

# 利用 DMSP/OLS 灯光影像提取 城镇建成区的聚类阈值法

邹进贵<sup>1</sup> 陈艳华<sup>1,2</sup> 丁 鸽<sup>1,3</sup> 宣 伟<sup>1</sup>

1 武汉大学测绘学院, 湖北 武汉, 430079

2 济南市房产测绘研究院, 山东 济南, 250014

3 济南市勘察测绘研究院, 山东 济南, 250013

**摘 要:**提出了便捷、高效地提取城镇建成区的聚类阈值法。该方法以连通区域循环标识算法所识别的建成区对象为基本空间单元,以对象的规模和地理重心处像元灰度(DN)值的大小为不同城镇化发展等级的衡量指标,进行空间聚类;然后,借助统计数据确定各聚类区域的最佳灯光阈值序列,提取城镇建成区,并对各建成区对象进行几何形态优化,利用消除运算去除细小碎片,填充运算填充内部空洞,平滑运算消除边缘锯齿。将提取结果与历年的《中国统计年鉴》数据和利用 Google Earth 提取的影像数据进行对比分析,结果表明,聚类阈值法能较好地提取城镇建成区的面积总量信息和空间格局特征,在数量尺度和空间格局上均有较高的有效性及可靠性。

**关键词:**DMSP/OLS; 空间聚类; 回归模型; PAEK 指数平滑算法; 几何形态优化

**中图法分类号:**P208

**文献标志码:**A

DMSP/OLS 夜间灯光影像可在较大尺度上分析城市发展的空间变化,是从事大尺度城镇化研究的有效数据手段<sup>[1]</sup>。该影像为灰度影像,表达城镇的灯光强度信息和空间分布关系,像元灰度(DN)值代表灯光强度,灰度值范围为 0~63,灰度值越大表示城镇建成区的可能性越大<sup>[2]</sup>。

目前,利用 DMSP/OLS 夜间灯光影像进行城镇建成区提取的方法主要包括两大类:基于像元的图像分类法<sup>[3,4]</sup>和基于区域阈值的图像分割法。第一类方法,以像元为单位,将大于某一阈值的像元识别为建成区像元,小于该阈值的像元识别为非建成区像元,大时空尺度应用的普适性没有得到有效证实。第二类方法,主要指统计数据比较法<sup>[5,6]</sup>,该方法以行政边界为划分依据,借助统计数据逐一确定各行政单元的阈值,当行政单元个数较多时,重复性操作多,效率低下。

由于我国的城市化水平和经济发展程度在单个行政单元内部存在明显的不均衡性<sup>[7]</sup>,必须打破行政边界的划分局限,对同一行政单元的不同区域确定不同的最佳阈值<sup>[8,9]</sup>。因此,本文提出

一套打破单个像元和行政边界的划分局限、兼顾城市化发展水平差异和城镇发展规模的聚类阈值法。该方法采用空间聚类思想,对识别的潜在建成区对象进行空间聚类,划分不同的集合,然后确定各集合的最佳阈值提取城镇建成区。

## 1 聚类阈值法

聚类阈值法最基本的空间单元为所识别的潜在建成区对象而不是单个像元或行政单元。

参考文献[2]的研究,DN 值大小与城市化强度成正比,且建成区规模越大,区域的城镇化发展等级越高。即研究对象灯光影像 DN 值的大小和建成区规模的大小可以作为不同城镇化发展等级的衡量指标。据此,聚类阈值法以各潜在建成区对象的地理重心处 DN 值大小和各对象所包含像元的数量多少为聚类依据,进行空间聚类,划分不同的城镇化发展等级。

本文对 NGDC 发布的 1992~2011 年的 F10 (1992~1994)、F12 (1994~1999)、F14 (1997~

收稿日期:2015-07-19

项目资助:国家自然科学基金(41074025)。

第一作者:邹进贵,博士,教授,主要从事精密工程与工业测量、变形监测和城市基础地理信息系统研究。jgzou@sgg.whu.edu.cn

通讯作者:陈艳华,硕士。cyh900902@163.com

2003)、F15(2000~2007)、F16(2004~2009)和 F18(2010~2011)共 32 期 DMSP/OLS (Version 4)非辐射标定夜间灯光强度影像进行处理,主要技术思路如下:① 数据预处理。利用一元二次回归模型对 DMSP/OLS 夜间灯光数据进行校正,获取 1992~2011 年的具有连续性和可比性的灯光影像;② 潜在建成区对象的识别。利用连通区域循环标识算法完成 2 728 个潜在建成区对象的识别;③ 空间聚类。根据地理重心处 DN 值大小和对象规模进行空间聚类,对潜在建成区对象划分不同的城镇化发展等级;④ 确定阈值。借助统计数据确定各发展等级的最佳阈值,提取城镇建成区;⑤ 建成区对象识别。再次利用连通区域标识算法对提取结果进行建成区对象识别。⑥ 几何形态优化。对所识别建成区对象进行几何形态优化,消除细小的城镇碎片、内部空洞和边缘锯齿;⑦ 精度评定。对得到的历年中国行政区划(本文中不含南海诸岛)内的城镇建成区信息,从数量尺度和空间格局两方面进行精度评定。

## 2 数据分析

### 2.1 影像数据校正

受传感器探测性能、卫星过境时间及成像环境不同的影响,未经校正的灯光影像不具有连续性和可比性<sup>[10]</sup>。本文应用相交分析和叠加分析,去除影像中的非稳定灯光像元,并应用一元二次回归模型对影像 DN 值进行校正:

$$DN_c = \begin{cases} DN_m, DN_m < 63 \\ 63, DN_m \geq 63 \end{cases} \quad (1)$$

式中,  $DN$ 、 $DN_c$  分别为校正前、后像元的灰度值;  $DN_m = a \times DN^2 + b \times DN + c$ ;  $a$ 、 $b$ 、 $c$  为回归参数。

参考文献[11]的研究:F16 传感器 2007 年的灯光数据具有最高累积灰度值;黑龙江省鸡西市历年灯光影像的扩展情况与城镇建成区的空间发展变化高度一致。故选取 F16(2007)灯光影像的鸡西市数据为标定数据集,确定式(1)中各回归参数,进行拟合回归<sup>[12]</sup>,然后,根据式(2)对回归校正后的影像 DN 值进行修正,完成 DMSP/OLS 影像数据的校正:

$$DN_{(n,i)} = \begin{cases} DN_{(n-1,i)}, DN_{(n-1,i)} > DN_{(n,i)} \\ DN_{(n,i)}, \text{其它} \end{cases} \quad (2)$$

式中,  $DN_{(n-1,i)}$ 、 $DN_{(n,i)}$  表示 DMSP/OLS 夜间灯光影像上  $i$  像元在第  $n-1$  年、 $n$  年的 DN 值。

### 2.2 潜在建成区对象识别

首先,利用较小阈值去除背景噪声。将灰度

值大于  $DN_T = 19$  的像元识别为前景像元,作为可能的城镇建成区。

然后,基于像元四邻域(上、下、左、右)的连通区域循环标识算法<sup>[13]</sup>,将提取的前景影像识别为由多个像元组成的、相互独立的潜在建成区对象并编号,完成 2 728 个对象( $C_1 \sim C_{2728}$ )的识别。过程为:从影像左上角开始逐行扫描,遇到第一个前景像元时,将之识别为具有一个单独像元的潜在建成区对象编号为  $C_1$ ,并作为  $C_1$  对象的种子像元标记为已识别;然后,对其四邻域进行扫描,若为前景像元,则编入  $C_1$  对象,并加入  $C_1$  对象的种子像元群,同时标记为已识别;递归此过程,直至所有与此对象相连通的前景像元被识别出来。继续逐行扫描影像像元,判断其是否为前景像元及是否被标记,若为前景像元且没有被标记,则识别为新的潜在建成区对象  $C_2$ ,并进一步完成  $C_2$  对象的标识。以此类推,直至将全部前景像元标识为潜在的建成区对象。

### 2.3 空间聚类

参考国家统计局发布的《中国城市发展研究报告》,本文将全国按城市化水平和经济发展程度,划分为 3 个等级进行空间聚类。

基于 ArcGIS 平台<sup>[14]</sup>,获取 2 728 个潜在建成区对象所包含的像元个数  $N_i$ 、地理重心点的空间位置和  $DN_i$  值大小。其中  $N_i$  表征建成区规模,  $DN_i$  表征城镇化强度,  $N_i$  和  $DN_i$  值越大,表示区域的城镇化发展等级越高。然后以  $N_i$  和  $DN_i$  值为约束条件,以对象地理重心点为节点,基于 Jenks 自然断点法构造 Voronoi 多边形,实现方差最小化聚类,聚类结果如图 1(国界数据来源:国家基础地理信息系统矢量数据)所示,其中,第一聚类区域具有最低的城镇化发展等级,对应最小的提取阈值;第三聚类区域具有最高的城镇化发展等级,对应最大的提取阈值。

然后,根据式(3),利用方差拟合优度  $R_{GVF}$  对聚类结果进行评定,一定范围内,聚类后的方差之和  $S_{DCM}$  越小,  $R_{GVF}$  越大,分类效果越好。

$$R_{GVF} = 1 - \frac{R_{SDCM}}{R_{SDAM}} \quad (3)$$

$$R_{SDCM} = \sum_{j=1}^k \frac{1}{N_j} \sum_{i=1}^{N_j} (z_{ij} - \bar{z}_j)^2 \quad (4)$$

$$R_{SDAM} = \frac{1}{N} \sum_{i=1}^N (z_i - \bar{z})^2 \quad (5)$$

式中,  $N$  表示对象的个数;  $z_i$  表示原始数据中第  $i$  个对象的 DN 值;  $\bar{z}$  表示原始数据的平均 DN 值;  $k$  表示聚类数,聚类后数据被分为  $k$  类;  $N_j$  表示第

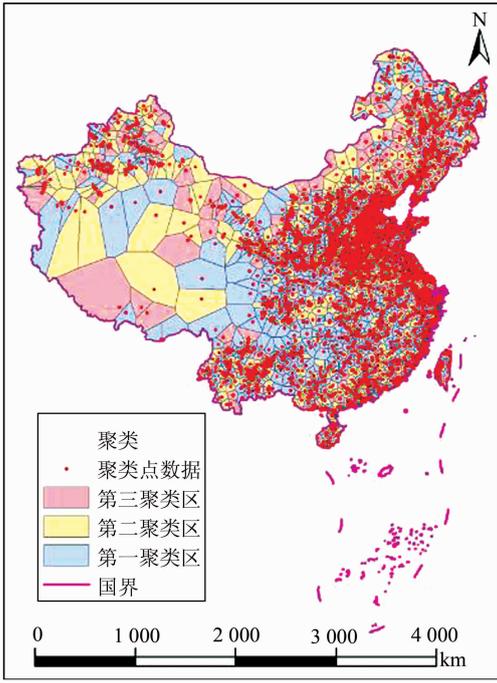


图1 不同城镇化发展水平空间聚类结果

Fig. 1 Spatial Clustered Results with Different Urban Development Levels

$j$  类对象的个数;  $z_{ij}$  表示第  $j$  类第  $i$  个对象的 DN 值, 表示第  $j$  类对象的平均 DN 值;  $R_{SDAM}$  表示原始数据的方差;  $R_{SDAM}$  表示聚类后的方差之和。本文中  $k = 3$ , 对聚类结果进行计算得到:  $R_{SDAM} = 15.88, R_{SDAM} = 165.17, R_{GVF} = 0.9039$ , 取得了较好的聚类效果。

### 2.4 阈值法提取建成区信息

借鉴基于统计数据提取城镇建成区的方法, 求取各聚类区域的最佳阈值。参考文献[15-16]对 DMSP/OLS 数据与人口、经济数据的相关分析, 设定各聚类区域的初始阈值分别为 35、50、57, 提取历年全国范围内的城镇建成区。用式(6)计算提取结果与同期统计数据间的差值, 并根据差值修正各聚类区域的阈值, 直到提取结果与统计数据充分接近为止<sup>[17]</sup>, 根据式(7), 得到各聚类区域的最佳阈值为 42、54、60, 并依据此阈值提取历年的建成区栅格影像:

$$\Delta(A_i) = S - E(A_i), i = 0, 1, 2 \dots \quad (6)$$

式中,  $A_i$  为设定的阈值序列;  $A_0$  为初始阈值序列;  $A_0 = \{35, 50, 57\}$ ,  $E(A_i)$  为根据  $A_i$  提取的城镇建成区面积;  $S$  为同期统计数据建成区面积;  $\Delta(A_i)$  为提取结果与统计数据间的差值; 若  $\Delta(A_i) < 0$ , 则增大阈值, 若  $\Delta(A_i) > 0$ , 则减小阈值。

$$|\Delta(A_{i-1})| \geq |\Delta(A_i)| \leq |\Delta(A_{i+1})| \quad (7)$$

式中,  $\Delta(A_{i-1})$ 、 $\Delta(A_i)$ 、 $\Delta(A_{i+1})$  分别表示在阈值

序列  $A_{i-1}$ 、 $A_i$ 、 $A_{i+1}$  下得到的统计数据与提取结果间的差值。

### 2.5 几何形态处理

再次利用连通区域循环标识算法对所提取结果进行建成区对象识别并编号, 同时, 对建成区对象的细小碎片、边缘锯齿和内部空洞等破损现象, 进行几何形态优化。

根据对象的面积属性, 应用消除运算删除面积小于  $30 \text{ km}^2$  的建成区对象; 应用填充运算对对象的内部空洞进行填充。由于建成区对象中的某些空洞是实际存在的, 需设定填充空洞的像元数阈值, 对小于此阈值的空洞进行填充, 大于此阈值的空洞则进行保留, 本文设定的填充像元数阈值为 20。最后, 应用 PAEK 指数多项式平滑算法, 设定多边形平滑容差值为  $0.5 \text{ km}$ , 对建成区对象计算出一个平滑的边界, 消除边缘锯齿、缺口, 实现对建成区对象边界的平滑。如图 2 所示为北京市 2008 年的建成区信息提取结果。

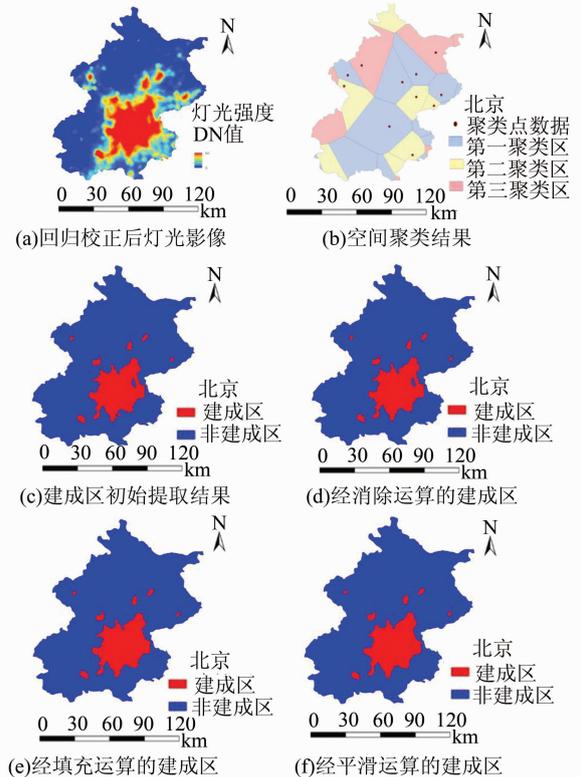


图2 聚类阈值法提取的北京市建成区空间信息

Fig. 2 Extraction of the Urban Construction Information of Beijing Based on Clustered Threshold Method

## 3 精度评定

### 3.1 数量尺度

利用历年《中国统计年鉴》中的城镇建成区统

计数据对 DMSP/OLS 数据提取的 1992~2011 年的建成区面积总量进行精度评定。如表 1 所示,历年提取结果的相对误差均在 10% 以内,误差最大的年份为 1996 年,相对误差为 9.30%,误

差最小的年份为 2003 年,相对误差仅为 -0.12%。由此可见,在数量尺度上,利用聚类阈值法提取 DMSP/OLS 灯光影像的城镇建成区信息取得了较好的精度。

表 1 利用统计数据对 DMSP/OLS 影像提取结果进行数量尺度的精度评定

Tab. 1 Accuracy Assessment on the Extraction Results of DMSP/OLS Data with the Statistical Data in Quantity Scale

| 年份   | 统计/km <sup>2</sup> | 提取/km <sup>2</sup> | 消除/平滑/km <sup>2</sup> | 填充/km <sup>2</sup> | 相对误差/% |
|------|--------------------|--------------------|-----------------------|--------------------|--------|
| 2011 | 43 603.2           | 44 010.57          | 42 308.73             | 43 255.54          | 0.80   |
| 2010 | 40 058.0           | 42 692.18          | 40 714.88             | 41 657.04          | -3.99  |
| 2009 | 38 107.3           | 40 441.74          | 37 300.02             | 38 244.97          | -0.36  |
| 2008 | 36 295.3           | 38 833.35          | 36 541.26             | 37 486.21          | -3.28  |
| 2007 | 35 469.7           | 36 575.71          | 33 997.80             | 34 877.59          | 1.67   |
| 2006 | 33 659.8           | 35 862.09          | 33 184.12             | 34 065.76          | -1.21  |
| 2005 | 32 520.7           | 33 575.71          | 31 084.74             | 31 912.39          | 1.87   |
| 2004 | 30 406.2           | 31 917.98          | 29 130.60             | 30 010.38          | 1.30   |
| 2003 | 28 308.0           | 30 138.86          | 27 403.62             | 28 342.98          | -0.12  |
| 2002 | 25 972.6           | 28 858.75          | 26 219.40             | 27 065.67          | -4.21  |
| 2001 | 24 026.6           | 27 287.13          | 23 277.48             | 24 144.23          | -0.49  |
| 2000 | 22 439.3           | 24 877.61          | 21 994.58             | 22 908.81          | -2.09  |
| 1999 | 21 524.5           | 23 748.20          | 21 899.62             | 22 813.85          | -5.99  |
| 1998 | 21 379.6           | 23 007.38          | 20 065.57             | 20 142.85          | 5.78   |
| 1997 | 20 791.3           | 22 480.56          | 19 268.65             | 19 347.78          | 6.94   |
| 1996 | 20 214.2           | 21 738.56          | 18 264.11             | 18 333.94          | 9.30   |
| 1995 | 19 264.2           | 20 215.22          | 18 049.99             | 18 096.36          | 6.06   |
| 1994 | 17 939.5           | 18 617.89          | 17 170.20             | 17 232.58          | 3.94   |
| 1993 | 16 588.3           | 17 232.46          | 16 875.08             | 16 932.80          | -2.08  |
| 1992 | 14 958.7           | 15 417.15          | 14 116.56             | 14 160.32          | 5.34   |

### 3.2 空间格局

利用 Google Earth<sup>[18]</sup> 提取 2011 年城镇建成区影像对 DMSP/OLS 数据的提取结果在空间格局上进行精度评定。具体地,选取北京、郑州和秦皇岛 3 个城市代表发展程度不同的“大城市”、“中等城市”和“小城市”,进行对比。

表 2 利用 Google Earth 提取结果对 DMSP/OLS 影像提取结果进行精度评定

Tab. 2 Accuracy Assessment on the Extraction Results of DMSP/OLS Data with the Google Earth Extraction Data

| 城市  | Google Earth /km <sup>2</sup> | DMSP/OLS /km <sup>2</sup> | 绝对差值 /km <sup>2</sup> | 重叠度 /% |
|-----|-------------------------------|---------------------------|-----------------------|--------|
| 北京  | 1 268                         | 1 345.89                  | -77.89                | 93.86  |
| 郑州  | 297                           | 320.04                    | -33.04                | 92.24  |
| 秦皇岛 | 80.8                          | 73.37                     | 7.43                  | 90.80  |

表 2 和图 3 显示,以 Google Earth 提取的城镇建成区信息作为参考,3 个城市 DMSP/OLS 提取结果的精度均在 92% 以上,DMSP/OLS 提取的建成区空间格局与 Google Earth 提取的建成区格局特征基本吻合。

### 4 结语

本文提出的聚类阈值法,以所识别的建成区对象为基本空间单元,既消除了以单一阈值提取

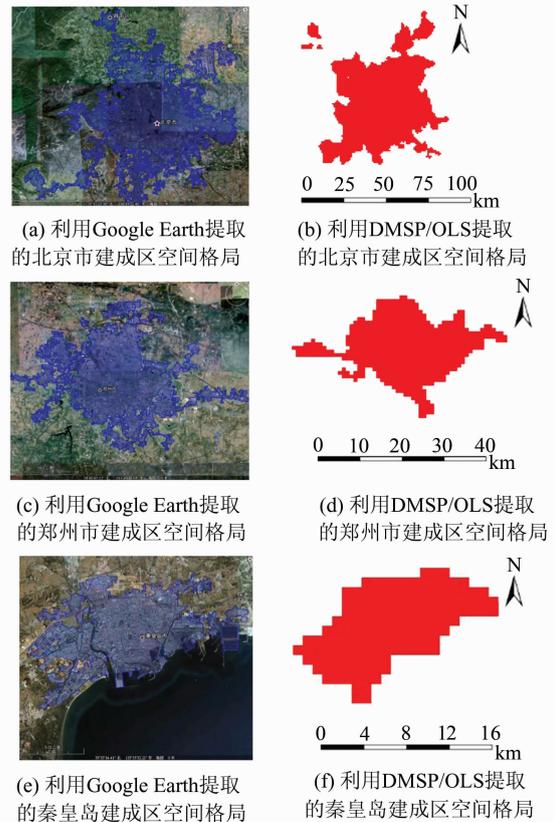


图 3 Google Earth 提取结果与 DMSP/OLS 影像提取结果对比分析

Fig. 3 Comparative Analysis on the Extraction Results of DMSP/OLS Data and Google Earth Data in Spatial Scale

同一行政单元内部不同城镇化发展等级区域的不合理性,又顾及了建成区对象的规模与发展等级划分。以该方法为基础,本文以 DMSP/OLS (Version 4)的32期非辐射标定夜间灯光强度影像为实验数据,提取历年中国区域(本文中不含南海诸岛)的城镇建成区信息。通过与统计数据 and Google Earth 提取结果进行对比分析,验证该方法在数量尺度和空间格局上的有效性及可靠性。该方法可在一定程度上弥补现有统计资料空间信息不足、逐一确定各行政单元阈值效率低下的缺陷,实现快速、有效、低成本地更新城镇建成区数据,为 DMSP/OLS 灯光影像充分发挥其时效性和经济性起到积极的推动作用。

### 参 考 文 献

- [1] Zhuo Li, Shi Peijun, Chen Jin, et al. Application of Compound Night Light Index Derived from DMSP/OLS Data to Urbanization Analysis in China in the 1990s [J]. *Acta Geographica Sinica*, 2003, 58(6): 893-902(卓莉, 史培军, 陈晋, 等. 20世纪90年代中国内地城市时空变化特征——基于灯光指数 CNLI 方法的探讨[J]. 地理学报, 2003, 58(6): 893-902)
- [2] Chen Jin, Zhuo Li, Shi Peijun, et al. The Study on Urbanization Process in China Based on DMSP/OLS Data: Development of Light Index for Urbanization Level Estimation[J]. *Journal of Remote Sensing*, 2003, 7(3): 168-175(陈晋, 卓莉, 史培军, 等. 基于 DMSP/OLS 数据的中国内地城市化过程研究: 反映区域城市化水平的灯光指数的构建[J]. 遥感学报, 2003, 7(3): 168-175)
- [3] Cao X, Chen J, Imura H, et al. A SVM-based Method to Extract Urban Areas from DMSP/OLS and SPOT VGT Data[J]. *Remote Sensing of Environment*, 2009, 113(10): 2 205-2 209
- [4] Lu D S, Tian H Q, Zhou G M, et al. Regional Mapping of Human Settlements in Southeastern China with Multisensor Remotely Sensed Data[J]. *Remote Sensing of Environment*, 2008, 112(9): 3 668-3 679
- [5] Wang Y Y, Xu D, Zhu X G, et al. The Spatio-temporal Characteristics of the Urban and Town Construction Land Expansion in Jiangsu Province from DMSP/OLS Nighttime Images[J]. *Modern Urban Research*, 2010, 2: 67-73(王跃云, 徐昀, 朱喜钢. 江苏省城镇建设用地扩展时空格局演化——基于 1993、1998、2003 年度夜间灯光数据的分析[J]. 现代城市研究, 2010, 2: 67-73)
- [6] Li Jinggang, He Chunyang, Shi Peijun, et al. The Use of Multisource Satellite and Geospatial Data to Study the Ecological Effects of Urbanization: A Case of the Urban Agglomerations in Bohai Tin [J]. *Journal of Remote Sensing*, 2007, 11(1): 115-126(李景刚, 何春阳, 史培军, 等. 基于 DMSP/OLS 灯光数据的快速城市化过程的生态效应评价研究——以环渤海城市群地区为例[J]. 遥感学报, 2007, 11(1): 115-126)
- [7] Wang Xiaohui. Urban Expansion in China over the Past 30 Years Detected Using DMSP/OLS Night-time Light Data[D]. Nanjing: Nanjing University, 2013(王晓慧. 基于 DMSP/OLS 夜间灯光数据的中国近 30 年城镇扩展研究[D]. 南京: 南京大学, 2013)
- [8] Ma T, Zhou C H, Pei T, et al. Quantitative Estimation of Urbanization Dynamics Using Time Series of DMSP/OLS Nighttime Light Data: A Comparative Case Study from China's Cities [J]. *Remote sensing of Environment*, 2012, 124:99-107
- [9] Henderson M, Yeh E T, Gong P, et al. Validation of Urban Boundaries Derived from Global Nighttime Satellite Imagery [J]. *International Journal of Remote Sensing*, 2003, 24(3): 595-609
- [10] National Oceanic and Atmospheric Administration (NOAA). Version 4 DMSP/OLS Nighttime Lights Time Series[OL]. <http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>, 2013
- [11] Liu Z F, He C Y, Zhang Q F, et al. Extracting the Dynamics of Urban Expansion in China Using DMSP/OLS Nighttime Light Data from 1992 to 2008 [J]. *Landscape and Urban Planning*, 2012, 106: 62-72
- [12] Elvidge C D, Ziskin D, Baugh K E, et al. A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data [J]. *Energies*, 2009, 2(3):595-622
- [13] Shu Song. Spatial Cluster Analysis of Urban Landscape Pattern Using Stable Nighttime Light Satellite Images[D]. Shanghai: East China Normal University, 2013(舒松. 基于稳定夜间灯光遥感数据的城市群空间模式识别方法研究[D]. 上海: 华东师范大学, 2013)
- [14] Zhang Qiaoying. Spatial Expansion of New First-tier Cities Using the DMSP/OLS Nighttime Light Images [J]. *Journal of Geomatics*, 2015, 40(1):72-75(张巧英. 基于 DMSP/OLS 灯光数据的新一线城市空间扩展研究[J]. 测绘地理信息, 2015, 40(1): 72-75)
- [15] Amaral S, Monteiro A M V, Camara G, et al. DMSP/OLS Night-time Light Imagery for Urban Population Estimates in the Brazilian Amazon [J]. *Inter-*

*national Journal of Remote Sensing*, 2006, 27(5/6): 855-870

- [16] Florida R, Mellander C, Gulden T. Global Metropolis: Assessing Economic Activity in Urban Centers Based on Nighttime Satellite Images [J]. *Professional Geographer*, 2012, 64(2):178-187
- [17] Yang Y, He C Y, Zhao Y Y, et al. Research on the Layered Threshold Method for Extracting Urban Land Using the DMSP/OLS Stable Nighttime Light Data [J]. *Journal of Image and Graphics*, 2011, 16(4):666-673 (杨洋, 何春阳, 赵媛媛, 等.

利用 DMSP/OLS 稳定夜间灯光数据提取城镇信息的分层阈值法研究[J]. *中国图像图形学报*, 2011, 16(4): 666-673)

- [18] Song Jinchao, Li Xinhua, Lin Tao, et al. A Method of Extracting Urban Builtup Area Based on DMSP/OLS Nighttime Data and Google Earth [J]. *Journal of Geo-information Science*, 2015, 17(6):750-756 (宋金超, 李新虎, 吝涛, 等. 基于夜晚灯光数据和 Google Earth 的城市建成区提取分析[J]. *地球信息科学*, 2015, 17(6):750-756)

## A Clustered Threshold Method for Extracting Urban Built-up Area Using the DMSP/OLS Nighttime Light Images

ZOU Jingui<sup>1</sup> CHEN Yanhua<sup>1,2</sup> DING Ge<sup>1,3</sup> XUAN Wei<sup>1</sup>

1 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

2 Jinan Real Estate Measuring Institute, Jinan 250014, China

3 Jinan Institute of Geotechnical Investigation Surveying, Jinan 250013, China

**Abstract:** A clustered threshold method for conveniently and efficiently extracting urban built-up areas has been proposed in this study. This method has broken through the limitations created by single pixel analysis and the administrative boundaries by using built-up objects identified by the recursive connected-region labeling algorithm as basic spatial units. It classifies urban development levels using spatial clustering on the basis of the size of the objects and the DN value of the geographic center of objects and extracts built-up areas based on the optimal threshold sequence as determined from statistical data. It optimizes the geometric morphology of built-up objects by removing small scraps, stuffing internal holes, and smoothing the jagged edges. Extraction results were analyzed and compared to the statistical data found in the *China Statistical Yearbook* over the years and images derived from Google Earth. These results show that the clustered threshold method can effectively obtain the total acreage and spatial pattern of urban built-up areas, with high validity and reliability in both quantity scale and spatial pattern.

**Key words:** DMSP/OLS; spatial clustering; regression model; PAEK exponential smoothing algorithm; geometry morphology optimization

**First author:** ZOU Jingui, PhD, professor, specializes in precise engineering and industrial surveying, deformation monitoring and urban geographic information system research. E-mail: jgzou@sgg.whu.edu.cn

**Corresponding author:** CHEN Yanhua, master. E-mail: cyh900902@163.com

**Foundation support:** The National Natural Science Foundation of China, No. 41074025.