

基于时空因素聚类算法的研究与应用

闫双舰^{1,2}, 罗泽¹, 阎保平¹

(1. 中国科学院计算机网络信息中心 e-Science 应用推进总体组, 北京 100190; 2. 中国科学院大学 信息学院, 北京 100049)

摘要: 为了发现移动对象的迁徙轨迹和经停地, 提出结合经停地检测算法和单链接聚类算法的方法。通过青海湖鸟类的历史位置信息验证该方法的准确性和有效性, 并应用于本领域的其他方法进行分析比较, 如 DBSCAN 聚类算法、减聚类及模糊聚类算法。结果显示提出的方法能够克服对比算法仅考虑迁徙数据空间位置信息的缺点, 准确有效地挖掘出鸟类经停地和迁徙轨迹。

关键词: 聚类分析; 经停地检测算法; 单链接算法; 迁徙轨迹; 经停地发现

中图分类号: TP311.1 **文献标志码:** A **文章编号:** 1001-3695(2013)12-3690-04

doi:10.3969/j.issn.1001-3695.2013.12.044

Research and application of clustering algorithm based on factor of time and space

YAN Shuang-jian^{1,2}, LUO Ze¹, YAN Bao-ping¹

(1. e-Science application advance Zongtizu, Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; 2. Information School, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In order to discover the migratory trajectory and stopover, this paper proposed a method combining stopover detection algorithm and single-linkage cluster algorithm. It verified the accuracy and effectiveness of the above method by the history location information of the birds in Qinghai Lake, and other methods applied to the field was analyzed and compared, such as DBSCAN clustering algorithm, subtractive clustering and fuzzy clustering algorithm. The result shows that the proposed method can overcome the shortcomings of conventional approaches that only consider the space location information of the migratory data, and discover the bird stopover and migratory trajectory accurately and effectively.

Key words: clustering analysis; stopover detection algorithm; single linkage algorithm; migratory trajectory; stopover discovery

0 引言

生物遥测(biotlemetry)是指针对自由生活的动物个体状态变量的远程测量。生物遥测技术的出现使人们能够远程地跟踪候鸟的位置和移动状态。特别是基于全球定位系统(global positioning system, GPS)的卫星遥测, 具有每天 24 h 都能获取候鸟在地面(或空中)的高精度位置信息、候鸟的位置信息更新速度快且连续等优势。基于 GPS 的遥测技术使人们能够监测和绘制候鸟移动和迁徙的细节。候鸟的位置信息提供了移动路线的基本单位, 展示了动物个体与其周围生态系统的交互。掌握动物的移动模式能够帮助科学家更好地理解动物生态学中的一些关键概念, 包括资源使用(resource use)、活动范围(home range)以及分布情况和种群动态。

基于 GPS 的生物遥测技术与动物生态学(animal ecology)的结合, 在提供新的研究机会和研究方向的同时, 也对技术和分析能力提出了挑战。本文研究的数据来自中国科学院计算机网络信息中心、动物研究所与美国地质勘探局西部生态研究中心、Patuxent 野生动物研究中心, 以及联合国粮农组织的合

作项目(项目网站: <http://www.werc.usgs.gov/Project.aspx?ProjectID=40>)。

基于获得的数据, 开展候鸟行为研究的一项重要研究内容是发现候鸟迁徙的迁徙路线及迁徙过程中的经停地, 研究分析迁徙路线及经停地的地面生态特征和环境特征。为支持领域科研人员有效地发现候鸟迁徙路线和经停地, 本文提出并实现了一个结合 Stopover 检测算法和 Single-linkage 聚类算法的方法。

本文的贡献包括以下三个方面:

a) 本文提出并实现了结合 Stopover 检测算法和 Single-linkage 聚类算法的移动轨迹数据分析方法, 能够有效地提取出移动对象的经停地和移动路线。将该方法应用到了迁徙候鸟的栖息地和经停地发现中, 获得了斑头雁迁徙过程中栖息地和经停地选择的知识。

b) 利用斑头雁真实的位置历史数据对提出的方法的准确性和有效性进行了验证。

c) 将本文提出的方法与应用于本领域的两种典型聚类算法(DBSCAN 算法、Subtractive Cluster 和 Fuzzy C-means 算法)进行对比分析, 总结了不同方法的优缺点及适用场景。

收稿日期: 2013-04-10; 修回日期: 2013-05-16

作者简介: 闫双舰(1987-), 男, 河南开封人, 硕士研究生, 主要研究方向为数据挖掘、海量数据存储(659382132@qq.com); 罗泽, 男, 云南人, 研究员, 硕士, 博士, 主要研究方向为数据网格、面向服务架构、e-Science; 阎保平(1950-), 女, 山东人, 研究员, 博导, 博士(后), 主要研究方向为大规模科学数据共享技术、数据网格、下一代互联网技术。

1 相关工作

近来在卫星和跟踪设备上的进展,使得获取移动对象大量轨迹数据成为可能,例如包括车辆位置数据、飓风轨迹数据、动物移动数据。在这些轨迹数据上执行数据分析引起了越来越多研究人员的兴趣。本文主要介绍在研究候鸟迁徙领域的相关工作。

面向候鸟栖息地发现、栖息地连接性分析、迁徙路线发现等科学问题,唐明杰等人^[1-4]提出并实现了一个数据挖掘系统。他们观察到,栖息地是候鸟停留时间比较长的地方,在数学上对应于点密集的区域,因此提出了一种层次化的聚类算法,在不同的密度层面上发现可能的栖息地。基于聚类的结果计算候鸟准确的活动范围和持续的时间。文献[3]提出的层次化聚类算法的基础是基于密度的聚类算法 DBSCAN,通过不同的参数设置支持从大尺度到小尺度地理空间的层次化聚类分析。但是文献[3]提出的分层聚类算法存在的一个问题是没有考虑时间维度,仅从地理空间相邻的程度(通过计算欧式空间两点间的距离)判断候鸟的接近程度,可能产生不准确的聚类结果。

Subtractive Cluster 和 Fuzzy C-means 聚类联合使用的方法已经在文献中出现。文献[5]将该方法应用到白鹳的研究,白鹳每年秋季从北半球(欧洲)迁徙到南半球(非洲)。文献[5]把整个迁徙数据按照不同角度(如正向迁徙和反向迁徙)分片,联合 Subtractive Cluster 和 Fuzzy C-means^[6]分析方法,显示移动对象的行为特征。该方法应用于迁徙的不同阶段,能够检测出移动对象在运动过程中的经停地,同时计算移动对象在某个给定位置出现的概率,进而反映移动过程中不同区域的重要程度。

与上面两种方法相比,本文提出的方法考虑了时间因素,且该方法对噪声数据不敏感。

2 改进聚类算法

本文提出联合 Stopover 检测算法和 Single-linkage 聚类算法的方法,对鸟类迁徙数据进行先后处理。Stopover 检测算法考虑迁徙数据的时间和空间信息,按时间序列对数据集检测,当数据集的空间距离小于等于空间阈值,且停留时间大于等于时间阈值,把该数据集当成一个聚类中心点。鸟类迁徙数据随时间变化,空间呈现反复性,即会出现两个聚类中心重叠或者相距很近的情况。因此使用 Single-linkage 处理 Stopover 检测算法的聚类中心点,能够得到比较理想的聚类结果。

图 1 描述了联合 Stopover 检测算法和 Single-linkage 聚类算法的方法。 S 、 $T1$ 、 $T2$ 、 $T3$ 、 E 在 Stopover 检测算法中位于空间阈值(图中圆半径)范围内,且 E 和 S 之间的时间间隔大于等于时间阈值,则 S 、 $T1$ 、 $T2$ 、 $T3$ 、 E 构成了一个 Stopover。 $C1$ 、 $C2$ 是 Stopover 算法的聚类中心点,对 $C1$ 和 $C2$ 基于 Single-linkage 算法进行聚类,SC 是 Single-linkage 聚类算法的结果。

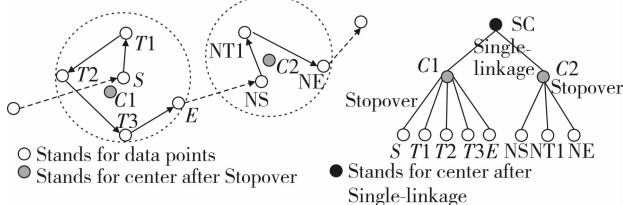


图 1 联合方法处理示意图

设预处理后的数据集为 D , Stopover 函数实现 Stopover 检测算法, SingLinkage 函数实现 Single-linkage 聚类算法,联合方法的伪代码如下:

```
CombineMethods( $D$ , StDist, StTime, MaxCenters);
clusterCenters = Stopover( $D$ , StDist, StTime);
//Stopover 函数返回检测算法的聚类结果
newCenters = SingleLinkage(clusterCenters, MaxCenters);
//SingleLinkage 函数返回 single-linkage 聚类算法的聚类结果
return newCenters;
```

2.1 Stopover 检测算法

Stopover^[7-9]算法的主要思想是:按时间序列,依次处理所有数据点,发现在一定空间范围阈值内,把停留时间大于某一时间阈值的数据子集作为停留地。停留地中心点位置坐标取该子集的平均值。

算法输入:距离阈值(StDist),时间阈值(StTime)。

算法输出:包含所有聚类中心点的数据集合 clusterCenters,有效数据 ValidDataNum(参与计算聚类中心点的数据个数,作为衡量算法好坏及参数是否合适的一个标准)。每个 center 的属性值组成如下:纬度均值(Lat_A)、经度均值(Lng_A)、停留时间(Time_D)、开始时间(Time_S)、结束时间(Time_E)。

算法伪代码如下:

```
clusterCenters = [];
ValidDataNum = 0;
Stopover( $D$ , StDist, StTime)
clusterCenters = [];
for index1 = 1 to Length( $D$ ) - 1
    for index2 = index1 + 1 to Length( $D$ )
        if(dist(index1, index2) >= StDist)
            if(timeDist( $i$ ,  $j$ ) >= StTime)
                //  $k$ .lg 和  $k$ .lt 分别对应数据点的经纬度
                Lat_A = ( $\sum_{k=i}^j k$ .lt)/( $j-i+1$ )
                Lng_A = ( $\sum_{k=i}^j k$ .lg)/( $j-i+1$ )
                Add(clusterCenters, Lat_A, Lng_A, timeDist( $i$ ,  $j$ ), time( $i$ ), time( $j$ ));
                //Time_S = time( $i$ );
                //Time_E = time( $j$ ); Time_D = time( $j$ ) - time( $i$ )
                ValidData = ValidDataNum +  $j-i+1$ ;
                else  $i=j+1$ ; break;
            else  $j=j+1$ ;
    return clusterCenters;
end;
```

2.2 Single-linkage 算法

Single-linkage^[10](单连接)算法是一种凝聚层次聚类算法。算法思想是:将每一个对象作为聚类中心,然后合并这些原子聚类中心为越来越大的聚类中心点,直到所有的对象都成为一个聚类中心点,或者满足某个终止条件。绝大多数层次聚类方法属于这一类型,只是在相似度的定义上有所不同。

使用树状图(dendrogram)表示层次聚类的过程。如图 2 所示,横坐标标号是聚类对象,纵坐标是簇间的相似度。假设初始有五个数据信息点,其中,最下面 $L=0$ 层都作为单元簇。在 $L=1$,数据点 4、5 的欧式距离最小,聚在一起形成一个簇,同时数据点 1、3 也聚在一起形成另一簇,这些聚集的数据点在以后各层仍保留在同一个簇中。当簇间数据点距离为 2.5 时,五个数据点形成一个簇。

算法定义簇间的距离如下: $D_{\min}(C_i, C_j) = \min|P_1 - P_2|$, P_1 属于 C_i , P_2 属于 C_j , P_1 、 P_2 是两个不同数据点, C_i 、 C_j 是两个

包含若干数据点的不同簇。

算法输入:数据集 $D1$ (Stopover 算法的聚类结果)。

算法输出:聚类中心点集合。

Single-linkage 算法伪代码如下:

```
SingLinkage(  $D1$ , MaxDist)
do
( $c_i, c_j$ ) = MinDist(  $C$  ); // 每个对象被初始化一个聚类中心点
 $C = C - c_i - c_j + \text{combine}(c_i, c_j)$ ; // 更新聚类中心点集合
Until MinDist(  $C$  ) < = MaxDist;
end;
MinDist(  $D1$  )
// 返回两个类中相似度
 $D \min(c_i, c_j) = \min p \in c_i, p \in c_j | p - p' |$ ;
combine( $c_i, c_j$ )
return  $c_i \cup c_j$ ; // 返回  $c_i$  和  $c_j$  的并集
```

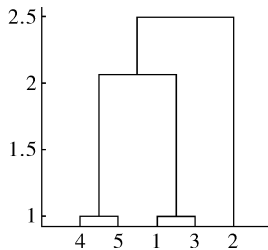


图2 数据对象{1,2,3,4,5}层次聚类的树状

联合方法对 Stopover 检测算法,选择合适的空间阈值和时间阈值,然后设置 Single-linkage 聚类算法的终止条件,从而达到避免 Stopover 检测算法和 Single-linkage 聚类算法各自缺点的目的。

3 实验评估及结果分析

实验以一只斑头雁的数据进行分析。该斑头雁共有 5 053 组历史数据。记录了 GPS 设备从 2007 年 3 月份到 2009 年 3 月的迁徙过程。在文献[11]中记录了斑头雁自身飞行速度为 16~21 m/s,但斑头雁飞行过程中易受高空气流影响(高空气流速度为 8~20 m/s),因此斑头雁迁徙过程中理论极限飞行速度为 41 m/s(147.6 km/h)。因此,迁徙路径中的数据点,如果两点间的速度大于 160 km/h(略大于 147.6 km/h),可以肯定两点必有其一是噪声数据。实验前首先去除噪声数据,去噪后剩下 5 001 个位置历史数据,去噪操作能够增加迁徙数据的有效利用率。实验证明,去噪前 Stopover 检测算法数据利用率为 69.38%,去噪后该算法数据利用率达 76.38%。

实验环境如下:台式机,Intel® 32 位双核处理器,内存 2.0 GB;操作系统 Windows XP。主要开发环境:MATLAB R2008a, DEV-C++ 5(开发语言 C++),Google Earth。

在联合 Stopover 和 Single-linkage 方法实验基础上,对比 DBSCAN 算法和 Subtractive Cluster 及 Fuzzy C-means 方法的实验结果,并从适用场景、算法优缺点方面比较方法的优劣。

3.1 联合方法实验分析

3.1.1 Stopover 检测算法实验结果

Stopover 算法输入参数:距离阈值(StDist)为 20 km,时间阈值(StTime)为 48 h。算法输出为 94 聚类中心点,分别用 Marker1, Marker2, ..., Marker94 表示。

Marker 1;36.741 99.792 25-Mar-2007 11:40:00 09-Apr-2007 06:10:00 14.771

Marker 2;36.742 99.795 10-Apr-2007 19:00:00 30-Apr-2007 21:00:00 20.083

Marker 3;36.801 99.793 02-May-2007 09:00:00 04-May-2007 09:00:00 2

...

Marker 1 聚类中心点的纬度为 36.741,经度为 99.792;经停地停留的开始时间为 25-Mar-2007 11:40:00;经停地停留的结束时间为 09-Apr-2007 06:10:00;经停地停留时间为 14.771 天(大于时间阈值 2 天)。其他聚类中心点代表的意义相同。Stopover 检测算法实验结果如图 3 所示。

Stopover 算法存在的缺点:空间阈值和时间阈值的设定和应用领域相关,需要领域知识才能确定恰当的参数。若设置阈值很大时,聚类效果不准确;若设置阈值很小时,聚类中心点较多。图 3 中 Stopover 检测算法得到了 94 个聚类中心点,有些聚类点相距很近,不能作为有效聚类中心,但这为下一步分析提供了有参考意义的中间数据。

3.1.2 Single-linkage 聚类算法实验结果

算法输入:Stopover 算法聚类点(94 个聚类中心),算法用分层二叉聚类树的形式表现如图 4 所示。

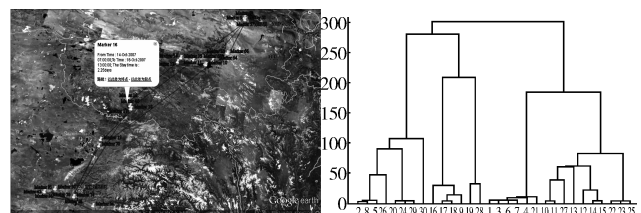


图3 移动对象聚类

中心点及迁徙路径图

图4 Single-linkage 算法

聚类结果

图 4 是 Single-linkage 算法的输出结果。横坐标代表 30 个叶子节点,纵坐标 0~300 km 指对象间的距离。如果输入数据点少于等于 30 个,每个叶子节点对应一个原数据点;如果数据点多于 30,低层数据点被压缩,以满足 30 个叶子节点。

Single-linkage 算法直接处理原数据(5 001 组数据)存在以下不足:a)衡量 Single-linkage 算法聚类结果的好坏,主要参考共性分类相关系数(cophenetic correlation coefficient),共性分类相关系数越接近 1,说明聚类方法越能反应数据本身,聚类效果越好,处理原数据的共性分类相关系数为 0.854 6,若先经 Stopover 检测算法后,再使用 Single-linkage 聚类算法的共性分类相关系数为 0.959 7;b)原数据对应更多的叶子节点,可视化效果较差;c)Single-linkage 算法对噪声数据过分敏感,对原数据处理产生较大的误差,Stopover 检测算法具有较强抗干扰性,故首先选择 Stopover 检测算法处理原数据。

根据 Stopover 检测算法发现的 94 个聚类中心点所处的位置,设 Single-linkage 算法的对象间的相似度阈值为 200 km,实验结果如图 5 所示,该图描绘具有时间属性的聚类中心点序列。

3.2 DBSCAN 实验结果

利用 DBSCAN 聚类方法处理数据。邻域 ε 和 MinPts 阈值是 DBSCAN 算法的输入参数,且 ε 和 MinPts 的大小需要根据经验来确定。从实验结果分析可知,当 ε 固定时,MinPts 越大,聚类数目越少,被判定为噪声数据的点就越多;而当 MinPts 固定不变时, ε 越大,聚类数目越少,被判定为噪声数据的样本数据就减少。

经过比较参数取值,选取 0.03 作为 ε 领域的取值,对应的空间范围在 46~63 km,可以得到如表 1 所示的实验数据。

表 1 DBSCAN 实验结果

input		output	
ε	MinPts	聚类数	噪声数
0.03	10	5	57
0.03	20	5	68
0.03	80	6	79
0.03	150	6	127
0.03	200	4	515

对应的输入参数 ε 为 0.03, MinPts 为 20, 输出聚类中心点数目为 5, 噪声数据个数为 68。

聚类中心坐标信息如下: cluste0(99.8129, 36.7508), cluster4(96.7890, 35.1127), cluster1(93.4320, 33.0821), cluster2(91.7623, 31.5118), cluster3(89.6618, 29.2080)。

图 6 所示是 DBSCAN 算法的聚类在 Google Earth 中的可视化效果, cluster0、cluster4、cluster1、cluster2、cluster3 聚类中心点依次如图所示, 黄色路径是对象的迁徙轨迹, 绿色地标指聚类中心点, 蓝色路径连接聚类中心(见电子版)。

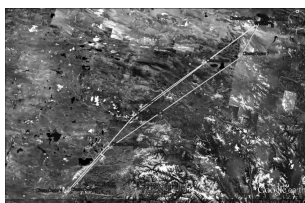


图 5 相似度为 200 km 的聚类结果

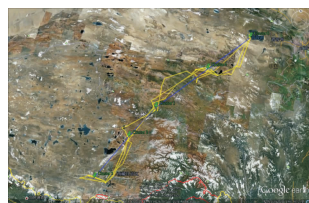


图 6 DBSCAN 聚类结果

3.3 Subtractive Cluster 和 Fuzzy C-means 实验结果

Subtractive Cluster 和 Fuzzy C-means 方法处理该数据集^[5,6]。输入参数包括 radii、xBounds 和 options。参变量 radii 取值范围是 0~1, 用来指定一个聚类中心在每一纬度上的影响范围, 首先把数据集的数据映射到 0~1, 当 radii 越小时, 越容易生成较多的聚类中心。radii 是一个向量值, 对于不同的纬度, 可以选择不同的影响因子。

结合 Subtractive Cluster、Fuzzy C-means 算法的聚类实验。输入条件为 radii = 0.005, others = default; 输出为聚类数目 6。Subtractive Cluster 算法和 Fuzzy C-means 聚类算法的参数意义及选择可参考文献^[5]。聚类结果在 Google Earth 可视化如图 7 所示。

聚类中心依次为: cluster1(99.7342, 36.6321)、cluster2(96.2859, 35.0929)、cluster3(93.4394, 33.0777)、cluster4(91.7725, 31.4908)、cluster5(90.4063, 29.0353) 和 cluster6(89.2520, 29.3084)。红色地标指聚类中心点, 红色路径连接聚类中心点(见电子版)。

3.4 实验结果对比

以 BH07_67695 的迁徙数据, 对每一种方法选择合理的参数, 三种方法的聚类结果如图 8 所示。

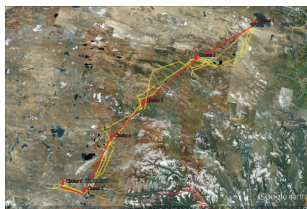


图 7 结合减聚类模糊聚类的聚类结果



图 8 三种聚类方法的比较

绿色地标和路径为 Stopover 算法、Single-linkage 聚类结果

图, 迁徙路径具有周期性, 能够准确地反映出迁徙特征^[12]; 红色地标和路径对应减聚类模糊聚类的结果; 蓝色地标和路径是 DBSCAN 的效果图(见电子版)。可以看出 Stopover 和 Single-linkage 算法能够得出与时间相关的迁徙序列^[13], 另外两种算法只能得到具有空间属性的聚类中心点。

3.5 算法性能分析

从时间复杂度、使用场景、优缺点对比三种方法, 结果如表 2 所示。

表 2 聚类算法比较

算法	时间复杂度	适应场景	优点	缺点
DBSCAN	$O(n^2)$, 最优情况下 $O(n \lg n)$	发现任意形状的簇	抗噪声	未考虑时间因素
减聚类 + 模糊聚类	$O(n^3 \times \lg 2n)$	更能反映现实世界	减聚类的输出作为模糊聚类的输入, 达到全局最优	处理复杂, 未考虑时间因素
Stopover + 单链接算法	$O(n^2)$	适应时空数据的聚类分析	考虑时空因素, 避免单链接算法易受噪声数据影响, 同时又避免 Stopover 受输入参数的影响	暂无

a) Stopover 算法从空间和时间两个角度分析移动对象数据, 把对象复杂的迁徙过程生成一个与时间相关的序列; b) 避免了 Single-linkage 基于距离且受噪声数据影响的缺陷; c) Single-linkage 弥补了 Stopover 生成太多聚类中心点的不足。相比 DBSCAN、Subtractive 和 Fuzzy C-means, Stopover 和 Single-linkage 是一种较理想的聚类方法。

4 结束语

本文以青海湖鸟类迁徙数据为对象, 进行聚类方法的分析研究。提出并实现 Stopover 检测算法和 Single-linkage 的联合方法, 通过与 DBSCAN 算法、Subtractive Cluster 及 Fuzzy C-means 算法的比较, 体现联合算法的优势。

移动对象的数据挖掘面临很多挑战, 但也极具研究意义。鸟类数据的处理分析对鸟类学家及生态学家有很大的研究和应用价值: a) 生态保护, 聚类中心点的发现, 可对生态环境及其周围野生动物实施保护措施, 增强目的性; b) 异常检测^[14], 鸟类迁徙受自然条件影响, 分析鸟类迁徙的异常, 对自然灾害的预测和防范也至关重要; c) 通过实验, 从性能、适用场景及优缺点比较对现有算法提出改进, 为以后研究提供参考。

下一步的研究方向。首先, 从多角度分析数据, 如迁徙方向、年份、对象类别; 其次, 聚类算法参数值(即输入参数)的选择对实验结果有很大的影响, 如何把经验因素的影响降到最低, 应该是以后研究的重点之一; 最后, 还将进一步探寻比较算法优劣的标准。

参考文献:

- [1] TANG Ming-jie, ZHOU Yuan-chun. Discovery of migration habitats and routes of wild bird species by clustering and association analysis [C]//Proc of the 5th International Conference on Advanced Data Mining and Application. [S. l.]: Springer-Verlag, 2009: 288-301.
- [2] CUI Peng, HOU Yuan-sheng. Bird migration and risk for H5N1 transmission into Qinghai lake [J]. China: Vector-Borne and Zoonotic Diseases, 2010, 11(5): 567-576.
- [3] TANG Ming-jie, ZHOU Yuan-chun, LI Jin-yan. Exploring the wild birds' migration data for the disease spread study (下转第 3697 页)

4 结束语

本文在CS理论的基础上对AIC进行了研究。阐述了CS理论的三个核心问题,分析了AIC的结构与功能,然后在MATLAB/Simulink平台上建立了AIC系统模型,并进行了相关实验与仿真。实验结果符合CS理论要求,验证了建模的可行性。因此,在处理物联网海量数据时,用AIC器件代替传统的ADC能够有效解决ADC采样速率不足等问题,实现对宽带信号的低速率采样、高概率重构。

参考文献:

- [1] 周开乐,丁帅,胡小建. 面向海量数据应用的物联网信息服务系统研究综述[J]. 计算机应用研究, 2012, 29(1): 8-11.
- [2] CANDES E, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Trans on Information Theory, 2006, 52(2): 489-509.
- [3] 石光明,刘丹华,高大化,等. 压缩感知理论及其研究进展[J]. 电子学报, 2009, 37(5): 1070-1081.
- [4] 戴琼海,付长军,李向阳. 压缩感知研究[J]. 计算机学报, 2011, 34(3): 425-434.
- [5] 李卓凡,闫敬文. 压缩感知及应用[J]. 微计算机应用, 2010, 31(3): 12-16.
- [6] ANITORI L, MALEKI A, OTTEN M, et al. Design and analysis of compressed sensing radar detectors[J]. Signal Processing, 2013, 61(4): 813-827.
- [7] 刘记红,徐少坤,高勋章,等. 压缩感知雷达成像技术综述[J]. 信号处理, 2011, 27(2): 251-260.
- [8] MENG J, LI H, HAN Z. Sparse event detection in wireless sensor networks using compressive sensing[C]//Proc of the 43rd Annual Conference on Information Sciences and Systems. 2009.
- [9] LING Q, TIAN Z. Decentralized sparse signal recovery for compressive sleeping wireless sensor networks[J]. IEEE Trans on Signal Processing, 2012, 58(7): 3816-3827.
- [10] BOURQUARD A. Binary compressed imaging[J]. IEEE Trans on Signal Processing, 2013, 22(3): 1042-1055.
- [11] LI Y, SHA X, WANG K, et al. The seight-block compressed sensing and its application to image reconstruction[C]//Proc of the 2nd International Conference on Instrumentation, Measurement, Computer, Communication and Control. 2012.
- [12] HUO C, ZHANG R, YIN D. Compression technique for compressed sensing hyperspectral images[J]. International Journal of Remote Sensing, 2012, 33(5): 1586-1604.
- [13] RIVENSON Y, STERN A, JAVIDI B. Overview of compressive sensing techniques applied in holography[J]. Applied Optics, 2013, 52(1): 423-432.
- [14] 焦李成,杨淑媛,刘芳,等. 压缩感知回顾与展望[J]. 电子学报, 2011, 39(7): 1651-1662.
- [15] 邵文泽,韦志辉. 压缩感知基本理论: 回顾与展望[J]. 中国图象图形学报, 2012, 17(1): 1-12.
- [16] CANDES E. The restricted isometry property and its implications for compressed sensing[J]. Académie des Sciences, 2006, 346(1): 592-598.
- [17] PROVOST J, LESAGE F. The application of compressed sensing for photoacoustic yomography[J]. IEEE Trans on Medical Imaging, 2009, 28(4): 585-594.
- [18] CANDES E, ROMBERG J. Practical signal recovery from random projections[C]//Proc of SPIE. 2005: 76-86.
- [19] CHEN S, DONOHO D, SAUNDERS M. Atomic decomposition by basis pursuit[J]. SIAM Rev, 2001, 43(1): 129-159.
- [20] ELAD M. Optimized projections for compressed sensing[J]. IEEE Trans on Signal Processing, 2007, 55(12): 5695-5702.
- [21] GILBERT A, MUTHUKRISHNAN S, STRAUSS M. Improved time bounds for near optimal sparse fourier representation via sampling[C]//Proc of SPIE. 2005.
- [22] LASKA J. Random sampling for analog-to-information conversion of wideband signals[J]. IEEE Trans on Information, 2006, 58(6): 457-579.
- [23] BARANIUK R, CEVHER V, DUARTE M, et al. Model-based compressive sensing[J]. IEEE Trans on Information of Theory, 2010, 56(4): 1982-2001.
- [1] 周开乐,丁帅,胡小建. 面向海量数据应用的物联网信息服务系统研究综述[J]. 计算机应用研究, 2012, 29(1): 8-11.
- [2] CANDES E, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Trans on Information Theory, 2006, 52(2): 489-509.
- [3] 石光明,刘丹华,高大化,等. 压缩感知理论及其研究进展[J]. 电子学报, 2009, 37(5): 1070-1081.
- [4] 戴琼海,付长军,李向阳. 压缩感知研究[J]. 计算机学报, 2011, 34(3): 425-434.
- [5] 李卓凡,闫敬文. 压缩感知及应用[J]. 微计算机应用, 2010, 31(3): 12-16.
- [6] ANITORI L, MALEKI A, OTTEN M, et al. Design and analysis of compressed sensing radar detectors[J]. Signal Processing, 2013, 61(4): 813-827.
- [7] 刘记红,徐少坤,高勋章,等. 压缩感知雷达成像技术综述[J]. 信号处理, 2011, 27(2): 251-260.
- [8] MENG J, LI H, HAN Z. Sparse event detection in wireless sensor networks using compressive sensing[C]//Proc of the 43rd Annual Conference on Information Sciences and Systems. 2009.
- [9] LING Q, TIAN Z. Decentralized sparse signal recovery for compressive sleeping wireless sensor networks[J]. IEEE Trans on Signal Processing, 2012, 58(7): 3816-3827.
- [10] BOURQUARD A. Binary compressed imaging[J]. IEEE Trans on Signal Processing, 2013, 22(3): 1042-1055.
- [11] LI Y, SHA X, WANG K, et al. The seight-block compressed sensing and its application to image reconstruction[C]//Proc of the 2nd International Conference on Instrumentation, Measurement, Computer, Communication and Control. 2012.
- [12] HUO C, ZHANG R, YIN D. Compression technique for compressed sensing hyperspectral images[J]. International Journal of Remote Sensing, 2012, 33(5): 1586-1604.
- [13] RIVENSON Y, STERN A, JAVIDI B. Overview of compressive sensing techniques applied in holography[J]. Applied Optics, 2013, 52(1): 423-432.
- [14] 焦李成,杨淑媛,刘芳,等. 压缩感知回顾与展望[J]. 电子学报, 2011, 39(7): 1651-1662.
- [15] 邵文泽,韦志辉. 压缩感知基本理论: 回顾与展望[J]. 中国图象图形学报, 2012, 17(1): 1-12.
- [16] CANDES E. The restricted isometry property and its implications for compressed sensing[J]. Académie des Sciences, 2006, 346(1): 592-598.
- [17] PROVOST J, LESAGE F. The application of compressed sensing for photoacoustic yomography[J]. IEEE Trans on Medical Imaging, 2009, 28(4): 585-594.
- [18] CANDES E, ROMBERG J. Practical signal recovery from random projections[C]//Proc of SPIE. 2005: 76-86.
- [19] CHEN S, DONOHO D, SAUNDERS M. Atomic decomposition by basis pursuit[J]. SIAM Rev, 2001, 43(1): 129-159.
- [20] ELAD M. Optimized projections for compressed sensing[J]. IEEE Trans on Signal Processing, 2007, 55(12): 5695-5702.
- [21] GILBERT A, MUTHUKRISHNAN S, STRAUSS M. Improved time bounds for near optimal sparse fourier representation via sampling[C]//Proc of SPIE. 2005.
- [22] LASKA J. Random sampling for analog-to-information conversion of wideband signals[J]. IEEE Trans on Information, 2006, 58(6): 457-579.
- [23] BARANIUK R, CEVHER V, DUARTE M, et al. Model-based compressive sensing[J]. IEEE Trans on Information of Theory, 2010, 56(4): 1982-2001.

(上接第3693页)

- of H5N1: a clustering and association approach[J]. Knowledge and Information Systems, 2011, 27(2): 227-251.
- [4] LI Sha-sha, ZHOU Yuan-chun, KOU Zheng. A wavelet packet based approach for the research of the Avian influenza virus cross-species infection[C]//Proc of the 1st International Conference on Networking and Distributed Computing. 2010: 301-304.
- [5] CARNEIRO C, ALP A, MACEDO J. Advanced data mining method for discovering regions and trajectories of moving objects: "Ciconia Ciconia" scenario[C]//Lecture Notes in Geoinformation and Cartography. Berlin: Springer-Verlag, 2008: 201-224.
- [6] PRIYONO A, RIDWAN M, ALIAS J A, et al. Generation of fuzzy rules with subtractive clustering[J]. Journal Teknologi, 2012, 43(D): 143-153.
- [7] ZHENG Yu, XIE Xing. Learning location correlation from GPS trajectories[C]//Proc of the 11th International Conference on Mobile Data Management. 2010: 27-32.
- [8] LI Quan-nan, ZHENG Yu, XIE Xing, et al. Mining user similarity based on location history [C]// Proc of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM Press, 2008: 1-10.
- [9] ZHENG Yu, ZHANG Li-zhu, MA Zheng-xin. Recommending friends and locations based on individual location history[J]. ACM Transaction on the Web, 2011, 5(1): 1-44.
- [10] HAN Jia-wei, KAMBER M, PEI Jian. Data mining concepts and techniques[M]. 3rd ed. [S.l.]: Morgan Kaufmann, 2011.
- [11] WARD S, BISHOP C M, WOAKES A J, et al. Heart rate and the rate of oxygen consumption of flying and walking barnacle geese and bar-headed geese [J]. Journal of Experimental Biology, 2002, 205(21): 3347-3356.
- [12] GIANNOTTI F, NANNI M, PINELLI F, et al. Trajectory pattern mining[C]//Proc of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2007: 330-339.
- [13] VERHEIN F, CHAWLA S. Mining spatio-temporal patterns in object mobility databases [J]. Data Mining and Knowledge Discovery, 2008, 16(1): 5-38.
- [14] LI Xiao-lei, HAN Jia-wei, KIM S, et al. ROAM: rule-and motif-based anomaly detection in massive moving object data sets[C]//Proc of the 7th SIAM International Conference on Data Mining. 2007.