

基于 GIS 数据库的田野考古地层剖面空间数据挖掘 ——以陕西临潼姜寨遗址为例

杨 林¹, 闫国年¹, 毕硕本¹, 裴安平², 盛业华¹, 陈济民¹, 孙懿青¹

(1. 南京师范大学地理信息科学江苏省重点实验室, 江苏 南京 210097; 2. 南京师范大学文物与博物馆系, 江苏 南京 210097)

摘要: 介绍一种基于地理信息系统(GIS)数据库进行田野考古地层剖面空间数据挖掘的方法。即利用虚拟钻孔扫描剖面地层获得相关信息, 从中提取田野考古中所需的各文化层及遗迹的厚度、顶界和底界深度等空间数据。实验结果表明, 该方法能自动获取地层信息且数据可信度较高。

关键词: 地理信息系统; 田野考古; 数据挖掘; 虚拟钻孔

中图分类号: TP311; P208 **文献标识码:** A **文章编号:** 1672- 0504(2005) 02- 0028- 04

GIS 应用于考古学研究始于 20 世纪 80 年代初欧洲和北美一些国家的遗址预测及景观考古, GIS 技术(与遥感技术结合)在我国考古领域的应用集中在地面和地下文物普查, 如 2002 年华东师范大学城市与环境遥感考古开放研究实验室完成上海地区遥感考古调查, 发现 200 余处^[1]有古遗址特征的地点。上述研究针对具体考古点的很少, 本文结合田野考古地理信息系统(FAGIS)^[2], 介绍基于 GIS 数据库的田野考古地层剖面空间数据挖掘方法^[3]。

传统的田野考古地层数据获取基于手工量测, 费时费力且精度难以保证。利用虚拟钻孔(虚拟洛阳铲)扫描数字化并经过配准的地层剖面矢量图形获得相关信息, 对文化层进行空间矢量数据挖掘, 提取考古所需的各文化层的厚度、顶界和底界深度及遗迹的厚度、口径、底径、顶界和底界深度等空间数据。本文基于陕西临潼姜寨遗址的发掘报告^[4], 以报告中的 11 个地层剖面图为实验对象, 探讨基于 GIS 数据库对文化层进行空间数据挖掘的方法。实验结果表明, 这种方法能自动获取可信度更高的地层信息, 是 GIS 方法应用于田野考古的一次成功尝试。

1 遗址文化层的分布

姜寨遗址位于陕西省临潼县, 西南紧靠临河, 南依骊山, 北望渭水, 是新石器时代聚落遗址, 属于仰韶文化半坡类型^[4], 其探方分布如图 1。遗址地层堆积比较复杂, 原因有二: 一是延续时间较长, 从仰韶早期到龙山晚期都有人居住, 大约延续两千多年; 二是各

时期在此居住的人数不同, 形成的村落大小各异。故有些地段只有一个时期的堆积, 有些地段则有两个甚至三四个时期的叠压堆积。在叠压堆积的地段内, 有的紧紧相接, 有的只有间隔的叠压关系。具有这种地层叠压关系的遗址, 是考古工作中比较理想的遗址^[4]。

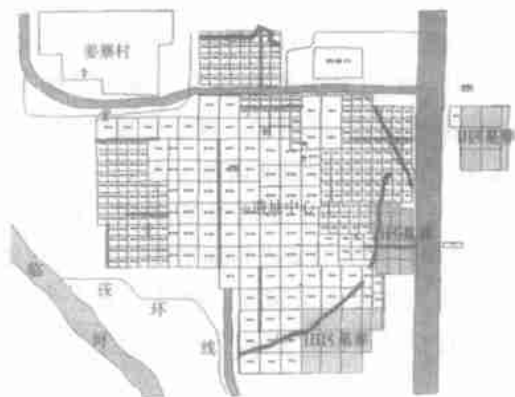


图 1 姜寨遗址探方分布
Fig. 1 Distribution of plots in Jiangzhai Site

根据文献[4]对其中 11 个剖面文化层土质、土色、包含物与出土遗迹、遗物文化特征的描述加以综合整理、分析, 得出姜寨遗址 5 个时期文化层由上至下的垂直分布情况(以遗址西部 T 198 东壁剖面为例说明): 农耕土, 厚 0.20~ 0.28 m; 灰褐色土, 质松, 厚 0.32~ 0.56 m, 出土遗物特征与半坡晚期、西王村仰韶晚期相同, 属姜寨第四期文化遗存; 深灰色土, 质松软, 厚 0.32~ 0.52 m, 遗物较少, 出土遗物特征与庙底沟仰韶遗存相似, 属姜寨第三期文化遗存; 浅灰色土, 夹有大量红烧土块, 质较硬, 厚约 0.32~ 0.52 m, 出土遗物的文化特征均与半坡早期相同, 属姜寨

收稿日期: 2004- 10- 18; 修订日期: 2005- 01- 11

基金项目: 国家“863”计划项目“虚拟地理环境系统的研究与开发”(2001AA135130); 南京师范大学重点项目“虚拟地理环境下三维地理空间数据模型与数据处理方法研究”(2004105X GQ2B46)

作者简介: 杨林(1976-), 女, 博士研究生, 主要研究方向为数字摄影测量和地理信息系统。E-mail: yangejus@126.com

一期文化遗存^[4]。

2 文化层的标准层与剖面图形库结构制定

本文首先制定包含全部文化层的标准文化层。在此基础上,对文献[4]中的 11 个剖面地层图进行数字化、配准、建立拓扑关系和属性赋值等处理,对姜寨遗址各文化层进行界定,最后对剖面上文化层、窑穴等遗迹进行空间数据挖掘。

2.1 标准文化层及其界定

通过对姜寨遗址各时期文化层分布情况的总结可知,该遗址标准文化层可划分为 9 层(表 1)。基于标准文化层对 11 个剖面进行文化层界定(表 2)。表 2 中“一”等层号为文献[4]中地层的原始编号,“生”表示该剖面中的生土层。

表 1 姜寨遗址标准文化层
Table 1 Standard cultural stratum of Jiangzhai Site

标准层编号	原始文化层
壹	农耕土层
贰	扰土层
叁	客省庄二期(姜寨五期)文化层
肆	半坡晚期文化(西王村仰韶文化或姜寨四期)层
伍	庙底沟仰韶文化(姜寨三期)层
陆	史家仰韶文化(姜寨二期)层
柒	半坡早期(姜寨一期)文化层
捌	腐殖质土层
玖	生土层

表 2 各剖面地层文化层分布
Table 2 Cultural stratum distribution of every section

文化层与各剖面	壹	贰	叁	肆	伍	陆	柒	捌	玖
东部 T 84~ T 106 北剖面	—	—	—	—	—	—	三	—	生
T 1~ T 4 北剖面	—	—	—	—	—	—	二、三	—	生
南部 T 263~ T 265 西剖面	—	—	—	—	—	—	三 A、三 B、四	—	生
T 198 东剖面	—	—	—	二	三	—	四	—	生
西部 T 135~ T 138 西剖面	—	—	—	—	—	—	二、三、四	—	生
T 205~ T 214 南剖面	—	—	—	二	三	四	五	—	生
西北 T 246~ T 248 南剖面	—	—	—	—	—	—	三	四	生
北部 T 180~ T 184 东剖面	—	—	—	—	—	—	二	三	生
T 143~ T 153 北剖面	—	—	—	—	—	—	四	—	生
中部 ZHT 2~ ZHT 7 西剖面	—	—	—	二	三	四	五 A、五 B	六	生

2.2 剖面图 GIS 数据库逻辑结构

与剖面文化层数据挖掘相关的数据库均以中文拼音缩写命名,各数据库的逻辑结构简要说明如下。
1) 地层空间信息(DCKJXX):文化层编号、最大厚度、最小厚度、平均厚度、顶界最大深度、顶界最小深度、顶界平均深度、底界最大深度、底界最小深度、底界平均深度、IP(流水号);
2) 地层累计空间信息(DCLJKJXX):结构同 1);
3) 地层遗迹空间信息(DCYJKJXX):文化层编号、最大深度、最小深度、平均深度、最大口径、最小口径、平均口径、最大底径、最小底径、平均底径、顶界最大深度、顶界最小深度、顶界平均深度、底界最大深度、底界最小深度、底界

平均深度、IP; 4) 地层累计遗迹空间信息(DCLJYJKJXX):结构同 3);
5) 虚拟钻孔表(XNZKB):虚拟钻孔编号、钻孔与地层交点的横坐标、钻孔与地层交点的纵坐标、地层编号、地层厚度、流水号、虚拟钻孔是否与地层只有一个交点、土结构、土颜色、土中杂质含量、块粒大小、坚硬程度、地层最大深度、地层最小深度、地层最大厚度、地层最小厚度、出土陶器种类、出土陶器数量、其它遗物种类、其它遗物数量、遗迹编号;
6) 特征点属性表(TZDSXB):流水号、特征点坐标、特征点与左边最近虚拟钻孔的距离、特征点与右边最近虚拟钻孔的距离、特征点左边最近虚拟钻孔编号、特征点右边最近虚拟钻孔编号、特征点编号。

3 文化层的空间数据挖掘

基于 GIS 数据库对文化层的空间数据挖掘,主要是基于矢量图形的空间数据挖掘,但包含少量必不可少的文化层属性数据,如“所属文化层”等。主要挖掘内容为姜寨遗址中 11 个剖面的文化层最大厚度、最小厚度和平均厚度,顶界和底界距地表的深度、最小深度和平均深度及文化层的缺失情况。这对于田野考古地层学分析文化层层位关系及预测地层走势等具有重要意义。

3.1 文化层空间数据挖掘过程

对文化层进行空间数据挖掘时,虚拟钻孔按一定的间隔扫描整个剖面,虚拟钻孔在和文化层相交的过程中,每个文化层有上、下界两个相交点,通过记录这两点的纵向数据值,再求差,得到该文化层在这条虚拟钻孔位置处的厚度(扫描间隔越小,虚拟钻孔越多,所求厚度越准确)。通过对多个虚拟钻孔的统计,可以获得当前剖面上各文化层的最大厚度、最小厚度和平均厚度($\bar{L}_j = \frac{1}{n} \sum_{i=1}^n (L_i)$, L_i 为各钻孔在该文化层的厚度(纵向数据差), $\bar{L}_j$ 为某文化层在第 j 个剖面的平均厚度)。如果某一层平均厚度为 0,说明这个文化层在该剖面缺失。一个剖面计算完成后,可以存到地层累计空间信息表中,具体方法如下:
1) 对相同的极值参数项:对应项进行比较,大者取更大,小者取更小,再计算、累加其它剖面。
2) 对相同的平均值参数项:当累计到第 j 个剖面时,平均厚度 = $(L \times (j-1) + L_j) / j$ (L 为当前累计的平均厚度)。顶界和底界距地表深度的计算方法与厚度相同。
3) 通过设置虚拟钻孔的扫描间隔,提取虚拟钻孔的文化层厚度、深度等数据,然后把当前剖面的各文化层

厚度、深度等数据添加到文化层累计库中。

3.2 文化层空间数据挖掘的结果

以虚拟钻孔扫描间隔 0.5 m 为例,对姜寨遗址西部 T198 东壁进行文化层厚度、深度等空间数据挖掘,其结果如表 3 所示。对比“遗址文化层的分布”中给出的属性数据,可以看出基于 GIS 数据库的空间数据挖掘准确度更高,信息更全面。以虚拟钻孔扫描间隔 2 m 为例(根据精度要求不同,可调整钻孔扫描间隔),对姜寨遗址的 11 个剖面进行文化层厚度、深度等的空间数据挖掘(表 4)。

通过姜寨一期文化层的分层、厚度等资料和对 11 幅剖面的对比分析,推断姜寨遗址的原中心广场为一个深度大于 2 m(相对于一期文化层的顶界而言)的圆形洼地^[5]。通过对遗址不同断面地层图(已配准的地层剖面矢量图形)进行虚拟钻孔扫描,可以获得不同断面上各相同地层表面的若干离散点(钻孔与地层交点)的三维坐标,利用带约束条件的格网算法(TIN 或 GRID)生成各地层的表面模型,进而生

成整个遗址地层的数字高程模型(DEM),实现数据挖掘描述和可视化的功能。这为再现远古地层的空间层位关系奠定了良好的数据基础,更重要的是可以预测未挖掘遗址的地层走势与大致形态,实现数据挖掘估值功能。如果考虑到各遗物(器物)和遗迹在地层的原始空间位置,将更全面的展示原始地层的丰富信息。随着断面和钻孔间隔的减小,所展现的地层 DEM 将更真实。

4 地层剖面上遗迹的空间数据挖掘

基于 GIS 数据库对剖面上窑穴等遗迹的空间数据挖掘,主要是基于矢量图形的空间数据挖掘,也需要少量的文化层属性数据,如所属文化层、遗迹名等。挖掘的主要内容是姜寨遗址中 11 个剖面上各文化层中遗迹的最大厚度、最小厚度、平均厚度、最大口径、最小口径、平均口径、最大底径、最小底径和平均底径,以及该遗迹顶界和底界距地表的最大深度、最小深度、平均深度等。

表 3 T 205~ T214 南壁文化层矢量数据挖掘结果 (单位:m)
Table 3 Data mining of south cultural sections of T205~ T214

文化层	最大厚度	最小厚度	平均厚度	顶界最大深度	顶界最小深度	顶界平均深度	底界最大深度	底界最小深度	底界平均深度
壹	0.321	0.186	0.237	0	0	0	0.321	0.186	0.237
贰、叁									
肆	0.629	0.226	0.428	0.321	0.186	0.237	0.846	0.492	0.673
伍	0.536	0.427	0.477	0.846	0.492	0.673	1.309	1.034	1.150
陆									
柒	0.599	0.273	0.432	1.309	1.034	1.150	1.736	1.472	1.582
捌									
玖	0.328	0.177	0.268	1.736	1.472	1.582	1.918	1.735	1.850

表 4 文化层矢量数据挖掘结果 (单位:m)
Table 4 Data mining of 11 sections

文化层	最大厚度	最小厚度	平均厚度	顶界最大深度	顶界最小深度	顶界平均深度	底界最大深度	底界最小深度	底界平均深度
壹	0.713	0.086	0.223	0	0	0	0.713	0.086	0.223
贰	1.309	0.148	0.359	0.376	0.086	0.259	1.521	0.307	0.787
叁	2.362	0.209	0.613	1.521	0.181	0.682	3.670	0.536	1.435
肆	2.415	0.002	0.590	3.670	0.133	1.030	4.996	0.284	1.620
伍	0.542	0.391	0.476	0.782	0.492	0.664	1.243	1.034	1.140
陆	0.846	0.118	0.507	3.299	1.098	2.739	3.801	1.393	3.246
柒	1.632	0.046	0.581	4.223	0.122	1.467	4.658	0.368	2.035
捌	0.628	0.189	0.354	1.213	0.631	0.793	1.642	1.045	1.180
玖	2.823	0.084	0.779	4.996	0.149	2.467	6.087	1.408	3.467

4.1 遗迹空间数据挖掘的基本过程

对剖面上遗迹的厚度、深度进行空间数据挖掘与对文化层厚度、深度的空间数据挖掘过程和方法基本相同。关于窑穴等遗迹的口径、底径、顶界和底界距地表深度,由于遗迹形状的复杂性,采用任意设置虚拟钻孔的方法,即利用人机交互的方式,人工确定遗迹口径、底径的边界位置。该遗迹的口径(或底

径)分别为遗迹左右边界点配准后的坐标,即每个遗迹上左右虚拟钻孔点的斜距。计算完剖面上所有文化层的全部遗迹的口径(或底径)后,对同一文化层上同类遗迹比较其最大、最小口径和最大、最小底径,然后计算该剖面上各文化层中同类遗迹的平均口径、底径,一个剖面计算完成后,可以存到地层累计遗迹空间信息表中。遗迹顶界和底界距地表深度的计算方法与厚度相同。图 2 为对剖面上窑穴进行空间数据挖掘的界面。

4.2 遗迹空间数据挖掘的结果

T84~ T106 北剖面上的窑穴进行空间数据挖掘, 结果以虚拟钻孔扫描间隔 0.5 m 为例, 对遗址东部 见表 5, 其中空白部分表示该文化层没有窑穴存在。

表 5 对窑穴进行矢量数据挖掘的结果(单位: m)
Table 5 Data mining of stoves

文化层	最大深度	最小深度	平均深度	最大口径	最小口径	平均口径	最大底径	最小底径	平均底径	顶界最大深度	顶界最小深度	顶界平均深度	底界最大深度	底界最小深度	底界平均深度
壹~叁肆伍陆柒捌、玖	1.056	0.582	0.718	1.312	0.566	0.939	0.949	0	0.475	1.301	0.206	0.537	2.725	1.603	2.366
	1.185	0.079	0.787	2.564	0.445	1.064	1.918	0.912	1.049	1.079	0.441	0.597	2.093	1.356	1.453

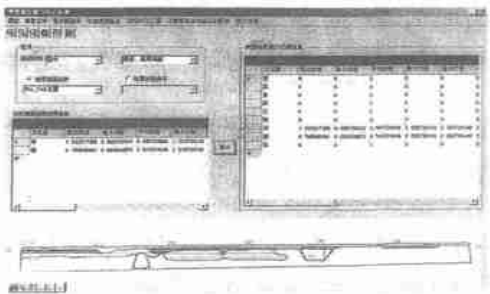


图 2 利用虚拟钻孔对窑穴进行数据挖掘
Fig. 2 Data mining of stove with virtual bore

5 结语

田野考古是数据挖掘的一个新的应用领域, 在对姜寨遗址各时期文化层分布情况全面分析及制定姜寨遗址标准文化层的基础上, 对各时期的文化层进行基于 GIS 数据库的空间矢量数据挖掘分析。实验结果表明, 不仅改进了传统的数据获取方式, 而且

地层数据准确度更高, 信息更全面, 为定量分析田野考古地层相对早晚关系^[6], 建立整个遗址地层的数字高程模型, 恢复远古生活场景奠定了很好的基础。

参考文献:

[1] 陈德超, 刘树人, 陈中原, 等. 上海地区古遗址遥感初步研究[J]. 华东师范大学学报(自然科学版), 2002(3): 83- 87.
[2] 李安波, 毕硕本, 裴安平, 等. 田野考古地理信息系统研究与建设[J]. 地理与地理信息科学, 2004, 19(1): 39- 42.
[3] 邱凯昌. 空间数据挖掘与知识发现[M]. 武汉: 武汉大学出版社, 2001.
[4] 西安半坡博物馆. 姜寨——新石器时代遗址发掘报告(上、下)[M]. 北京: 文物出版社, 1988.
[5] 毕硕本. 聚落考古中空间数据挖掘与知识发现的研究——以史前聚落半坡类型姜寨遗址为例[D]. 南京: 南京师范大学出版社, 2004. 33- 57.
[6] 冯恩学. 田野考古学[M]. 长春: 吉林大学出版社, 1993.

Data Mining of Field Archaeology Stratum Based on Geographical Information System Database:
A Case Study of Jiangzhai Site

YANG Lin¹, LV Guo- nian¹, BI Shuo- ben¹, PEI An- ping², SHENG Ye- hua¹, CHEN Ji- min¹, SUN Yi- qing¹

(1. Jiangsu Provincial Key Lab. of Geographical Information Science, Nanjing Normal University, Nanjing 210097;

2. Department of Cultural Relic & Museum, Nanjing Normal University, Nanjing 210097, China)

Abstract: In order to improve traditional data- acquiring method which is based on toilsome manual survey and the data is not accurate enough to be used in precise analysis, the article introduces a method of stratum data- mining of field archaeology based on geographical information system(GIS) database. There are 11 sections in Jiangzhai Site, Lintong, Shanxi in China, which are instanced for testing the method. The pretreatment is a series of GIS progress including framing standard cultural stratum of Jiangzhai Site and demarcating the different stratum of the 11 sections, digitizing, transforming and building topology of the 11 sections, and then setting down the framework of the GIS database and evaluating the attributive result. After that, spatial data- mining is done, namely, with virtual bore scanning section stratum maps which have been digitized and transformed, to distill spatial data from the information of the intersection of virtual bore lines and section stratum(lines). The spatial data involve the width, top length and bottom length of stratum and relic, the top and bottom caliber of relic, which are necessary in field archaeology. The scanning interval could be altered and the less interval value, the more accurate the data. In addition, according to some simple arithmetic that data could be made new information to meet some other need in field archaeology analysis, even describing and visualizing the stratum. At the same time, the data of the same cultural stratum in different section could be added up to acquire more information of some larger region, predicting and evaluating the distribution of the relic. It is proved that not only kinds of stratum information can be gotten automatically through the method, but also the information is more believable and abundant than the original data in related documents.

Key words: geographical information system; field archaeology; data- mining; virtual bore