

文章编号: 1672-8785(2016)03-0025-07

一种基于视觉注意模型的可移动 目标检测方法

吕云波 张志龙 李吉成 李树新

(国防科学技术大学电子科学与工程学院 ATR 国家重点实验室, 湖南长沙 410073)

摘 要: 提出了一种基于视觉显著性的目标检测算法, 用于对位于地面上的车辆、飞机等地面可移动目标进行检测和定位。针对地面可移动目标在场景中较小的特点, 设计了一种用于对目标检测进行引导的基于视觉注意机制的目标显著模型。首先, 提取图像目标的颜色特征、强度特征和方向 Gabor 特征, 并将其结合起来用于计算显著图。然后以超像素为单位对显著值进行计算, 并结合人眼视觉敏感度对不同距离的超像素之间的差异进行加权处理。实验结果表明, 本文算法可以有效地检测并定位出复杂背景中的地面可移动目标。

关键词: 视觉显著性; 地面可移动目标; 人眼视觉敏感度; Gabor 特征; 超像素

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2016.03.006

A Ground Mobile Target Detection Method Based on Visual Attention Model

LV Yun-bo, ZHANG Zhi-long, LI Ji-cheng, LI Shu-xin

(ATR National Key Laboratory, College of Electronic Science and Engineering,
National University of Defense Technology, Changsha 410073, China)

Abstract: A target detection method for detection and location of mobile targets such as vehicles and airplanes on the ground based on a visual attention model is proposed. According to the factor that the mobile targets in natural scenes are small, a target saliency model for guiding target detection based on visual attention mechanism is designed. First, the features such as color, identity and orientation in a target image are extracted and are combined with each other to calculate the saliency map. Then, the saliency value is calculated in the unit of superpixel and a weight is provided for the difference among the superpixels at different distances by combining with the human visual sensitivity. The experimental results show that the algorithm can detect and locate mobile targets against complicated background on the ground effectively.

Key words: visual saliency; ground mobile target; human visual sensitivity; Gabor feature; superpixel

收稿日期: 2016-02-26

基金项目: 国家自然科学基金项目 (61101185)

作者简介: 吕云波 (1985-), 男, 云南昆明人, 硕士研究生, 主要研究方向为目标检测与识别。

E-mail: 570938695@qq.com

0 引言

地面可移动目标是指位于地面或海面上有可能改变位置的目标,比如飞机、车辆和舰船等。可移动目标是航空航天军事侦察和精确打击密切关注的热点目标。无论是高分辨率遥感图像的自动判读处理,还是各种精确制导武器的自动目标识别,都需要使用针对各种可移动目标的检测识别技术。可移动目标检测识别面临的主要困难如下:目标种类繁多;同一类目标会由于姿态、观测距离和观察角度的不同而表现出不同的外观;复杂的战场环境及人为干扰也会大大增加可移动目标检测识别的难度。

视觉是人类最重要的一种感知觉。大脑所接收的外界信息绝大多数都来自人类的视觉。人类视觉系统对接收到的外界信息具有很强的筛选能力,即拒绝无用信息,保留有用信息,并及时作出反应。这种主动的、有选择性的筛选信息的活动称为注意。在视觉感知过程中,视觉注意系统通过视觉注意机制完成对信息的筛选,并把注意力集中在少数感兴趣的区域上以待后续处理^[1]。正是由于这种选择性,人类视觉系统才能够有效地完成复杂信息的处理任务。同时,人类视觉系统在处理视觉信息时,能够综合颜色、亮度、饱和度、方向和纹理等多种特征,并具有较高的鲁棒性。如果能够在目标检测识别时参照和模仿人类视觉系统的这种选择注意过程,使计算机只对可迅速引起观察者注意的显著区域进行及时处理,那么所得到的目标检测识别方法将会具备人类视觉的主动性、选择性、鲁棒性和高效性。

本文研究了一种基于视觉注意的可移动目标检测方法。其基本思路是,首先通过模仿人类视觉系统的多种信息处理过程,得到一种更为完善的视觉注意模型;然后利用视觉注意模型计算输入图像的显著图,接着根据显著图完成目标检测。具体说来,本文的研究思路主要体现在以下三个方面:

(1) 通过局部对比提取目标差异。图像中目标与背景的差异在本质上是目标与背景的特征

差异。目标在颜色、亮度、饱和度、纹理和方向等特征上的独特性,使其在与背景比较时可产生较强的显著值。

(2) 根据距离改变敏感度。根据生理心理学领域对于视觉注意机制的研究成果,眼睛在观察场景图像时会选择一个最显著的区域作为注视点;离注视点越近的地方,视觉注意的敏感度越高;离注视点越远的地方,视觉注意的敏感度越低。因此,有必要根据距离改变敏感度。

(3) 采用 Gabor 滤波器组提取图像特征。研究人员通过模拟生物视觉发现^[2], Gabor 函数的特性与视觉神经细胞的工作机理比较相近。一维 Gabor 函数可以很好地描述简单细胞感受野的剖面线,同时二维 Gabor 函数在模拟感受野时也可以得到不错的效果。另外, Gabor 函数还具有独特的时频特性,它不仅可以达到时频测不准关系的下界,而且还可以更好地兼顾信号在时域和频域中的分辨能力。

综上所述,本文在设计视觉注意模型时,从人类视觉系统的生理学过程出发,选取了更加符合人眼视觉特性的 Gabor 特征,并考虑了视觉敏感度的变换,使得视觉注意模型用于可移动目标检测时更准确并具有更强的鲁棒性。

1 基于视觉注意的目标显著性检测

假设待处理的场景图像为 $f(i, j)$ 。图 1 所示为位置 (i, j) 的显著性 $S(i, j)$ 的计算流程。首先,按照某种方式对图像进行分块;其次,对于以 (i, j) 为中心的中心图像块 C 和参与对比的多个周围图像块 S_k ($k = 1, 2, \dots, K$), 分别提取 Gabor 特征 $F_1 \sim F_{24}$ 、亮度特征 F_{25} 和标准方差特征 F_{26} ; 然后根据图像块的特征计算局部反差;接着根据局部反差得到显著度;最后,通过综合显著度得到位置 (i, j) 的显著性 $S(i, j)$ 。

这个计算过程涉及图像分块方法、特征提取方法、视觉对比敏感度模型、图像块的差异性函数、显著图的生成以及多视距显著性的融合处理。下面对以上内容进行逐一介绍。

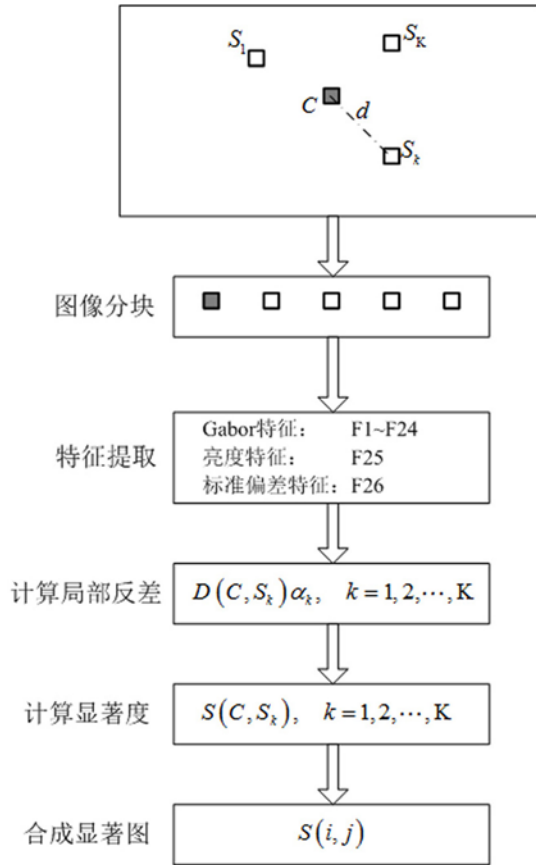


图 1 显著性的计算流程

1.1 图像分块和特征提取

1.1.1 图像分块

根据生物学领域的研究成果, 人类视觉在观察场景时以图像区域为基本单位。根据这一特点, 本文将待处理图像分为 $p \times p$ 的图像块, 其中每个图像块相当于一个巨大的像素 (称为超级像素)。以超级像素为单位来计算显著图可以极大地减少计算量。同时, 为了保障小目标的显著性不因图像块过大而降低, 在图像分块时应使不同图像块之间具有较大的重叠性。根据相关经验, 本文取图像块的大小为 $p = 8$ 。

1.1.2 Gabor 特征提取

Gabor 变换是时频信号分析的重要工具。Gabor 滤波器具有良好的方向和频率选择特性, 能在空域和频域同时达到最优的联合分辨率。二维 Gabor 函数可以很好地描述哺乳动物视觉皮层细胞的感受野特性^[3], 因此在图像处理、模式识别和计算机视觉领域有着广泛应用。

二维 Gabor 滤波器的脉冲响应函数可以表示为一个正弦平面波与一个二维高斯函数的乘积。该函数由相互正交的实部和虚部组成, 即

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \times \left[\cos\left(2\pi\frac{x'}{\lambda} + \psi\right) + i \cdot \sin\left(2\pi\frac{x'}{\lambda} + \psi\right)\right] \quad (1)$$

式中, $x' = x \cos \theta + y \sin \theta$; $y' = -x \sin \theta + y \cos \theta$ 。

二维 Gabor 滤波器具有以下 5 个关键参数:

(1) λ 为波长, 其单位为像素, 通常应不小于 2 但不能大于图像大小的 $1/5$ 。

(2) θ 为方向, 用于确定 Gabor 滤波器平行条纹的方向, 其取值范围为 $[0^\circ, 360^\circ]$ 。

(3) ψ 为相位偏移, 其取值范围为 $[-180^\circ, 180^\circ]$ 。 $\psi = 0^\circ$ 对应于 Center-On 函数; $\psi = 180^\circ$ 对应于 Center-Off 函数; $\psi = \pm 90^\circ$ 对应于反对称函数。

(4) γ 为长宽比, 用于确定 Gabor 函数形状的椭圆率。当 $\gamma = 1$ 时, Gabor 函数的形状为圆形; 当 $\gamma < 1$ 时, Gabor 函数的形状会沿着平行条纹方向拉长; 通常 $\gamma = 0.5$ 。

(5) σ 为高斯因子的标准差, 其值无法直接设置, 而是随着带宽 b 变化。而 b 又是 Gabor 滤波器的半响应空间频率带宽, 与 σ/λ 有关。

在我们的显著性计算模型中, 二维 Gabor 滤波器用于提取图像块的细节特征。图像块的 Gabor 特征通过图像块与 Gabor 滤波器组逐一卷积获得。设 $I(i, j)$ 表示待分析的图像块, 它与 Gabor 滤波器 $g(x, y; \lambda, \theta, \psi, \sigma, \gamma)$ 的卷积结果 $o_{\theta, \sigma}(i, j)$ 为

$$o_{\theta, \sigma}(i, j) = I(i, j) \otimes g(i, j; \lambda, \theta, \psi, \sigma, \gamma) \quad (2)$$

式中, $\otimes$ 为卷积算子; θ 和 σ 分别为 Gabor 滤波器的方向和尺度。 $o_{\theta, \sigma}(i, j)$ 表示对应方向取 θ 和尺度取 σ 时 Gabor 滤波器与输入图像的卷积结果, 即经滤波得到的特征向量。

当 θ 和 σ 取不同值时, 就可得到不同方向和尺度上的 Gabor 滤波器。在选择尺度 σ 时, 尺度越小, 则越能提取出更多的图像细节特征; 尺度越大, 所提取的内容更侧重于图像的轮廓特征。在选择方向 θ 时, θ 越多, 则越能得到更丰富的图像细节。但受运算速度的限制, 往往

需要根据图像特点选择几个最能反映图像大部分特征的方向,以便减少计算量和提高处理速度。因此,本文的显著性计算模型选择 4 个方向 ($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$) 和 3 个尺度 ($\sigma = 2, 4, 8$) 共计 24 个 Gabor 滤波器对图像进行了卷积,得到了 24 个 Gabor 特征,构成了特征向量的 $F_1 \sim F_{24}$ 。

1.1.3 其他特征的提取

明暗视觉也是人类视觉系统的一项重要功能,它对于视觉显著性的形成具有至关重要的作用。亮度特征是指光线作用于人眼时所引起的反映明亮程度的感觉。亮度特征可以使我们在比较均匀的背景中更容易发现亮度的突然变化,从而找到显著区域或显著目标。因此,我们建立的模型还引入了图像亮度特征,使图像中一些纹理或颜色对比不是很明显的目标得以凸显,以提高显著性检测结果的准确性。

我们用 F_{25} 表示图像块的亮度特征,其定义为图像块的灰度平均值:

$$F_{25} = \bar{I} = \frac{1}{N} \sum_{(i,j) \in \Omega} I(i,j) \quad (3)$$

式中, $I(i,j)$ 为图像位置 (i,j) 处的灰度值; Ω 代表图像块的区域; N 为 Ω 的像素数量。

另外,图像块的灰度标准偏差有助于发现图像复杂程度突变的区域,所以是另一个重要的亮度特征。因此,我们也将引入显著性计算模型作为图像块的第 26 个特征,并用 F_{26} 表示,其计算方法为

$$F_{26} = \left\{ \frac{1}{N} \sum_{(i,j) \in \Omega} [I(i,j) - \bar{I}]^2 \right\}^{1/2} \quad (4)$$

1.2 对比敏感度模型

生理视觉研究表明^[4-5],人眼的视网膜上有锥细胞和杆细胞两种视觉细胞。其中,锥细胞对场景中的颜色信息敏感,而杆细胞则对场景中的亮度信息敏感。与杆细胞相比,锥细胞具有更强的细节分辨能力,能够以更小的视角去观察相同距离的景物。根据人类视觉显著性的原理^[4],锥细胞在人眼视网膜上的不同区域内具

有不同的分布密度,因此人眼在观看图像时也有不同的视觉灵敏度。中央凹具有最大的锥细胞密度,其视觉灵敏度也最高;距离中央凹越远,锥细胞密度越小,视觉灵敏度也越低。因此,当人眼的视线固定在场景中的某个位置时,人眼对该位置的视觉分辨率最高。随着相对于该位置的距离不断增加,人眼的视觉分辨率会逐渐降低。

图 2 所示为人眼观察距离与视网膜离心距离之间的关系。其中, v 为人眼与图像之间的距离, e 为图像中位置 (x,y) 的视线与图像中心位置 (x_0,y_0) 的视线之间的夹角, d 为 (x,y) 与 (x_0,y_0) 的欧氏距离。模型中,根据人眼视觉的这种敏感度特性来确定图像块之间的差异权重。根据文献^[4],人眼的对比敏感度 $C_s(f,e)$ 被定义为对比度阈值 $C_t(f,e)$ 的倒数:

$$C_s(f,e) = \frac{1}{C_t(f,e)} \quad (5)$$

对比度阈值的定义为

$$C_t(f,e) = C_0 \exp \left(\alpha f \frac{e + e_2}{e_2} \right) \quad (6)$$

式中, f 为空间频率(周期/度); e 为视网膜偏心距(度); C_0 为最小对比度阈值; α 为空间频率衰减常数; e_2 为半分辨率偏心距。根据文献^[4]中的实验,三个参数分别选择为 $C_0 = 1/64$, $\alpha = 0.106$, $e_2 = 2.3$ 。

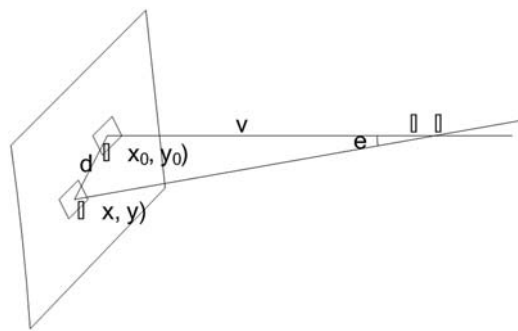


图 2 观看距离与视网膜离心距离之间的关系图

假设关注点为图像块的中心 (x_0,y_0) , 位置 (x,y) 的视网膜偏心距 e 可按照式 (7) 计算:

$$e = \tan^{-1}(d/v) \quad (7)$$

式中, d 为 (x, y) 与 (x_0, y_0) 之间的欧氏距离; v 为眼睛到图像的视距。视距与图像高度的比值通常为 $3 \sim 6$ 。

综上所述, 我们可以得到图像块之间差异的权重参数 $a_{(i,j)}$, 其计算公式为

$$a_{(i,j)} = \frac{1}{C_0 \exp\left(\alpha f \frac{e + e_2}{e_2}\right)} \quad (8)$$

1.3 目标图像块的差异性函数

图像中每个图像块的显著性可通过该图像块与图像中其余所有图像块之间的特征差异之和进行计算。一个图像块与其周围图像块的特征差异度越大, 则该图像块的显著性越强; 与其周围图像块的特征差异度越小, 则该图像块的显著性越弱。这里, 我们使用 Gabor 特征、亮度特征和方差特征的欧几里得距离来表示每个图像块与相邻图像块之间的对比度差异。其计算表达式如下:

$$D^{(v)}(c, s) = \sum_{i=1}^{26} \left| f_i^{(c)} - f_i^{(s)} \right| \quad (9)$$

式中, $D^{(v)}(c, s)$ 为图像块的显著度; $f_i^{(c)}$ 为中心图像特征; $f_i^{(s)}$ 为周围图像块特征; i 为特征标号, 即在前面所求的图像块特征 $F_1 \sim F_{26}$ 。

1.4 显著图的生成

这里, 我们将所有图像块的显著度与视觉灵敏度权重 α_k 加权后进行求和, 最终得到了总显著图。

$$S^{(v)}(c) = \sum_{k=1}^n \alpha_k \cdot D_k^{(v)}(c, s) \quad (10)$$



(a) 图像1



(b) 图像2



(c) 图像3

图 3 地面车辆和飞机目标图像示例

式中, α 为图像块之间的显著度差异权重, 即式(8)中计算的 $a_{(i,j)}$ 。

1.5 显著图引导的目标检测

在获得图像的显著图之后, 可根据显著图对地面可移动目标进行检测。本文采用固定阈值法进行了目标检测。阈值计算方法如下:

$$T = \frac{3}{W * H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} S(x, y) \quad (11)$$

式中, W 和 H 分别为图像的宽和高; $S(x, y)$ 为 (x, y) 位置上的显著值。

经过阈值分割后, 对从图像中获得的显著区域进行了鉴别, 并将长宽比异常的去掉, 获得了最终的目标检测结果。

2 实验结果

为了验证本文方法在地面可移动目标检测中的性能, 我们采用该方法对从网上搜集的地面车辆和飞机目标可见光图像进行了检测。图 3 所示为实验图像示例。

首先, 将只采用单个特征计算的显著图与综合使用全部特征计算的显著图进行了对比。图 4 中, 第一列是采用颜色特征计算得到的显著图, 第二列是采用文献 [6] 中的 QS 算法所获得的结果, 第三列是采用本文方法所获得的结果, 第四列是真实标记图像。从中可以看出, 仅采用颜色特征虽然对地面目标有一定的突出效果, 但对于地面背景没有很好地加以抑制, 造成了强烈的干扰。文献 [6] 中的 QS 方法则较好地解决了这个问题, 但受到第一幅图像中烟雾的干扰, 其附近的目标被掩埋。而采用本文方法

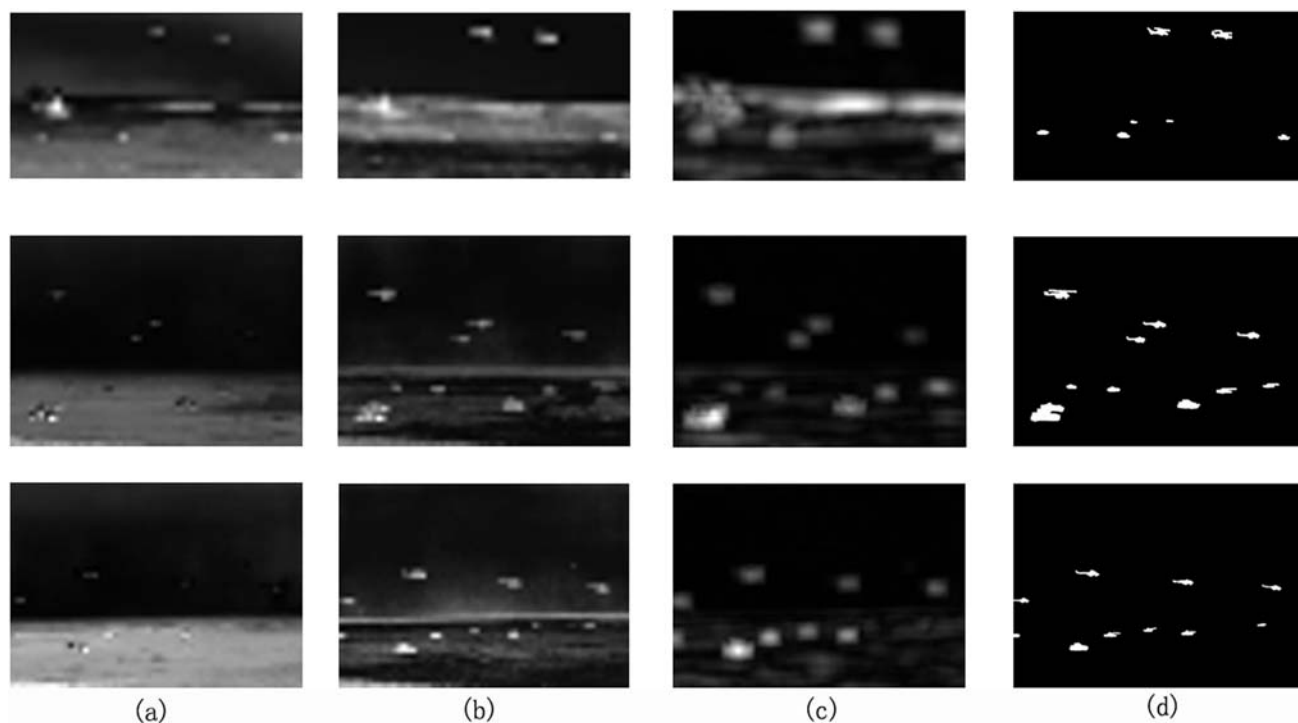


图 4 采用不同方法进行显著图计算的结果: (a) 通过颜色特征差异计算得到的显著图; (b) 采用 QS 方法所获得的显著图; (c) 采用本文方法所获得的显著图; (d) 目标位置的手工标记图

所得到的结果不仅能够较好地突出显著地面目标, 而且还能够有效抑制地面背景的干扰和较好地提取烟雾干扰下的目标。表 1 所示为通过计算显著图与真实标记图之间的 KL 散度 (Kullback-Leibler Divergence, KLD) 值^[7]所反映的显著图与真实标记图之间的接近程度。当显著图与真实标记图越接近时, KLD 值越小; 反之, 则 KLD 值越大。从表 1 中也可以看出, 采用本文方法所得到的显著图的 KLD 值均小于采用颜色特征的方法和 QS 方法。因此, 本文方法可以获得与真实标记更为接近的显著图。

表 1 采用不同方法所得的显著图的 KLD 值

KLD 值	图像 1	图像 2	图像 3
颜色方法	0.2595	0.1889	0.3617
QS 方法	0.2333	0.1150	0.1594
本文方法	0.1337	0.0533	0.0445

获得图像的显著图后, 以显著图为指导, 进

行阈值分割 (其结果见图 5)。与图 5 中第三列的用 QS 方法所获的结果相比, 本文方法可以有效提取出烟雾干扰下的目标 (如图像 1 左下角的目标), 同时还可以产生较少的虚警检测 (如图像 2 中的结果)。另外, 本文方法的检测效果也比 QS 方法更加稳定, 证明了前者对地面可移动目标进行检测的有效性。

3 结束语

本文综合利用图像中的 Gabor 特征、亮度特征、局部和全局特征等信息, 并结合人眼视觉灵敏度, 提出了一种新的以显著性引导的目标检测方法。该模型首先将输入图像转换成小的图像块, 然后使用 Gabor 滤波器提取图像块的特征, 其中还融合了图像块的亮度特征和标准差特征。每个图像块的显著性值是通过计算图像中该图像块与所有其他图像块的欧氏距离差异, 并加权人的视觉灵敏度对显著性值的影响获得的。实验结果表明, 本文方法可以获得与真实目

标标记图像更加接近的显著图, 并可获得更加稳健的地面可移动目标检测结果。该方法还可以推广应用到针对红外图像中的可移动目标的

检测中。下一步, 我们将结合红外图像的特点, 采用本文方法研究红外图像中的可移动目标检测问题。

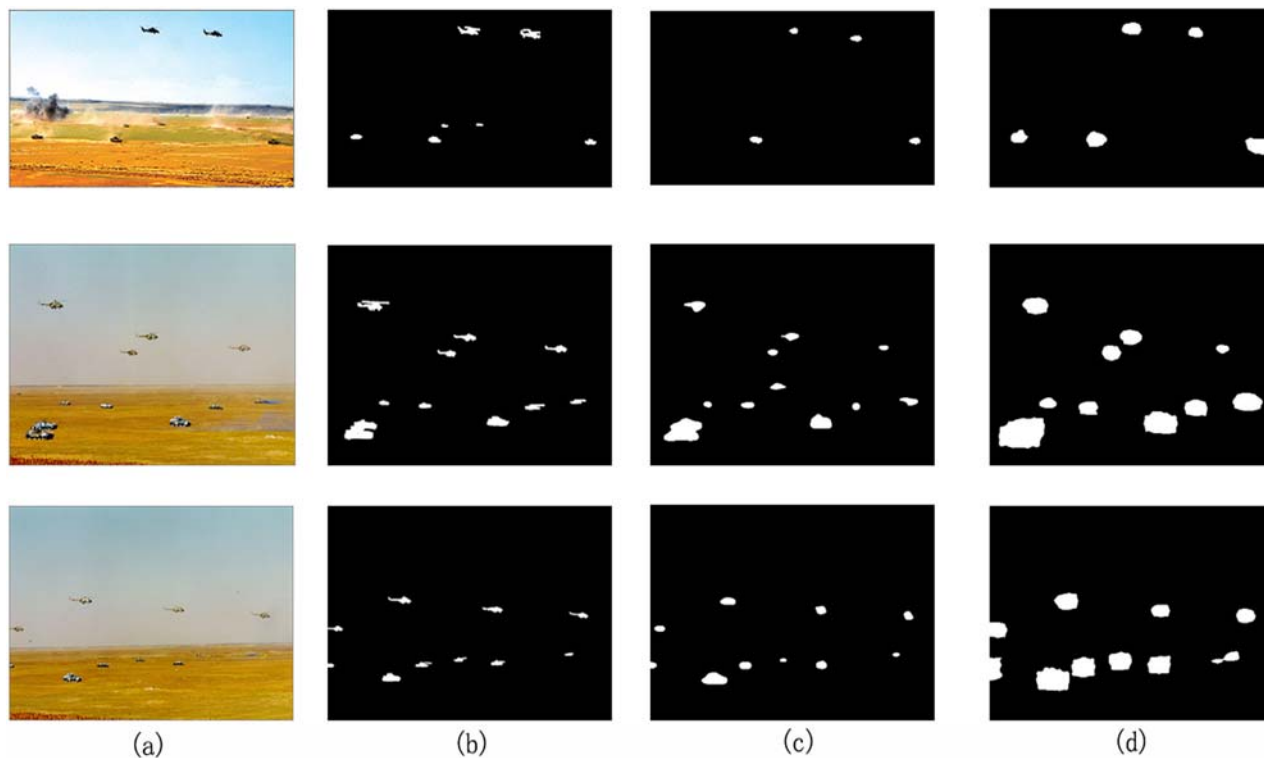


图 5 采用不同方法进行显著目标检测的结果: (a) 原始图像; (b) 目标位置的手工标记图; (c) 采用 QS 方法所获得的目标检测结果; (d) 采用本文方法所获得的目标检测结果

参考文献

- [1] 范娜. 视觉注意模型及其在目标检测中的应用 [D]. 西安: 西安电子科技大学, 2012.
- [2] Gonzalez R C, Woods R E. Digital Image Processing [M]. Englewood Cliffs: Prentice-Hall, 1992.
- [3] Greenspan H, Belongie S, Goodman R, et al. Over Complete Steerable Pyramid Filters and Rotation Invariance [C]. Seattle: IEEE Int Conf Computer Vision and Pattern Recognition, 1994: 222-228.
- [4] Wandell B A. Foundations of Vision [M]. Sunderland: Sinauer Associates, 1995.
- [5] Geisler W S, Perry J S. A Real-time Foveated Multi-resolution System for Low-bandwidth Video Communication [C]. SPIE, 1998, **3299**: 294-305.
- [6] Fang W M, Lin W S, Lee B S, et al. Bottom-Up Saliency Detection Model Based on Human Visual Sensitivity and Amplitude Spectrum [J]. IEEE Trans Multimedia, 2012, **14**(1): 187-198.
- [7] 仇媛媛. 基于视觉显著性的物体检测方法研究 [D]. 上海: 上海交通大学, 2013.