

基于滑动近似熵突变检测的广佛城乡过渡带提取

周财霖

中山大学 地理科学与规划学院, 广州 510000

摘要: 突变检测法是基于遥感影像提取城市城乡过渡带的一种重要方法: 采用最大似然地物分类后, 基于滑动窗口获得城乡用地比率图像, 提取从中心向四周射出的射线组成城乡用地比率的数据序列, 基于滑动近似熵方法提取城乡突变点, 并通过显著性检验, 从而组成过渡带。利用广佛 2004 年和 2017 年 Landsat TM 影像考虑方法的可行性, 观察广佛的城市格局动态变化和城乡发展。

关键词: 突变检测 近似熵 城乡过渡带

1 引言

随着社会经济的发展和城市扩张的进程, 不同于传统的双层城市格局: 城市-乡村, 在城乡之间出现了新的圈层——城乡过渡带, 并成为了城市格局中的重要组成部分, 是城市向外扩张的直观显现, 与核心和乡村发生直接互动和作用, 在城市中发挥着独特作用。对城镇过渡带在城市研究中具有重要意义。城镇过渡带的提取是其中的关键环节。

城镇过渡带的提取可以基于生态景观指数, 社会经济变量, 行政边界等因素进行提取, 但往往是以区域作为单元, 精度较低。而基于遥感影像的提取具有其他方式不可比拟的时效性和科学性, 能够以像素为单元进行提取, 精度较高, 更利于进行多年尺度的动态变化的。

广州与佛山由于长期的协同发展与一体化推进, 具有罕见的双城一体的城市结构, 因此对于广佛城乡过渡带具有独特的城市研究意义。这种非典型的圈层城市结构, 对于传统的过渡带提取方法也提出了新挑战。

2 研究方法

2.1 突变检测法

突变检测法 (abrupt change detection) 是时间序列中检测异常点的经典方法。抛开应用情景, 单纯从统计学的角度, 突变就是一串按顺序排列的数据中, 从一种统计特性到另一种统计特性剧烈变化。因此关键的检测原理就是考察统计特征量(如平均

值和方差)的变化, 得到变化剧烈的数据位置。

由于数据序列的不同来源, 以及对突变点的不同成因解释, 突变检测在学术研究上具有广泛的应用。最主要的突变检测应用在气象学范畴中: 在天气过程变化中的跳跃性转变现象, 如多年范围内的由于厄尔尼诺或拉尼娜现象导致的异常暖年或冷年。

章文波 (2001) 首次将突变检测方法应用到空间尺度上, 利用滑动 T 检验基于 TM 影像提取北京时城乡过渡带。本文基于这种思想探索新突变检测提取城乡过渡带的算法, 开创空间突变检测新思路。

2.2 城市用地比率

突变检测法仅能对 1 条数据序列进行分析, 对于多条数据序列需要更复杂和高级的工具, 因此对于遥感影像的多波段进行处理时较困难的。另外, DN 数值或者地表反射率变化时非常剧烈的, 基于一条像素的 DN 值进行突变检测效果不好。因此需要对遥感影像进行处理, 实现: (1) 多波段转换成单波段; (2) 以一种能代替城镇变化的指标代替原始的 DN 值; (3) 通过类似滤波的操作使单像元值能表示像元附近整个区域的值, 实现图像的平滑和去噪。而城市用地比率能有效解决以上问题。

城市用地比率是一定范围内城镇用地面积占区域总面积的百分比。可以看到离核心区越近的城市区域值越高, 离城市越远值越低。因此本文语境下的城乡过渡地区为: 城市用地比率快速变化的区域。

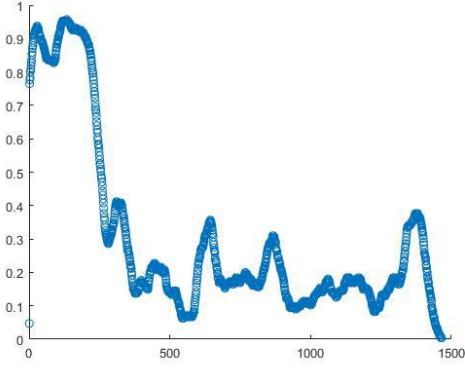


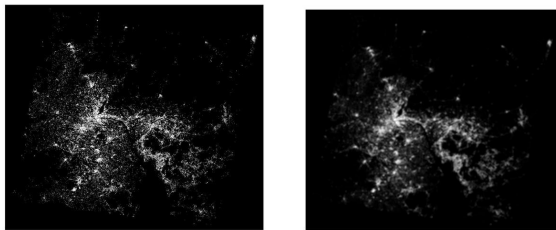
图 1 城市用地比率变化（南向一断面线）

Fig.1 Change of Urban Land-use Ratio (One cut-line toward S.)

对影像进行最大似然分类后，提取出表示城镇用地和非城镇用地的二值图像，在二值图像上以每个像元为中心开边长为 N 的正方形窗口，统计该窗口内城市用地面积占整个窗口面积的比率，作为该像元的值。滑动比率窗口计算式为：

$$D(i, j) = \frac{\sum_{i=1}^N \sum_{j=1}^N g(i, j)}{N \times N}$$

式中， N 为统计窗口大小， $g(i, j)$ 为窗口内像元 i, j 的像元值（1 或 0）， D 为计算出来的城市用地比率值。对整幅二值图像的每一像元点遍历，得到最终的城市用地比率图。窗口大小的 N 值需要进行多次实验，选择最适宜的值。 N 值过大时会引起图像边缘的模糊， N 值过小时则不足以剔除小斑块和随机变化的影响。另外由于城市核心区内和附近的图斑分布比较密集，还需要保证 N 值足够小以保持一定的精度。经过实验，最终选择 N 为 59。



(a) 城市用地 (b) 城市用地比率
(a) Urban land-use (b) Urban land-use ratio

图 2 城市用地比率影像对比

Fig.2 Compare of urban land-use ratio process

2.3 近似熵

传统的突变检测方法，包括 Mann-Kendall，滑动 F 检验，滑动 T 检验等，都是基于统计理论，从

线性统计的角度出发。这类方法对于大数据量的数据集以及高度复杂、动态、非线性的数据序列效果较差，并且对于数据集有严苛的要求。因此近年来非线性分析方法成为突变检测的主要发展方向。

	近似熵	线性统计
考察核心	熵	统计量
抗噪声能力	强	无
数据长度	不依赖	较小
数据要求	几乎无	较多
多突变点	可行	不适宜

表 1 突变检测方法对比

Table 1 Comparison of Detect Methods

近似熵(Approximate Entropy, $ApEn$)由 Pincus 在 1991 年提出，有效测量时间序列的复杂性和规律性的一种无需粗粒化的复杂性测度分析方法。 $ApEn$ 是一个非负标量，表达时间序列的复杂性。对于任意一段数据样本 $\{u(i); i = 1, 2, \dots, n\}$ ，给定维数 m 和容许偏差 r ，公式推导如下：

1) 构建一组维数为 m 的向量：

$$X(i) = [u(i), u(i+1), \dots, u(i+m-1)] \\ (i = 1, 2, \dots, n-m+1)$$

2) 定义向量 $X(i)$ 和 $X(j)$ 之间的欧式距离：

$$d[X(i), X(j)] = \max[|X(i+k) - X(j+k)|] \\ (k = 0, 1, 2, \dots, m-1)$$

3) 计算 i 各值下的 $C_i^m(r)$ ：

$$C_i^m(r) = \{\text{数目} \mid d[X(i), X(j)] \leq r\} / (N-m+1)$$

4) 对 $C_i^m(r)$ 取对数，并对所有 i 取平均值得到 ϕ^m

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r)$$

5) 将维数 m 加 1，重复 (1) ~ (4) 的步骤，计算 ϕ^{m+1}

6) 最终得到数据序列的近似熵：

$$ApEn(m, r) = \phi^m(r) - \phi^{m+1}(r)$$

由推导过程可以看出， $ApEn$ 反映的是序列重构后 m 维情况下两点的自相似程度大小，以及序列在向量重构后，维数 m 增加时产生新模式的概率大小，熵值越大，产生新模式的几率就越大，因而序

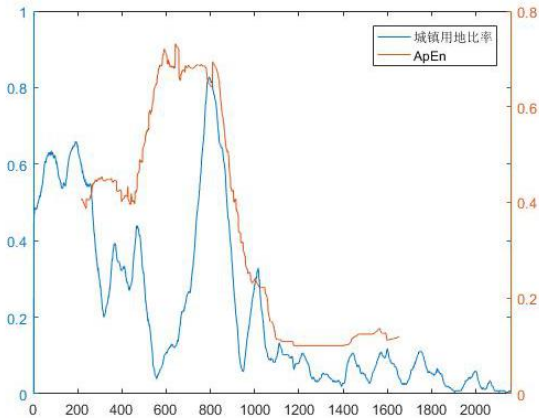
列就越复杂。

2.4 滑动近似熵

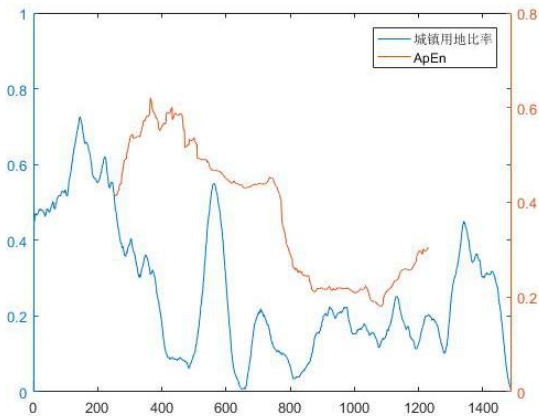
近似熵仅能表达一条数据序列的复杂性，而要实现突变点的提取，还需要结合滑动方法进行子序列的提取和分析，具体步骤如下：

- ①根据样本的长度初步选择滑动窗口长度 h ，维数 m 和容许偏差 r ；
- ②由样本的第 i 个数据为中心，选取长度为 h 的子序列；
- ③计算子序列的 $ApEn$ 值作为 i 点值；
- ④计算出大多数 i 点值，得到一个 $ApEn$ 序列，一部分 i 点由于处在样本序列的头或尾端，因此没有对应的子序列。

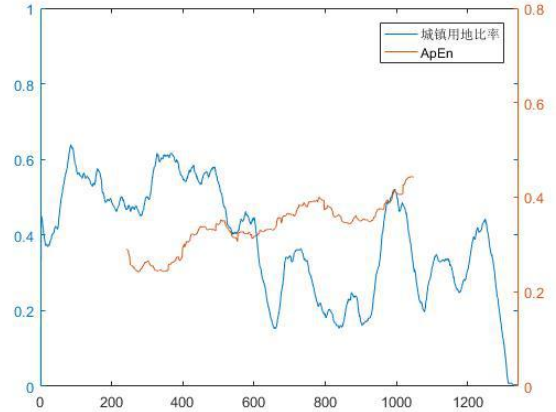
简单的来说，就是以每个样本点作为中心提取固定长度的子序列，以该子序列的 $ApEn$ 值作为 i 点是否为突变点的判断依据。如果子序列 $ApEn$ 值较高，说明该点所处周围的样本点剧烈变化，因此该点可以考虑为突变点。



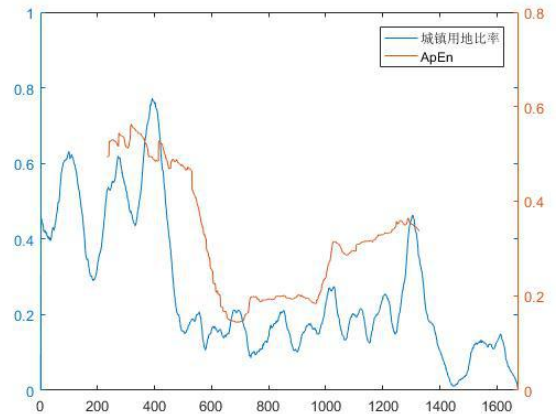
(a)东断面线
(a) Cut-line (E)



(b)西断面线
(b) Cut-line (W)



(c)北断面线
(c) Cut-line (N)



(d)南断面线
(d) Cut-line (S)

图 3 滑动近似熵结果
Fig.3 Result of ApEn

上图以本次对广州 2017 年 Landsat 遥感影像的研究结果为例，城镇用地比率变化发生剧烈变化的区间内，往往对应的 $ApEn$ 值较高，可以看出其对于突变点检测有效性。但与普通的线性统计相比，并非简单的对应关系，因为近似熵是更为复杂的系统，相较于突变点更加不敏感，因此如上文所述，更有利于大数据尺度下的多突变点检测。

3 结果处理与分析

3.1 城乡过渡带提取

首先根据城市用地二值影像提取的点格局，根

据点格局分析中的质心选定影像中的核心点，并得到影像中城市中心点的行列号。然后根据参数方程的思想提取从城市中心点出发的，每隔 1° 形成的切割射线。每条射线上的城市用地比率值作为一个数据序列。

每条数据序列包含许多 0 值点（图像中的外围黑色区域），会影响突变检测的结果，因此应该手动对影像进行掩膜切割，剔除多余的 0 值点。

对处理后提取的每条数据序列进行滑动近似熵计算。在参数的选择上，维数 m 通常为 2 或 3，此处取 2；容许偏差 r 一般在 $0.1\sigma \sim 0.2\sigma$ （ σ 为子序列的标准差）之间，此处取 0.15σ ； h 一般视样本长度在 0~2000 之间决定。经过多次尝试，此处 h 取 600，既能保证不丢失过多的首尾信息，又能保证突变检测的准确。

仅仅依靠近似熵进行突变点的判断缺乏统计学依据，因此需要对得到准突变点进行 T 检验，实现二次筛选。序列数据下的突变点 T 检验公式如下：

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{1/n_1 + 1/n_2}} \sim t(n_1 + n_2 - 2)$$

其中：

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

式中， N_1 为基准点前的样本组即前序列的样本数， S_1 和为前序列的方差和均值。类似的， N_2 ， S_2 和为后序列的样本数，方差和均值。简单来说，就是通过断点的前后数据列的统计量的差异，来判断断点是否为突变点。

本研究中，T 检验值大于置信度 $\alpha = 0.01$ 水平下的 T 值作为判断依据，进行二次筛选得到空间上的所有突变点。

本研究所称的的城乡过渡带只考虑城镇用地和农田。而位于城镇中的林地，公园，水体由于属于非城镇区域，因此周边的像素点会很大程度的被误判为突变点。所以需要城镇用地的分类图像进行掩膜，剔除突变点结果中的属于林地，水体，绿地的干扰点。

对得到的点格局进行 1km 的缓冲区处理，得到面要素组成的城乡过渡带。

总体实验流程如下：

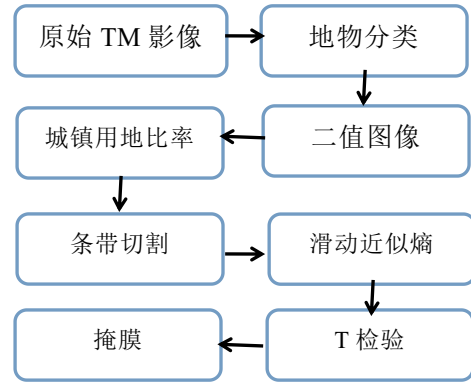


图 4 操作流程

Fig.4 technical process

最终提取结果如下：

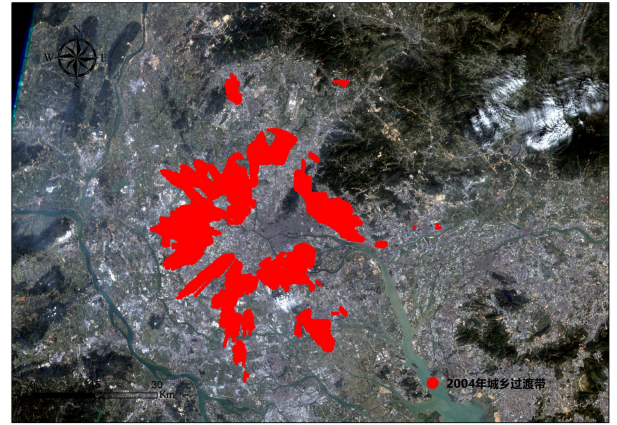


图 5 2004 年城乡过渡带结果

Fig.5 2004 Urban-rural transition zone

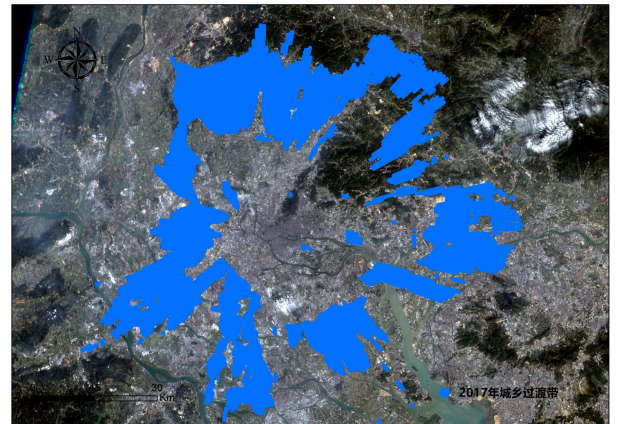


图 6 2017 年城乡过渡带结果

Fig.5 2017 Urban-rural transition zone

从提取的结果来看，基本符合客观实际：城乡过渡带呈圈层状环绕广佛城镇中心，并且 2017 年结果与 2004 年结果相比发生了较大的变化，符合近年

来广佛城镇格局快速发展的状况。初步确定提取结果的有效性,作为进一步分析的依据。

3.2 结果分析

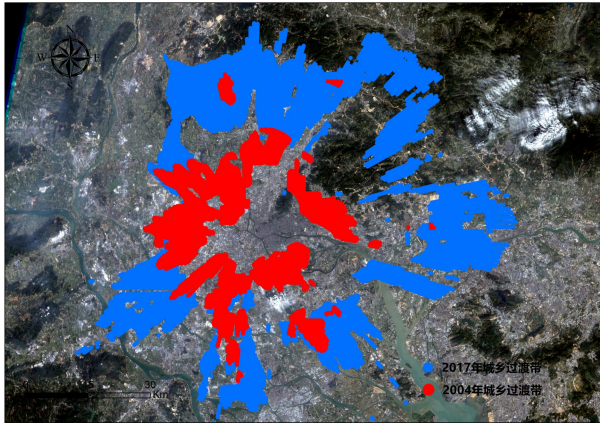


图 8 2004-2017 年城乡过渡带变化
Fig.8 Change of Urban-rural transition zone

将 2004 年和 2017 年的广州城乡过渡带提取结果叠置比较,以此分析十年来广州的城市格局变化。

(1) 广佛地区城乡过渡带均呈现破碎的圈层环绕分布,厚薄不均且地区差异较大。并且随着时间发展和城镇扩张,破碎程度进一步增加。

这主要归因于广佛所在区域独特的地势地貌:珠三角下游河网密布,河流对于城镇的切割作用明显;丘陵山台对城镇扩张的限制明显,例如广州东北部山区,天河北-从化-增城一带,城乡过渡带明显较其他地区薄弱;南方地区与北方地区相比,由于资源丰富和环境良好,城镇和聚居点更趋于沿河网平原离散分布,长久形成松散的城市格局。而随着时间发展,城市向外扩张,城乡过渡带覆盖区域上存在更多的河网和山地丘陵,因此过渡带破碎程度随外扩而增加。

(2) 广佛城镇中心紧密连接,呈现明显的双城合体。历史渊源上广佛长期属于同一城市,由于地理接壤和经济发展的协同效应和互补优势,两地的整体性得以延续。2009 年后“广佛同城”政策的不断推进,两地在城市格局上的融合在提取结果上得到证实。

由结果还可发现,佛山南海-广州荔湾的广佛通道基本不存在城乡过渡带,直接将广佛各自中心城区联结起来。因为处在广佛边界的南海大沥和广州芳村上的城镇区不仅紧密连接,而且由于长期的工

商业区位优势,经过多年的发展,城镇化水平非常高。因此广佛地区的城乡过渡带围绕的是“双核”,两地的过渡带圈层有机结合。

(3) 城乡过渡带在近十年内快速变化,城市扩张明显。由叠置对比发现,2004 年和 2017 年的城乡过渡带重叠区域非常小,说明以前属于过渡带的区域很多都发展成了建成区域,而以前属于乡村的地区发展成了城乡过渡地区。

2004 年的城乡过渡带区域主要有:东部的天河北-黄埔萝岗一带,北部的白云-花都-官窑一带,西部的三水-丹灶-西樵一带,南部的海珠南-三山-北滘一带。而到 2017 年,主要的城乡过渡带区域为:东部的增城区和从化区,北部的花都和清远广州边界区域,西部的三水区和高明区,南部的番禺-南沙-顺德南一带。

总而言之,从基于遥感影像的城乡过渡带提取能够很好的分辨出城市的分布格局和发展差异,与客观实际的城市规划和扩张发展吻合度较高,两者能够很好的相互解释,并为广州的城乡研究提供客观全面的资料。

4 结 论

本文基于 TM 遥感影像的地物分类影像,通过滑动窗口获取城市用地比率影像,通过滑动近似熵的突变检测获取影像中的城市用地比率突变点,经后处理后得到最终的城镇过渡带。本文的主要创新点有:

(1) 首次在在空间尺度上使用滑动近似熵方法,为空间突变点的探测提供新思路。并且对算法进行了适应性的改进,且结合传统的线性统计方法,保证结果的显著性和提取质量。

(2) 首次仅基于遥感影像提取广佛的城乡过渡带,尽管基于遥感影像没有考虑人口经济等因素,但具有时效性和客观性的特点,对于城市格局的动态检测有重要意义。

(3) 首次对十年尺度范围内的广佛城乡结构进行对比分析,为城市研究提供可靠的资料和初步观点。

但是本研究还有很多局限性:

(1) 突变检测尽管能够很好的分析出序列的突

变点,但并不能解释导致突变的原因和机制,也不能比较各种因子的贡献。城乡过渡带的形成机理是复杂和多元的,需要结合社会经济等更多因素进行考虑和分析,单纯的区域提取需要进一步的研究。

(2) 尽管目前对城乡过渡带提取的研究非常多,但不同研究对城乡过渡带的定义不同,提取方法不同,所获得的结果差异也较大。因此简单的以其他研究提取成果作为精度评价依据的意义不大,需要更可靠的精度评价方法。

(3) 突变点尽管在统计值上具有显著性,但离真正的城乡过渡带需要更多的实际分析和讨论,因为存在很多的干扰,如城市核心区中的绿地园林前的区域由于城镇用地比率的剧变容易被误判成城乡过渡带。

不同于一般的地物类型提取,在不同语境和研究背景下,城镇过渡带可以有不同的定义和内涵,相应的提取方法也不尽一。因此不同城乡过渡带提

取的未来的研究可以基于多种突变检测的方法,并结合经济,人口等除遥感影像外的指标进行更全面的提取。

参考文献(References)

- Zhang W B, Fang X Q, Zhang L S. 1999. Study on the method of dividing Urban-Rural Transition Zone by using remote sensing image. *Journal of remote sensing* 1999, 3(3):199-202 (章文波, 方修琦, 张兰生. 利用遥感影像划分城乡过渡带方法的研究[J]. *遥感学报*, 1999, 3(3):199-202.)
- He W, He T, Cheng H Y, et al. A new mutation detection method based on approximate entropy [J]. *Acta physica Sinica*, 2011, 060 (004): 813-821.(何文平, 何涛, 成海英,等. 基于近似熵的突变检测新方法[J]. *物理学报*, 2011, 060(004):813-821..)
- Sun D Y, Huang Q, Wang Y M, et al. Application of sliding approximate entropy in the analysis of runoff sequence mutation [J]. *Journal of hydropower*, 2014, 33 (4): 1-6.(孙东永, 黄强, 王义民,等. 滑动近似熵在径流序列突变性分析中的应用[J]. *水力发电学报*, 2014, 33(4):1-6.)

Extraction of Urban-rural transition zone of Guangzhou and Foshan Based on ApEn Abrupt Change detection

ZHOU Cailin

Geography and Planning School of Sun Yat-sen University, Guangzhou 510000, China

Abstract: Mutation detection method is an important method to extract Urban-Rural Transition Zone Based on Remote Sensing Image: after the maximum likelihood feature classification, the urban-rural land ratio image is obtained based on the sliding window, the ray emitted from the center to the surrounding is extracted to form the data sequence of urban-rural land ratio, and the urban-rural mutation point is extracted based on the sliding approximate entropy method, and t-test is carried out to form the transition zone. Based on the feasibility of Landsat TM image in 2004 and 2017, the dynamic changes of urban pattern and urban-rural development of Guangzhou & Foshan were observed.

Key words: ApEn, Urban-rural transition zone, abrupt change