

文章编号:1673-8411(2018)03-0083-03

## 雷达回波特征向量应用初探

卓健,廖胜石,陈少斌,韩宇龙,奉意杰  
(崇左市气象局,崇左 532200)

**摘要:**在雷达产品相似度比较中提出雷达回波特征向量这一概念,将雷达回波产品用一组有限长度的数字进行描述,使得雷达回波相似度的比较转化为线性代数计算。通过受干扰回波自动识别,雷达回波的聚类应用展示雷达回波特征向量的应用前景。

**关键词:**雷达回波;特征向量;研究

**中图分类号:**P412.25

**文献标识码:**A

## A primary study on the application of radar echo eigenvectors

Zhou Jian, Liao Shengshi, Chen Shaobin, Han Yulong, Feng Yijie  
(Chongzuo Meteorological Bureau, Chongzuo Guangxi 532200)

**Abstract:** In radar product similarity comparison, the concept of radar echo eigenvector is proposed. The radar echo product is described by a set of finite-length numbers, which convert the comparison of radar echo similarity into a linear algebraic calculation. Through the automatic identification of disturbed echoes, the application of radar echo clustering shows the application prospects of radar echo eigenvectors.

**Keywords:** radar echo; eigenvector; search

### 1 引言

雷达资料大量应用于短时临近预报业务中,对天气雷达资料的应用已积累了大量经验,但是在雷达资料自动识别领域研究相对较少,使用的方法较复杂,不适用于对海量雷达数据进行处理。例如在雷达回波外推业务中,需要对外推结果进行评估,不少学者在这方面开展了研究。曾小团等<sup>[1]</sup>将外推得到的雷达产品与实况产品采用格点对格点逐一检验,计算出各点外推结果与实况的偏离程度。曹春燕等<sup>[2]</sup>方法基本相同,标准简化到是否出现回波。张蕾等<sup>[3]</sup>将出现阈值以上为活跃点,某点预报与实况均为活跃则记为正确。这一类方法我们称之为重叠区域比较法。重叠区域比较法的优点是定义明确,计算

结果清晰,对于大面积少变化的回波容易获得较高的评分。重叠区域比较法也有不便之处,例如对同一时次的同一块回波,在不同雷达的产品中由于作为观测者的雷达站与回波相对位置不同,若只根据产品图片,重叠区域比较法无法将其识别出来。对于不同过程,但是相似度非常高的回波,若仅有旋转角度的差异,重叠区域比较法即使能将其判别出来计算量也非常巨大。在查询雷达回波与历史过程的相似应用中,使用重叠区域比较法每次查阅比对都需要把要比对的历史数据重新读取计算一次。以上种种不便,需要我们引入一种高效、快速的计算手段。

对回波外推结果的判别和查找历史相似,都可以看作回波相似度的判别。在计算机对新闻的自动分类时<sup>[4]</sup>,根据稿件中主题词的使用频率作为特征向

收稿日期:2018-01-21

基金项目:广西区气象局面上项目(桂气科2017M22)资助

作者简介:卓健(1972-),男,广西宾阳人,大学本科,工程师,主要从事网络与雷达维护保障工作,E-mail:261817252@qq.com。

量,使用两个特征向量的夹角定量衡量它们之间的相似性,本文参照这一思路,根据雷达回波特点,提出雷达回波特征向量这一概念。

雷达回波特征向量使得两两雷达回波图像比较可以使用计算的方法进行,在业务实际运用中,这一方法可以用来寻找海量雷达回波产品中受干扰的样本,也可以在实时业务中用于自动监控,发现受干扰的雷达回波,降低人工劳动强度。

雷达回波特征向量还可以应用在雷达资料的聚类分析上,聚类分析是数据挖掘中解决其他问题的起点,特征分析与聚类技术在气象系统内多有研究<sup>[5-7]</sup>,但是将聚类分析应用在雷达资料应用领域则比较少见,雷达回波特征向量的提出,使雷达数据可以利用聚类分析方法划分成有意义和有用的簇,本文采用k-means算法进行雷达回波图像的聚类研究。

2 雷达回波特征向量

假设有如图1(见彩页)所示一组雷达组合反射率产品,对于源图(图1,见彩页)和用于比较的三幅对比图像(图1b,c,d,见彩页),人们可以轻易的区分出图1d有明显区别,但是图1b与图1c则很难区分出谁更接近源图。当有两个不同的方法外推出的结果分别是图1b和图1c,需要研究一种客观统一的标准评价标准,这种评价方法还要具有快速、节省空间的特点,以适应数据挖掘技术的需求。

2.1 雷达回波特征向量

根据经典Z-I关系公式 $Z=aI^b$ ,降水只与回波强度有关,所以不同强度的回波可以看成对降水贡献度不一致的因素,把每个体扫雷达产品中各个强度回波出现的次数分别统计并依次排列,就得到这个雷达产品的一个向量,我们用这个向量代表这个时次的回波,称为雷达回波的特征向量(Feature Vector),向量中每一个维度的大小代表这个等级强度的回波对降水贡献度。图1为雷达组合反射率产品(单位为dBZ),数据包括无回波,从-5dB开始每级间隔5dBZ一直到65dBZ及以上,共分为16级。较强的回波才是对天气变化的主要影响因素,所以我们只统计强度大于10dBZ的回波,由此可以得到组合反射率产品一个12维的向量。每一时次雷达回波都可以对应这样一个特征向量。当雷达回波用特征向量表示,就可以通过计算得到雷达回波相似度。

2.2 雷达回波特征向量夹角与尺度无关

假设这样一种情况,某次雷达产品回波面积只

有1个像素,回波强度为10dBZ,另一次同等强度回波面积为10像素×10像素,计算特征向量夹角,发现两块回波向量夹角余弦等于1,从这个例子可以看出特征向量夹角是一种无尺度的度量,它只是表明了要度量的两个目标组成成分的比例差异。虽然雷达回波特征向量夹角与尺度无关,但是本身的信息已经包含了部分尺度信息,如图形图像学中的面积,对于成分比例基本一致,但是尺度差异较大的回波,可以在不增加信息量的情况进行区分。

对于雷达回波外推比较这种应用,由于是进行短时间外推,可以默认为两个时次回波的尺度差异不大,所以对于这种应用直接使用特征向量距离进行判断。

3 雷达回波特征向量的应用

雷达产品受干扰产生突变或者雷达资料的聚类分析,都可以归类为雷达资料相似度的比较,雷达回波特征向量在这一应用场景可以得到发挥。

3.1 资料预处理

本文使用雷达组合反射率产品来源全国综合气象信息共享平台(CIMISS),使用气象数据统一服务接口(MUSIC: Meteorological Unified Service Interface Community)获取的RADA\_L3\_PUP\_CREF产品,选用南宁雷达站产品(站号Z9771),时间跨度从2016年8月至2017年9月,共57228个体扫文件。为后面计算方便,先对资料进行预处理,从开始时间逐个读取雷达组合反射率产品,按照特征向量定义将编号、雷达产品时间和12个特征向量维度统计值存入特征向量表(表1)。

表1 组合反射率回波特征向量

编号	时间	维度1	维度2	维度3	...	维度12
1	201608010100	2079	524	189	...	0
2	201608010106	2101	508	201	...	0
3	201608010111	2063	557	196	...	0
...	...	...	...	...	...	...
57228	201708080909	4237	2392	1017	...	0

3.2 有干扰回波图像的识别

雷达产品有些时候会受到各种干扰,比如表现为图片出现“麻点”,由于k-means算法对噪声点比较敏感,所以首先需要对噪声进行识别剔除,若需人工手段从海量的数据中找出受干扰的产品是很难实

现的问题,机器识别技术在这一方面就可以派上用场。

### 3.2.1 降维

人类比较容易理解二、三维的图形,我们将雷达回波特征向量的12个维度进行降维处理,转换为2个维度以便观察,用雷达扫描范围内笛卡尔坐标系上方最强回波反射率因子的合计值,以及这些回波占扫描范围的百分比,这可以看作使用反射率因子总量和面积比例对雷达回波产品的一种全局性描述方法。

### 3.2.2 结果分析

将特征转换得到的结果使用散点图表示(图略),发现右下红圈位置有一明显的离群点,根据编号查看对应的产品(图略),发现这一体扫时次的产品包含大量强度大于45dBZ的麻点,最强点回波强度达81dBZ,对比前后时次的同样产品都不存在如此强度的回波,再使用相邻时次向量余弦夹角时间序列变化观察,发现在这一时次向量余弦夹角出现突变(图略),其后恢复正常,查看强回波的分布形状散乱无序,综合以上因素,判断这张图片受到干扰,在下一步的试验中做剔除处理。从近6万张的产品图片找出仅有的一张受干扰图片,由于已提前做过数据预处理,用特征转换方法找出受干扰图片几乎不费时间雷达回波与降水的研究中,认为不同的天气系统造成的降雨使用的Z-I关系法的a,b系数不同<sup>[8,9]</sup>,但是如何自动识别出不同的天气系统是个难题,本文尝试研究一种方法,将回波分成几个大的类型,这种方法可以为机器识别雷达回波对应的天气系统打下基础。

### 3.2.3 k-means 聚类

虽然对于雷达回波与天气系统的关系已经有一定的研究,但是这些研究一般仅使用少量天气过程的样本,并没有建立起一套完善的雷达回波分类定义,所以很难使用KNN(K-Nearest Neighbor)等算法将海量的雷达数据进行分类。若使用人工手段,对所有雷达资料进行类别标注的成本太高,这时候使用聚类方法是一个好的选择,通过聚类,可以达到以下目的:

(1)从庞大的样本集合中选出一些具有代表性的由专家加以标注;

(2)在无类别信息情况下,寻找好的特征。

聚类属于机器学习的无监督学习,传统的聚类算法有划分方法(partitioning method)、层次方法(hierarchical methods)、基于密度的方法(density-based

methods)、基于网格的方法(grid-based methods)、基于模型的方法(model-based methods)。本文使用无监督学习的k-means方法进行研究,k-means方法属于划分方法,它把n个对象根据他们的属性分为k个聚类,使得所获得的聚类满足:同一聚类中的对象相似度较高;而不同聚类中的对象相似度较小。

### 3.2.4 试验结果

试验取k=10,随机选择10张雷达回波产品作为初始质心,按照k-means方法迭代5000次,得到分为10簇的回波,按聚类得到各簇的数量降序排列(图略),取迭代结束后得到的簇心距离最近的一个时次图像作为该簇代表图像(图略),可以看出各簇的代表图像区别程度较大。试验进行了多次,每次试验所分得的簇代表图像也接近,各轮次同类型簇大小基本一致,各簇簇内数量差异度<2%,可见的雷达回波特征向量适合k-means聚类方法。

按聚类得到各簇的回波类型约占原始数据的一半,这类回波即弱切面积小,对应的天气现象也基本一致,在进一步应用中,这类回波无须进一步分类,简单的归为一类即可。如对应的回波类型面积虽然比较大,但是回波并不强,可以考虑合并为同一类回波,如回波均有较大面积的强回波,两类图像总量6086张,虽然数量还是比较大,但是已经比原始数据小了很多,假如需要有效区分出更细致的强回波的类型,可以再引入其他特征向量进一步细分,同时可以使用人机结合的方法,将细分的小类交由专家给其打上标签,值得注意的是,由于k-means算法的特点,比代表图像强回波面积更大、强度更强的回波都被归类到这两类回波中,这两类回波都是对强对流天气深入研究的重点。如对应的回波的特点是有小范围强度较大的回波,由于地面观测站点比较稀疏,这类回波造成的降水等天气现象不一定能被自动站观测到,所以这类回波是否会造成局地强降水也是需要重点研究的地方。如回波面积不大,强度也不强,可以考虑合并为同一类回波。

## 3.3 程序运行环境

程序使用一台普通图形工作站,主要配置:CPU Intel i5-4460S 2.90GHz,内存12G。数据预处理部分共运算32min,使用预处理好的数据进行k-means聚类,每轮迭代5000次约耗时70min。

## 4 小结和讨论

本文从雷达资料相似度判断角度出发,为处理海量雷达数据,提出雷达回波特征 (下转第95页)



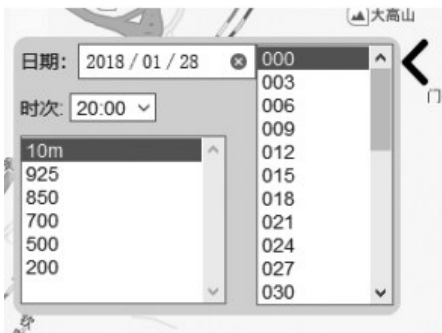


图7 大风数据控件

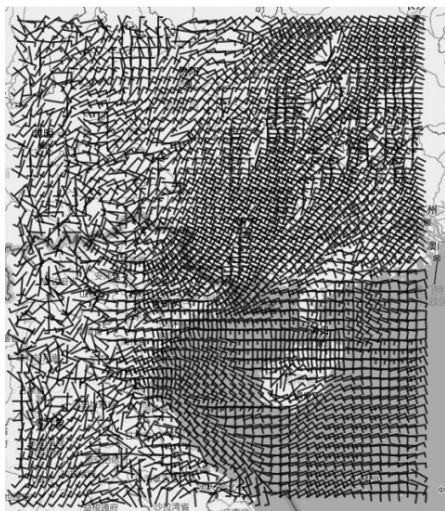


图8 大风预报风杆显示界面

### 3 结语

采用 OpenLayers 等 JS 类库搭建的北部湾气象可视化系统,结构清晰,能对实况天气信息进行直观丰富的可视化显示,还可通过雷达与卫星云图对钦州的短期临近预报提供指导意义,对北部湾大风的研究与预报具有重要意义。系统可扩展性强,对将来建立覆盖北部湾地区的实况、预报、预警系统奠定了基础。

#### 参考文献:

- [1] 郭明强,黄颖,谢忠,等. WebGIS 之 OpenLayers 全面解析[M]. 北京:电子工业出版社,2016:16-50.
- [2] 郑永泉,陈峰蓉,邢维东,等. 广西气象短信网页系统二次开发本地化的实现[J]. 气象研究与应用,2018,39(1):99-101.
- [3] 蒋礼珍,符永兴,徐一晖. 市级区域自动站信息处理系统的开发和应用[J]. 气象研究与应用,2010,31(1):77-79.
- [4] 黄伟萱. 新版 Micaps 资料的本地化处理与维护[J]. 气象研究与应用,2014,35(2):81-83.
- [5] 叶凯,蒲智,林思成. 基于 WebGIS 技术的塔里木河流域河湖信息系统[J]. 计算机系统应用,2018,27(2):91-96.
- [6] 李新庆,胡文东,王凡,等. Openlayers 在宁夏区域自动气象站显示系统中的应用与研究[J]. 宁夏工程技术,2015,14(2):111-114.
- [7] 赵文芳,刘旭林,聂凯. 基于 WebGIS 的气象综合显示系统改进与实现[J]. 应用气象学报. 2015,26(3):378-384.

(上接第 85 页)

向量这一概念,通过数学建模的方法,将复杂的相似度计算简化成矢量计算。通过对特征向量的提取,将复杂图像的特征信息简化成一串数值串,便于存储与计算。

在受干扰图像自动识别实验中,根据特征向量进行特征转换可快速发现离群点,通过噪声图像产生前后雷达回波特征向量余弦夹角变化的时间变化序列,可以看出雷达回波特征向量余弦夹角可以应用在实时业务系统中,用于监控回波的突变。

在雷达产品的聚类试验中,对五万多张雷达产品的多次聚类试验表明,由于雷达回波特征向量的使用,使得雷达回波图像的聚类分析可用计算的方法进行,计算快速,聚类效果良好。

本文的试验表明雷达回波特征向量的引入使得处理海量雷达产品有了简明、高效和可行的方法,是一种有益的尝试,具有良好应用前景。

#### 参考文献:

- [1] 曾小团,梁巧倩,农孟松,等. 交叉相关算法在强对流天气临近预报中的应用[J]. 气象,2010,36(1):31-40.
- [2] 曹春燕,陈元昭,刘东华,等. 光流法及其在临近预报中的应用[J]. 气象学报 2015,73(3):471-480.
- [3] 张蕾,魏鸣,李南,等. 改进的光流法在回波外推预报中的应用[J]. 科学技术与工程,2014,14(32):133-148.
- [4] 吴军,数学之美[M]. 北京:人民邮电出版社,2016:127-135.
- [5] 潘江,张培昌. 不同类型降水回波的自动识别方法[J]. 南京气象学院学报,1999,22(3):398-402.
- [6] 王庆国,黄归兰,黄增俊,等. “回南天”的客观分析方法研究[J]. 气象研究与应用,2014,35(2):1-6.
- [7] 罗思泽,黄庆国,杨兰,等. 梧州前汛期强对流天气多普勒雷达特征分析[J]. 气象研究与应用,2016,37(4):18-21.
- [8] 东高红,吕江津. 不同校准方法检验雷达定量估测降水的效果对比[J]. 气象与环境学报,2012,28(4):38-42.
- [9] 郑媛媛,谢亦峰,吴林林,等. 多普勒雷达定量估测降水的三种方法比较试验[J]. 热带气象学报,2004,20(3):192-197.