

## 基于旋转角预估的红外指定目标快速捕获

薛松<sup>1,2</sup>, 韩广良<sup>1</sup>

(1. 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033;  
2. 中国科学院大学, 北京 100049)

**摘要:** 针对抗旋转的红外目标捕获问题, 提出了一种基于旋转角预估的目标快速识别方法。该方法首先对图像进行局部自适应阈值的快速分割和连通域标记, 采用积分图像加速计算; 将每个连通域缩放至标准尺寸, 以缩放后的二值图像的像素值作为特征, 用神经网络估计目标的旋转角, 并在角度估计过程中滤除一些非目标; 最后, 用神经网络进行目标识别。针对实际应用中样本量往往较小, 模板与实际识别时可能存在差异的问题, 提出了一种边缘随机生长和消去的样本生成方法。实验表明, 提出的方法计算量小, 在模板形状有少量偏差的情况下, 仍能有较高的准确率。

**关键词:** 红外目标识别; 抗旋转; 自适应分割; 神经网络

**中图分类号:** TP751.1 **文献标志码:** A **文章编号:** 1007-2276(2013)11-2907-06

## Fast capture of appointed infrared targets based on estimation of rotation angle

Xue Song<sup>1,2</sup>, Han Guangliang<sup>1</sup>

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;  
2. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** For the problem of recognizing infrared targets, a method based on estimation of rotation angle was proposed. The method first segmented the image by local adaptive threshold and mark connected areas. Integral image was used to accelerate the computation of the local threshold. The connected areas were resized to the same size. The pixel values of the resized images were used as features. Then the rotation angle of the target was estimated by a neural network. Some areas were filtered during the estimation. At last, the area was recognized by another neural network. For some applications, the number of sample was not sufficient and a little difference existed between the samples and targets. For this problem, a method based on random growth and erosion was proposed to generate samples. Experiments show the method is effective and has a high recognizing rate even when the shape of samples is not exact.

**Key words:** infrared target recognition; rotation invariant; adaptive segment; neural network

收稿日期: 2013-03-11; 修订日期: 2013-04-07

基金项目: 国家自然科学基金(61172111)

作者简介: 薛松(1988-), 男, 硕士生, 主要从事视频目标捕获方面的研究。Email: [sxuemail@foxmail.com](mailto:sxuemail@foxmail.com)

导师简介: 韩广良(1968-), 男, 研究员, 硕士生导师, 博士, 主要从事机器视觉及视频目标识别跟踪方面的研究。  
Email: [hangl@ciomp.ac.cn](mailto:hangl@ciomp.ac.cn)

## 0 引言

红外目标的自动捕获是工程应用中的一个关键问题,该技术在军事领域如红外制导中,应用需求尤其紧迫,因而得到广泛的研究<sup>[1]</sup>,红外图像与可见图像相比,目标在图像中往往灰度比较单一,内部纹理信息不丰富,呈现的主要为轮廓形状信息,一般无法通过通常的纹理识别方法,而只能利用目标的形状信息来自动识别。空对地固定目标的自动捕获,一般有其他途径获取的目标图像作为模板来作为基准,同时目标姿态可获取,所以此类问题往往通过匹配的方法来解决,而在红外导引头对地机动目标自动识别应用中,如针对坦克、机场上位置不固定的飞机等,预期目标没有确定的姿态信息能够提供,即目标存在随机的旋转、缩放等变化,同时其灰度信息也不一定是图像中最显著的,这时就需要一种能够根据目标的典型先验知识来自动识别的方法,即抗目标旋转缩放的基于知识的识别算法。

针对抗旋转的红外目标自动识别问题,目前已取得较多的研究成果。其中使用最广泛的方法是先提取一些抗旋转的特征,然后使用一些分类器,如决策树、支持向量机、神经网络等进行识别<sup>[1]</sup>。对于指定形状的识别,提取的特征可以是 Hu 矩和 Zernike 矩<sup>[2-3]</sup>等旋转不变特征,也有使用圆谐展开以达到旋转不变的方法。梯度特征有时也被使用<sup>[4]</sup>。但特征的选取具有较大的人为性和经验性。Hu 矩虽然具有旋转不变性,但由于提取的特征较为简单,容易使具有相似矩特征的干扰物被误识为目标。圆谐展开等相对复杂的特征一般计算量较大,实时性不强。神经网络在识别问题中也很常用,但一般都要求先提取维数不能过多的特征<sup>[5]</sup>。

文中采取另一种思路:先用神经网络估计出目标的旋转角度,再进行识别。只要角度估计得比较准确,识别过程就不需要抵抗大角度旋转,从而不需要旋转不变的特征。这样,就把旋转不变性的问题转化为角度估计和无需抵抗旋转的识别这两个相对容易解决的问题。

目前已有少量研究采用的类似的预估角度再识别的方法,参考文献[6]采用了这一思路,但它面向的是具有比较显著的边缘、纹理等特征的可见图像,

对于目标几乎不含纹理特征的红外图像适应性不好。并且由于需要复杂的反复迭代,运算量较大,因此适用于对实时性要求不高的应用。

文中总体识别过程依次为自适应分割、连通域标记、旋转角估计和目标识别。同时,针对模板可能的偏差,文中提出了一种随机生长消去的样本生成方法。

## 1 局部自适应快速分割

图像自适应分割最常用的方法为双峰法和 OTSU 法。但红外目标捕获的应用中,目标在全图中所占区域往往较小,且不一定为全局最亮或最暗。故需要用局部自适应阈值方法。参考文献[7]使用局部 OTSU 和插值的方法对光照不均的图像进行分割,分割效果好,但对多块图像使用 OTSU 分割比较耗时。一些红外小目标的分割方法<sup>[8-9]</sup>也可借鉴。但考虑到误分割的区域可以在识别过程中滤除,且文中后续的算法识别一个形状是否为目标所需的时间很短,所以文中采用了一种会造成一些误分割但速度很快的方法。其基本思想为:对于图像的一个局部,如果其含有目标,则它应该有较好的可分性,因而具有较大的方差。如果目标区域在局部所占的比例不是很小,那么,使用像素均值作为阈值即可对目标区域进行较好的分割。使用积分图像,可以快速地计算出每个像素周围一个窗口中的像素均值和方差。二值分割后,再进行连通域标记,标出的连通域中就会包含目标连通域。具体算法如下:

STEP1 输入图像为  $f(i, j)$  ( $0 \leq i < h, 0 \leq j < w$ ) 计算输入图像的积分图像  $f_s$  和平方积分图像  $f_{sq}$ , 即

$$f_s(i, j) = \sum_{0 \leq i_0 < i} \sum_{0 \leq j_0 < j} f(i_0, j_0) \quad (0 \leq i < h+1, 0 \leq j < w+1) \quad (1)$$

$$f_{sq}(i, j) = \sum_{0 \leq i_0 < i} \sum_{0 \leq j_0 < j} f^2(i_0, j_0) \quad (0 \leq i < h+1, 0 \leq j < w+1) \quad (2)$$

STEP2 取以每个像素为中心的  $(2w_{win}+1) \times (2w_{win}+1)$  大小的窗口,其中  $2w_{win}+1$  取为目标大小的 1-3 倍。对于以像素  $f(i, j)$  为中心的窗口,计算其均值和方差,计算方法如下。

$$s = f_s(i+w_{win}+1, j+w_{win}+1) + f_s(i-w_{win}, j-w_{win}) - f_s(i-w_{win}, j+w_{win}+1) - f_s(i+w_{win}+1, j-w_{win}) \quad (3)$$

$$s_{sq} = f_{sq}(i+w_{win}+1, j+w_{win}+1) + f_{sq}(i-w_{win}, j-w_{win}) - f_{sq}(i-w_{win}, j+w_{win}+1) - f_{sq}(i+w_{win}+1, j-w_{win}) \quad (4)$$

$$\mu = s / (2w_{win}+1)^2 \quad (5)$$

$$\sigma^2 = [s_{sq}(2w_{win}+1)^2 - s^2] / (2w_{win}+1)^2 \quad (6)$$

式中:  $s$  为窗口中的像素和;  $s_{sq}$  为窗口中像素的平方和。  $\mu$  为像素均值,  $\sigma^2$  为方差。计算分割图像的像素值,如下:

$$f_b(i, j) = \begin{cases} 1, & \sigma^2 > T_\sigma \text{ and } f(i, j) > \mu \\ 0, & \text{else} \end{cases} \quad (7)$$

式中:  $f_b$  为分割图像;  $T_\sigma$  为设定的阈值。

STEP3 对  $f_b$  做连通域标记,记录每个连通域外接矩形窗的窗和高,滤除窗口过小和过大的连通域。其余连通域则作为文中系统下一步角度预估的输入。

图 1 为一图像经 STEP2 后的分割结果,其中图(a)为输入图像,图(b)为结果,可以看到,背景灰度不均一时,这种分割方法会引入很多背景中的形状。因此它对很多算法的预处理是不适用的,但对于文中的系统,这对于后续的计算影响较小。而且这种方法的计算量之小,可以大大减小目标捕获的总时间。

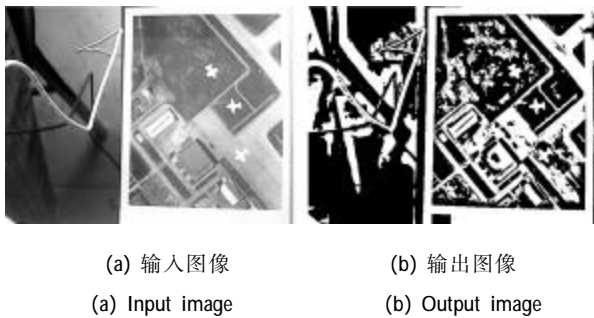


图 1 自适应分割结果示例  
Fig.1 Example of adaptive segment

## 2 目标旋转角预估

要在目标具有旋转的情况下对目标进行识别,常用的方法是提取其旋转不变特征。文中采用另一种方法,在识别之前先对目标的旋转角度进行预估。此节的目的在于:设计一个角度预估器,该预估器的输入为一个具有一定形状的连通域,输出为角度。对于目标连通域的输入,输出的角度为其转角的预估,允许一定的偏差,对于非目标连通域的输入,输出角度不限。

### 2.1 角度预估神经网络的结构

估计角度的一种方法是计算外接矩形中心到形状重心的向量方向,但一些形状的重心和其外接矩形的中心距离很近,向量方向的计算受噪声影响很大。文中采用一种更加通用的方法,通过神经网络的训练,得到一个角度预估器。

文中所设计的预估计的网络的结构使用单隐层 BP 网,输入层和隐层的转移函数为 tansig 函数。为了抵抗一定缩放的影响,同时也为了使网络地输入维数不至于太高,文中将连通域的二值图像缩放到标准尺寸,将缩放后图像的每个像素值作为网络的输入,经实验,文中取 25×25 像素为标准尺寸。网络的输入为 625 维,网络的输出为目标旋转角的 sin 和 cos 值。

### 2.2 边缘随机生长和消去的训练样本生成方法

实际应用时,样本的获取也是一个问题。在很多应用,如红外制导中,执行任务之前,经常只能获取少量的,甚至只有一幅目标的实际图像。有时只能获得目标的可见图像,红外图像的目标模板需要凭经验人为地制备。同时,执行任务时还存在天气、成像系统噪声和其他干扰。这些都会使执行任务时获得的目标形状和模板有一定的区别。虽然神经网络具有一定的容错能力,但以上问题仍会导致网络的识别率下降。因此,文中设计了一种训练样本生成方法,由目标的一个形状模板随机地生成相似的模式,使网络对以上的这些干扰有更好的适应性。

样本生成的具体为:把模板图像每隔一定角度旋转一次,在旋转后的图像中找出目标的外接矩形,将外接矩形内的图像缩放至标准尺寸。然后在形状的边缘上随机消去和生长出一些目标像素,具体步骤如图 2 所示。

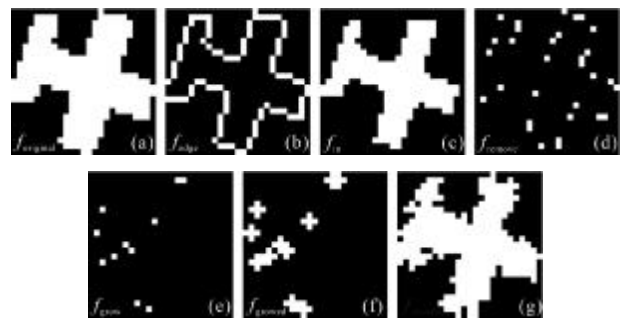


图 2 样本生成步骤  
Fig.2 Step of sample generation

STEP1 记目标模板图像为  $f_{original}$ 。分别提取目标轮廓和内部像素,生成两幅图像  $f_{edge}$  和  $f_{in}$ 。

STEP2 将  $f_{edge}$  的非零值像素,按照一个随机的比例随机分成三组,其中两组生成图像  $f_{grow}$  和  $f_{remove}$ 。

STEP3 将  $f_{grow}$  用 3×3 的十字模板膨胀,生成图像  $f_{growed}$ 。

#### STEP4 生成样本图像

$$f_{\text{sample}} = (f_{\text{grewed}} \cup f_{\text{edge}} - f_{\text{remove}}) \cup f_{\text{in}} \quad (8)$$

然后将样本图像的各个像素值作为样本输入, 旋转角度的  $\sin$  和  $\cos$  值作为输出。

由于此节的预估器只需对目标形状具有正确的角度输出, 其他形状的输出不影响识别。因此, 样本非常符合 BP 网相似输入产生相似输出的要求。虽然由于未进行降维, 网络输入维数较高, 但由于输入的情况相对单一, 同时, 后续识别允许角度预估有一定偏差, 因此, 网络很容易收敛到符合要求的精度。另外, 虽然网络的大维数增加了从输入到输出的运算量, 但由于省去了降维的计算, 总的角度预估时间减少了。

### 3 小角度旋转的目标识别

由于上一节的方法已经完成了目标旋转角的估计, 因此本节中的识别只需要抵抗较小的角度旋转。

#### 3.1 角度预估的负例预滤除

文中通过实验发现, 利用上一节的角度识别网络, 可以预先滤除大部分的负例。其滤除负例的原理为: 对于目标, 当其旋转一个角度时, 角度预估网络的输出也会旋转一个相近的角度, 而对于非目标, 由于训练时未将其作为样本, 相应的输出角度变化不一定, 且往往与其旋转角度有较大的偏差。由于将图像旋转  $180^\circ$  是一个很简单的运算, 因此, 文中采取以下方法: 将第二节中生成的标准尺寸图像旋转  $180^\circ$ , 再计算角度预估网络的输出。设旋转前输出为  $(x_1, y_1)$ , 旋转后输出为  $(x_2, y_2)$ , 若

$$(x_1 x_2 + y_1 y_2)^2 / [(x_1^2 + y_1^2)(x_2^2 + y_2^2)] < T_r \quad (9)$$

$$\text{或 } x_1 x_2 + y_1 y_2 < 0 \quad (10)$$

则直接判断该连通域不是目标, 不再进行后续识别。其中  $T_r$  为一设定阈值, 公式(10)的左边即为两向量的点积, 而公式(9)左边即为两向量夹角余弦值的平方。两式的意义即为: 旋转前后角度预估网络的输出向量夹角小于一个阈值。

#### 3.2 基于神经网络的目标识别

对于小角度旋转的形状, 现有的识别方法较多, 文中使用神经网络的方法。

##### 3.2.1 识别网络的结构

网络结构与角度预估器中使用的结构相同。将

连通域二值图像按预估角度反向旋转, 然后缩放到标准尺寸, 得到图像的各像素值作为网络的输入。网络的输出为连通域的为目标的置信度, 目标为 1, 非目标为 0。

##### 3.2.2 正负例样本的生成

正例样本的生成与角度预估网络的方法类似, 不同的是, 旋转角度的范围仅选择一个较小的范围。样本的输出为 1。

对于负例样本, 一种方法是使用随机二值向量, 但实验发现效果不甚理想。因此文中使用了从图库中获得负例的方法。具体方法为: 选择一个包含大量不同复杂场景的图库, 文中选择 PASCAL VOC2007 的图库。从图库中随机抽取一些图片, 按前述方法生成输入识别网络的向量, 作为负例。

##### 3.2.3 逐次添加负例的网络训练方法

由于负例需要尽可能全面地包含各种情况, 因此需要样本的数量较大。但过多的负例会使训练变得很慢, 内存消耗过大。而且, 从图库中随机地抽取样本, 当样本数量较多时, 实际上其中会包含大量相似的模式, 造成不必要的样本量增多。因此文中采用多次训练, 每次增加误识别负例的方法。具体为: 先随机抽取图片, 生成一些样本作为负例, 训练网络。然后, 再随机抽取一些图片, 用训练好的网络识别其中的连通域, 把其中误识为正例的样本加入样本集中, 再次训练网络。如此重复 3~4 次后, 网络就具有较好的负例拒绝能力。

### 4 实验结果和分析

该节中使用模拟空对地红外目标捕获的视频作为测试视频, 视频通过目标模型和场景的手工制作, 并使用含有一定噪声的摄像头拍摄得到, 模拟地面飞机捕获的任务。视频分辨率为  $640 \times 480$ 。为了模拟不同的情况, 测试视频分为三段, 第一段为简单背景, 第二段为复杂背景, 第三段在背景复杂的基础上, 通过调节摄像头焦距使图像模糊, 使目标外形与模板有少量差异。视频中包含两种飞机形状的多个目标。并且设置了一些干扰物。为了和真实应用时的情况相近, 用于训练网络的目标图像分为两组, 第一组仅包含了两种目标(约定称为目标 1 和目标 2)的模板二值图像, 如图 3 所示, 即模拟执行任务前未预料到目标图像与模板的差异。第二组在第一组的基础

上,加入目标模糊情况下的二值图像。

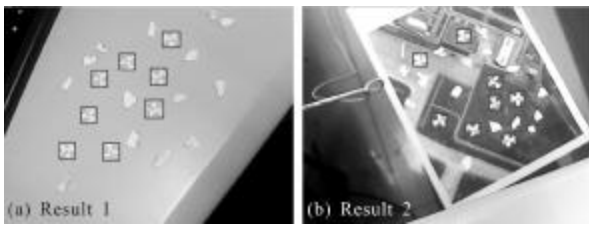


图 3 样本生成步骤

Fig.3 Templates of targets

实验平台是 CPU 为 Intel i3-2102,内存 2Gb 的 PC 机。程序只使用一个 CPU 核中的一个线程。

图像分割时,取  $T_e=100$ 。角度预估网络的样本生成时,每隔  $0.2^\circ$  旋转一次模板,即每个模板生成 1 800 个样本。识别网络训练时,从在正负  $30^\circ$  的范围内,每隔  $0.05^\circ$  旋转一次模板,即每个模板生成 1 200 个正例样本。负例样本的初始数量取 5 000,重复训练时,每次增加误识样本的数量取初始负例样本数量的 3%,增加 4 次。训练网络时,随机取样本的 70% 为训练集,20% 为验证集,10% 为测试集。经过测试,对于实验中的目标模板,两个网络的隐层节点数均取 10 较为合适。表 1 为使用两组模板生成网络的总时间(包括样本生成和网络训练),以及得到网络在测试集上的精度。

表 1 网络准确度和生成时间

Tab.1 Accuracy of network and time for generating network

No.	MSE of angle estimation	Recognition rate of positive samples/%	Recognition rate of negative samples/%	Time for training/s
1	0.004 17	100	98.9	348
2	0.007 51	100	99.5	463

可以看到,角度的估计有较小的误差,误差在识别网络的适应范围内。识别网络在测试集上有很高的准确率,说明网络对正负例样本可以有很好的区分度。部分图库中得到的样本可能与目标很相近而无法区分,造成一定的误识率。

然后在模拟场景的视频中测试网络的性能,并与经典的 Hu 矩识别方法以及用单个神经网络识别点阵图像的传统方法对比。角度预估滤除负例的阈值  $T_r$  取 0.5,结果如表 2 所示,图 4 为两种场景中一帧

的识别结果。其中,网络 1 由第一组模板生成,即模糊目标未知,网络 2 由第二组模板生成,即模糊目标已知。

表 2 识别结果

Tab.2 Result of recognition

No.	Method	Net	Frame count	Recognition rate /%	False alarm rate/%	Time /ms · frame <sup>-1</sup>
1	Proposed	1	90	97.0	0	24.8
	Hu	-		85.0	10.9	26.2
	Single net	-		90.6	2.8	22.3
2	Proposed	1	297	94.3	0	25.8
	Hu	-		83.3	11.4	28.2
	Single net	-		88.2	3.1	23.2
3	Proposed	1	390	89.0	0	26.8
	Hu	-		96.2	0.006	26.9
	Single net	-		74.8	11.2	28.8
	Proposed	2		80.2	2.3	23.9

可见,对于模板制备比较准确的情况,文中的方法具有很高的准确度。对于模板制备有一定偏差的情况,识别率同样较高。同时,文中的方法计算量小,具有较好的实时性。与传统方法中实时性较好的 Hu 矩和单神经网络方法相比,文中方法在处理时间基本不变的情况下,提高了正例的识别率,并大大降低了负例的误识率。并且,在得到目标位置信息的同时,也能得到目标的旋转角。

另外,实验中还发现,使用角度预估器的负例预滤除,可以滤去大约 60% 的负例,减少了后续计算量和负例的误识机率。

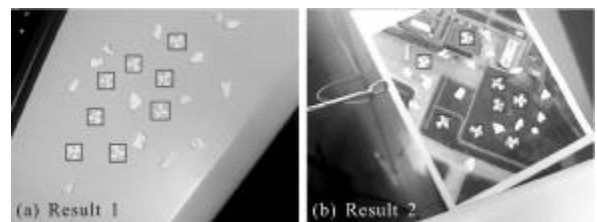


图 4 单帧捕获结果

Fig.4 Result for capture on single frame

## 5 结 论

文中提出了一种基于角度预估的抗旋转红外目标快速捕获方法,该方法识别准确度高,计算量小。文中的主要贡献在于:(1)提出了角度预估神经网络和识别网络的双网络识别方法,使单个网络完成的任务简单化,精度提高,同时避免了复杂的旋转不变特征提取;(2)提出了使用角度预估网络预先滤除识别负例的方法,提高了负例的拒绝率,同时减少了运算量;(3)提出了一种边缘随机生长和消去的训练样本生成方法,使得在模板较少且不准确的情况下,识别精度仍能保持较高。

文中算法实现了地面红外目标的自动捕获,后续进一步的工作是研究一种可以替代文中图像分割算法的方法,在计算量没有很大增加的情况下,减少误分割的目标,以减少识别过程的计算量和虚警数量。

### 参考文献:

- [1] Fan Jinxiang, Zhang Yuan, Wang Sheyang. Analysis of application status of ATR technologies in infrared imaging guided missiles [J]. *Infrared and Laser Engineering*, 2007, 36(6): 778-781. (in Chinese)  
范晋祥, 张渊, 王社阳. 红外成像制导弹自动目标识别应用现状的分析[J]. *红外与激光工程*, 2007, 36(6): 778-781.
- [2] Fu Yan, Wang Mei, Zhang Chunqin. SAR image target recognition based on Hu invariant Moments and SVM[C]// Fifth International Conference on Information Assurance and Security, 2009: 585-588.
- [3] Singh C, Walia E, Mittal N. Rotation invariant complex Zernike moments features and their applications to human face and character recognition [C]//The Institution of Engineering and Technology, 2011: 255-266.
- [4] Wan Lei, Zeng Wenjing, Zhang Tiedong, et al. Real-time detection of marine infrared objects based on the fusion of gradient information [J]. *Infrared and Laser Engineering*, 2013, 42(1): 41-45. (in Chinese)  
万磊, 曾文静, 张铁栋, 等. 基于梯度信息融合的海面红外目标实时检测[J]. *红外与激光工程*, 2013, 42(1): 41-45.
- [5] Li Junmei, Hu Yihua, Tao Xiaohong. Recognition method based on principal component analysis and back-propagation neural network [J]. *Infrared and Laser Engineering*, 2005, 34(6): 719-723. (in Chinese)  
李军梅, 胡以华, 陶小红. 基于主成分分析与 BP 神经网络的识别方法研究[J]. *红外与激光工程*, 2005, 34(6): 719-723.
- [6] Minoru Fukumi, Sigeru Omatu. Rotation-invariant neural pattern recognition system estimating a rotation angle [C]// IEEE Transactions on Neural Networks, 1997: 568-581.
- [7] Shen Han. Segmentation for uneven lighting document image [D]. Shanghai: East China Normal University, 2006. (in Chinese)  
沈晗. 非均匀光照文本图像的分割[D]. 上海: 华东师范大学, 2006.
- [8] Gao Jing, Sun Jiying, Liu Jing, et al. FLIR target recognition based on local fuzzy threshold [J]. *Optics and Precision Engineering*, 2011, 19(12): 3056-3063. (in Chinese)  
高晶, 孙继银, 刘婧, 等. 基于区域模糊阈值的前视红外目标识别[J]. *光学精密工程*, 2011, 19(12): 3056-3063.
- [9] Jin Yongliang, Wang Yanjie, Liu Yanying. Pre-detection method for small infrared target [J]. *Optics and Precision Engineering*, 2012, 20(1): 171-178.  
靳永亮, 王延杰, 刘艳滢, 等. 红外弱小目标的分割预检测[J]. *光学精密工程*, 2012, 20(1): 171-178.