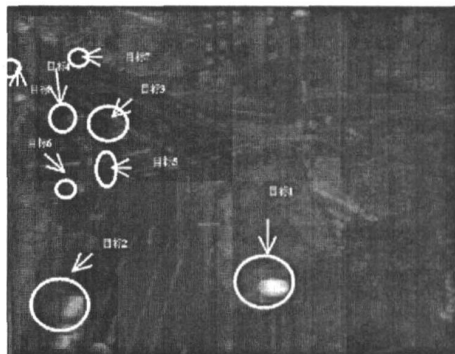


图 1 修正因子作用示意图

Fig. 1 The effect of  $\theta_M$



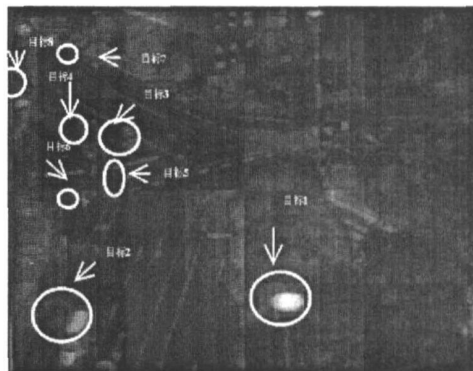
(c) 第 32 帧原始图像



(d) 第 32 帧静态显著图



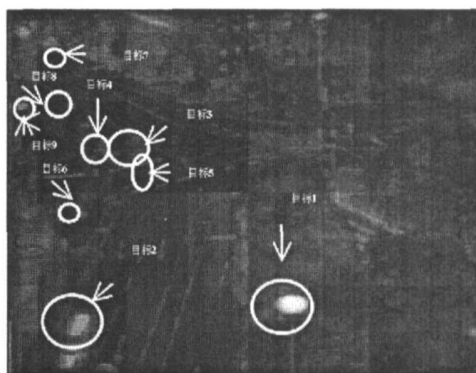
(e) 第 45 帧原始图像



(f) 第 45 帧静态显著图



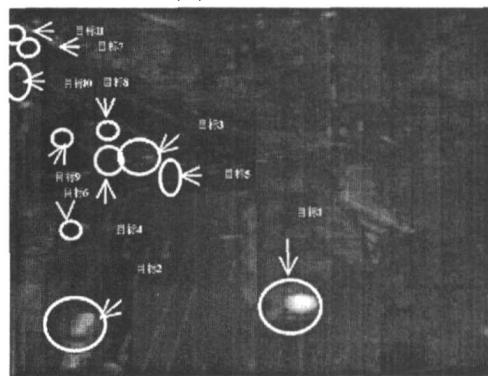
(g) 第 89 帧原始图像



(h) 第 89 帧静态显著图



(i) 第 147 帧原始图像



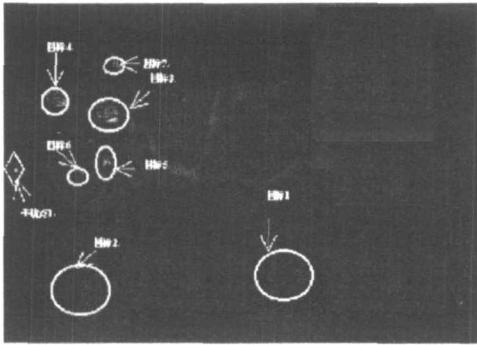
(j) 第 147 帧静态显著图

图 2 原始图像序列及其对应的静态显著图

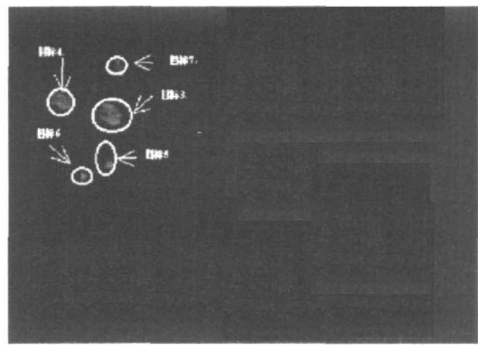
Fig. 2 The original images and their static state salient images

实验二 如图3所示,图3(a)、(c)、(e)、(g)、(i)是利用灰度投影法得到的第5、32、45、89、147帧动态显著图。在图3(a)中可以看出,在静态显著图中非常显著的目标1、2由于其不具有运动特性,因此在动态显著图中十分不显著,而目标3、4、5、6、7由于具有运动特性,故而在动态显著图中较为显著。在第32、89、147帧中新出现的目标8、9、10、11在动态显著图上均较为显著;在第89帧中,目标3不再运动,因此其显著性降低;在第147帧中目标4也不再运动,故而在这一帧中目标3和目标4几乎不见。同时在图3(a)、(c)、(e)、(g)、(i)中,可以看出由于相机微小运动或是其他因素所引起的运动区域,例如干扰点1、2以及干扰区域,在动态显著图中也呈现出来了,但在静态显著图中它们则完全不具有显著性。

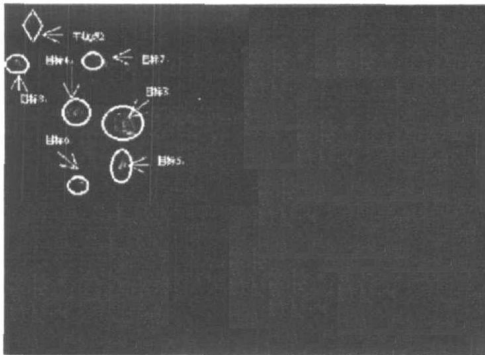
图3(b)、(d)、(f)、(j)是静态显著图和动态显著图的融合结果图,将其称为融入运动特性的目标显著性特征提取图,在这组融合图像中,既具有静态显著性又具有动态显著性的目标呈现出较强的显著性特征,而在静态显著图中显著的目标1、2由于不具有动态显著性,因此在图3(b)中几乎消失了,而目标3、4由于逐渐停止运动,故而分别在第89、147帧消失,而干扰点1和2等干扰区域由于在静态显著图中显著性很差,因此在图3(b)、(d)、(f)、(j)中也难见踪影了。可以看出,通过融合处理后,图像中能够引起人类视觉注意的运动目标得到很好的加强,而其他区域则受到明显的抑制。



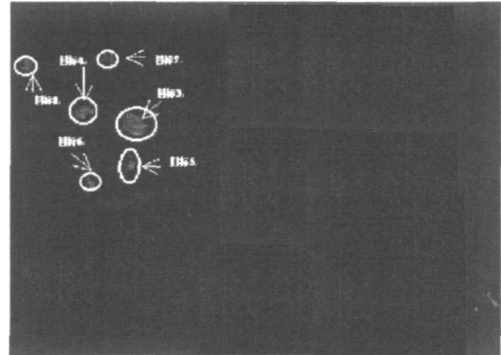
(a) 第5帧运动显著性特征提取图



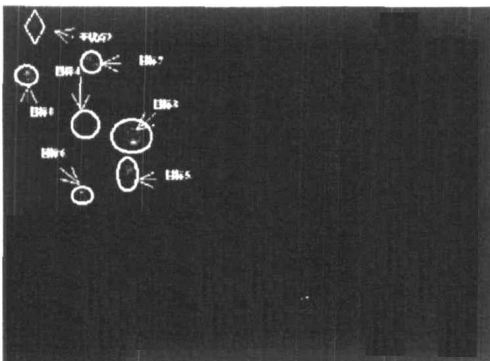
(b) 第5帧静态-动态显著性特征融合图



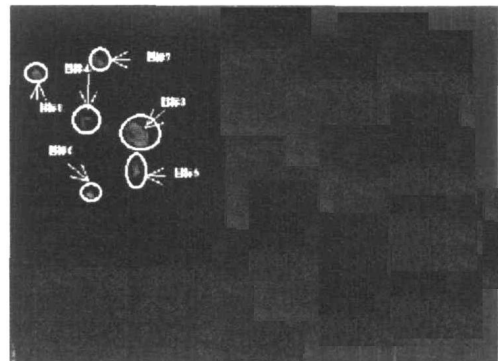
(c) 第32帧运动显著性特征提取图



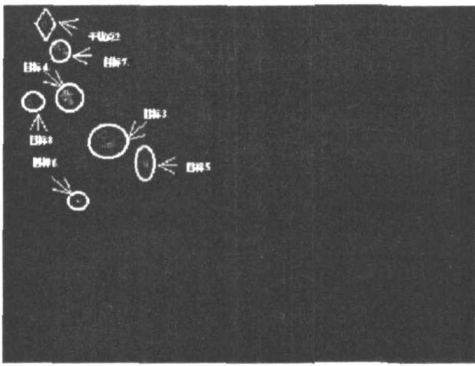
(d) 第32帧静态-动态显著性特征融合图



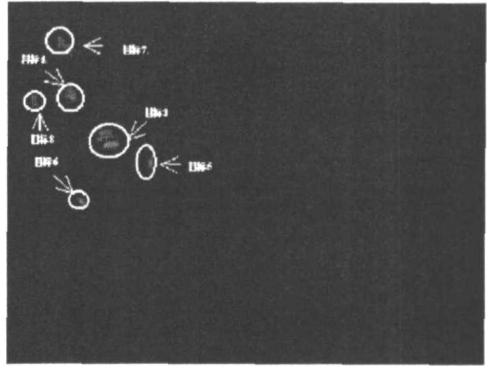
(e) 第45帧运动显著性特征提取图



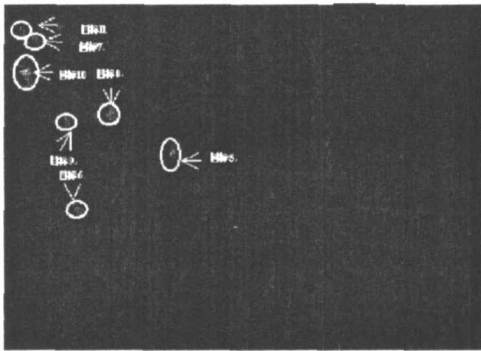
(f) 第45帧静态-动态显著性特征融合图



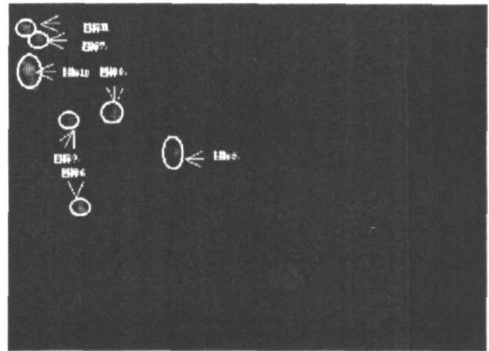
(g) 第 89 帧运动显著性特征提取图



(h) 第 89 帧静态- 动态显著性特征融合图



(i) 第 147 帧运动显著性特征提取图



(j) 第 147 帧静态- 动态显著性特征融合图

图 3 图像序列运动显著图及其对应的静态- 动态显著性特征融合图

Fig. 3 The motion saliency images and the images in which the static saliency fusing motion characteristic

### 3 小结

综上所述, 本文根据灰度图像序列特点, 利用图像底层的静态特征- 灰度、细节等以及目标的运动特性, 建立起的一种融入运动特性的显著性特征提取方法, 对于能够引起人类视觉注意的运动目标具有较好的检测能力。实验证明, 该方法具有较强的稳定性和实用性, 抗干扰能力强, 从视觉效果的角度出发, 能够较大地提高运动目标在复杂背景中的信杂比, 为实现复杂背景中运动目标的跟踪前检测方法提供了一种新的思路。

### 参考文献:

- [1] Itti L, Koch C, Niebur E. A Model of Saliency-based Visual Attention for Rapid Scene Analysis[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254- 1259.
- [2] Itti L, Koch C. Computational Modeling of Visual Attention[J]. Nature Review Neuroscience, 2001, 2(3): 194- 230.
- [3] Itti L, Koch C. Feature Combination Strategies for Saliency-based Visual Attention Systems[J]. Journal of Electronic Imaging, 2001, 10(1): 161 - 169.
- [4] 张志龙. 基于遥感图像的重要目标特征提取与识别方法研究[D]. 长沙: 国防科技大学, 2005.
- [5] Stentford FWM. An Evolutionary Programming Approach to the Simulation of Visual Attention[C]//Proc. of the IEEE Congress on Evolutionary Computation, Seoul, 2001: 851- 858.
- [6] Dan W, Riad H, Ronald S. Improved Likelihood Function in Particle-based IR Eye Tracking[C]//Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [7] 张鹏, 王润生. 基于视点转移和视区追踪的图像显著区域检测[J]. 软件学报, 2004, 15(6): 891- 898.
- [8] Itti L, Koch C, Niebur E. Rapid Biologically Inspired Scene Classification Using Features Share with Visual Attention[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(2): 300- 312.