GIS 支持的 Eigenface 法
——适于只有一个模型单元时的矿产定量预测方法

张振飞 ,高凤亮 ,马智民 ,姜常义
(长安大学资源学院 ,西安　710054)

[摘　要] 在研究和开发程度较低地区开展矿产预测 ,往往只能有很少矿产地用以建立定量预测模型 ,此时大多数统计分析方法不适用。一种 GIS 支持的基于“单元簇”概念的多源地学信息综合分析方法即 Eigenface 法 ,比较适合于只有一个模型单元时的建模。单元簇是相邻若干网格单元的空间定量组合 Eigenface 法先求出这种高维组合变量空间的某个低维特征子空间 ,然后计算未知单元簇和已知单元簇在特征子空间上投影点的距离来评价未知单元的找矿有利性。Eigenface 法与 GIS 集成是将单元作为区图元 ,利用 GIS 空间 - 属性分析功能提取预测信息并对单元和单元簇进行操作。以新疆乌斯拉台 — 亦格尔达坂地区铜铅锌银多金属矿产预测为例说明了其应用效果。

[关键词]矿产预测　Eigenface 法　GIS　新疆
[中图分类号]P628　[文献标识码]A　[文章编号]0495 - 5331(2001)06 - 0051 - 04

0　引 言

在研究和开发程度较低地区开展矿产预测 ,往往只能有很少的矿产地用以建立定量预测模型 ,此时大多数传统的统计分析方法应用效果不好[1]。如果利用其它地区已建立的成矿模型 ,则可能会有过多主观因素介入 ,影响预测可靠性[2]。因此 ,研究只有少量已知矿产地情况下的预测建模方法是很有必要的。本文介绍一种新的矿产预测方法 ,即 Eigenface 法 ,它适于只有一个已知矿产地情况下的预测。该方法与其他统计预测方法一样也划分基本单元[1,3] ,不同的是以网格状单元的空间组合(“单元簇”)为定量分析对象 ,从而可能捕捉到地质变量组合所隐含的有关成矿的局部空间结构信息。将 Eigenface 法与 GIS 加以集成 ,对单元和单元簇的管理操作及变量数据提取都可利用 GIS 的空间 - 属性分析功能来实现。论述了该方法的原理、算法 ,并以新疆乌斯拉台 —亦格尔达坂地区铜铅锌银多金属矿产预测为例说明其应用效果。使用了 MAPGIS v. 5. 32(武汉华地公司 ,1997) 。

1　Eigenface 法原理及其与 GIS 的集成

1. 1　“单元簇”概念

作者曾对“单元簇”概念作过初步探讨[4] ,这里作一些补充说明。设矩形网格状基本单元划分方案已确定 ,每个单元可用 p 个地质变量来描述。一个单元及其周围 n 圈[最多 $(2n+1)^2$ 个]单元构成一个单元簇。最中间的单元称中心单元 ,周围的称外围单元。构成单元簇的每个单元称该单元簇的组成单元。单元簇组成单元的个数叫做它的体积(V)。显然 ,一个单元簇各组成单元的相对空间关系是惟一确定的。以控制单元和未知单元为中心的单元簇分别称已知簇和未知簇。称第 k 个组成单元与它所在单元簇之间非空间关系的密切程度(即重要程度)为该组成单元的簇权。计算各组成单元中地质变量定量组合的过程称为标度。通过标度 ,单元簇可以反映任意复杂的局部地质结构。本文所述 Eigenface 法是单元簇标度的途径之一。

1. 2　基于单元簇概念的 Eigenface 法的原理与算法

Eigenface 法是近几年发展起来的一种图像识别方法 ,如 Turk 等用之于人脸的识别[5] , Essa 及赵力庄等用该方法识别人脸的表情[6,7] ; Sclarroff 用该方法进行三维物体的模板匹配、跟踪[8]。该方法可看作是一种包含空间分析的多变量统计分析 ,可以用于多源地学信息的综合分析与建模预测 ,但目前在矿产预测中还未见应用的报道。赵力庄等详细介绍了人脸图象分析中 , Eigenface 法的计算过程[7]。其要点是将人脸 $N \times N$ 个点构成的图像看成作 N^2 维的变量空间(每个点为一个变量)。选择不同表情的多个典型人脸 ,计算出一个 $N^2 \times N^2$ 维的协方差矩阵 C。计算 C 的特征值 v_i 及相应的特征向量 u_i。可以取最大的一个或多个 v_i 所对应的 u_i ,构成人脸特征

[收稿日期]2001 - 01 - 14 ;[修订日期]2001 - 05 - 14 ;[责任编辑]曲丽莉。

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

子空间,每个具体人脸在特征子空间上投影位置不同,根据该位置可以对表情进行判别。

在用于矿产预测时,将单元簇当作"脸"图象,但处理方法需改造。本文把该方法改造成一种适于单目标预测的方法。因只涉及常规和简单计算,故只将主要作法简述如下:

1)单元簇不同组成单元中的相同变量视为不同变量。因为地质因素控矿、示矿往往有显著的方向性,例如断层上盘和下盘常常有不同的含矿性[5],故对于与一个矿产地有不同空间关系的同一类地质因素或标志,可以区别对待。

2)所有变量的平均值视为 0,用一个已知簇计算协方差矩阵。即用 pV(变量数 ×单元簇体积) 维向量与其转置相乘,得 $pV \times pV$ 维的协方差矩阵。

3)已知簇特征子空间的维数主要参考几个最大特征值的和占所有特征值总和的比例来定,例如,如果最大的一个特征值已占到特征值总和的 60 %以上,取 1 维;如果两个最大特征值的和才可以占到 60 %以上,取 2 维,余类推。

4)通过计算和比较未知簇在特征子空间的投影点与已知簇投影点的距离(欧氏距离,以下称有利性指标),对未知簇中心单元的找矿有利性作出评价。距离越小,反映未知簇与已知簇在对应的组成单元中变量组合越相似。

1.3 基于单元簇的 Eigenface 法与 GIS 的集成

本文将所述方法与 MAPGIS 进行集成,利用 MAPGIS 的开发库编写了包括 Eigenface 法在内的矿产预测应用程序。将单元作为区实体,是集成的关键。单元划分形成两个文件("单元.wp"和"单元.wl"),纳入由地、物、化、遥资料构成的空间数据库(在本文的实例中化探和遥感资料缺),利用 GIS 作空间 - 属性分析。通过建立单元簇类实现对单元簇的操作。单元簇类封装其组成单元号数组指针、变量数据数组指针以及获取数据的成员函数。通过 ado(active data object)接口对属性数据进行操作。建立一个 C ++ 类库封装不同计算方法(包括 Eigenface 法)。还有一些C ++ 类用于完成预测的其它步骤如预测范围确定、网格单元划分、预测方法选择、变量提取、数据表生成及成果图的生成、显示等。

2 在新疆乌斯拉台—亦格尔达坂地区铜铅锌银多金属矿产预测中的应用

2.1 研究区概况及地质矿产空间 - 属性数据库

研究区位于新疆东部,东经 86°～89°,北纬 42°～43°之间,面积约 14 000 km^2。横跨北天山、中天山、南天山 3 个大地构造单元,地质条件比较复杂。区内铜矿产有中型矿床 1 个(彩华沟),小型 2 个,矿(化)点一批,均为火山岩型,产于早古生代火山岩系中;铅锌矿产有小型矿床 1 处,矿(化)点若干,为层控型,产于泥盆系碎屑岩系中;银矿产仅有矿点 1 处,为层控型,产于下石炭统碎屑岩中。其他矿产如砂金、铁等本文未考虑。根据资料情况建立多个图层图形文件及相应的属性表,作为矿产预测依据。资料平均精度可以达到 1∶200000。

2.2 Eigenface 法的应用

2.2.1 应用的思路

以地质成矿规律为基础是矿产预测的一条重要原则,但由于地质成矿过程的长期复杂性和多因素特点,在预测过程中需要估计到有一些地质研究难以察觉的间接、隐蔽、抽象和综合信息的存在。包括 Eigenface 法在内的矿床统计预测通过对大量地质变量的综合分析来挖掘这些信息。一般来说难以对这些信息作出清楚具体的描述,但它们可能已存在于所建立的数学模型中而被利用了。在模型区稀少情况下(如本区银矿产预测),这一点尤应强调。因此在本区 Eigenface 法应用中重视两点:(1)经网格单元划分将研究区充分离散化之后,根据资料情况提取较多的变量。有些变量从地质观点来看未必显示出明确的成矿或找矿意义,但它们有可能携带有用信息,仍参与建模计算。(2)本区铜、铅锌矿床(点)数较多,其地质成矿特征相似,但分布于全区不同地段。故建立多个互相独立的铜矿(或铅锌矿)预测模型,比较它们的预测结果并与已知矿产分布情况进行比较,如有较好的吻合性,说明模型可靠。如果可以确认模型可靠,就可以确认模型中包含了充分的有用信息,所使用的方法能够利用极少的已知样品来挖掘较多的信息。这一点很重要,因为本区的银仅有一个已知矿点,预测结果无从检验,需要根据方法的可靠性来保证结果的可靠性。

2.2.2 预测范围确定、网格单元划分及变量提取

根据资料覆盖情况确定矿产预测的空间范围(形成"研究范围.wp")。在该范围内划分近正方形的网格单元 1236 个(单元面积约 1 km^2),如图 1。

从空间 - 属性库中提取多个变量,包括各种面状地质体的单元面积比(不考虑中生代及以后的地层和岩体)、单元岩性熵、断层数、断层交点数、断层优势走向、单元中心到各类地质体(岩体和断裂)的

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

最小距离、由物探等值线图提取的单元平均场值、最大梯度、等值线形态指标(定义为单元边界上 8 个点的场值升降变化次数)等。剔除部分分布单调、明显无分析意义的变量,得到 42 个基本变量。

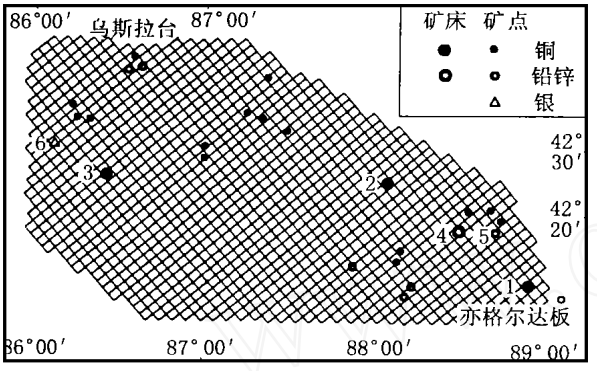

图 1　研究范围、基本单元划分及矿床(点)分布图

1 —彩华沟中型铜矿床;2 —可可乃克小型铜矿床;3 —巴仑台小型铜矿床;4 —马鞍桥小型铅锌矿床;5 —5 号铅锌矿点;6 —莫托萨拉银矿点

2.2.3　多金属找矿有利性计算结果及讨论

取单元簇体积为 9,特征子空间维数为 1,对铜、铅锌矿分别以不同矿床(点)所在单元为中心建立了多个 Eigenface 模型;对银矿以 6 号矿点所在单元为中心建立了一个模型。找矿有利性计算结果如图 2～6(铜和铅锌各只列出 2 个结果)。

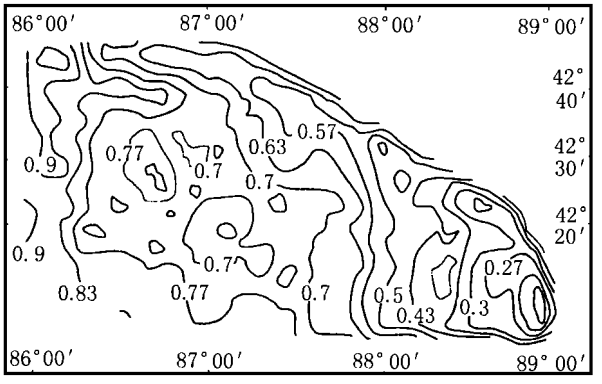

图 2　以彩华沟铜矿床所在单元为中心的已知簇所计算的铜矿找矿有利性指标等值线图

比较图 1、2、3 可以看出,铜矿找矿有利性的变化趋势与已知矿点的分布基本一致。以 1 号和 3 号矿床为模型单元,虽然这两个矿床空间上相距甚远,但预测结果有明显的一致性。绝大部分矿点的位置与有利性指标等值线的相对低凹处相吻合。铜矿找矿有利地段主要靠近研究区北东边部分布。

比较图 1、4、5,可以看出类似的特点。尤其在研究区南中偏东部,有若干铅锌矿点分布,该地段铅锌找矿有利性指标等值线坡度相对变缓。

以上结果可以说明所用方法的有效性,从而提供了一种解决本区银矿产预测中模型单元过少问题的办法。图 6 是银矿产找矿有利性指标计算结果。可看出除莫托萨拉银矿点一带以外,在研究区中偏南部也有等值线图相对低凹地段分布。参照铜和铅锌找矿有利性指标空间分布的规律性,可以推断本区莫托萨拉类型的银矿找矿方向,除该矿点外围就矿找矿,也不能忽视在研究区中部下石炭统碎屑岩分布区开展工作。

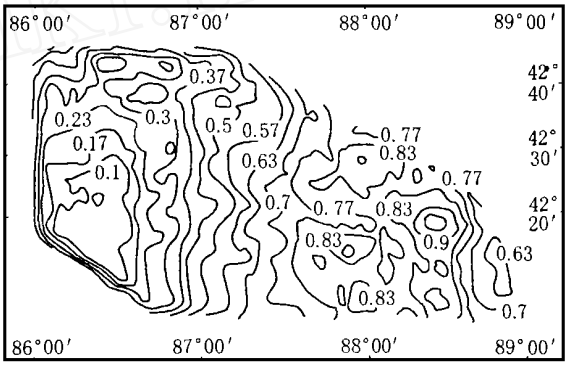

图 3　以巴仑台铜矿床所在单元为中心的已知簇所计算的铜矿找矿有利性指标等值线图

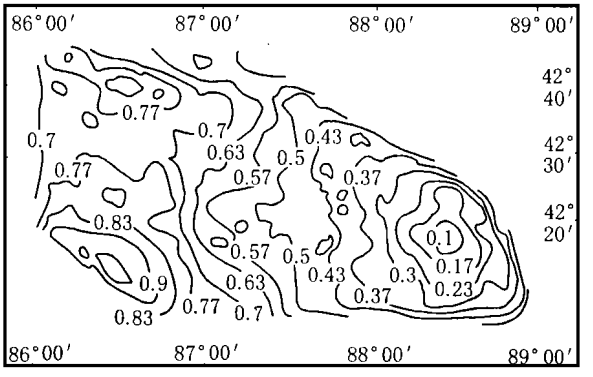

图 4　以马鞍桥铅锌矿床所在单元为中心的已知簇所计算的铅锌矿找矿有利性指标等值线图

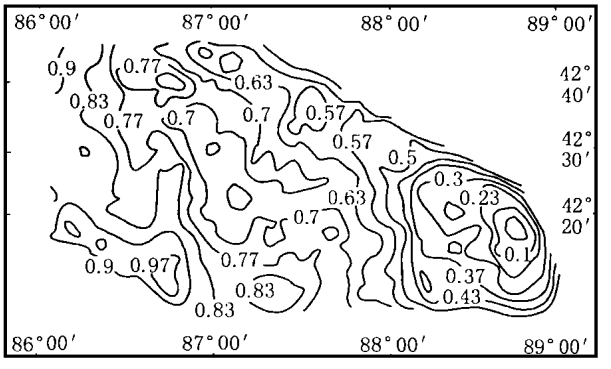

图 5　以 5 号铅锌矿点所在单元为中心的已知簇所计算的铅锌矿找矿有利性指标等值线图

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

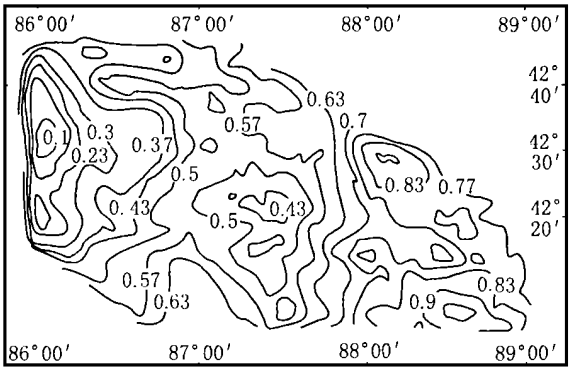

图 6　以莫托萨拉银矿所在单元为中心的已知簇所计算的银找矿有利性指标等值线图

以上实例说明,GIS 支持的 Eigenface 法可以用于只有一个模型单元情况下的区域矿产定量预测。进一步研究应注意根据实际合理确定单元簇体积、剔除找矿有利性指标的全区性趋势变化、提高预测精度。此外,研究区边部因单元簇不完整而有边界效应,在预测结果使用和地质解释时应予注意。

本文蒙中国地质大学李紫金、胡光道教授提出宝贵意见,谨表示衷心感谢!

[参考文献]

[1] 赵鹏大,胡旺亮,李紫金. 矿床统计预测[M]. 北京:地质出版社,1994,70～82.

[2] Wyborn L and Mernagh T. Using GIS for Mineral potential evaluation in areas with few known mineral occurrences[M]. In: Second National forum on GIS in the geoscience - forum proceedings, AGSO, 1995: 199～211.

[3] 王于天. 成矿预测单元的基本概念及其划分方法[J]. 地质论评,1990,36(6):24～29.

[4] 张振飞,胡光道,曾章仁. 矿产预测中空间地质结构的定量类比 - 单元簇的概念及应用[J]. 地球科学,1999,24(6):661～665.

[5] Turk M. Pentland A. Face recognition using eigenfaces[M]. In: Proc Computer Vision and pattern Recognition Conference, 1991: 586～591.

[6] Essa I Pentland A. Facial expresion recognition using a dynamic model and motion energy[M]. In: Proc International Conference on Computer Vision '95, Cambridge, MA, 1995: 20～23.

[7] 赵力庄,高文,陈熙霖. Eigenface 的变维分类方法及其在表情识别中的应用[J]. 计算机学报,1999,22(6):627～632.

[8] Sclarroff S Pentland A. Model matching for correspondence and recognition [M]. MIT Media Laboratory: Perceptual Computing Section Technical Report 304, 1995.

[9] 翟裕生. 矿田构造学概论[M]. 北京:冶金工业出版社,1982:153～195.

A MINERAL RESOURCE PREDICTION METHOD USING IN AREAS WITH ONLY ONE MODEL CELL: GIS - SUPPORTED EIGENFACES

ZHANG Zhen - fei, GAO Feng - liang, MA Zhi - ming, JIANG Chang - yi

Abstract: In lowly developed regions, there are often few known mineral occurrences for modeling for statistical mineral resources prediction. A GIS - supported modeling method, the cell - cluster based eigenfaces, can be used in areas with only one model cell. A cell cluster is a spatial association of adjacent raster cells, each described by p geovariables. The model is a subset of the eigenvectors of the covariance matrix of the known cluster. An unknown cell is taken as the center of a cluster and then evaluated by the Euclid distance from the projection of the cluster on the eigenvector subset to the projection of the known cluster on the same subset. Cells are taken as polygon elements thus cells and cell clusters can be manipulated by GIS. A case study of Cu, Pb - Zn and Ag resource prediction in Uslatai - Ygerdaban region, Xinjiang, is presented.

Key words: Mineral resource prediction, Eigenfaces, GIS, Xinjiang

[第一作者简介]

张振飞(1961 年 -),男,1982 年毕业于武汉地质学院矿产系,1985 年在武汉地质学院北京研究生部矿产普查勘探专业获硕士学位,2001 年在中国地质大学地球探测与信息技术专业获博士学位。现为长安大学资源学院副教授,主要从事数学地质和矿产勘查方面的科研、教学。

通讯地址:西安市雁塔路南段　长安大学雁塔校区资源学院　邮政编码:710054

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net