

# 基于数据挖掘的雷达探测目标误差测量技术

张彤

(成都理工大学 工程技术学院, 四川 乐山 614000)

**摘要:** 传统技术缺少对数据处理的部分, 导致测量误差较大, 为了避免该技术对雷达测量带来的弊端, 提出了基于数据挖掘的雷达探测目标误差测量技术研究; 在小波变换下对雷达采集到的图像进行去噪处理, 通过构建网络层识别模型进行图像切割、归一化处理, 由此设计具体识别流程; 为剔除冗余多余且空洞无用数据, 对其进行相应的转换、集成和匹配, 将数据处理成一致形式, 以此读取雷达数据文件, 完成数据预处理; 依靠决策树在 SQL Server 2005 中进行误差数据挖掘, 由此完成雷达探测目标误差测量; 假设 4 种雷达运动轨迹目标类型, 如此进行雷达仿真验证分析, 由实验结果可知, 该技术误差测量精度与实际值相差较小, 最大误差为 2.0 km, 在实际误差允许范围内, 为雷达精准探测奠定基础。

**关键词:** 数据挖掘; 雷达探测; 目标误差; 测量

## Data Mining Based Radar Detection Target Error Measurement Technology

Zhang Tong

(Engineering & Technical College, Chengdu University of Technology, Leshan 614000, China)

**Abstract:** Under the wavelet transform, the image collected by the radar is measured and denoised. The image is cut by establishing a network layer recognition model, normalized, and the specific identification process is transformed and designed. In order to eliminate scattered and redundant and empty and useless data, corresponding conversion, integration and matching are performed, and the data is processed in the decision tree for error data mining in SQL Server 2005 to complete the radar detection target error measurement. Assuming the target type of the radar moving trajectory, the radar simulation verification analysis is performed in this way. From the experimental results, it can be known that the technical error measurement accuracy is acceptable and the maximum error is 2.0 kilometers, within the actual error allowable range, it is extended for the accurate detection of radar basis.

**Keywords:** data mining; radar detection; target error; measurement

## 0 引言

随着雷达技术的不断发展与进步, 不同领域都将该技术进行了很好的应用, 如军事领域、农业畜牧等领域, 他们通过该技术得到了具有丰富信息且非常清晰的图像<sup>[1]</sup>。雷达图像中存在着许许多多的目标, 其中不乏一些复杂的, 识别这些复杂的图像是非常困难的, 雷达技术最基础的功能和应用就是对所需目标进行识别, 所以对如何利用雷达技术识别所需目标是一个备受瞩目的焦点<sup>[2]</sup>。目标的不同特性和变化多端的天气对雷达在进行探测时的效能有着非常大的影响, 雷达所探测的信息中夹杂着许多杂波信息, 很难得到具有一定影响价值的规律和模式, 也就导致了“数据丰富, 知识贫乏”的出现<sup>[3]</sup>。

雷达探测目标的误差测量技术不仅非常深奥复杂, 而且是一种具有系统性的工作, 主要是由“白箱”建模和“黑箱”建模两种基本的方法组成, 所谓“白箱”就是通过理论分析推导, 进而得出机理模型, 也就是对之前已经经过验证的知识进行推导; “黑箱”是, 通过测试雷达系统所输入/出的数据信息, 并作出相应的判断的方法。当今主要

是用传统的方法进行“黑箱”建模, 即数理统计的方法进行仿真雷达探测, 但是这种方式不能反映出真实的误差的特性, 而是只能进行误差总体上的规律的反映, 对整体测量技术的真实性有比较大的影响, 因为该方法得到的系统误差和标准差只是通过分析历史上存有的误差数据资料和对通过对比模仿得到的, 比对模型是高斯白噪声模型。数据挖掘是指对隐含且未知的数据进行提取的过程, 所选数据是随机的、数量庞大的且不清晰的, 但这些信息确认是潜在的且有用的, 所以提出了建立于该方法之上的雷达探测目标误差测量技术。

## 1 基于数据挖掘的雷达探测目标图像识别

数据挖掘技术的激光雷达图像识别思想为: 最开始要进行图像的去噪, 在小波变换下进行去噪处理, 接着提取图像的 Zernike 不变矩阵的特征进行, 最后建立深度的网络层进行识别模型<sup>[4]</sup>。

### 1.1 目标图像的特征提取

图像切割: 如果把具有极其复杂性的图像直接输入到分类器里面, 分类器会随之变得非常复杂。所以, 需要切割整个完整的图像, 切割步骤为: 1) 确定图像的阈值, 即从众多阈值中选择适合的数值; 2) 对图像做二值化处理和平滑处理<sup>[5]</sup>。从水平的方向开始扫描图像, 计算从第一行起的图像中白色像素的数量, 如果数量大于  $N$  (已设阈值)

收稿日期: 2020-02-28; 修回日期: 2020-04-07。

作者简介: 张彤(1975-), 男, 陕西宝鸡人, 硕士, 讲师, 主要从事计算机软件开发、大数据技术及应用方向的研究。

时，需要对此进行特殊的标记。若与它相邻的两行仍然出现这种情况，即白色像素的数量大于已设阈值，那么视为之前的标记为合理标记，否则的话，就要从头开始再次操作<sup>[6]</sup>。若出现合理的标记时，就要开始改为向下的扫描。接着计算以最后一行起的每行中白色像素的数量，如果数量大于  $N$ ，就要将这一整行进行特殊标记。然后接着做向上的扫描，若相邻两行仍是白色像素的数量大于  $N$ ，同上需要视之前的标注为合理的，且重新进行该操作<sup>[7]</sup>。若过程中有合理标记出现，就要停下水平方向的扫描。垂直方向同理，需在水平方向的扫描区进行<sup>[8]</sup>。

### 1.2 图像归一化

一般需要在分类器中输入图像之前对其进行归一化处理，包括位置、大小等的归一化处理。

1) 位置归一化。主要是将目标移动到所规定的位置，以达到位置偏差的消除目的。有下面两种比较简单的归一化方法：(1) 对目标的质心进行计算，随后将其移动到指定的位置，该方法被称为建立于质心之上的归一化方法；(2) 对目标的外边框进行计算，然后寻找重心，再将其移动到指定位置，该方法称之为建立于目标外边框之上的位置归一化方法<sup>[9]</sup>。

2) 大小归一化。进行变换大小不一的目标，使其大小相同，该变换过程为大小归一化。主要有下面两种方法：(1) 为了使目标符合规定，需对目标外框进行放大或者缩小，缩放过程要按照比例进行；(2) 根据目标在水平和垂直两个方向的分布进行大小归一化<sup>[10]</sup>。

### 1.3 图像识别

如何在向量机中进行目标图像特征的输入，并对其进行分类和预测，主要有下面 3 个步骤：

1) 建模，也就是向量机模型的建立。模型中的核函数为径向基函数。该函数在模型选择时由以下两个参数决定：惩罚因子和核参数。分类器的性质和预测能力受到该参数的直接影响<sup>[11]</sup>。

2) 学习，通过样本进行训练以对向量机提出支持。用于目标图像识别的支持向量机的学习过程如图 1 所示。

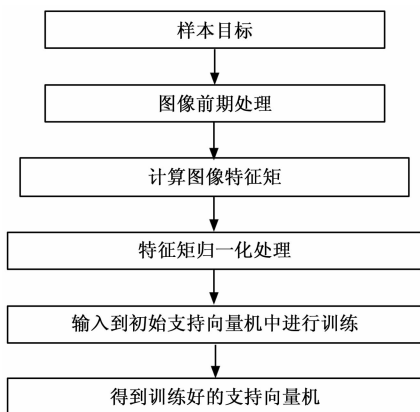


图 1 支持向量机的学习过程

像进行前期处理后，计算目标识别图像的特征矩，并对特征矩进行归一化处理，再输入到初始支持向量机中进行训练，得到训练后的支持向量机。

3) 预测，在向量机中输入未经分类的图像特征，并对此做出分类和预测。

根据上述内容，设计具体识别流程如图 2 所示。

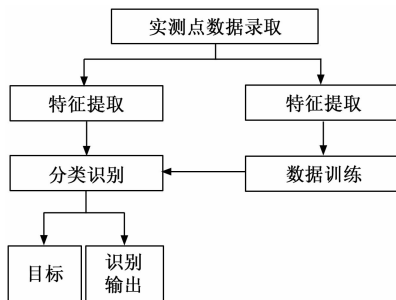


图 2 雷达图像的识别流程

由图 2 可知，雷达图像的识别流程为录入实际测量点的数据，分别提取实际测量点的数据特征，再对实际测量点的数据进行训练后，分类识别目标图像样本，并输出目标图像识别结果。

## 2 数据预处理

在进行数据挖掘时，必不可少的一步为数据的预处理。因为被挖掘的原始数据有着冗杂多余且空洞无用的特点，在进行数据挖掘时要对其进行加工，处理其中的噪声，这就是所谓的数据预处理<sup>[12]</sup>。数据预处理流程如图 3 所示。

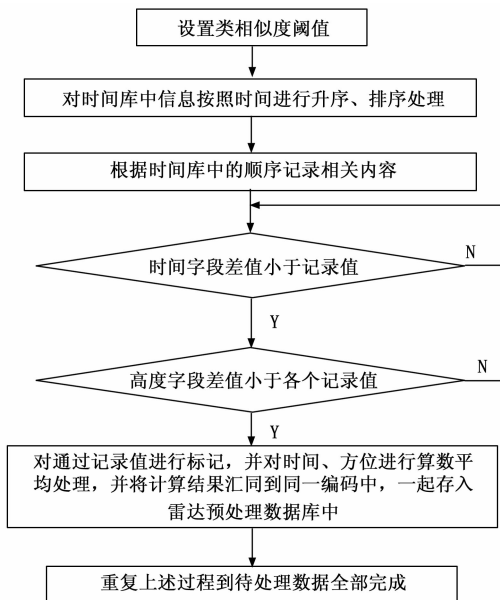


图 3 数据预处理流程

预处理的主要流程为：

1) 将所需处理的数据进行相应的转换、集成和匹配。要想实现数据挖掘，就要将目标雷达所得到的实际测量数据和真实航迹数据进行对比，并获得差值。但是上述两种数据在采集时的步长是不同的，并且被存放在不同的文件

由图 1 可知，通过目标图像识别样本，对目标识别图

中，所以对二者做航迹匹配处理，并且为了二者步长变得一样，还需要对数据进行三点插值处理。

2) 将数据处理成一致的。目的是保证挖掘结果的精确度和干净整洁，且具备一致性的特点。

3) 读取雷达数据文件，并将其装入数组之中，指针指向数组初始位置，数组是否为计算机保留的值？如果是，则结束程序；如果不是，则查看数据是否为模式数据？如果是，则需解码，并将结果存储到计算机指令之中，否则直接查看解码时间信息；将解码结果存储到模式数据库之中，查看解码时间信息，并将结果存储到时间表之中，解码方位信息，并将结果存储成 AZ 格式。不断重复上述步骤，直到获取所需数据。

### 3 雷达探测目标误差测量

雷达探测误差数据挖掘需要输入和预测属性，前者为真值距离和真值方位，后者为距离和方位误差，而误差模型的建立还要依靠决策树来完成，评估树的准确性还要依据保留测试法来实现。保留测试法对样本数据集进行划分，划分为训练和测试数据集，并且二者是互不相交的。

在 SQL Server 2005 中进行误差数据的挖掘，为了选择适合且可用的决策树算法，需要选择挖掘模型参数。决策树一即树形结构的流程图，属性的测试由内部节点表示，测试结果由分支表示，类别则由叶节点表示，根节点即树的最高层，如图 4 所示。

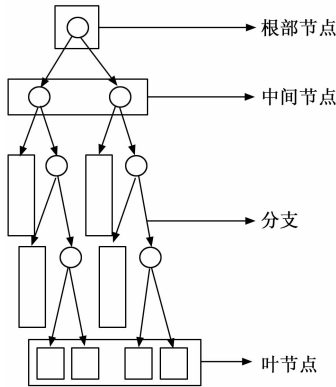


图 4 决策树树形结构图

为了形成决策树，需要进行训练样本子集的选择，若该树不能给出任何对象正确答案，则要将没有得到正确答案的特殊情况添加到该树中，一直重复上述操作，直到正确决定集的出现为止。完成一个符合要求决策树，即每个叶子都可以代表一个类别，每一个节点都代表一个属性，每一个分支都代表一个属性所对应可能值。

进行熵计算方法下的拆分数计算，和节点的拆分，拆分方法为二分法。通过以上操作，挖掘模型会分别生成距离和方位误差的决策树。通过对包含在叶节点中的最少的样本数的设定进行决策树的剪枝和修剪，并对其进行评估，评估需要根据测试数据集完成，确定了距离误差决策树在样本数最少时生成预测性能比较好。

基于数据挖掘的雷达探测目标误差测量技术实现流程如图 5 所示。

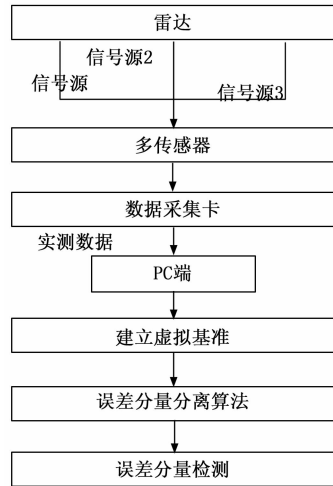


图 5 雷达探测目标误差测量流程

由于雷达探测目标通常是通过相关指令实现目标探测的，相应获取的雷达运行轨迹误差图像真实反应了从雷达指令发布的完整信息。通过信号源为被探测目标提供所需要的信号，利用多传感器融合相关数据参数，再通过数据采集卡，将模拟信号转化为数字信号，传输到 PC 端，以此建立雷达探测目标误差测量的虚拟基准。利用数据挖掘方法分离出各个误差分量，并且分析出产生误差的根本原因。根据探测结果，还可以优化雷达探测位置，提高测量精度。

### 4 雷达仿真结果验证

针对基于数据挖掘的雷达探测目标误差测量技术研究，进行仿真实验验证。

#### 4.1 数据挖掘处理

雷达信号中存在的孤立点需通过挖掘分离处理，使用数据挖掘技术在后期处理时主要通过以下两种方法得以应用。

1) 进行原始的雷达信号和已经被处理过且孤立点已经被删除的信号之间的比较，将其送到雷达终端，雷达终端会将目标点迹图显示出来，以达到有效性确定的目的。

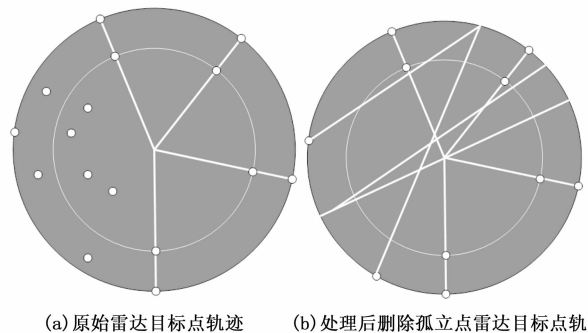


图 6 雷达目标点迹

从以上点迹图可以很明显地看出经过处理和未经处理的雷达信号的区别。两幅图由于是在不同的时间段采集的，会

有略微的区别，但是前后两幅图因采集数据的时段不同略有差异，但是可以看出原始点击图中显示的航线附近非常多的干扰点在处理过的图中得到了明显的解决，可以断定该处理有效过滤且去除了价目表的信号。

2) 进行提取孤立点信号，并将其送到终端显示器上，接着对其进行观察，确定其规律后，做出合理的处理。通过以上方法，提取了 4 种类型：

(1) 出现在 25 km 的半径范围内，该范围是以雷达站为圆心计算的，除此之外，25~50 km 之间偶尔也会有假目标出现。真假目标距雷达站有着相同的距离；假目标数量比较多，一般在 2~3 个，情况很糟糕时甚至会出现 4~5 个之多，分布为扇形，并且会出现在真目标的顺/逆时针方向。

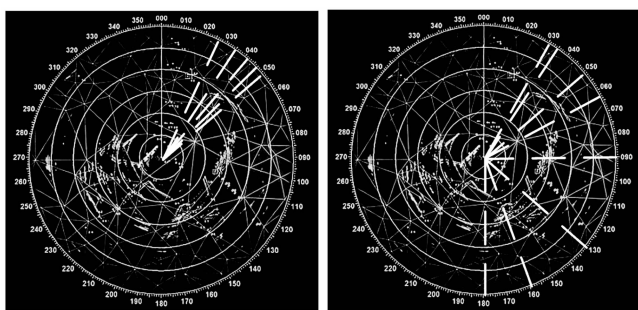
(2) 出现在半径 50~120 km 之内，出现更加频繁的位置是 65 km 和 100 km 附近，而且 30~50 km、120~180 km 的范围之间也会有假目标的出现。真假目标距雷达站有着相同的距离；多数情况下是一个，恶劣时会出现两个。若假目标在真目标的逆时针方向出现，从雷达矩阵引两条夹角为 40° 的射线到真目标和假目标上面。反之，如果假目标在真目标的顺时针方向出现，则所需射线的夹角为 30°。

(3) 出现在 270~280 km 方位一般情况下是一个假目标，不好的情况会出现两个，二者之间离的非常近，且夹角也比较小。

(4) 假目标很偶尔的会出现，且没有规律，不会长时间的出现。

#### 4.2 雷达运动轨迹研究及结果分析

根据上述 4 种假设目标类型，对雷达运动轨迹展开分析，如图 7 所示。



(a) 第一类

(b) 第二类

(c) 第三类

(d) 第四类

图 7 4 种假设目标类型雷达目标点轨迹

用图 7 所示的目标点轨迹，将传统方法与基于数据挖掘方法的雷达探测目标误差测量精度进行对比分析，结果如表 1

~4 所示。

表 1 第一类假设情况下两种方法测量精度对比结果

轨迹	实际情况/km	传统方法/km	基于数据挖掘方法/km
1	26.0	22.0	25.5
2	34.5	31.0	34.0
3	41.0	39.5	40.0
4	46.0	43.5	45.5
5	50.0	45.0	51.5

表 2 第二类假设情况下两种方法测量精度对比结果

轨迹	实际情况/km	传统方法/km	基于数据挖掘方法/km
1	30.0	26.0	30.5
2	35.0	30.0	33.5
3	51.0	45.0	49.0
4	60.0	65.5	59.0
5	90.0	89.5	90.8
6	130.0	125.5	128.0
7	160.0	176.8	159.5
8	180.0	186.2	180.9

表 3 第三类假设情况下两种方法测量精度对比结果

轨迹	实际情况/km	传统方法/km	基于数据挖掘方法/km
1	271.5	265.0	272.5
2	274.0	265.0	272.5
3	276.0	282.0	275.0
4	278.0	280.0	278.5

表 4 第四类假设情况下两种方法测量精度对比结果

轨迹	实际情况/km	传统方法/km	基于数据挖掘方法/km
1	78.0	75.0	78.4
2	145.0	140.2	144.5
3	233.0	235.8	232.3
4	254.5	250.0	255.8
5	273.5	269.5	276.9
6	315.5	310.0	314.8

由上述表内容可知，采用传统方法雷达探测目标误差测量精度与实际值相差较大，最大误差为 16.8 km；而基于数据挖掘方法雷达探测目标误差测量精度与实际值相差较小，最大误差为 2.0 km。

#### 5 结束语

建立于数据挖掘之上的雷达探测方法可以有效地分析和测量目标误差，提出了在这方面研究的新路径。在以后的研究中，为了能够得到更加具有真实性的雷达探测误差的结果，会对与误差探测有关的机制、环境等影响因素进行研究。

有几个问题在做研究的时候发现依然需要想出解决办法，

(下转第 24 页)