

# 基于 AHP 的跟踪算法性能评价研究\*

樊祥<sup>1</sup>, 方义强<sup>1,2</sup>, 程正东<sup>1</sup>, 朱斌<sup>1</sup>, 施展<sup>1,2</sup>

(1 解放军电子工程学院, 合肥 230037; 2 脉冲功率激光技术国家重点实验室, 合肥 230037)

**摘要:**针对跟踪算法评价方法不够健全的问题,文中提出了基于 AHP 法的跟踪算法评价方法。主要从跟踪算法的跟踪准确度、算法的实时性以及算法的硬件可实现性这三个指标考虑了算法性能的评价问题,讨论了三个指标量化和规范的方法,运用 AHP 法给出了计算各个指标的权重和综合评价分数的方法。最后采用 4 种典型的跟踪算法和实际拍摄的红外图像序列进行了小目标的跟踪实验,并根据实验得到的数据计算了算法的评价分数,结果表明了文中跟踪算法评价方法的合理性。

**关键词:**跟踪算法;性能评价;层次分析法

**中图分类号:**TP391.4 **文献标志码:**A

## Performance Evaluation for Tracking Algorithm Based on AHP

FAN Xiang<sup>1</sup>, FANG Yiqiang<sup>1,2</sup>, CHENG Zhengdong<sup>1</sup>, ZHU Bin<sup>1</sup>, SHI Zhan<sup>1,2</sup>

(1 Electronic Engineering Institute of PLA, Hefei 230037, China;

2 State Key Laboratory of Pulsed Power Laser Technology, Hefei 230037, China)

**Abstract:** According to imperfectness of performance evaluation for tracking algorithm, a method of evaluation for tracking algorithm based on AHP was proposed. As for the evaluation, three parameters including veracity, real-time demand and the implementation on hardware of the tracking algorithm were mainly considered, and then the method for quantizing and regulating the three parameters was also discussed here. AHP was adopted to calculate the weight of the three parameters and the complete score of the evaluation. Finally, the small target tracking experiments were made by 4 typical small target tracking algorithms and 3 real IR images series, and the evaluation score of the algorithms were also calculated according to the experiment results. The results verify the rationality of the method of evaluation.

**Keywords:** tracking algorithm; performance evaluation; analytic hierarchy process

## 0 引言

对目标的检测跟踪算法在近几年得到了很大的发展。为了能够对算法有一个合适、客观的评价,更好的把握算法改进和研究的方向,促进对算法的优化,需要一个科学的算法性能评价方法。对于检测算法的性能评价研究得比较多,很多文献都对这方面做了比较深入的讨论研究<sup>[1]</sup>;而现行文献资料对于目标跟踪算法的评价方法研究较少,只分别从一些侧面对跟踪算法进行过初步的研究。文献[2]采用了有效跟踪评价和有效跟踪精度评价的方法对跟踪算法进行评价;综述文献[3]提到了目前国际上在智能视频算法中通过精确度、完整性来进行跟踪算法的性能评价;文献[4-6]用算法具有低的跟踪误差来说明算法性能优越性的问题;文献[7]用算法的跟踪有效性和精度来说明算法性能;文献[8-9]通过算法的实时性

来说明算法的优越性。但是这些文献都没有系统的对跟踪算法的性能进行评价,主要表现在两个方面:一是采用的性能指标不够全面;二是没有对各性能指标进行综合。

因此文中主要针对跟踪算法的评价方法不够健全的问题,利用层次分析法(AHP)提出了一种评价跟踪算法的方法,AHP法为运筹学中进行系统效能评价的一种方法,文中是第一次用AHP法来进行跟踪算法的性能评价。

## 1 主要性能指标的分析

要对跟踪算法进行评价,首先需选取合适的指标作为评价的依据。通过对相关文献和算法的研究,一个完整的跟踪算法主要涉及到的指标有跟踪准确度、算法的实时性和算法的硬件实现等,因此文中主要从算法的这几个方面来考虑跟踪算法的性能评价问题。

\* 收稿日期:2012-07-05

作者简介:樊祥(1963-),男,安徽滁州人,教授,研究方向:信号处理,光电对抗。

### 1.1 跟踪准确度

跟踪准确度是最直观的反映算法性能的一个方面。文献[10]提出了一个跟踪率的概念,跟踪率定义为,  $\alpha = n/N$ , 其中  $n$  为持续跟踪上的图像帧数,  $N$  为跟踪图像的总帧数。该方法可以在一定程度上表征算法对目标的跟踪准确度, 但该方法只是从是否跟踪上目标来说明跟踪的准确度问题, 是一个比较粗糙的方法, 不能精确的描述一个算法的跟踪准确度, 尤其是对于对红外小目标的跟踪描述。而多数文献用目标的算法预测位置与实际位置的差, 即跟踪误差来说明跟踪的准确度问题<sup>[3-4,7]</sup>, 该方法精确的描述了每帧图像的误差, 但同时该方法不对跟丢的情况进行判断, 使得跟踪误差在遇到目标跟丢的情况下失去了意义。因此文中将两者结合起来描述跟踪的准确度。

由于存在目标跟丢的情况, 因此需要对跟踪误差做进一步的定义。假设对于某个算法, 算法得到的有效预测坐标与实际坐标的平均跟踪误差为  $E(\mu)$ , 其定义如下:

$$E(\mu) = \frac{1}{n} \sum_{i=1}^n \mu_i \quad (1)$$

$$\mu_i = \sqrt{(x_{ei} - x_{ii})^2 + (y_{ei} - y_{ii})^2} \quad (2)$$

式中:  $x_{ii}$ 、 $y_{ii}$  是真实目标的坐标值;  $x_{ei}$ 、 $y_{ei}$  是算法跟踪有效的目标坐标值。这里计算平均跟踪误差时, 只计算了算法对目标有效跟踪部分的坐标值, 而没有考虑跟丢时误差的计算, 显然这样不够合理。因此在讨论误差时, 文中提出采用  $E(\mu)/S_{11} = E(\mu) \cdot N/n$  来衡量算法的跟踪误差。

### 1.2 跟踪算法的实时性

对目标的自动检测跟踪对实时性的要求很高, 在算法评价时必须要考虑其实时性的问题。现行的很多算法结构比较复杂, 时间开销很大, 虽然这些算法在一些特定的场合在某些方面可以取得相对较好的效果, 但是这样的算法实时性差, 仍然没有实际应用的价值。因此, 在考虑跟踪算法的性能评价时, 必须要把实时性考虑在内。

设  $v$  为算法处理速度, 单位为 fps(帧/s), 文中算法的处理速度用计算机仿真得到。算法的处理速度应该有一个合适的值, 这与有些文献上认为的处理速度越快越好的观点是不一样的。因为算法首先应该满足实时性的要求, 但是当算法达到的速度过快时, 已经远超出了实时性的要求, 超出的部分对于算法性能的提高没有影响或影响不大, 而且很多情况下处理速度的加快是以牺牲其它方面的性能为代价的。因此文中

在考虑实时性的时候提出处理速度上限和下限的概念, 分别用  $v_H$  和  $v_L$  表示, 用一个分段函数来表述处理速度对算法性能的影响, 表达形式如下:

$$\eta = \begin{cases} v_H & \text{if } v \geq v_H \\ v & \text{if } v_L \leq v \leq v_H \\ 0 & \text{if } v \leq v_L \end{cases} \quad (3)$$

当  $v < v_L$  时, 认为算法过慢, 为无效算法; 当  $v_L \leq v \leq v_H$  时, 算法性能随着速度的增加线性提高; 当  $v > v_H$  时, 认为算法处理速度已超过了需求的处理速度。

### 1.3 算法的硬件可实现性

好的算法同时也应该适合于硬件的实现<sup>[11]</sup>, 而很多算法在设计时没有考虑到其硬件实现的问题, 没有考虑算法所需要的硬件资源情况以及现阶段的可实现性问题, 使得这些算法只能停留在研究层面。因此, 文中提出把算法的硬件可实现性也作为评价算法的一个重要的依据。

对于一个算法, 其越容易通过硬件来实现, 对硬件资源的需求越少, 则认为该算法的可实现性越好。与前两者不同的是, 算法硬件的可实现性为定性指标, 因此文中采用 5 个等级来量化算法的可实现性, 用  $\delta$  表示,  $\delta$  的值越大说明其硬件可实现性越好, 其取值情况如表 1 所示<sup>[12]</sup>。

表 1  $\delta$  的取值与等级的对应关系

等级	1	2	3	4	5
$\delta$	0.1	0.3	0.5	0.7	0.9

其中, 硬件可实现性的权值可以通过算法实现对硬件资源的需求、硬件开发难易程度等方面来综合衡量, 通过评价者的经验或相关专家讨论来确定。也可以把硬件可实现性分成更详细的下级指标参数, 由评价者或者组织各领域专家进行打分确定, 文中对此不做更深入的研究。

## 2 算法性能的综合计算

通过前面的论述可知, 对于小目标跟踪算法评价的内容如图 1 所示。

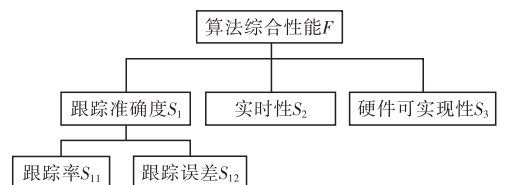


图 1 跟踪算法性能评价内容

在确定各个指标后, 需要按照实际的需求情况,

对各个指标进行加权综合,形成对跟踪算法性能评价的综合指标。常用的可用于指标加权的方法有 AHP 法、模糊综合评价法等,而从跟踪算法评价的内容可以看到评价指标体系具有明显的层次性,而且涉及的指标参数较少,因此文中采用 AHP 法来计算综合指标。

AHP 法是美国匹兹堡大学运筹学专家 T. L. Saaty 于 20 世纪 70 年代提出的一种系统分析方法。1982 年天津大学许树柏等将该方法引入我国,研究内容主要集中在判断矩阵、比例标度、一致性问题、可信度上。AHP 法是一种实用的多准则决策方法,该方法以其定性与定量相结合处理各种决策因素的特点,以及系统、灵活、简洁的优点,在我国得到了广泛的应用<sup>[12]</sup>。

### 2.1 构造权重判断矩阵

根据 AHP 法,首先需要构造权重判断矩阵,以区分各个指标的权重。在构造判断矩阵时,要根据实际情况选择合适的标度法,由决策者直接通过对指标两两比较得到,常见的标度法有 1 ~ 9 标度法,9/9 ~ 9/1 分数标度法、90/9 ~ 99/9 指数标度法等。如表 2 为应用比较多的 1 ~ 9 标度法<sup>[12-13]</sup>。

表 2 1 ~ 9 标度法等量级化表

值	意义
1	指标 1 与指标 2 同等重要
3	指标 1 略微比指标 2 重要
5	指标 1 明显比指标 2 重要
7	指标 1 当然比指标 2 重要
9	指标 1 绝对比指标 2 重要

Note:2,4,6,8 是可以使用的其他中间值;

若指标 1 不如指标 2 重要,则取值为 1/u, u 为 1 ~ 9

如对于第一层指标,假设其判断矩阵为 A,则其形式如下:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

式中,  $a_{ij}$  表示指标  $a_i$  相对于指标  $a_j$  的相对权重,指标  $a_1$ 、 $a_2$ 、 $a_3$  分别对应于跟踪准确度、实时性和硬件可实现性。

### 2.2 指标权重计算

根据 AHP 法,得到判断矩阵后,对指标权重的计算分 3 个步骤,以第一层指标为例。

第一步,计算判断矩阵 A 的每一行元素的乘积:

$$M_i = \prod_{j=1}^n a_{i,j} = \prod_{j=1}^3 a_{i,j}, \quad i = 1, 2, 3 \quad (4)$$

式中  $n$  为判断矩阵 A 的维数。

第二步,计算  $M_i$  的  $n$  次方根:

$$\bar{\omega}_i = (M_i)^{1/n} = (M_i)^{1/3}, \quad i = 1, 2, 3 \quad (5)$$

第三步,对  $\bar{\omega}_i$  进行归一化处理:

$$\omega_i = \bar{\omega}_i / \sum_{i=1}^n \bar{\omega}_i = \bar{\omega}_i / \sum_{i=1}^3 \bar{\omega}_i, \quad i = 1, 2, 3 \quad (6)$$

这样就得到了所需求的权向量  $\omega = [\omega_1, \omega_2, \omega_3]$ ,  $\omega_1$ 、 $\omega_2$ 、 $\omega_3$  分别为跟踪准确度、实时性和硬件可实现性。

### 2.3 一致性检验

对于判断矩阵,由于人的主观度量可能存在一定的偏差,因此为了提高权重评价的可靠性,最后需要对判断矩阵做一致性检验。

判断矩阵的最大特征值  $\lambda_{\max}$  为:

$$\lambda_{\max} = \sum_{i=1}^3 \frac{[A\omega]_i}{3\omega_i} \quad (7)$$

式中,  $[A\omega]_i$  为  $A\omega$  向量中的第  $i$  个元素。

一致性检验的公式为:

$$CI = \frac{\lambda_{\max} - n}{n - 1} = \frac{\lambda_{\max} - 3}{2} \quad (8)$$

$$CR = \frac{CI}{RI} \quad (9)$$

其中:  $CI$  为一致性指标,  $CR$  为一致性比例,  $RI$  为修正因子。通常情况下,当  $CR < 0.1$  时,认为该判断矩阵满足一致性要求,对于维数小于 3 的不需要进行判断。 $RI$  的取值和维数相关,其取值和维数的对应关系如表 3 所示<sup>[12]</sup>。

表 3 RI 的取值表

dimension	RI	dimension	RI	dimension	RI
1	0	4	0.96	7	1.32
2	0	5	1.12	8	1.41
3	0.58	6	1.24	9	1.45

### 2.4 指标的规范化

对上述的用于评价的各种指标,其量纲不同,各指标值的量级也不一样,很可能造成评价的不合理性,因此在进行综合评价前,有必要把各指标值转换为可以综合处理的量化值,一般都规范到  $[0, 1]$  的范围内。

跟踪率  $S_{11} = \alpha = n/N$  和硬件可实现性  $S_3 = \delta$  的值已经在  $[0, 1]$  内,且为无量纲的量,满足规范要求;跟踪实时性  $S_2$ , 通过  $S_2 = \eta/v_H$  来规范;跟踪误差  $S_{12}$  属于成本型指标,即误差越小,跟踪效果越好,根据成本型指标的线性变换方法<sup>[12]</sup>,首先得到最小误差  $E_{\min}(\mu)$ ,然后通过线性变换  $S_{12} = \frac{n}{N} \frac{E_{\min}(\mu)}{E(\mu)}$  把跟踪

误差的值限定在  $[0, 1]$  的范围内,而理论上的最小误差值为零,因此  $E_{\min}(\mu)$  必须为一个非零值,对于小于该值的误差按照  $E_{\min}(\mu)$  处理。

### 2.5 计算综合评价分数

在获得各指标的权重后,通过与对应评价值的乘积就可以得到最后的综合评价分数。对于不同的跟踪算法,得分最高者性能好。计算算子为:

$$F = [\omega_1, \omega_2, \omega_3] [S_1, S_2, S_3]^T \quad (10)$$

$S_1, S_2, S_3$  分别为跟踪准确度、实时性和硬件可实现性的评价分数。其中跟踪准确度的分数  $S_1$  同样可以通过 AHP 法由跟踪率和跟踪误差获得,即  $S_1 = [\omega_{11}, \omega_{12}] [S_{11}, S_{12}]^T$ 。

通过前面的讨论可以得到基于 AHP 的小目标跟踪算法的评价流程如图 2 所示。

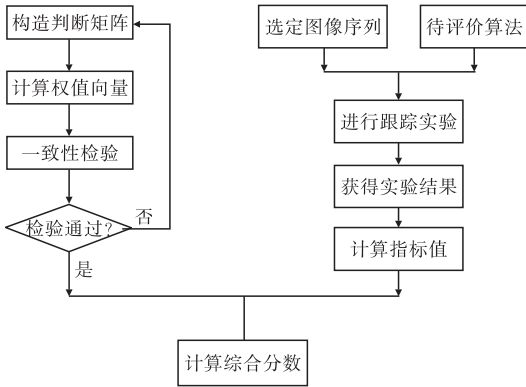


图 2 基于 AHP 的小目标跟踪算法的评价流程图

### 3 实验验证

为了验证文中方法的合理性以及对一些参数的确定和优化,文中采用流行的几种跟踪算法进行了实验,并通过上述评价方法来进行跟踪算法的性能评价,得到各个算法的综合评价分数。

在实验之前,文中通过综合考虑构造了判断矩阵,用于实验后的性能讨论,由于该评价系统中涉及的指标参数少,容易进行两两比较判断,因此文中采用了 1 ~ 9 标度法来构造权重判断矩阵。对于跟踪率和跟踪误差采用的判断矩阵为  $A_1$ ,对于跟踪准确度、实时性和硬件可实现性的判断矩阵为  $A_2$ ,则其取值如下:

$$A_1 = \begin{bmatrix} 1 & 2 & 4 \\ 1/3 & 1 & 2 \\ 1/4 & 1/2 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 2 & 4 \\ 1/2 & 1 & 2 \\ 1/4 & 1/2 & 1 \end{bmatrix}$$

求得权向量分别为  $[\omega_{11}, \omega_{12}] = [0.75, 0.25]$ ;  $[\omega_1, \omega_2, \omega_3] = [0.57, 0.29, 0.14]$ 。对于  $A_2$ , 其一致

性比例  $RI = 0$ ,符合一致性要求。

由于文中采用的跟踪算法是针对红外小目标的跟踪的,认为当跟踪误差超过 3 个像素时就认为目标开始丢失,并且取  $E_{\min}(\mu)$  取 0.2(个像素);算法处理速度的上限设为 50fps,下限设为 1fps。文中实验总共包含 3 组图像序列和 4 种不同的算法。图 3 为参与实验的 3 组图像序列,其中小目标用白色矩形框标识。

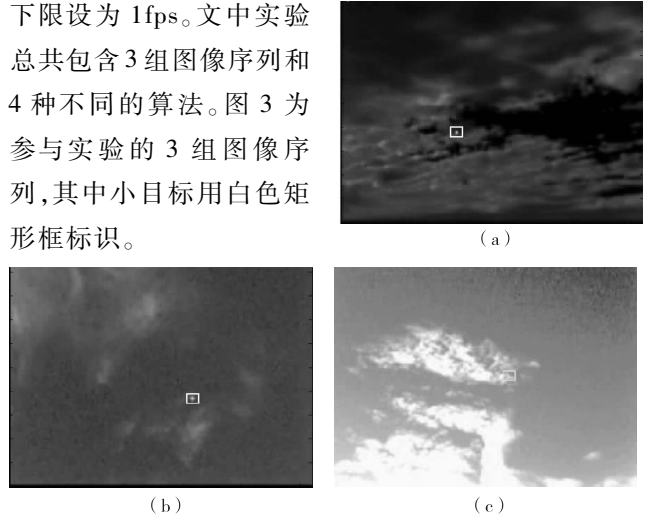


图 3 参与实验的 3 组图像序列

表 4 ~ 表 6 为分别采用粒子滤波算法<sup>[14]</sup>、均值位移算法<sup>[15]</sup>、自适应阈值分割算法<sup>[16]</sup>和基于数学形态学的小目标跟踪算法进行实验得到的数据;表 7 ~ 表 9 为由各个实验得到的数据根据本章提出的评价方法计算得到的指标分数。

表 4 采用图像序列 1 实验后得到的数据

实验结果	粒子滤波 算法	均值位移 算法	自适应阈值 分割算法	数学形态 学算法
$E(\mu)/\text{pixel}$	1.8253	2.159	0.134	0.4732
$n/N$	400/400	124/400	400/400	400/400
$v/\text{fps}$	29.7	6.1	28.2	31.3
$\delta$	0.5	0.5	0.9	0.7

表 5 采用图像序列 2 实验后得到的数据

实验结果	粒子滤波 算法	均值位移 算法	自适应阈值 分割算法	数学形态 学算法
$E(\mu)/\text{pixel}$	0.5809	1.6924	0.007	0.4298
$n/N$	180/180	180/180	180/180	180/180
$v/\text{fps}$	31.2	5.9	28.0	31.4
$\delta$	0.5	0.5	0.9	0.7

表 6 采用图像序列 3 实验后得到的数据

实验结果	粒子滤波 算法	均值位移 算法	自适应阈值 分割算法	数学形态 学算法
$E(\mu)/\text{pixel}$	1.9773	2.454	0.3037	0.6849
$n/N$	115/400	90/400	99/400	400/400
$v/\text{fps}$	30.8	5.4	28.3	31.3
$\delta$	0.5	0.5	0.9	0.7

表 7 采用图像序列 1 实验数据计算得到的指标分数

性能指标	粒子滤波	均值位移	自适应阈值	数学形态
	算法	算法	分割算法	学算法
$S_{11}$	100%	31%	100%	100%
$S_{12}$	0.1096	0.0287	1	0.4227
$S_1$	0.7774	0.2397	1	0.8557
$S_2$	0.5940	0.1200	0.5640	0.6260
$S_3$	0.5	0.5	0.9	0.7
$F$	0.6854	0.2414	0.8596	0.7673

表 8 采用图像序列 2 实验数据计算得到的指标分数

性能指标	粒子滤波	均值位移	自适应阈值	数学形态
	算法	算法	分割算法	学算法
$S_{11}$	100%	100%	100%	100%
$S_{12}$	0.3443	0.1182	1	0.4653
$S_1$	0.8361	0.7795	1	0.8663
$S_2$	0.6240	0.1080	0.5600	0.6280
$S_3$	0.5	0.5	0.9	0.7
$F$	0.7275	0.5456	0.8584	0.7739

表 9 采用图像序列 3 实验数据计算得到的指标分数

性能指标	粒子滤波	均值位移	自适应阈值	数学形态
	算法	算法	分割算法	学算法
$S_{11}$	28.75%	22.50%	24.75%	100%
$S_{12}$	0.0291	0.0183	0.1630	0.2920
$S_1$	0.2229	0.1733	0.2264	0.8230
$S_2$	0.6160	0.1080	0.5660	0.6260
$S_3$	0.5	0.5	0.9	0.7
$F$	0.3757	0.2001	0.4192	0.7486

从计算结果可以看到,对于图像序列 1 和图像序列 2,由于目标所在的位置背景比较简单,信噪比较高,因此除均值位移算法跟踪效果相对较差外,其它的算法都能实现较好的跟踪。而对于图像序列 3,由于背景复杂,信噪比低,只有数学形态学算法实现了对小目标较好的跟踪,表现出来优良的跟踪能力。

### 4 小结

通过实验的计算结果可以看到,采用文中提出的评价方法通过把跟踪算法的各指标进行综合,实现对跟踪算法的性能评价。评价的计算方法灵活、合理,对跟踪算法的评价问题,尤其是对于单凭一项指标无法做出判断的算法的评价,具有较好的参考意义。

#### 参考文献:

[1] 高陈强,张天骐,李强,等. 几种典型红外弱小目标检测算法的性能评估[J]. 重庆邮电大学学报:自然科学版, 2010,22(3):386-391.

[2] 周进,吴钦章. 弱小目标跟踪算法性能评估的研究[J].

光电工程,2007,34(1):19-22.

[3] 李鹏飞,陈朝武,李晓峰. 智能视频算法评估综述[J]. 计算机辅助设计与图形学学报,2010,22(2):352-360.

[4] Andrew Cilia. Object tracking using cluster of elastically linked feature trackers[J]. Journal of Electronic Imaging 2008,17(2):023019-1-023019-11.

[5] Ruiming Liu, Erqi Liu, Jie Yang, et al. Automatically detect and track infrared small targets with kernel Fukunaga-Koontz transform and Kalman prediction[J]. APPLIED OPTICS, 2007 46(31):7780-7791.

[6] 袁广林,薛模根,韩裕生,等. 基于自适应多特征融合的 mean shift 目标跟踪[J]. 计算机研究与发展,2010,47(9):1663-1671.

[7] 王相海,方玲玲,丛志环,等. 卡尔曼粒子滤波的视频车辆跟踪算法研究[J]. 中国图象图形学报,2010,15(11):1615-1622.

[8] Cui-yun Li, Hong-bing Ji. Marginalized particle filter based track-before-detect algorithm for small dim infrared target [C]//IEEE. AICI 2009:321-325.

[9] 刘献如,蔡自兴,唐王进. 基于 SAD 与 UKF-Mean Shift 的主动目标跟踪[J]. 模式识别与人工智能,2010,23(5):646-652.

[10] 艾斯卡尔·艾木都拉,王保柱. 恒虚警率 PDAF 的弱特点目标跟踪技术性能分析[J]. 计算机工程与应用,2009,45(3):168-171.

[11] HU Tao-tao, FAN Xiang, Zhang Yujin, et al. Infrared small target tracking based on SOPC [C]//SPIE. Parallel Processing for Imaging Applications, 2011, Vol. 7872:78720U-1-78720U-9.

[12] 张杰,唐宏,苏凯,等. 效能评估方法研究[M]. 北京:国防工业出版社,2009.

[13] Lyle M, Dawley, Lenore A Marentette, Long A M. Developing a decision model for joint improvised explosive device defeat organization (JIEDDO) proposal selection, ADA 484369[R]. Department of The Air Force Air University, 2008.

[14] 任彪,樊祥,马东辉. 基于多特征融合与粒子滤波的红外弱小目标跟踪方法[J]. 弹箭与制导学报,2009,29(5):304-307.

[15] Dorin Comaniciu, Visvanathan Ramesh. Mean shift and optical prediction for efficient object tracking [C]//IEEE International Conference on Image Processing, 2000:70-73.

[16] 薛丰廷,彭鼎祥. 红外跟踪系统中的自适应阈值分割[J]. 激光与红外,2008,38(4):386-388.