大型电力网络数据库故障数据关联查询模型

大型电力网络数据库聚集和存储大量的分布式电力电网管理系统数据，常因为通信冲突和电网节点损耗产生故障数据，对故障进行有效的查询是实现电力数据调度和故障诊断的关键。传统的大型电力网络数据库故障数据查询模型采用平均互信息关联维提取算法，当关联维特征出现独立同分布状态时，对故障数据提取性能受限，查询准确度不高。提出一种基于差分累积函数特征挖掘的大型电力网络数据库故障数据查询算法。构建大型电力网络数据库的实体模型，分析故障数据产生的机理，提取故障数据的信号模型特征，基于相空间重构方法构建高维相空间。在高维相空间中，挖掘故障数据的差分累积函数特征，实现对故障数据查询算法改进。仿真实验表明，采用该模型进行大型电力网络数据库故障数据查询，对电网数据库的故障数据分类性能较好，提高对电力网络的故障诊断和检测能力。

王昌辉/贵州建设职业技术学院
罗兴春/贵州电网公司培训与评价中心

王昌辉/副教授

关键词/Keywords
差分累积函数·
特征挖掘·
数据库·
电力·

随着Web技术的不断发展，大型电力电网系统逐渐构建和完善电力数据库平台，实现电力和电网信息的智能管理。电力系统作为一种最重要的基础性设施，关乎人民生产生活的稳定，对电力数据进行智能网络数据库集成和管理，可以实现对大型电力数据的优化调度，对维护电力系统的稳定，确保供电的安全具有重要意义。大型电力网络数据库是通过局域网提供给Web查询接口，通过跨平台访问这些电力数据。随着时间和信息的累积，这些数据库蕴含的信息量倍级增长，对大型电力网络数据库的查询和管理成为难题。大型电力网络数据库聚集和存储大量的分布式电力电网管理系统数据，常因为通信冲突和电网节点损耗产生故障数据，对故障进行有效的查询是实现电力数据调度和故障诊断的关键。研究大型电力网络数据库的故障数据优化查询方法受到人们的广泛关注[1]。

传统方法中，对大型电力网络数据库的故障数据查询方法主要有信号处理方法、人工智能调度方法、模式识别方法和数据特征提取方法等。其中，数据特征提取算法较为典型，通过提取电力网络数据库的故障数据特征，实现对故障数据的信号特征分析和查询，达到电力网络故障识别和诊断的目的。这一过程设计需要对故障数据进行采集和分析，对大型电力网络数据库的故障数据的实时采集主要有时域分析故障数据采集法、频域分析故障数据采集法、统计分析故障数据采集法和信息论分析故障数据采集法等[2-4]。但综合分析传统方法可见，传统的大型电力网络数据库的故障数据查询模型受到非关联特征干扰较为严重，自适应态协调能力较差，对低信噪比的故障数据检测性能不好。例如在文献[5]中的大型电力网络数据库故障数据查询模型采用平均互信息关联维提取算法，当关联维特征出现独立同分布状态时，对故障数据提取性能受限，查询准确度不高[6-10]。

针对上述问题，本文对传统的方法的弊端进行改进，提出一种基于差分累积函数管理特征挖掘的大型电力网络数据库故障数据查询算法。首先构建大型电力网络数据库的实体模型，分析故障数据产生的机理，提取故障数据的信号模型特征，基于相空间重构方法构建高维相空间。在高维相空间中，提取电力

中国知网 https://www.cnki.net

网络数据库的故障数据的差分累积函数特征，实现对故障数据查询算法改进。最后进行故障数据的查询和分类仿真实验，展示了本文算法的优越性能。

大型电力网络数据库实体模型构建及故障数据产生机理

1. 大型电力网络数据库的实体模型构建

为了实现大型电力网络数据库故障数据查询，需要首先构建大型电力网络数据库的实体模型。大型电力网络数据库访问的主要步骤分为重采样、权值更新和数据访问三部分，大型电力网络数据库的实体模型构建具体实现描述如下：令R为大型电力网络数据库中包含有四元组（E_i，E_j，d，t）的信任关系，电力数据分类属性$A=\{A_1, A_2, \cdots, A_m\}$，在对电力数据最近时刻查询获得的信任值取最大值1，则间隔时间越长的信任值对当前的影响越小，计算时间减弱函数为

$$f(k)=\begin{cases} f(k-1)-\dfrac{1}{n} & 1\leqslant k<n \\ 1 & k=n \end{cases} \tag{1}$$

需要构建电力数据权重分配机制，进行分类属性权重评估，对于关系数据四元组（E_i，E_j，d，t），进行属性值IDF权重定义为

$$IDF_i(v)=\lg[n/F_i(v)] \tag{2}$$

式中，n为电力数据的元组数；$F_i(v)$为电力数据集中满足差异权重属性分配的元组数。假设$k-1$时刻的大型电力网络数据库集成的电网通信信息状态方程为

$$P_i(t)=\sum_{n=1}^{N}\frac{A}{r}e^{-\mathrm{j}kr}R_{in}\frac{1}{r}e^{-\mathrm{j}kr} \tag{3}$$

$$P_i(t)=\frac{A}{r^2}\sum_{n=1}^{N}e^{-\mathrm{j}2kr}a_{in}e^{\mathrm{j}\Psi_{in}} \tag{4}$$

式中，A为电力网络数据库查询指向性信号的信号强度；r为MA长度；k为波数；其中$k=2\pi f/c$；f为信号的瞬时频率；c为指向性数据库传递权重；得到$k-1$时刻的大型电力网络数据库集成的状态后验分布的粒子集为$\{x_{k-1}^i, w_{k-1}^i\}$。进行数值属性权重评估，假设B_j为电网系统各存储节点反馈权系数，电网用户客户端采用Send Chunk控制电力数据库存储系统中的数据故障特征，得到大型电力网络数据库查询指向性信息特征模型。其中，对大型电力网络数据库分层特征挖掘准确度概率函数密度表达式为

$$p(y|\alpha,\theta)=\sum_{k=1}^{K}\alpha_k p_k(y|\mu_k,\sum\nolimits_k) \tag{5}$$

式中，$\delta x_i=x_j-x_i$，x_j是x_i近邻的一组指数分离相量，呈现在多维矩阵构建的指数状态空间中，控制状态访问指向性分布，得到对大型电力网络数据库数据对象的层析集成粒度为

$$a_k=(\sum\nolimits_k+\varepsilon U)^{-1}(\frac{1}{|w|}\sum_{i\in w_k}I_i p_i-u_k\bar{p}_k) \tag{6}$$

式中，I_i为数据对象的分层数据；u_k为数据对象功率因数。

采用分布式代码执行方式，将任务代码分布到多个数据节点，实现分布式执行。执行完后，结果统一返回到客户端，从而有效避免大型电力网络数据库中的缓存数据的移动与传输。通过上述分析，实现了对大型电力网络数据库的实体模型构建，为进行故障特征分析和故障数据查询提供模型基础。

2. 大型电力网络数据库故障数据的产生及预处理

对大型电力网络数据库的智能故障诊断的第一步是进行故障数据的模型构建，大型电力网络数据库故障数据的产生主要来源是因为通信冲突和电网节点损耗产生故障数据，对故障进行有效的查询是实现电力数据调度和故障诊断的关键。在电网系统中，数据采用分布式激振方式存储于数据库中，进行状态估计，得到数据库访问输出为

$$\hat{x}_k=\sum_{i=1}^{N}w_k^i x_k^i \tag{7}$$

式中，w_k^i为第k个项目下的第i个故障特征；x_k^i为故障数据。

采用相空间重构和非线性时间序列分析方法分析大型电力网络数据库的故障特征，设大型电力网络数据库振动信号时间序列为$\{x_1, x_2, \cdots, x_N\}$，当嵌入维数为$m$，延迟时间间隔为$\tau$，故障数

中国知网 https://www.cnki.net

据的后验概率估计为$p(x_0)$，得到

$$P_{ij}(k)=\frac{[l_j(k)-l_i(k)]\eta_{ij}(k)}{\sum_{j\in N_i(k)}[l_j(k)-l_i(k)]\eta_{ij}(k)} \quad (8)$$

式中，$l_j(k)$为编号为j的第k个数据的故障概率；$n_{ij}(k)$为均衡概率。

考虑大型电力网络数据库故障特征相关性参数，其中

$$j\in N_i(k),N_i(k)=\{\|x_j(k)-x_i(k)\|<r_d(k)\} \quad (9)$$

采用轨迹协方差矩阵X^TX的特征值作为该观测时间序列的特征，通过预获取有限电力信息，推导出实体关联知识库，利用表征关联知识有效指导电力数据资源消重和整合，得到故障数据提取的简化公式为

$$E_{q\times r}=Y_{q\times u}W_{u\times u}Y_{u\times r}^{U} \quad (10)$$

式中，$Y_{q\times u}$为故障维度矩阵；$W_{u\times u}$为故障均衡概率；大型电力网络数据库的故障数据形成m维空间。从分析可知，上述大型电力网络数据库故障特征相关数据的数据量比较大，计算的复杂度比较高，无法对大型电力网络数据库故障征兆进行及时的挖掘。因此，需要去除其中的冗余数据和非关联数据的干扰，从而降低计算复杂度。

差分累积函数特征挖掘及故障数据查询模型改进实现

1. 故障数据差分累积函数特征挖掘算法

在上述进行数据库实体模型构建和故障数据特征分析预处理的基础上，进行大型电力网络数据库故障数据查询模型改进实现。从上述分析可知，传统的大型电力网络数据库故障数据查询模型采用平均互信息关联维提取算法，当关联维特征出现独立同分布状态时，查询滞后，准确度不高，提出一种基于差分累积函数特征挖掘的大型电力网络数据库故障数据查询算法。基于相空间重构方法构建高维相空间，基于F. Takens的嵌入定理，故障数据的序列$\{x_n\}_{n=1}^{N}$以如下向量在相空间中形成新的映射

$$x_n=[x_n,x_{n-\tau},\cdots,x_{n-(m-1)\tau}] \quad (11)$$

式中，n为输入的数据集；τ为对变量的序列数。

大型电力网络数据库故障数据的相空间重构轨迹为

$$X=[x(t_0),x(t_0+\Delta t),\cdots,x(t_0+(K-1)\Delta t)] \quad (12)$$

式中，$x(t)$为大型电力网络数据库故障数据嵌入的相空间中的一组空间状态矢量；m为嵌入维数；Δt为抽样时间间隔；$K=N-(m-1)J$。

在高维相空间中，挖掘故障数据的差分累积函数特征，可以得到数据库的差分累积函数特征

$$y[n]=\sum_{i=0}^{M}a_i x[n-i]+\sum_{i=1}^{M}b_i y[n-i] \quad (13)$$

式中，a_0，a_1，…，a_M为差分累积向量；$a_i(i=0,1,\cdots,M)$和$b_i(i=1,\cdots,M)$为差分累积函数的特征系数。根据差分累积函数的特征提取结果，利用表征关联知识有效指导电力数据资源消重和整合，在上述构建的实体模型中，进行故障数据关联查询和故障诊断，改进算法的设计思路如图1所示。

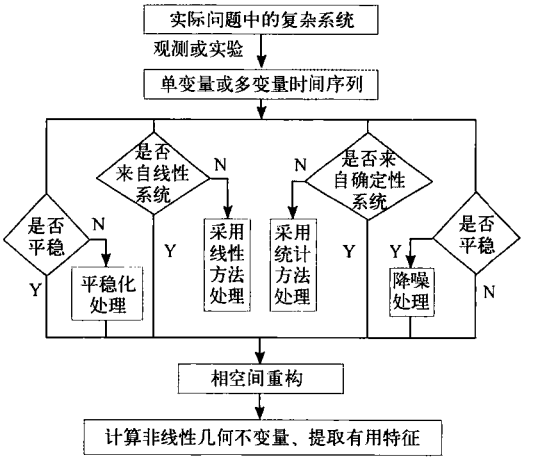


图1　大型电力网络数据看的故障数据查询表改进算法设计流程

2. 故障数据关联查询模型实现

基于差分累积函数特征挖掘和相空间重构，实现故障数据的关联查询，对故障特征矩阵$\boldsymbol{X}$进行奇异值（SVD）分解

$$\boldsymbol{X}=\boldsymbol{UDV}^{\mathrm{T}} \quad (14)$$

式中，$\boldsymbol{U}\in R^{m\times n}$正交矩阵，$\boldsymbol{V}\in R^{M\times N}$，且$\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{-1}$，$\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{V}^{-1}$；$\boldsymbol{D}\in R^{m\times M}$，且满足$\boldsymbol{D}=[\sum\ 0]$进行大型电力网络数据库的故障特征提取

$$E_n=\sum_{l=1}^{n}y_l\varphi_l \quad (15)$$

式中，y_l为故障数。

求出其l个特征值λ_1，λ_2，…，λ_l和特征向量矩

中国知网 https://www.cnki.net

阵$\boldsymbol{Y}=[y_1, y_2, \cdots, y_l]$，故障数据的关联特征估计$\hat{S}_w$为

$$\hat{S}_w=\sum_{i=1}^{c}p_i\frac{1}{n_i}\sum_{k=1}^{n_i}\left[\left(\vec{X}_k^{(i)}-\vec{m}_i\right)\left(\vec{X}_k^{(i)}-\vec{m}_i\right)^{\mathrm{T}}\right] \quad (16)$$

式中，X_k为对故障数的统计数；m_i为故障特征的最大值。

在关联数据查询中，查询跟踪轨迹对角线方向线段的关联相似点的确定性特征表示为

$$DET=\frac{Nl}{Nr}=\frac{\sum_{l=l_{\min}}^{N-1}lp(l)}{\sum_{i=1}^{N}\sum_{j=1}^{N}R_{ij}} \quad (17)$$

式中，l为跟踪轨迹对角线方向线段的长度；R_{ij}为查询中的特征数；Nl是包含在与查询跟踪轨迹对角线平行线段中的关联相似点的个数。大型电力网络数据库故障数据路径R_1为

$$R_1=\frac{\mu_w}{2\pi Kh}\frac{1}{m}\ln\frac{R_{eh}}{2(i+1)r_w} \quad (18)$$

式中，k为数据库故障中路径的个数；R_{eh}为故障数据路径。

通过上述分析，求出其l个特征值λ_1，λ_2，…，λ_l和特征向量矩阵$\boldsymbol{Y}=[y_1, y_2, \cdots, y_l]$，得到数据库查询查准系数为

$$R=\frac{\mu_w}{2\pi Kh}\left[\frac{1}{m}\ln\frac{R_{eh}}{2(i+1)r_w}+1.7\ln\frac{R_{eh}}{r_f}+12Z_f+25Z_f^2+\frac{\mu_o}{\mu_w}\ln\frac{r_w}{r_w'}\right] \quad (19)$$

通过上述设计，实现对大型电力网络数据库的故障数据关联查询算法改进。

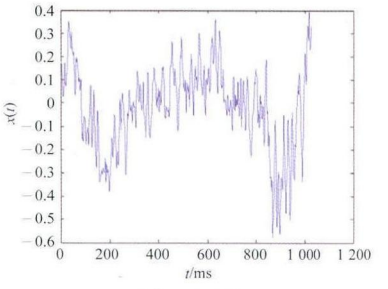
仿真实验及性能分析

为了测试本文算法在进行大型电力网络数据库故障数据采和关联查询中的性能，进行仿真实验。仿真实验的硬件环境描述如下：试验平台为通用计算机，CPU为Intel® Core™ i7－2600@3.40 GHz，内存为4×4 GB DDR3@1600 9－9－9－24，操作系统Windows7，开发工具为Matlab 7.0，并行处理使用OpenMP 2.0和MPICH NT 1.2.5，因此电力网络通信信号OFDM的物理层参数可参考采用IEEE 802.11a的标准，其他重要的仿真参数设定如下表所示。

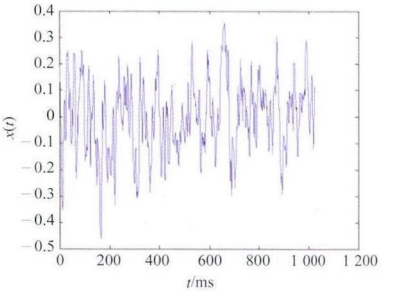
表 仿真参数设定

参数	参数值
工作频率/ MHz	1 800
系统带宽/ MHz	20
电网最大功率/kW	10 000
故障数据载波间隔	0.321 5 MHz（20 MHz/64）
有效符号周期	3.2 μs（1/0.312 5 MHz）

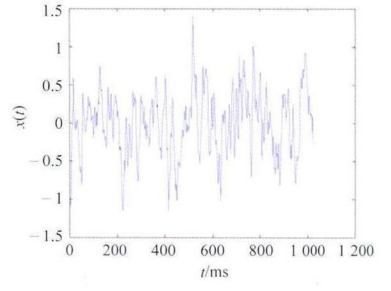
根据上述仿真环境和参数设定，构建大型网络数据库，故障数据信号采样频率为60 kHz，大型电力网络数据库中存储大量的分布式电力电网管理系统数据，故障数据产生来源与通信冲突和电网节点损耗，故障种类假设为三类，分别记为A、B和C类故障数据。首先进行故障信号模型构建，得到信号采样时域波形如图2所示。



（a）A类故障



（b）B类故障

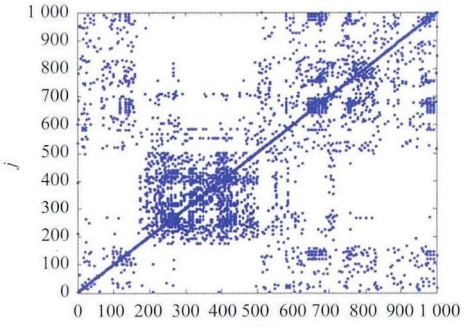


（c）C类故障

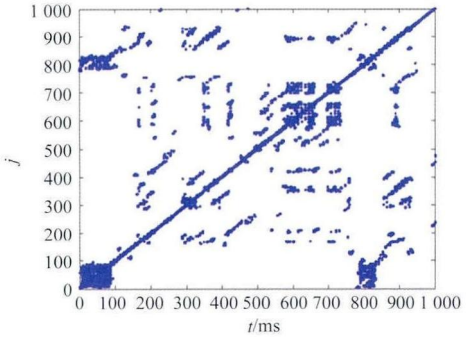
图2 大型电力网络数据库故障数据信号采样时域波形

中国知网 https://www.cnki.net

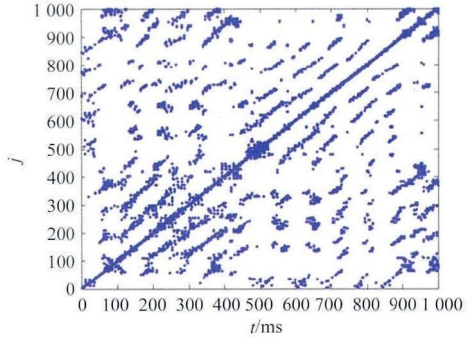
以其中一组样本为例，进行故障数据的特征分析，于相空间重构方法构建高维相空间，在高维相空间中，挖掘故障数据的差分累积函数特征，得到仿真结果如图3所示。



(a) A类故障



(b) B类故障



(c) C类故障

图3 高维相空间中的故障数据的差分累积函数特征提取结果

从图3可以得到高维相空间中故障数据的差分累积函数特征提取有规律性表现，点线分布具有规律性，在对角线周围有规律线段，通过该规律特征可以实现对故障数据管理查询挖掘，为实现故障诊断提供特征基础。进一步采用本文设计故障数据查询模型对电网数据库中故障的分类性能，实现最终的故障诊断，以1 024个故障样本数据为测试对象，进行故障数据分类查询，得到分类查询结果如图4所示。

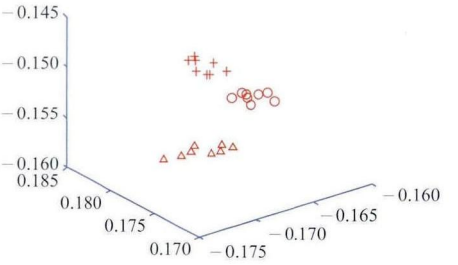


图4 故障数据分类查询

由图3可见这三类故障具有可分性，故障数据的聚类能力高，基于本文设计的大型电力网络数据库的故障数据关联查询模型，可以实现对大型电力网络数据库的故障准确诊断。

结束语

电力系统作为一种最重要的基础性设施，关乎人民生产生活的稳定，对电力数据进行智能网络数据库集成和管理，可以实现对大型电力数据的优化调度，对维护电力系统的稳定，确保供电的安全具有重要意义。大型电力网络数据库聚集和存储大量的分布式电力电网管理系统数据，常因为通信冲突和电网节点损耗产生故障数据，对故障进行有效的查询是实现电力数据调度和故障诊断的关键。本文提出一种基于差分累积函数特征挖掘的大型电力网络数据库故障数据查询算法。构建大型电力网络数据库的实体模型，分析故障数据产生的机理，提取故障数据的信号模型特征，基于相空间重构方法构建高维相空间。在高维相空间中，挖掘故障数据的差分累积函数特征，实现对故障数据查询算法改进。研究和实验分析得出，本文算法对电网数据库的故障数据分类查询性能较好，特征分析能力较好，提高对电力网络的故障诊断和检测能力，在电网数据库集成构建和故障诊断等领域具有较好的应用价值。

参考文献

[1] 孟祥福，严丽，马宗民，等．基于语义相似度的数据库自适应查询松弛方法 [J]．计算机学报，

中国知网 https://www.cnki.net

2011，34 (5)：812-824.

[2] 寇月，申德荣，李冬，等．一种基于语义及统计分析的Deep Web实体识别机制 [J]．软件学报，2008，19 (2)：194-208.

[3] 李红升．基于粒子群算法改进电力信息系统的安全研究 [J]．科技通报，2013，29 (4)：158-161.

[4] 张文亮，刘壮志，王明俊．智能电网的研究进展及发展趋势 [J]．电网技术，2009，33 (13)：1-11.

[5] 王晓波，樊纪元．电力调度中心统一数据平台的设计 [J]．电力系统自动化，2006，30 (22)：89-92.

[6] 蒋海波，王晓京，范明钰，等．基于水平纠删码的云存储数据布局方法 [J]．四川大学学报：工程科学版，2013，45 (2)：103-109.

[7] 覃雄派，王会举，李芙蓉，等．数据管理技术的新格局 [J]．软件学报，2013，24 (2)：175-197.

[8] 刘冬宁，汤庸，滕少华，等．基于时态数据库的极小子结构逻辑系统 [J]．计算机学报，2013，36 (8)：1592-1601.

[9] 谭鹏许，陈越，兰巨龙，等．用于云存储的安全容错编码 [J]．通信学报，2014，35 (3)：109-114.

[10] 柳虹，周根贵，傅培华．分层供应链复杂网络局部演化模型 [J]．计算机科学，2013，40 (2)：270-273.

（收稿日期：2014－10－20）EA

(上接第34页)

实现了一种基于嵌入式的大型建筑火情监控电气系统。系统采用CC2430芯片的无线传输模块将摄像机检测到的大型建筑火情数据，直接传给地面站，在计算机上显示及存档。通过图像采集和控制模块对摄像机检测到的大型建筑火情进行操作，将图像的处置结果，发送云台得到控制参数，实现机载云台的及时调整，通过S3C2440处理器的自动监控模块，帮助摄像机自动追踪对地目标。软件设计过程中，对基于背景差分的大型建筑火情监控算法进行了详细分析，并给出了大型建筑火情监控的程序改进代码。仿真实验结果说明，所设计系统可对大型建筑火情进行准确监控，并具有较高的准确度和效率。

参考文献

[1] 淳于江民，张晰．微型无人直升机技术研究现状与展望 [J]．机器人技术与应用，2004 (6)：6-11.

[2] 董国忠，王省书，胡春生．无人机的应用及发展趋势 [J]．国防科技，2006 (10)：34-38.

[3] 余静，游志胜．自动目标识别与跟踪技术研究综述 [J]．计算机应用研究，2005，22 (1)：12-15.

[4] 刘义海，张效民，赵延安，等．一种低功耗水下目标识别系统设计 [J]．计算机测量与控制，2012，20 (2)：470-473.

[5] 徐贵力，倪立学，程月华．基于合作目标和视觉的无人飞行器全天候自动着陆导引关键技术 [J]．航空学报，2008，29 (2)：437-442.

[6] 王林，彭辉，朱华勇，等．应用无人机跟踪地面目标最新研究进展 [J]．系统仿真学报，2011，22 (1)：172-177.

[7] 温佩芝．红外成像制导自动目标检测与识别方法研究 [D]．沈阳：中国科学院沈阳自动化研究所，2004：46-47.

[8] Yizong Cheng. Mean shift, mode seeking, and clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17 (8)：790-799.

[9] 刘松涛．红外成像制导舰船目标识别与跟踪实用技术研究 [D]．烟台：海军航空工程学院，2003：36-37.

[10] 李俊山，曹素平，谭园园．基于数学形态学的红外运动小目标识别算法研究 [J]．电光与控制，2006，13 (2)：11-14.

（收稿日期：2014－10－05）EA

中国知网 https://www.cnki.net