

文章编号: 1007-4619 (2004)04-0317-06

应用空间统计学理论解译遥感影像信息“缺失”区

冯益明, 雷相东, 陆元昌

(中国林业科学研究院 资源信息所, 北京 100091)

摘要: 借助图像处理软件 ERDAS、地理信息系统软件 ArcInfo 以及空间统计分析软件 ILWIS, 在对 TM 遥感影像进行分类的基础上, 运用空间统计学理论以及 Kriging 插值技术, 内插了影像真实信息“缺失”斑块的信息, 插值结果通过了精度检验。为解译影像信息“缺失”区, 提供了一种手段和方法。

关键词: 空间统计学; 遥感; GIS; 信息“缺失”区; Kriging 插值

中图分类号: TP79/O211.9 **文献标识码:** A

1 引言

我们在对遥感影像进行分类时, 经常会遇到由于受云、云阴影、建筑物或山体本影、烟尘等遮挡, 造成影像上部分区域真实信息的“缺失”的现象, 目前对信息“缺失”区这一问题的处理是把这些区域单划一类, 这样一来, 所作的分类结果, 就很难反映地面的真实情况。要解决这一问题, 有时可以通过采用另一时相的遥感资料来克服, 但需要花费一笔额外经费。借助信息“缺失”区周围影像信息, 内插估计“缺失”区信息, 是解决此类问题的一种手段。一般是采用传统的内插模型, 如最邻近法, 双线性重采样^[1,2], 但效果不好。由于地物在空间上的分布具有自相关性, 即在空间上越靠近的事物或现象就越相似, 如温度、水分、土壤特征、植被特征等在空间上的分布都反映了这种“近朱者赤, 近墨者黑”的现象, 地物的空间分布服从地学分布规律, 因此, 地学统计分析技术对于解决这类问题有独特的效果。

2 空间统计学简介

空间统计学是应用数理统计学的一个新分支, 空间统计学通过研究对象在空间上不同间隔的抽样点上的差异, 定量描述其空间变化规律。它主要包括两方面的研究内容: 第一为半方差图, 描述研究对

象的空间相关性的方法; 第二为空间局部内插理论, 主要为 Kriging 内插法, 通过空间上抽样点的调查数据, 对空间上未测点进行估计^[3-7]。将 Kriging 理论用于遥感影像分析时, 一般是用来描述影像纹理 (image texture), 空间结构 (spatial structure) 和插值 (interpolation)^[8,9]。在国外, Silvana Amaral 等^[10]曾应用空间统计内插方法, 消除遥感影像上的噪音。在此, 作者着重研究用空间统计 Kriging 插值技术, 预测解译影像信息“缺失”区。

2.1 Kriging 基本概念

Kriging 这一术语由法国 G. Matheron 于 1965 年提出的, 以纪念 D. G. Krige 在 20 世纪 40 年代末和 50 年代初首次提出并运用的一类内插技术^[3]。

Kriging 技术的核心是根据空间一些采样点的观测值, 通过加权平均来对未测点进行估计。其具体方法如下: 假设在一个区域内测量点坐标为 X_i , 变量观测值 $Z(X_i)$, $i=1, 2, 3, \dots, n$ 。则未测点 X_0 的估计值可用这 n 个样本点的线性组合来表示^[11], 如图 1。

$$\bar{Z}(X_0) = \sum_{i=1}^n \lambda_i Z(X_i) \quad (1)$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (2)$$

其中, λ_i 是有关已测量点权重, 其由无偏估计和方差最小两个假设条件来确定, 因此又称为最优内插法。

收稿日期: 2003-02-17; 修订日期: 2003-03-14

基金项目: 国家“十五”攻关课题“东北天然林景观合理配置研究与示范(2001BA510B0705)”部分内容

作者简介: 冯益明(1972—), 男, 助理研究员, 河南商城人, 博士, 主要从事统计、GIS 与 RS 应用等研究, 已发表论文 10 余篇。

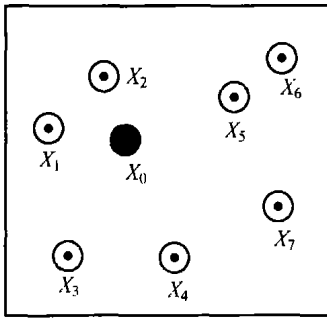


图 1 Kriging 理论

Fig. 1 Kriging theory

2.2 空间自相关理论

Kriging 实质就是根据空间已知采样点来推知未知点的值,这一方法除了服从数理统计法则外,主要依据是世间万物在空间上的分布均具有自相关性。

半方差分析是空间统计学中的一个重要组成部分,它以空间上任一距离分隔的两点上随机变量的差异为基础,分析随机变量的空间自相关性。以半方差函数来刻画随机变量在空间分布上的差异。其数学表述为:设 $Z(x)$ 为区域化随机变量, h 为两点空间分隔距离, $Z(x_i)$ 和 $Z(x_i + h)$ 分别是区域化变量 $Z(x)$ 在空间位置 x_i 和 $x_i + h$ 上的值 ($i = 1, 2, \dots, N(h)$), 则空间上具有相同间距 h 的 $N(h)$ 对观测值半方差公式为^[11]:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(X_i) - Z(X_i + h)]^2 \quad (3)$$

它是点对间差异的一半,因此将 $\gamma(h)$ 称为半方差。间距为 h 的点对取值越相似,半方差越小,而自相关越大。以半方差图直观刻画区域化变量的空间自相关性。以 $\gamma(h)$ 为纵轴, h 为横轴,绘制出 $\gamma(h)$ 随 h 增加的变化曲线为半方差图^[5]。从作出的半方差图(图 2)可以得到半方差图的 3 个基本参数。(1)块金方差 C_0 (nugget variance): 当 $h=0$ 时,在纵轴上的截距;(2)阈值(sill): 最大的 $\gamma(h)$ 值,是空间上随机变量的先验方差,即 $C(0) = \gamma(\infty)$ 。阈值与块金方差之差 C : 表示由于调查数据中存在的空间自相关性引起的方差变化范围。(3)变程(range): 实际中,将 $|h| \geq \alpha$, $C(h) = 0$ 时的间距 α 定义为变程,它给出了随机变量在空间上自相关性的尺度,在变程距离内,空间上越靠近在一起的点之间的相关性越大,相隔距离大于 α 的点间不具有自相关性,同时,变程也表示了空间内插的极限距离,在 α 内的内插是有意义的^[12]。

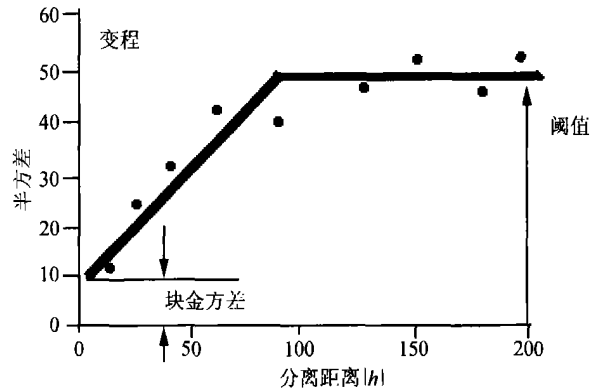


图 2 半方差图 3 个基本参数

Fig. 2 Three basic parameters of semivariogram

由于空间自相关的广泛存在,传统的内插法大多不能准确地估计出不同空间位置的未测点值。Kriging 是根据半方差分析所提供的空间自相关程度的信息来进行插值,因此可以对未测点给出最优无偏估计,而且能同时提供估计值的误差和精确度。

从以上的理论分析可知,我们把遥感影像上的像素值作为区域化变量,像素值的空间变化规律就可以用半方差函数来模拟。只要信息“缺失”区直径不大于像素值空间内插极限距离,通过影像“缺失”区周围的像素值,应用 Kriging 插值技术,是完全可以内插恢复出“缺失”区像素值,本文以遥感影像的分类为例,来具体阐述解译影像信息“缺失”区过程。

3 实验研究

3.1 实验区数据

研究数据选择东北汪清县林业局金沟岭实验林场 ($130^{\circ}5'56'' - 130^{\circ}19'16''E$, $43^{\circ}17'26'' - 43^{\circ}25'26''N$) 1997-7-11 Landsat TM ($30m \times 30m$ 分辨率) 遥感影像 (673 行, 542 列, 面积约 $328km^2$) 为数据源进行植被分类。从金沟岭林场影像中选择一块有云层遮挡影像 ($130^{\circ}6'3'' - 130^{\circ}7'31''E$, $43^{\circ}18'10'' - 43^{\circ}19'13''N$, 58 行, 56 列, 面积约 $3km^2$, 其中信息“缺失”区最大直径为 25 个像素), 作为进行 Kriging 插值估计的实验数据。对分类和插值估计精度, 均以 1997 年森林资源二类调查数据 (林相图及相应属性数据库) 为基底材料, 进行检验。

3.2 方法

3.2.1 遥感影像分类

影像分类整个工作借助图像处理软件 Erdas8.5

完成。

图像预处理 选择 Landsat TM 数据 4, 5, 3 的波段组合, 在 ERDAS 8.5 中, 将其按顺序配以红、绿、蓝三原色, 形成假彩色图像; 以 1:5 万比例尺的地形图作为底图, 选择 TM 图像和地形图上的同名地物点作为控制点, 对图像进行几何精校正。

分类 东北汪清县林业局金沟岭实验林场属国有林区, 土地利用类型主要以森林为主, 依此实际情况结合影像特点, 根据中华人民共和国 1982 年颁发的《森林资源调查主要技术规定》, 确定将影像按分类目标分为 5 类, 分别是: (A) 针叶林; (B) 阔叶林; (C) 混交林; (D) 林业其它用地; (E) 云及阴影。根据影像色调、纹理等判读标志, 从影像上进行植被类别的特征提取, 选取理想的训练样本, 对图像进行最大似然法监督分类, 生成植被类型分类图。

精度检验 用林相图以及相应的属性数据作为分类精度检验数据。在影像上, 每一类选取大约 30 个参考样点进行检验, 将林相图中相应位置的植被类型与影像分类类型进行比较, 由误差矩阵得出精度报告, 分类总精度达到 80.25%, 满足分类精度要求。

3.2.2 Kriging 插值

由于受云的影响, 该遥感影像上有多处不能直接判读其真实信息, 这也正满足实验需求。以下研究如何运用空间统计学 kriging 插值技术对原始影像进行分析, 来恢复原始影像被遮挡的斑块信息。整个工作过程在遥感图像处理软件 Erdas 8.5、GIS 软件环境 ArcInfo 8.10 以及空间统计学软件 ILWIS 3.0 中进行。

(1) 数据准备

利用 Erdas 软件, 从金沟岭实验林场遥感影像中分割出含有云层和云层阴影的斑块, 以此作为插值研究数据源。以 ASCII 形式输出影像像素值。在 Arc/Info 中, 导入 ASCII 文件, 并以图形方式显示。

对导入的数据进行编辑处理, 删除云层(图 3 中 P_2 , P_4 多边形)与云层阴影遮挡部分(图 3 中 P_1 , P_3 多边形)像素点, 图形编辑处理结果如图 3。作为进行空间统计分析 with 内插图。

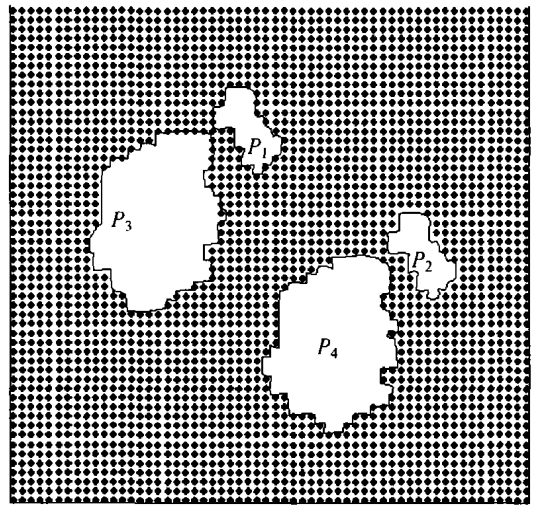


图 3 空间统计分析与内插图

Fig. 3 Spatial statistics analysis and interpolation

(2) 半方差函数理论模型模拟与插值

借助 Arc/Info 扩展模块 (geostatistical analyst) 分红、绿、蓝三个波段分别对其像素值进行半方差函数理论模型模拟与 Kriging 插值。

Kriging 方法是根据半方差函数分析所提供的变量空间自相关程度的信息来进行插值, 因此, 插值准确程度与半方差函数拟合的好坏直接相关。半方差函数理论模型的拟合过程主要包括确定曲线类型、参数最优估计和回归模型的检验。在模型拟合过程中, 红、绿、蓝三波段像素值半方差各向趋于同性^[10]。比较几个不同理论模型的拟合程度, 表现以指数模型拟合的半方差函数较好。表 1 是红、绿、蓝三波段像素值半方差函数理论模型得出的相应参数。

表 1 各波段像素半方差函数理论模型相关参数

Table 1 The best-fitted semivariogram models of image pixels of each band

波段类型	理论模型	块金值	基台值	块金值/基台值	变程/m
红	指数模型	0.94794	2.7804	34.1	867.15
绿	指数模型	150	185	27.0	829.73
蓝	指数模型	20	120	16.7	829.73

从表 1 可看出, 由于三波段像素值均表现出一定的基台值, 反映出三波段像素值在研究区内具有

平稳特性或近平稳特征。红波段与绿波段块金值与基台值的比例介于 25%—75%, 表明红、绿波段像

素值,具中等空间相关性;蓝波段块金值与基台值的比例小于 25%,表明蓝波段像素值具有强烈的空间相关性。因此,利用各波段已知像素值去预估未知像素值是合理的。

图 4、图 5、图 6 分别是红、绿、蓝三波段像素值以指数模型拟合的半方差函数图,检验表明,模型拟合符合要求。图 7 是绿波段像素值,依据拟合的半方差模型进行 Kriging 内插结果图。

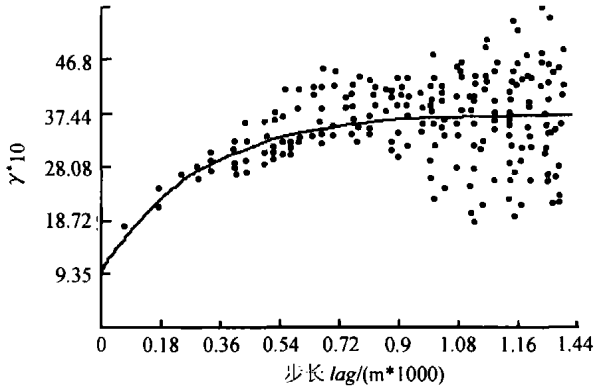


图 4 红波段像素值半方差函数图

Fig.4 The semivariogram model of image pixels of red band

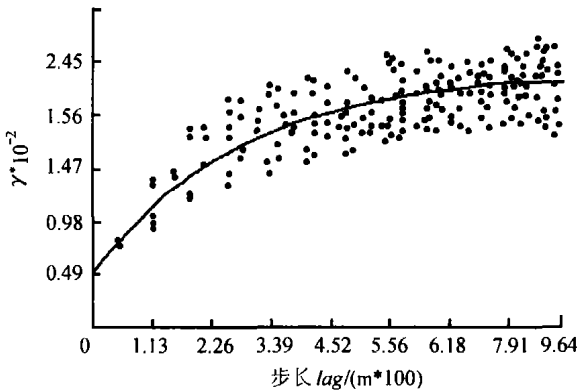


图 5 绿波段像素值半方差函数图

Fig.5 The semivariogram model of image pixels of green band

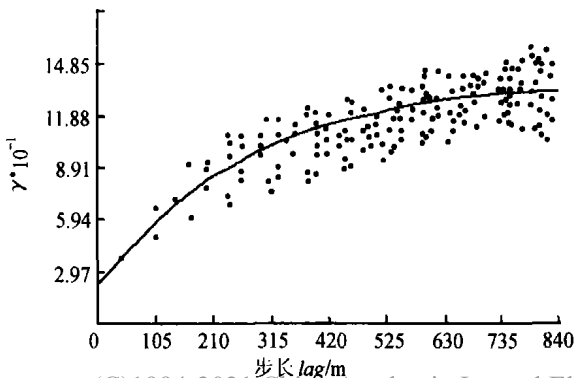


图 6 蓝波段像素值半方差函数图

Fig.6 Semivariogram model of image of blue band

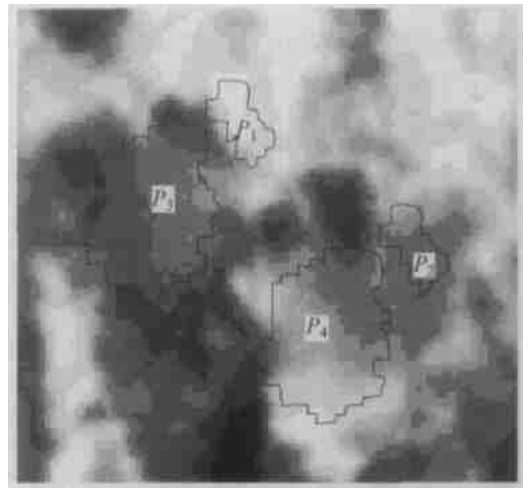


图 7 绿波段像素值 kriging 插值结果图

Fig.7 Result of Kriging interpolation of image pixels of green band

3.2.3 信息“缺失”区像元插值信息的再分类

在分别对红、绿、蓝各波段进行 Kriging 插值处理后,在 Arc/Info 8.10 中,把插值结果转换为栅格数据(*.img 格式)。红、绿、蓝各波段插值数据在 Erdas 中进行波段重组,根据前面 3.2.1 所讲,借助 3.2.1 中建立的分类因子,对插值重组影像图进行再分类,得分类结果图如图 8。对信息“缺失”区再分类结果与二类资源调查林相图比对,进行精度检验, P_1 、 P_2 、 P_3 、 P_4 四个多边形信息“缺失”插值分类精度分别为:80%,76.7%,80%,76.7%。总体反映出插值区域尺度越大,插值精度越低。

由图 8 可看出:影像信息“缺失”区,经空间统计学 Kriging 插值后,对“缺失”区像素信息进行了恢复(四个多边形区域即为“缺失”区信息恢复后信息)。从中可看出:由于受云层及阴影的遮挡,信息“缺失”区内含多种植被类型,我们不能简单地将被遮挡区归为一类。

4 结论与讨论

(1) 遥感影像解译一直是遥感信息处理中的一个关键问题,有许多方法,如考虑影像的纹理特征、空间结构,引入知识、建立图像解译专家系统等等。本文用到的 Kriging 内插模型,从方法上看属数理统计范畴,但理论上却抓住了地物在空间分布上具有连续性与自相关性这一事实,因而估计效果较好。

(2) 在实际应用中,首先要保证已经对“缺失”区周围信息准确掌握,才能应用 Kriging 方法。

(3) 如果影像信息“缺失”区面积过大,超过像

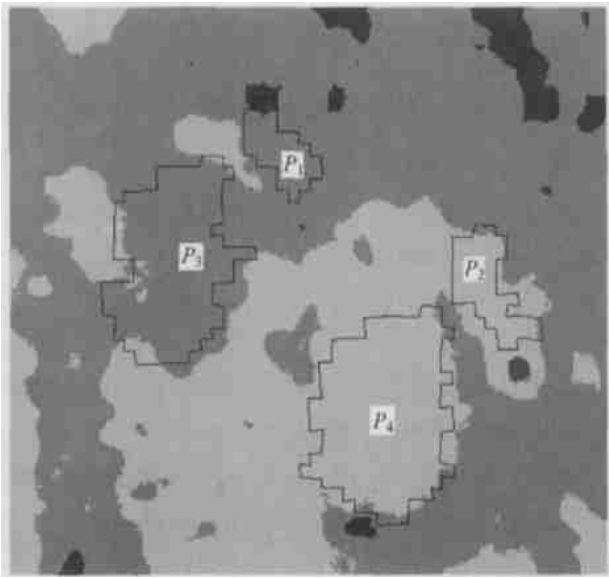


图8 插值分类结果图

Fig.8 Classification result of interpolation

素值空间内插极限距离(即变程),用 Kriging 内插没有实际意义。

(4) 如果“缺失”区含有与其周围像素不相似信息,或者信息“缺失”区地表特征复杂多变,则 Kriging 内插精度会明显下降。这时只能依靠掌握的一些基底材料,如地形以及对当地的实际了解,再结合 Kriging 内插,才能够得到好的效果。

致谢 在论文完成过程中,得到中国林业科学院资源信息研究所洪玲霞、陈尔学、庞勇,中国科学院生态中心黄志林博士等帮助,在此一并致谢。

参考文献(References)

[1] YOU Hong-jian, LIU Tong, LIU Shao-chuang *et al.* Fast Generating Geo-referenced Image Based on Airborne 3D Imager[J]. *Acta Geodaetica et Cartographica Sinica*, 2000, **29**(4):324-328. [尤红建,

刘彤,刘少创等.利用3维成像仪快速生成遥感地学编码图像[J].*测绘学报*,2000,**29**(4):324~328.]

[2] Atkinson PM, Curran PJ, Webster R. Sampling Remotely Sensed Imagery for Storage, Retrieval, and Reconstruction [J]. *The Professional Geographer*, 1990, **37**:345-353

[3] HOU Jing-ru, YIN Zhen-nan, Li Wei-ming *et al.* Practical Geostatistics (Spatial-Information Statistics) [M]. Beijing: Geological Publishing House, 1998. [侯景儒,尹镇南,李维明等.实用地质统计学(空间信息统计学)[M].北京:地质出版社,1998.]

[4] WANG Jia-hua, GAO Hai-yu, Zhou Ye. Kriging Geological Mapping Technique-Modeling and Algorithm of Computer[M]. Beijing: Oil Industry Press, 1999. [王家华,高海余,周叶.克里金地质绘图技术——计算机的模型和算法[M].北京:石油工业出版社,1999.]

[5] Burgess TM, Webster R. Optimal Interpolation and Isarithmic Mapping of Soil Properties I: The Semivariogram and Punctual Kriging [J]. *Journal of Soil Science*, 1980, **31**:315-331.

[6] Issaks EH, Srivastava RM. An Introduction to Applied Geostatistics [M]. New York: Oxford University Press, 1989.

[7] Robertson GP. Geostatistics in Ecology: Interpolating with Known Variance [J]. *Ecology*, 1987, **68**(3):744-748.

[8] Curran, PJ. The Semivariogram in Remote Sensing: An introduction [J]. *Remote Sensing of Environment*, 1988, **24**: 493-507

[9] Chica-Olmo M. Computing Geostatistical Image Texture for Remotely Sensed Data Classification [J]. *Computers and Geosciences*, 2000, **26**(4):373-383.

[10] Silvana Amaral, Shimabukuru YE, Ahern FJ *et al.* Pre-processing and Semivariogram Analysis of Radarsat Fine Mode Images for Forest Application: Tapajos National Forest, Brazilian Amazon [C]. In: *Geomatics in the era of RADARSAT*, Ottawa, Canada, May 25-30, 1997.

[11] Webster R. Quantitative Spatial Analysis of Soil in Field [J]. *Advances in Soil Science*, 1985, **3**:1-70.

[12] Wu JG, Gao W, Tueller PT. Effects of Changing Spatial Scale on the Results of Statistical Analysis with Landscape Data: A Case Study [J]. *Geographic Information Sciences*, 1997, **3**:30-41.

Interpretation of Pixel-missing Patch of Remote Sensing Image with Kriging Interpolation of Spatial Statistics

FENG Yi-ming, LEI Xiang-dong, LU Yuan-chang

(*Institute of Forest Resource Information Techniques, CAF, Beijing 100091*)

Abstract: The purpose of this paper was mainly to interpret the pixel-missing patch of image by using the kriging interpolation technology of spatial statistics. The TM image (a resolution of $30\text{m} \times 30\text{m}$) of the Jingouling Forest Farm of the Wangqing Forestry Bureau of Northeast China on 1 July, 1997 was selected as data resource, and some patches with cloud and cloud shade from the TM image were adopted as experimental data in this paper. Based on the classification for the TM remote sensing image, the pixel information of real pixel-missing patches of image (the real pixel-missing patch is the patch with the unknown real information of partial regions on the image, which is caused by the shelters of cloud, cloud shade, soot and others) was restored by the kriging interpolation technology of spatial statistics theory under the image processing software—ERDAS, the geographic information system software—ArcInfo and the spatial statistics analysis software—ILWIS. The restored pixel information was reclassified, and then compared with the results of forest resource survey, thus the results (i.e. the classified precision of the restored information of each patch was more than 75%), was obtained. Moreover, the kriging interpolation technology used in this paper grasped the fact that the objects on the global surface possessed continuity and self-correlation, and therefore it had better estimation effects. Additionally, the applicable scopes and the applicable conditions of the kriging interpolation were discussed in detail from different aspects. In summary, this paper would provide a method and means for interpreting the pixel-missing patch of image.

Key words: spatial statistics; remote sensing; GIS; pixel-missing patch; Kriging interpolation