密度聚类算法在连续分布点云去噪中的应用

张巧英，陈　浩，朱　爽
（浙江省测绘大队，浙江 杭州 310007）

摘　要：在原始测量获取的点云数据中，除了目标数据外，还有大量的噪声数据。噪声往往无规律地分布在目标物体周围，难以用统一数学模型区分。基于密度的聚类算法将簇定义为密度相连的点的最大集合，能发现任意形状、大小的类簇，将该算法应用在点云去噪中，能将密度分布连续点进行聚类，从中提取出目标点云。

关键词：基于密度的聚类；点云密度分布；点云去噪
中图分类号：P237　　**文献标志码**：B　　**文章编号**：1672-4623（2011）06-0101-04

在点云数据的获取中，由于受人为因素、仪器因素、环境因素、测量方法和随机因素的影响和干扰，我们所获取的点云并不能完全精确地表现被测目标物体的表面信息[1]。在这种含有噪声点的数据上进行数据的后续处理（如数据简化、建模），会影响各阶段的数据处理的精度与质量。因此需要在获取点云数据后的数据预处理阶段进行数据的去噪处理。由于噪声点的产生不受人控制，噪声点的分布、噪声点形成的类簇的形状、大小等是随机产生。要用一个统一的数学模型来表达和区分目标点与噪声点不大容易。我们通常在点云处理软件中手动删除这个噪声点，不仅效率低，而且容易误删了目标点。考虑到当被测物体表面是连续时，通过扫描仪获取的点云也是沿表面连续分布；而噪声则在目标点云外随机分布。本文将数据挖掘领域中的聚类思想引入点云去噪中，利用聚类的思想将点云数据进行类别划分，然后在聚类的结果上选择出目标主体实现去噪处理。

1　基于密度的聚类算法

聚类分析的目的是在大规模的数据集中发现未知对象，是数据挖掘领域的重要研究课题。聚类也被称作非监督分类，是指在没有任何先验知识的情况下，根据数据的不同特征，将其划分为不同的数据类或簇。聚类结果是使得属于同一个簇的个体之间具有较大的相似性，而不同簇上的个体间的具有较大的相异性。

DBSCAN（Density-Based Spatial Clustering of Application with Noise）是一个比较有代表性的基于密度的聚类算法，它将簇定义为密度相连的点的最大集合，并可在有“噪声”的空间数据库中发现任意形状的簇类[2]。由于该算法不是基于各种各样的距离进行分类，这样就克服了基于距离判别准则只能发现“类圆形”聚类的限制。图 1 为一个实际三维数据点的聚类效果图。

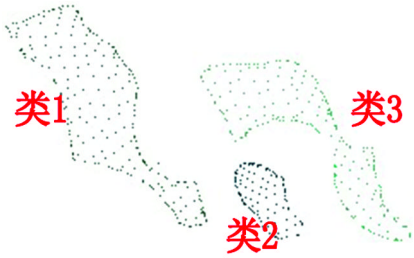


图 1　聚类结果

DBSCAN 通过遍历数据集中的每个对象的ε邻域[1]来寻找聚类。其算法描述如下：从任意点 p 开始，判断其是否已归属某类，如果已归类则继续下一点；若未归类，判断其是否为核心对象[1]，然后对该核心对象的邻域进行查询，搜索所有从 p 点密度可达的点，形成一个密度联通集；将上述过程重复执行，得到所有的密度联通集。

DBSCAN 算法不需要初始定义聚类数目，即可将具有足够高密度的区域划分为簇，并且对噪声不敏感。在点云数据的处理中，由于噪声的数目、分布均未知，所以 DBSCAN 算法能够较好地应用在此类问题中。但是该算法对初始定义的参数ε、Minpts 较为敏感，不同的参数设置会使得聚类结果有很大差别。故在点云去噪中需要根据实际数据的分辨率、被测目标的形状等因素设置相应的参数，获取最佳聚类结果。

2　点云的空间密度分布

点云的空间密度在本文中是指：以任意点为球心，半径为 r 的球体范围内的总点数，属于体密度；与常见的以分辨率表示的面密度有较大区别。在不考虑物体的反射率、折射率等情况下，点云的空间密度分布主要由 2 个因素决定：扫描仪的扫描原理、被测物体自身的几何形状。

收稿日期：2011-09-22

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

2.1 扫描仪的扫描原理对点云密度的影响

高分辨率扫描仪的扫描原理主要有 4 种[3]，一种是激光测距（LIDAR），利用三维激光扫描仪获取目标点与扫描仪的相对距离、角度，解算出目标点的三维坐标，常见的地面激光三维扫描仪和机载激光三维扫描仪属于此类；另一种是摄影测量（photogrammetry），利用相机获取目标点的立体相对，根据三角测量的原理反算出目标点的三维坐标，手持式激光扫描仪 Handyscan 即属于此种原理；此外还有结构光（structured light）和全息摄影（holography）2 种。

本文用来检验密度聚类算法在点云去噪应用中的数据，是采用加拿大 Creaform 公司的 HANDYSAN- Maxscan 手持式三维扫描仪获取，扫描仪如图 2 所示[4]。

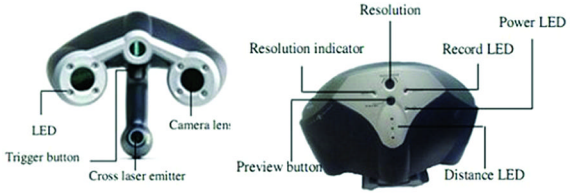


图 2 HANDYSAN- Maxscan 手持式三维扫描仪

该扫描仪的数据获取装置包括 2 个固定定向的数码相机和 1 个十字激光发射器。它的一个主要特点就是使用反射靶实现自定位：如果扫描仪能到同时扫描到被测物体上 4 个或以上的反射靶，就可以解算出扫描仪在由靶标确定的物方坐标系内的坐标，然后根据 2 个摄像机获取的相片建立物体表面点云[5]。扫描过程中，反射靶的影像是实时获取，并且实时解算坐标，目标物的其他点坐标就是通过这些点来解算。

手持式扫描仪的原理比较简单——摄影测量原理，但是它所采用的解决方案却是一种全新的思路。不仅可以获取高精度的点云数据，而且避免了传统摄影测量中的控制点、立体相对的定向、自相关等问题。

手持扫描仪在数据采集阶段可以对目标物进行多次重复的测量，软件自动处理这些数据产生一个优化后的结果。获取的点云分辨率（点间距）根据用户设置由系统软件自动处理。

2.2 被测物体自身形状对点云密度的影响

由于点云是沿被测物体表面分布，因此被测物体自身的几何形状会对点云密度分布产生很大影响。如目标物体中间区域与边缘区域、平坦区域与形状变化较大区域的密度分布有较大区别。同时噪声区域与目标体区域的点云密度分布同样有较大区别。

图 3 为高度 150 mm、最大直径 80 mm 的圆柱形笔筒，使用 HANDYSAN- Maxscan 手持式扫描仪获取分辨率为 0.7 mm 的含有噪声的物体点云。

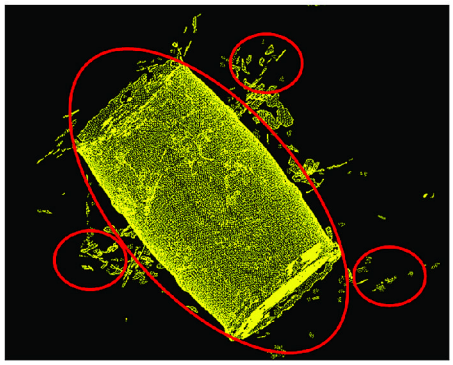


图 3 含有噪声的笔筒点云

其中目标区的密度分布均匀且连续，而噪声分布在目标点云的周围。虽然噪声也聚集成大大小小的类簇，但类簇的大小和位置都是随机分布且不连续。

对目标体的不同位置做密度分析：以筒口边缘一点为中心，半径为 3 的邻域内点数为 45；以筒表面上 1 点为中心，半径为 3 的邻域内点数为 82，具体位置如图 4 所示。由于笔筒表面连续，且无较大的形状变化在扫描仪的设置的扫描参数不变的情况下，筒身上不同位置点的密度基本一致；筒口边缘的点密度则相对为筒身的一半左右。

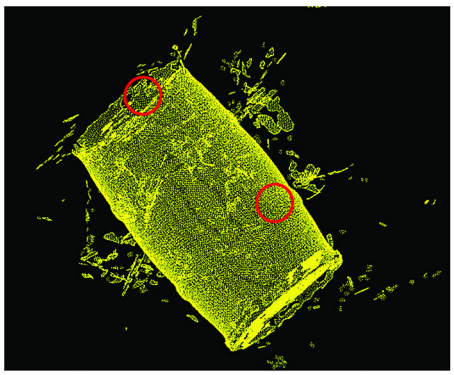


图 4 圆柱筒不同位置密度对比

3 对点云数据进行密度聚类

由于基于密度的聚类算法需要 2 个用户自定义参数：ε和 Minpts，且聚类结果对这 2 个参数都比较敏感。所以聚类前需对点云数据的密度信息作一个统计分析，然后根据结果选择合适的初始参数。

整个算法过程分为以下几个步骤：

1）选定一个半径ε_i对点云数据作统计分析。

2）根据统计结果选择初始ε_i和 Minpts（i)。

3）开始聚类（由于聚类过程中程序的主要开销是进行ε邻域内点的搜索，为提高程序效率本文中的程序使用了 kd 树数据结构来对点云数据进行管理）。

4）判断聚类结果质量，若能分离出目标则结束，否则返回 1）选择ε_{i+1}继续。

图 5 是该聚类算法的程序流程图。

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

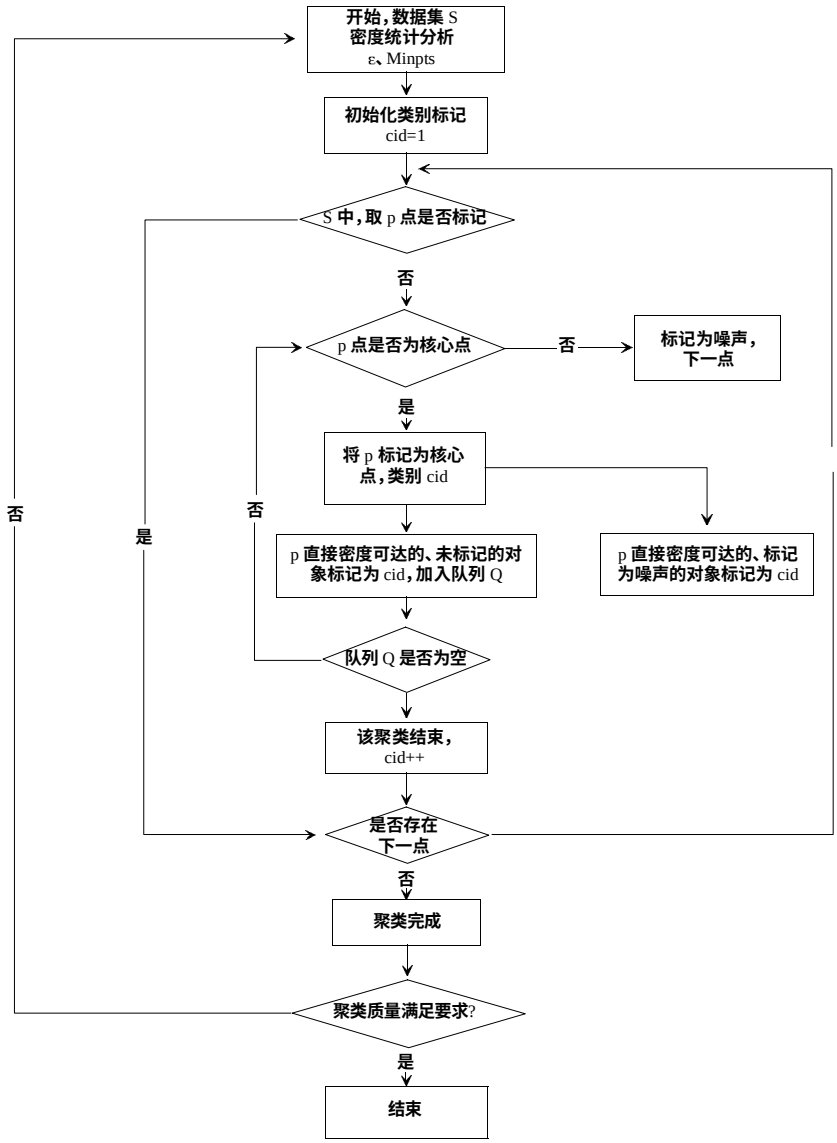
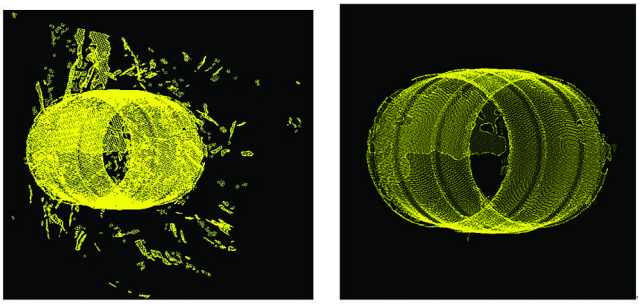


图 5　聚类算法的程序流程图

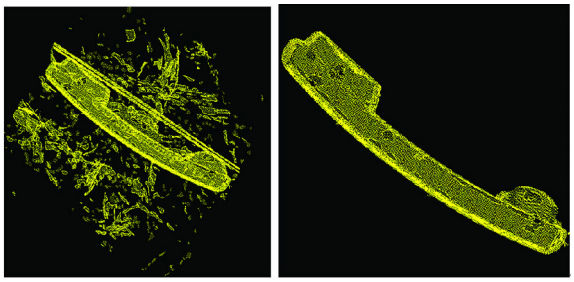
4　算法应用分析

本文采用了 3 组数据进行去噪分析实验，效果如图 6-8 所示。



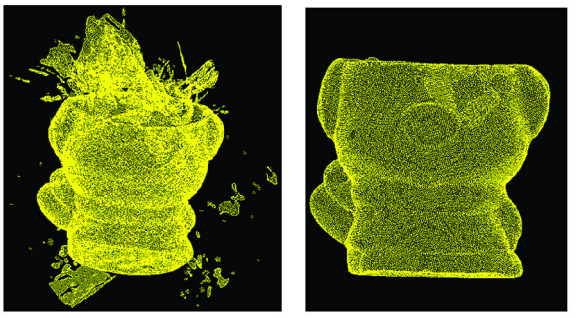
a)含噪声的原始数据　　b)聚类后分离的目标数据

图 6　圆柱笔筒聚类结果图



a)含噪声的原始数据　　b)聚类后分离的目标数据

图 7　电话听筒聚类结果图



a)含噪声的原始数据　　b)聚类后分离的目标数据

图 8　笔筒猪聚类结果图

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

下面以表格的形式给出聚类后的目标点数与实际点数的关系。

表 1 聚类结果显示

	原始点数	提取目标点数	目标点数提取比例
圆柱笔筒	93 984	68 908	73.318 9 %
电话听筒	124 516	40 894	32.842 4 %
笔筒猪	133 663	82 222	61.514 4 %

由图 6、7、8 及表 1 可以明显地看出，经过该算法的处理，在保持目标体数据特征的同时，有效减少了数据冗余。

5 结 语

本文将基于密度的聚类算法应用到点云去噪中，很好地实现了目标物与噪声点的分离。算法处理过程中没有对原始目标点云做重采样、删除或转换，最大限度地保留了目标体的原始扫描信息。但是本文中的应用还仅局限于小范围、表面连续的物体点云；并且初始参数的选择需要对点云进行密度统计再分析，过程中可能需要多次人工干预；聚类质量也是由人眼观察评价。所以下一步的工作应更多地考虑这三方面内容，使得基于该算法在点云去噪中的应用范围更广泛，人工干预过程更少。

参考文献

[1] 徐慧朴. 逆向工程中基于离散点云的噪声预处理及模型重构技术研究[D]. 大连：大连海事大学, 2009

[2] 毛国君, 段立娟, 王实, 等. 数据挖掘原理与算法[M]. 北京：清华大学出版社, 2005

[3] Leandro BORNAZ, Chiara PORPORATO. 3D High Accuracy Survey and Modeling of One of AOSTA'S Anthropomorphic Stelae [J]. XXI International CIPA Symposium, 2007, (10): 176-181

[4] Huipu Xu, Ying Hu. Unilateral Left-side Quartile Algorithm Based Data Processing Scheme for 3D Scattered Point Data[J]. International Journal of the Physical Sciences, 2010, (5)：1641-1646

[5] A.Adami, L.Fregonese, F.Guerra,et al. Digital Representations and Analysis of Deformations Induced in Map Supporting Materials[J]. XXI International CIPA Symposium, 2007, (10):7-11

[6] O.Etxaniz, E.Solaberrieta. Digital Modelling of a Human Skull [J]. Journal of Achievements in Materials and Manufacturing Engineering, 2008, (27)：55-58

[7] 刘艳丰. 基于 kd-tree 的点云数据空间管理理论与方法 [D]. 长沙：中南大学, 2009

第一作者简介：张巧英，高级工程师，主要从事地理信息及土地勘测定界工作。

（上接第 100 页）

参考文献

[1] 毛克彪, 唐华俊, 陈仲新, 等. 一个从 ASTER 数据中反演地表温度的劈窗算法[J].遥感信息, 2006(5):7-11

[2] 甘甫平, 陈伟涛.热红外遥感反演陆地表面温度研究进展[J]. 国土资源遥感, 2006, 67(1)：5-11

[3] 姚云军, 南鹏.分裂窗算法在热红外遥感陆面温度反演中的应用[J].兰州理工大学学报, 2007, 33(6)：89-92

[4] 毛克彪.用于 MODIS 数据的地表温度反演方法研究[D]. 南京:南京大学, 2004

[5] 赵英时.遥感应用分析原理与方法[M].北京：科学出版社, 2004

[6] 毛克彪, 覃志豪. 针对 MODIS 数据的大气水汽含量反演及 31 和 32 波段透过率计算[J]. 国土资源遥感, 2005(1)：26-30

[7] Price J C. Land Surface Temperature Measurements from the Split Window Channels of the NOAA 7 Advanced Very High Resolution Radiometer [J]. Journal of Geophysical Research, 1984,89:7231-7237

[8] Sobrino J A, Jimenez-Muiioz J C, and aolini L. Land Surface Temperature Retrieval from LANDSAT TM 5[J].Remote Sensing of Environment, 2004, 90:434-440

[9] GAO B C. AlexanderFH Geotz.Column Atmospheric Water Vapor and Vegetation Liquid Water Retrievals from Airborne Imaging Spectrometer Data [J]. Journal of Geophysical Research, 1990(95):3 549-3 564

[10] 毛克彪, 覃志豪.用 MODIS 影像反演环渤海地区的大气水汽含量[J].遥感信息, 2004(4):47-49

第一作者简介：王丽美，硕士，主要从事摄影测量与遥感方面的研究。

下期论文导读

陈智尧 邱儒琼：大比例尺地形图分幅统一编号的构想

在分析比较现行国家基本比例尺地形图和大比例尺地形图分幅编号的相关规定的基础上，提出了一种全国范围内大比例尺地形图的分幅编号规则，并结合分幅编号与地理坐标的关系，以实例解析了 2 个通用的应用模型。

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Image Fusion Based on Steerable Pyramid and IHS Transformation by LI Bo

Abstract This paper proposed a new image fusion method based on steerable pyramid and IHS Transformation. The intensity component of multispectral image(MS) and panchromatic image(PAN) was decomposed respectively by steerable pyramid transformation. Different rules are used to fuse sub-images of low frequency, high frequency, and band pass, and then the fused multispectral image could be reconstructed by inverse steerable pyramid and IHS transformation. The experiment results show that the new method is better than IHS fusion on sharping information of details and maintain spectral characteristics of multispectral image.

Key words image fusion; steerable pyramid; IHS transform (Page:83)

Research on Wetland Landscape Change in Wuhan City Based on 3S Technology by SHAO Jun

Abstract This thesis had taken 3S technology as its research area for Wuhan City. It had analyzed the wetland landscapes pattern changes in Wuhan City in the recent eighteen years by using the Landsat-TM images in year 1991 as well as in 2003, and CBERS02B-CCD images of year 2009 as its data source, and by extracting wetlands information through the method of multi-scale image segmentation and visual interpretation. The results show that the city nearly eighteen years total area of wetland reduced ,the number of patches increased, fragmentation level increased, mean patch shape index reduction, wetland landscape diversity index decreased , combined with the decrease of human activities on wetland landscape disturbance increased, urban expansion has greatly affected the wetland landscape pattern changes, urban space makes wetlands shrinking trend, wetland protection is not optimistic.

Key words 3S Technology, dynamic change,wetland landscape pattern,Wuhan City (Page:85)

Landscape Pattern Changes of Rural Residential Areas in Hilly Country of Southern Hunan by WANG Binwu

Abstract Based on GIS, the theory of landscape ecology, and the method of quantitative analysis, the landscape pattern changes of rural residential area were analyzed in Beihu District in Chenzhou City in the recent 5 years. Many landscape indices were calculated which expresses the rural settlements amount, the scale of the rural residential, spacial formation and spacial distribution, while analyzed the dynamic changes of the landscape pattern at different altitude rural residential. The results showed that rural residential areas generally gathered in gentle slope zone. The area of rural settlements increased by 189.27 hm2 in Beihu District from 2004 to 2009. The rural residential mean patch size (MPS) increased by 5.57% in hilly area of Beihu District within these 5 years. On the other hand, about spacial formation, the Patch fractal dimension and shape of the index increased respectively 0.031and 0.189 form 2004 to 2009, the patch of rural residential area was becoming smaller and smaller. For spatial distribution, the blocks neighbor distances became continuously short, the mean nearest neighbor distance(MNND) decreased by 13.29%, showing the characters of agglomerate distribution.

Key words GIS,rual residential area, landscape pattern, Beihu District (Page:89)

Coastal DEM Construction and Accuracy Assessment Based on ArcGIS by REN Yuesen

Abstract Taking the building of Fujian Jiulong River coastal DEM data for example, the paper introduced a method of Coastal DEM construction and accuracy assessment based on ArcGIS through integrated the land, marine and foreshore flat areas data. The accuracy assessment result shows that the Jiulong River area Coastal DEM of Fujian meets national standards of DEM production. So, this paper provides a technical support for the application of GIS in the coastal resource integrated management based on information.

Key words ArcGIS, Coastal, DEM construction, Accuracy assessment (Page:93)

Research on Land Use/ Cover Dynamic Changes with Remote Sense in Guang'an City by Chen Nan

Abstract As society to speed up the process of industrialization, the contradiction between economic development and ecological environment become more and more prominent, so it is an important way that carry out the typical area ecological environment remote sensing dynamic monitor analysis to recognize the present situation of the nation land utilization changes. Based on the remote sensing and GIS technology, this paper analyzed land use/cover dynamic degree changes and transfer changes of the region between 2000-2007 through establishing the Guangan city land utilization database, and established the land use/cover of the initial state matrix and transition probability matrix, and with Markov model to predict the area between 2014-2021 land use/cover the evolution of trends. The results show that the farming and the lawn area of this region reduces year by year, the rate of rise of the forest land area is quick. Although the Guangan returning farmland to forest project has made the remarkable progress, but has actually created the farming and the lawn fierce reduction, does not favor the ecological environment the balanced development, therefore should protect the farming and the lawn, emphasize the economy and environment sustainable development.

Key words LUCC,transfer matrix,Markov forecast model,Guang'an (Page:96)

Inversion of the Land Temperature Using Split-window Algorithm by WANG Limei

Abstract On the basis of analyzing the split window algorithm, some commonly employed algorithms were introduced. The Qin etal split window algorithm was analyzed and selected as the best algorithm for temperature retrieval. Radiometric calibration, atmospheric and geometric correction for MODIS data were performed. At the same time, Valor algorithm sensitive to the reflection of the vegetation was used to e-valuate the specific radioactivity of land surface and the land surface temperature of Jingjintang region was hence retrieved. The result showed that the temperature of urban area was the highest, and the temperature of water body the lowest. These retrieval results were more ideal. Besides, Qin etal split window algorithm and Valor algorithm could improve the accuracy of temperature retrieval, offering a good example of application of temperature retrieval in complicated areas.

Keywords split window algorithm; thermal infrared remote sensing; temperature retrieval; JingJinTang area (Page:99)

Application of Density-based Clustering Algorithms in Noise Removing of Continuous Distributed Point Clouds by ZHANG Qiaoying

Abstract There are lots of noise data in the raw data except the target data. And the noise data always distribute around the target object irregularly, it is impossible to build a math model to make a distinction between the noise data and target data. The cluster is defined as the maximum set of density-collected in Density-based clustering algorithms, it can discover arbitrary shaped or sized cluster. To apply this algorithm it the noise removing of point clouds, can make the continuous

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

distributed points as a cluster, and then extract the target point clouds.
Key words Density-based clustering algorithms, the density distribution of point cloud, noise remove (Page:101)

Method of Basic Geographical Information Module Implementation Based on Google Earth by YIN Qiang

Abstract This paper introduced the mentality and the implementation method of kinds application subsystem general modul redevelopment based on Googel Earth platform. It elaborated the implementation method of control and browse module and geographical information module, and implemented format conversion of the shp to kml file in the system.
Key words GE, basic geographical information modul, kml, format conversion (Page:105)

Establishment of the Geographical Name Inquiring System of Fuxin City Based on MO by REN Dongfeng

Abstract This paper talked about establishing the geographical name inquiring system of Fuxin city based on MapObjects controlling and VB language and in the foundation of the geographical name geodatabase of Fuxin city. The system achieved the function of brose, layers management, drawing, the geographical name inquiring, the buffer analysis, the shortest path analysis.
Key words geographical name inquiring system; system design;buffer analysis; path analysis (Page:107)

Formulas of Calculation of Road Horizontal Curve Coordinates in the Route Plane Control Survey Coordinate System by ZHEN Dengchun

Abstract A method of direct calculation of road horizontal curve coordinates in the route plane control survey coordinate system is introduced, and the related formulas, compact and practical, can be referenced for setting out of road horizontal curve, are derived.
Key words road;horizontal curve;coordinate calculation (Page:111)

Calculation and Application of Various Area in the Second Land Investigation by ZHANG Hui

Abstract This paper analyzed working method and mathematical models of line and sporadic feature, summary the advantages and disadvantages of various area proposed the improving and using direction, by Comparison between calculation formula of ellipsoid area and working method and precision assessment in the first land using status investigation.
Key words land investigation; area mature; precision assessment (Page:115)

Investigation about the Subdivision of the Digital Estate Figure of Wuhan by CHEN Zhen

Abstract The subdivision of the estate figure is apart of the plan of the real estate framing, and it's the basic figure of drawing and issuing the figure of the license of the estate. According to the provision of the property management at Wuhan, there are two ways of surveying and mapping the subdivision the figure. Framing is the basic unit of surveying and checking of the estate which is a very important code at surveying and management, and it is also the major index at the management of the records. The standardization of surveying and mapping the subdivision the estate figure is benefit for the department of the estate management, which can also support the service of the department. This thesis showed us some research about the surveying of boundary points, the coordination of the corner of the buildings, the serial number of the buildings and so on.
Key words the subdivision of the estate figure, express content, investigate of the technique (Page:118)

Role of Detection of Underground Pipeline in Municipal Engineering Design by XIAO Shun

Abstract Underground pipeline survey before carrying out municipal engineering is very important. This issue illustrated this significance by explaining the important role detailed municipal pipeline survey plays in municipal engineering, comparing between detailed municipal pipeline survey and underground pipeline survey and their pre- and follow-up services. Several illustrative cases were provided to enhance the conclusion.
Key words municipal engineering design, detailed municipal pipeline survey,detection of underground pipeline (Page:121)

Design and Analysis of the Deformation Monitoring Program about a Foundation Ditch in Chengdu by LI Yong

Abstract This paper summarized the foundation excavation monitor need pay attention to in the basic problems and general principles,and combining QingyangQu red east street in a Chengdu deformation observation projects,analyzed the project operation processes involved with some typical problems,including project profiles, benchmark layout,observation period and so on contents,and combined with actual situation corresponding conclusion.
Key words foundation ditch,benchmark,observation period (Page:125)

Optimum Design of CPIII Plane Control Network for High Speed Railway by XIAO Daiwen

Abstract By doing the simulation optimum design, the positional accurary,relative positional accurary and reliability of CPIII networks was analysed, and the result showed the reliability of CPIII network was bader. The optimum scheme of CPIII network was presented. And frequency of repeatable measurement of this CPIII network may was reduced.
Key words CPIII plane control network,positional accurary,reliability,optimum design,ballastless track (Page:127)

Thoughts of Surveying and Mapping Engineering Supervision by PENG Songlin

Abstract This paper starts with the analyzing the origin of relation and distinction of engineering supervision and project supervision, to discuss the need for the implementation of mapping and project supervision, and how could it be practiced. The focus is on how important the organization, legal system, market construction and other work are in promoting mapping and project supervision.
Key words supervision; engineering supervision of surveying and mapping;organization construction;legal system construction;market construction (Page:130)

Design and Practice of Deformation Monitoring of Building by FU Hai'ou

Abstract This paper expounded the design of the building's settlement monitoring process to Chengdu general tablet research building structural template Co., LTD as an example, the level of the stability analysis, combining results point on the watch for observation data statistics and analysis, and a detailed corresponding conclusion.
Key words subsidence monitoring,baseline point, stability (Page:133)

Application of Regession Analysis Model in Dam Deformation Monitor by YANG Yongchao

Abstract This article focused on a regression analysis to monitor dam

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net