Remote Sensing of Impervious Surfaces
Tropical and Subtropical Areas
Hongsheng Zhang, Hui Lin, Yuanzhi Zhang, Qihao Weng

Methodology of Combining Optical and SAR Data

Publication details
Hongsheng Zhang, Hui Lin, Yuanzhi Zhang, Qihao Weng
Published online on: 01 Sep 2015

How to cite: Hongsheng Zhang, Hui Lin, Yuanzhi Zhang, Qihao Weng. 01 Sep 2015, Methodology of Combining Optical and SAR Data from: Remote Sensing of Impervious Surfaces, Tropical and Subtropical Areas CRC Press
Accessed on: 29 Dec 2018

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Methodology of Combining Optical and SAR Data

3.1 Study Area

Six cities from four countries located on the tropical and subtropical areas (Figure 3.1) were carefully chosen as the study sites of this research. These cities are Guangzhou, Shenzhen, Hong Kong, Sao Paulo, Mumbai, and Cape Town. A basic description about these six cities is given in the following sections including their geography, climate, population, economy, and urban planning.

3.1.1 Site A: Guangzhou

Guangzhou, the capital city of the Guangdong province, China, is located at the center of the PRD metropolitan. The PRD is located downstream of the Pearl River, and is known as the third largest metropolitan in China, enjoying a tremendously fast development during the past 30 years. The region has had a quick urbanization process, with a population of over 19 million and an area of over 21,000 km² (Fan et al. 2008). However, due to strong interactions between human activities and the environment, serious environmental issues have also emerged and are causing a series of environmental problems, including air pollution and water pollution (Zhang et al. 2008). Nevertheless, unlike many other metropolitans in the world, the PRD is located in a subtropical humid area, with long period of cloudy weather throughout the whole year and different characteristics of plant phenology (Fan et al. 2008). In the PRD, about 80% of precipitation occurs during April to September, which is known as the wet season, while only about 20% of rainwater occurs from October to the following March, which is known as the dry season (Cai et al. 2004). Therefore, the seasonal effects of ISE from remote sensing imagery are likely to be different from those found in previous research, which was conducted mainly in midlatitude regions. Three sites in the PRD, Guangzhou, Shenzhen, and Hong Kong, have been chosen as the study sites in this book, with detailed descriptions in the following sections.
Guangzhou has undergone a dramatic urbanization process and is the third largest city in China. The population of Guangzhou reached 12.78 million in 2010 (Brinkhoff 2011). The study site selected in this research is located in the Huangpu District of Guangzhou (Figure 3.2), which is a medium urbanized area. The land cover in this site is characterized by residential areas, small rivers, farmland, small hills, and small lakes and water pools with seasonal waters. Impervious surfaces can be seen from the Landsat ETM+ image with both high and low reflectance. Within this study, this site was selected to evaluate ISE using Landsat ETM+ and ENVISAT ASAR.
images, as well as to investigate the seasonal effects of the plant phenology and climate of the tropical and subtropical regions.

### 3.1.2 Site B: Shenzhen

Shenzhen is located in the southern part of the PRD and to the north of Hong Kong. Shenzhen is China’s first and most successful special economic zone (SEZ), which has been highly urbanized in the past three decades. Commercial and industrial areas are intensively distributed all over the city due to the rapid development of the economy (Figure 3.3). The selected site is on the boundary of Shenzhen and Hong Kong, which is highly urbanized and is characterized with mainly commercial, residential, and greening areas of the city. SPOT-5 and ENVISAT ASAR images in Shenzhen were selected to test the effectiveness of the methods proposed in Section 3.5.

### 3.1.3 Site C: Hong Kong

Hong Kong is situated in the southern part of the PRD on the coast of the South China Sea. It consists of three main parts, the New Territories, Kowloon, and Hong Kong Island. Even though Hong Kong has been intensively urbanized, it is a mountainous city with a large area of mountains distributed all over the city. The impervious surface percentage can reach up to 100% in urban areas such as Kowloon, but is only moderately urbanized with a moderate impervious surface percentage in the rural areas such as the New Territories. In this study, a site located in the Yuen Long in the northern part of the New Territories was selected (Figure 3.4). This study site is typical in Hong Kong

![FIGURE 3.3](image-url)

(a) SPOT-5 (RGB: 4-1-2) and (b) ENVISAT ASAR images of the study site located in Shenzhen.
as it is moderately urbanized and includes both plain and mountain areas. Land covers in this study site include residential areas, farmland, mountains, and coastal sea surface. SPOT-5 and TerraSAR-X images in this area are used to address the ISE and evaluate comprehensively the effectiveness and performance of the synergistic use of optical and SAR data.

### 3.1.4 Site D: Sao Paulo

Sao Paulo is located in southeastern Brazil and is the largest city in Brazil by population and gross domestic product (GDP). Sao Paulo has a humid subtropical climate and is significantly influenced by monsoons, and the average annual precipitation is about 1454 mm. The Sao Paulo metropolitan area
has been undergoing a rapid urbanization process since the twentieth century. However, urbanization has brought significant environmental impacts to the ecosystem, such as the deforestation of rainforest (Torres et al. 2007). Therefore, the estimation of impervious surfaces would be beneficial for urban planning and environmental management of the city. In this study, Landsat TM, ENVISAT ASAR, and TerraSAR-X are used synergistically to extract the impervious surfaces of Sao Paulo. Figure 3.5 shows the Landsat TM and ENVISAT ASAR data in this study area.

3.1.5 Site E: Mumbai

Mumbai is located in western India, with a total urban area of approximately 465 km². Mumbai has a tropical wet and dry climate, with an average annual precipitation of about 2167 mm and an average annual temperature of 27.2°C. It is the main city of western India and a leading economic and financial center (Bhagat 2011; Moghadam and Helbich 2013). The population of Mumbai has nearly doubled in the last four decades according to the Indian Census of 2011 (Moghadam and Helbich 2013), with an increasing trend to reach about 27 million by 2025 (United Nations 2012). However, there are many problems introduced by the rapid urbanization in Mumbai, such as urban fragmentation (Gandy 2008). Therefore, remote sensing of the urbanization process of Mumbai would be very helpful to monitor the urban sprawl in order to improve the urban planning and management of Mumbai. In this study, Landsat TM, ENVISAT ASAR, and TerraSAR-X are used to estimate the impervious surface distribution of Mumbai. Figure 3.6 shows the Landsat TM and ENVISAT ASAR data in this study area.

![FIGURE 3.6](a) Landsat TM (RGB: 5-4-3) and (b) ENVISAT ASAR images of the study site in Mumbai.)
3.1.6 Site F: Cape Town

The city of Cape Town is located in the southwestern part of South Africa, covering about 2460 km². It is located at approximately latitude 33.55°S and longitude 18.25°E, which is nearly on the boundary of the subtropical region in the Southern Hemisphere. Cape Town enjoys a Mediterranean climate with warm and dry summers and cool and wet winters. The population of Cape Town was about 3.5 million in 2008 (Rebelo et al. 2011). Cape Town has been undergoing a rapid urbanization process with significant land use/land cover changes (Rebelo et al. 2011). Consequently, this rapid urbanization has caused great environmental impacts, especially damage to the biodiversity (Rebelo et al. 2011). Satellite monitoring of the urban sprawl in Cape Town will be important in order to monitor these impacts in a timely manner. In this study, Landsat TM, ENVISAT ASAR, and TerraSAR-X will be used to estimate the impervious surface distribution of Cape Town. Figure 3.7 shows the Landsat TM and ENVISAT ASAR data in this study area.

3.2 Satellite Data

3.2.1 Landsat ETM+

The Landsat ETM+ images had one panchromatic band and six bands with an image pixel size of 30 m × 30 m. In this study, only the 30 m data was used. The ETM+ image was acquired on December 31, 2010. In order to preprocess ETM+ images, a process should first be applied to get rid of the stripes on
the eastern and western edges of each scene that are caused by the footprints (location and spatial extent) of each band due to scan line corrector (SLC) failure. For this reason, the study area was located in the middle of each scene where there were no stripes, and thus no stripe removal operation was applied. We assumed that the atmospheric conditions were clear and homogeneous and the small area of clouds would not significantly impact the whole scene of the image, and thus no atmospheric correction was performed (Wu and Murray 2003).

### 3.2.2 SPOT-5

SPOT is a high-resolution optical satellite family launched by France. SPOT-5 was launched on May 4, 2002, with a higher spatial resolution of 2.5 and 5 m in panchromatic mode, and 10 m in multispectral mode. The SPOT-5 image used in this study was in a multispectral mode image, at precision 2A level, and was obtained on November 21, 2008. Therefore, the pixel size of the SPOT-5 data in this study is $10 \times 10$ m. The multispectral SPOT-5 data has four image bands located in the green region (500–590 nm), red region (610–680 nm), near-infrared region (780–890 nm), and shortwave infrared region (1580–1750 nm). The image was projected under the coordinate system of World Geodetic System 1984 (WGS84) and Universal Transverse Mercator (UTM) (Zone 50N).

### 3.2.3 ENVISAT ASAR

ASAR is a radar instrument on the ENVISAT satellite operated by the European Space Agency (ESA). ENVISAT was launched on March 1, 2002 with a projected mission duration of 5 years and continued to work for 10 years. Even though it stopped operating on April 8, 2012, the archive data is still beneficial for this study due to the nature of the study and the selected time period. ASAR operates in the C band (4–8 GHz) and generally has five operation modes: alternating polarization (AP) mode, image (IM) mode, wave (WV) mode, suivi global (GM) mode, and wide swath (WS) mode, where AP, IM, WV, GM, and WS are the identity codes. Raw data from these operation modes is the Level 0 data, and they can be further processed to Level 1 or even higher levels of data product by different treatments. The ASAR data used in this study is the wide swath mode (WSM) and IM precision (IMP) data, which is a Level 1b data product. The data was received by the Satellite Remote Sensing Receiving Station at the Chinese University of Hong Kong. The ASAR WSM data was obtained on September 23, 2010, on the descending direction with vertical transmit/vertical receive (V/V) polarization and a pixel size of $75 \times 75$ m. The ASAR IMP data was obtained on November 19, 2008, on the ascending direction, Track-25 of ENVISAT, with V/V polarization and a pixel size of $12.5 \times 12.5$ m. Additionally, due to the uncertainty of speckle noises in SAR images, the enhanced Lee filter is selected to filter the speckle noises. The enhanced Lee filter is an improved version of the Lee
filter that was designed to better preserve texture information, edges, linear features, and point targets in SAR images (Lee 1983). The enhanced Lee filter is an adaptive filter that was proven to be more suitable for preserving radiometric and textural information than other speckle filters (Lopes et al. 1990; Xie et al. 2002). The ASAR image was then geocoded and projected with the georeference system of the WGS84 and UTM (Zone 50N).

3.2.4 TerraSAR-X

TerraSAR-X (TSX) is a German earth observation satellite launched on June 15, 2007 and is still in operation. TSX operates in the X band (9.6 GHz) and has three main imaging modes: SpotLight mode, StripMap mode, and ScanSAR mode. The TSX image used in this study is in StripMap mode, obtained on November 16, 2008, with a spatial resolution of 3 × 3 m, and the scene size is 30 km (width) × 50 km (length). The TSX image was geocoded with Next ESA SAR Toolbox (NEST) 4C-1.1 software developed by ESA under the coordinate system of WGS84 and UTM (Zone 50N). Geometric correction was also conducted by the Range-Doppler Terrain Correction in NEST with digital elevation model (DEM) data. Additionally, due to the uncertainty of speckle noises in SAR images, the enhanced Lee filter is selected to filter the speckle noises.

3.3 Digital Orthophoto Data

The orthophoto is derived from aerial photographs that were taken mainly at a flying height of 2400 m in Hong Kong, and were named the Digital Orthophoto DOP 5000 series. The whole land area of the Hong Kong Special Administrative Region (HKSAR) is covered by 190 tiles of DOP 5000 images with a specific tile number for each image. The original DOP 5000 data has a ground pixel size of 0.5 × 0.5 m and is georeferenced in the coordinate system of the Hong Kong 1980 Grid. The DOP 5000 photo used in this study was taken on November 12, 2008, and was located on the northwestern part of Hong Kong with a tile number of 6-NW-A. To use the DOP 5000 data in this study, a coordinate system transformation was conducted on the data to transform it from Hong Kong 1980 Grid to UTM50N with WGS84.

3.4 In Situ Data

Field work was conducted on January 7, 2013 (winter) to collect information about the spectral reflectance and spatial texture of different land covers.
The field work was carried out in the Yuen Long district located in northwestern Hong Kong (Figure 3.8).

A Global Positioning System (GPS) device (Leica Zeno) was employed to locate the geocoordinates of the fields (Figure 3.9). Using a Nokia cell phone (Nokia 5320) to connect to the local differential GPS (DGPS) reference station during the field work, the DGPS technique was used to improve the location accuracy, which was up to 0.4 m. A spectrometer produced by Analytical Spectral Devices, Inc. (ASD) was employed to collect the hyperspectral reflectance of each land cover. The field of view (FOV) of the ASD spectrometer is 25 degrees and the area to be measured was set as 0.5 × 0.5 m according to the resolution of the digital orthophoto data, and therefore, the height of the ASD spectrometer should be 1.13 m (= 0.25/tan12.5°) from the ground. Moreover, a digital camera (Canon-Digital IXUS 115 HS) was used to take

FIGURE 3.8
Image from Google Earth showing the area where field data was collected.

FIGURE 3.9
Devices employed for field data collection: (a) ASD spectrometer, (b) Leica Zeno, (c) Nokia cell phone, and (d) Canon camera.
photos of each land cover, which were used to help analyze the color, texture, and shape features. During the field work, six main land cover types were considered, including dark impervious surfaces (i.e., asphalt road and old concrete roads), bright impervious surfaces (i.e., new concrete roads), vegetation (i.e., shrubs and grasses), farmland (i.e., crop), water surfaces (i.e., rivers and water pools), and bare soil in a field under construction. In this book, the collected field data were used to validate the results of visual interpretation of the satellite images.

3.5 Framework of Methods

Figure 3.10 shows the framework of methodology of this study, illustrating the mechanism of land cover diversity, the responses of both optical and SAR remote sensing, and the ISE from remote sensing images. The main idea shown in Figure 3.10 is that the difficulties of ISE from satellite images are caused by the diversity of land covers and their reflectance in the images. Moreover, land cover diversity is caused by the phenology and the climatology of tropical and subtropical regions and the extreme human activities surrounding urbanization. Meanwhile, the phenology is affected by the climatology, which mainly includes the characteristics of temperature and precipitation.

Therefore, the methodology of this study includes three parts: (1) to investigate the effects of phenology and climatology on ISE, aiming at finding the most suitable season for ISE from satellite images, (2) to investigate the characteristics of urban land covers, which are the direct cause of the difficulties in the accurate estimation of impervious surfaces, and (3) to address the methods for synergizing optical and SAR images in order to improve the accuracy of ISE.

3.5.1 Per-Pixel Modeling of Impervious Surfaces

Impervious surface mapping at a per-pixel level is actually a classification process where impervious and nonimpervious surfaces are a combination of various land cover types. Conventional LULC includes vegetation, urban areas, and water, and each land cover type shares similar spectral and spatial characteristics. Therefore, they are often identified individually during the classification procedure. However, impervious and nonimpervious surfaces consist of various land cover materials. For instance, impervious surfaces can be made up of dark materials (e.g., asphalt and old concrete) and bright materials (e.g., new concrete and metal), while nonimpervious surfaces are also very diverse in materials (e.g., vegetation, water, and base soils). In this study, a two-step approach is employed to estimate the impervious surfaces. First, six land cover types: dark impervious surfaces, bright impervious
FIGURE 3.10
Framework of the methodology of this research.
surfaces, vegetation, water body, bare soil, and shaded areas, are identified with a classification procedure using RF. Second, a procedure is conducted to combine various land covers into impervious and nonimpervious surfaces.

In particular, shaded areas are treated as a single land cover type as they often have unique spectral and spatial characteristics. Moreover, since shaded areas may be impervious (e.g., roads and rooftops) or nonimpervious (e.g., greening areas), they are treated as nonimpervious surfaces in the second step of combination in this study. Therefore, dark impervious surfaces and bright impervious surfaces are combined as impervious surfaces and vegetation, water, bare soil, and shade are combined as nonimpervious surfaces. Additionally, because misclassification may happen not only between impervious and nonimpervious land cover types but also among different subtypes of impervious or nonimpervious types, the accuracy of classification before and after the combination operation may be different. Therefore, in this study, an accuracy assessment is conducted on the classification results before and after combining impervious surfaces and nonimpervious surfaces subtypes.

### 3.5.2 Investigation of Seasonal Effects

The motivation of this study is to prove or reject two hypotheses. First, seasonal changes of the landscape components should be less problematic in subtropical monsoon areas since vegetation and canopy change less among the seasons, while water may have an impact, as there are many variable source areas (VSAs). VSAs are those areas filled with water in rainy seasons, and bare soil is exposed in dry seasons (Frankenberger et al. 1999). Second, the sensitivity of the ISE may depend on different methods. In this book, the per-pixel approach is adopted and two popular classifiers are selected, including ANN and SVM.

This book investigates the variation of ISE from different seasons of satellite images and the seasonal sensitivity of different methods. Four Landsat ETM+ images of four different seasons are employed to estimate the impervious surfaces at the pixel level. Seven land use types are defined to conduct the classification procedure according to the landscape of the study area. Table 3.1 gives a brief description about each land cover type, including water, vegetation, bare soil, clouds, shade, dark impervious surfaces, and bright impervious surfaces. In particular, clouds and shade are treated as one type of land cover, since they have unique spectral and spatial characteristics compared with other types of land cover. As the region is undergoing a dramatic urbanization process, a lot of bare soils appear on the areas under construction. Further, numerous cool roofing materials, which are light blue or white in color, are used to build up rooftops. These rooftops are designed to highly reflect the solar radiation in order to reduce the urban heat island effect. Thus, they appear to be bright impervious surfaces in...
Methodology of Combining Optical and SAR Data

3.5.3 Feature Extraction

3.5.3.1 Conventional Feature Extraction

According to previous literature, segmentation methods are superior over pixel-by-pixel methods because segmentation methods take texture characteristics into account (Dell’Acqua and Gamba 2003; Stasolla and Gamba 2008). A texture characteristic is important for the interpretation of SAR data because the speckles in SAR data result in difficulties for the pixel-by-pixel approaches. Therefore, in order to extract complementary information for urban impervious surfaces from ASAR data, texture feature extraction is necessary and important. In this study, the popular GLCM approach (Haralick et al. 1973) is employed to analyze the texture features of the ASAR data. For the application of GLCM, the size of image block and the texture measures with GLCM have been a major issue (Marceau et al. 1990). In terms of the classification of remote sensing images in urban areas, it is reported that a window size of $7 \times 7$ pixels is suitable with a test on the resolutions from $2.5 \times 2.5$ m to $10 \times 10$ m (Puissant et al. 2005). Moreover, four texture measures: homogeneity (HOM), dissimilarity (DISS), entropy (ENT), and angular second moment (ASM) were identified as effective indicators for the optical remote sensing data. ANN and SVM are then applied to classify the four seasonal changes of ETM+ images. After seven land use classes are available, a combining operation is employed to reclassify the five land use types into two types: the impervious surfaces and nonimpervious surfaces. During this period, the water body, vegetation cover, and bare soils are combined together to form the nonimpervious surfaces, while the dark and bright impervious surfaces are combined into the impervious surfaces class.

### TABLE 3.1
Definition of the Land Covers Used in This Study

<table>
<thead>
<tr>
<th>Land Use Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>Rivers, lakes, and other freshwater bodies</td>
</tr>
<tr>
<td>Vegetation</td>
<td>Grain crops, vegetable crops, grass, and other agricultural land</td>
</tr>
<tr>
<td>Bare soil</td>
<td>Land under construction with bare soils exposed</td>
</tr>
<tr>
<td>Clouds</td>
<td>Small and fragmentary clouds that are difficult to remove</td>
</tr>
<tr>
<td>Shades</td>
<td>Topographical shades and shades from tall buildings, trees, etc.</td>
</tr>
<tr>
<td>Dark impervious surfaces</td>
<td>Rooftops, roads, and parking lots that are made of asphalt, concrete tile, and other materials with low spectral reflectance</td>
</tr>
<tr>
<td>Bright impervious surfaces</td>
<td>Cool rooftops and green rooftops that are made of cool materials, such as metal, which are designed to highly reflect solar radiation</td>
</tr>
</tbody>
</table>

Note: Clouds and shade are treated as land covers since they have unique spectral and spatial characteristics compared with other land covers in satellite images.
texture description of different urban land cover types (Puissant et al. 2005). Thus, in this study, the window size is set as 7 × 7 pixels, and four texture measures, HOM, DISS, ENT, and ASM, are employed.

3.5.3.2 SAN Feature Extraction

3.5.3.2.1 Concept of Shape-Adaptive Neighborhood

The neighborhood is a basic and key concept in image processing. However, previous feature extraction approaches using neighborhoods with regular shapes had some shortcomings that could lead to error because terrain objects may have different irregular shapes. As more attention has been paid to human cognition, the procedure of human vision has been considered and applied in image processing. Considering human vision, the color and shape of the target are very important. Human beings recognize different objects by their color characteristics first, then by their shape feature, and other features such as texture. This procedure of recognizing an object generally happens within a local neighborhood in the image. If the object is gray, then the shape characteristics will be the most important features for human eyes because there are no colors with which to identify different objects. Based on this observation, the concept of SAN was proposed to start the procedure of feature extraction. Prior to the feature extracting procedure, color characteristics will be analyzed to determine the neighborhood of each pixel with an adaptive boundary.

**Definition**

A SAN is the neighborhood of a pixel containing but not necessarily centered on the pixel, whose shape is determined by the terrain object it represents (Zhang et al. 2013).

Figure 3.11 shows the concept of the SAN of a pixel (a), where the view port is used to represent the local range to search the SAN or the object. The

![Figure 3.11](image)

**FIGURE 3.11**

Illustration of a SAN of point A (a) in an irregular shape object, (b) in a rectangular object, (c) in a square object and (d) in a complex environment (e.g., sports ground).
feature of a SAN only represents the feature of the central pixel (not always in the center). If the pixels in the SAN are of the same terrain object as the central pixel, the judgment is correct. In contrast, if there are some misjudged pixels in the SAN, this will affect only the classification of the central pixel without any impact on other pixels in the SAN.

After determining the SAN of a pixel, the feature of the SAN can be extracted, including the color, texture, and shape feature, which describe characteristics of the central pixel and will then be used in the classifying procedure. The mechanism of the determination of SAN is consistent with that of the on-off switch model in the attention mechanism of human vision (Solso et al. 2004). According to cognitive psychology research, in a local range containing an object of the on-off switch model, those parts that do not belong to the object will be filtered out and thus do not catch our attention. Only the object itself will be able to pass the on-off switch and then cause the so-called attention. This phenomenon has also been proven with some evidence in neural experiments both in animals and humans (Solso et al. 2004). It is important to note that the SAN here is only suitable in those objects with no significant variation in color, or in the case of remote sensing images, with no significant variation in multispectral reflectance. However, a certain extent of variation is allowed because there is little variance in spectral reflectance (color), and this is the common case in reality. This situation can be handled with a heterogeneity-based threshold, which will be discussed in the following section.

3.5.3.2.2 Determination of a SAN

3.5.3.2.2.1 Spectral Feature Transformation  Multiple spectral features are crucial for the interpretation of remote sensing images. However, the problem of how to understand the information contained in different spectral bands becomes apparent when using visual interpretation, since a visible color (for human eyes) consists of only three components in existing popular color space (e.g., red, green, blue [RGB] and HSI). In this case, three of the bands are often assigned to be the red, green, and blue values used to generate a false-color image for visual interpretation of the image. Moreover, different combinations of bands are attempted in order to discriminate between different objects. This RGB mode of mapping the spectral feature space is also called the color space, referred to as the color characteristics in visual interpretation.

As discussed above, determination of the SAN depends on the color characteristics. Heterogeneity is used to describe the color feature as follows. Conventionally, there are several color models, including RGB, hue, saturation, intensity (HSI), and hue, saturation, value (HSV). In the processing of remote sensing images, the images are often transformed to the false-color composition; that is, in the RGB color space. However, according to the existing literature, the RGB color space is not consistent with human vision (Herodotou et al. 1999). A color point in RGB color space cannot really
represent the color recognized by human eyes as it corresponds differently from that of the human perception of color. Of these color spaces, the one closest to the human perception of color is HSV (Herodotou et al. 1999). The transformation formula from RGB color to HSV color is shown in Equation 3.1 (Herodotou et al. 1999), where the value of $H$ would be in the range $[0, 360]$, and the values of $S$ and $V$ would be in $[0, 1]$.

$$
\begin{align*}
H &= \begin{cases} 
\cos^{-1} \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} & \text{if } (B \leq G) \\
360^\circ - \cos^{-1} \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} & \text{if } (B > G)
\end{cases} \\
S &= \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \\
V &= \frac{\max(R, G, B)}{255}
\end{align*}
$$

(3.1)

In order to treat the three components in the same way for the calculation of the color feature, the $H$ needs to be transformed in the range $[0, 1]$. For the hue component, the visible spectrum distributes over the whole range $[0, 360]$; that is, including 0 but excluding 360. For instance, red, green, and blue colors are separated by $120^\circ$ within this range. In this way, $H$ does not assume the mathematical meaning of an angle (i.e., $90^\circ$ and $270^\circ$ represent different colors). Thus, it is reasonable to normalize $H$ to $[0, 1]$, which includes 0 but excludes 1, with a linear transformation. After the transformation, the color feature of the pixel can be expressed as

$$
CF = \omega_1 \cdot H + \omega_2 \cdot S + \omega_3 \cdot V
$$

(3.2)

where $\omega_1$, $\omega_2$, and $\omega_3$ are the weights of the three components, and $\omega_1 + \omega_2 + \omega_3 = 1$. Therefore, the color feature $CF$ here will be a single value instead of a vector of the three components. There are two advantages of this. First, it is convenient to place different weights with different components. According to psychological research, hue is related mostly to the color we determine (see) an object to be. Thus, it is reasonable to place higher weight on the $H$ component; that is, $\omega_1$. However, which combination of the three weights is the best probably depends on various applications. The second advantage is that it is computationally better than representing the color feature as a vector. Since the color feature is used to calculate the heterogeneity for a large number of times in the following steps, the representation of the color feature in this way will save a lot of time.
3.5.3.2.2 Determine SAN  The SAN of a pixel is determined within a view port (Figure 3.2) centered on the central pixel using a given heterogeneity threshold. The heterogeneity between two pixels is defined to determine the SAN of one pixel using its color feature. Let $CF_0$ be the color feature of the central pixel and $CF_i$ represent the color feature of the pixel $i$, which is to be determined whether inside the SAN or not. Thus, a simple way of expressing the heterogeneity between the two pixels is $\text{diff} = |CF_0 - CF_i|$. Given a threshold $T$ and that the SAN of the central pixel is $SAN_0$, the rule could be $i \in SAN_0$ if and only if $\text{diff} < T$, where iff represents the term “if and only if.”

The threshold of heterogeneity between two pixels is a key factor influencing the size of the SAN. If the threshold is too small, most of the SANs contain only a few pixels, which can result in difficulties for the feature extraction from the SANs. Thus, an appropriate threshold is crucial for the feature extraction procedure in the steps to follow. In a simple way, the optimal threshold is found with a threshold search procedure by quantitatively testing various thresholds and analyzing the size of the SANs. In this study, a series of numbers is tested for threshold values and corresponding sizes of SANs are counted. Results are plotted in a curve using a spline interpolation method. Finally, the threshold curve is used to determine the optimal threshold.

3.5.3.2.3 Extracting Spatial Features

3.5.3.2.3.1 Texture Extraction Method  There is no accepted quantitative definition of texture (Bharati et al. 2004). Rather, it is left as an intuitively obvious but quantitatively undefined characteristic associated with a given pixel. Various attempts have been made to give it an appropriate quantitative definition, but none appear to have achieved widespread acceptance. In this study, a texture analysis is conducted on each SAN to represent the texture characteristic of each pixel. Texture features can be extracted with the SAN. There are many methods cited in the previous literature for carrying out texture analysis, such as co-occurrence-matrix-based approaches (Zhang 2001), random distribution models (Bruzzone and Prieto 2002), and geostatistical methods (Curran 1988). Since each SAN has an uncertain shape, and considering the case of remote sensing images, the geostatistical approach was used to describe the spatial autocorrelation, which is closely related to the texture characteristics (Jensen 2007). The geostatistical approach was reported to successfully represent the autocorrelation of spatial data (e.g., remote sensing images) (Fabbri et al. 1993; Jensen 2007). In geostatistics, the variogram is calculated first and is fitted with a theoretical model such as the spherical model, and then the parameters of the variogram, such as nugget, sill value, and variable-range, are used to describe the characteristics of spatial autocorrelation.

However, the calculation of a variogram and the fitting of the theoretical variogram are time-consuming processes. Since we do not employ the variogram to predict some unknown pixels, but only to describe the texture
feature, there is no need to calculate all the function values of the variogram at every step length $h$. Only some of the key steps are helpful to describe the texture feature, such as when $h = 1$, the function value $\gamma(h)$ will be the sill value of the variogram. Thus, a selected series of steps was used to compute the function values, which is a modified version of the variogram, shown as follows:

$$\gamma(H) = \frac{1}{N(H)} \sum_{i=1}^{N(H)} [Z(x_i) - Z(x_i + H)]^2$$

(3.3)

where $H = [h_1, h_2, \ldots, h_n]$ denotes the selected series of steps and $\gamma(H)$ the resampled variogram, which is treated as the extracted feature of the texture. In specific applications, $h$ should be selected according to both the spatial resolution of the data and the landscape characteristics of the land surface. Generally, the higher the resolution of the data and the larger the size of the targets, the higher the value of $h$ should be used. In this study case, $H$ is empirically set to be $[1, 2, 3]$.

3.5.3.2.3 Description of Geometric Features and Their Effectiveness

Geometric features include many types, such as shape features and topological features. Since this study focuses on modeling the early processing of visual perception, only the shape features are considered. Shape features of images have received attention since 1993, when Fabbri first introduced the shape feature into multispectral remote sensing images (Fabbri et al. 1993; Jensen 2007). Description of shape characteristics in the traditional shape analysis methods contained the compact expression (compactness), the complexity description, and the curvature description. In this study, two kinds of compact expressions are employed as the shape descriptors: the aspect ratio ($R$) and the form factor ($F$), illustrated by Equation 3.4, where $L$ and $W$ are the length and width of the minimum boundary rectangle (MBR) of the SAN, $B$ is the perimeter of the SAN, and $A$ is the area of the SAN.

$$R = \frac{L}{W}; \quad F = \frac{|B|^2}{4\pi A}$$

(3.4)

Another issue of the shape feature extraction is the effectiveness of the shape description. The shape characteristics may be meaningless for some terrain objects with random shapes, such as natural forests, residential areas, and farmland, but they are important for roads, buildings, sports fields, and other targets with a regular shape. Therefore, the effectiveness of the shape feature is very important. The effectiveness of the shape can be defined as

$$\text{eff} = [Re, Fe]$$

(3.5)
where $Re \in \{0, 1\}$ is the effectiveness of aspect ratio, $Fe \in \{0, 1\}$ is the effectiveness of the form factor, and if the shape feature is meaningful for classification, it is assigned to be 1; otherwise, it will be 0. The assignment of shape effectiveness can be done with a supervised classification procedure, which can be conducted via four steps: (1) visually select a set of samples containing pixels with both effective and ineffective shape features, (2) generate the SANs of these samples and calculate their shape features, (3) use these samples to train a classifier (e.g., minimum distance classifier or maximal likelihood classifier) to classify the shape features (the output of the classification is 0 or 1), and (4) apply the classifier to assign the effectiveness of the shape features for other SANs.

Finally, the shape features of a SAN can be expressed as the Equation 3.6 in vector form. However, some managed forests and balanced residential areas may also have regular shapes. In this case, the determination of shape effectiveness should be more complicated than that presented above. For instance, texture features should be taken into account to help define the shape effectiveness. In this book, we only consider natural forest and ordinary residential areas.

$$SHA = [R, F, Re, Fe]$$ (3.6)

### 3.5.3.2.3 Integration of All the Features

The feature of a SAN contains the color feature, the texture feature, and the shape feature. The values of all these features are normalized into the same range $[0, 1]$. Finally, all features should be integrated to express the general feature of each SAN. This feature integration procedure is a data fusion procedure on the feature level, which can be illustrated by the following general model.

$$SANF = fusion(CF(k), TF(m), SF(n))$$ (3.7)

where $CF(k)$ is the color feature, $TF(m)$ is the texture feature, and $SF(n)$ is the shape feature. Each feature is multidimensional and can be represented as a vector, and the variants $k, m, n$ are the number of vector elements for color, texture, and shape features, respectively. The function $fusion()$ is the fusion method used to integrate all the features. In Equation 3.7, the function $fusion()$ is only a concept model. Many existing approaches can be used to conduct this operation, such as principal component analysis (PCA), Fourier transform, and wavelet transform, or simply link all three vectors of features into a larger vector if there are only a small number of features. In the experiment conducted in this study, there is one color feature (Equation 3.2), three texture features (Equation 3.3 where $H = [1, 2, 3]$), and four shape features (Equation 3.6). Thus, only the simple vector-based approach is used as there are only a total of eight features ($= 1 + 3 + 4$) in the study case.
3.5.4 Fusing the Optical and SAR Data

A flowchart of fusing the optical and SAR images is shown in Figure 3.12, illustrating the details of the combination of optical and SAR images to classify impervious surfaces. Several issues are important for the fusion of optical and SAR data, including the coregistration of the two data sources,
the feature extraction methods, the comparison between a single use of optical or SAR data, the comparison of the differences of various levels of fusion strategies, and the fusion methods for combining the two data sources. Since the feature extraction methods have already been described in Section 3.5.3, this section will focus on the methods of image coregistration, the comparison of the single use of the two data sources, the comparison of different levels of fusion, and the fusion methods between optical and SAR images.

3.5.5 Result Validation and Accuracy Assessment

For supervised classifiers (e.g., MLP, SVM, and RF), both the size and quality of the training data are highly significant in determining the performance of classification, while testing data is imperative in assessing the accuracy of both supervised and unsupervised classifications. In order to quantitatively assess the accuracy of the ISE, an appropriate sampling framework should be used. Five main types of sampling schemas were summarized previously by Jensen (2007): the simple random schema, the system schema, the layer schema, the layer system schema, and the cluster schema, where different backgrounds represent different classes in the layer schema. Different schemas are suitable for different cases and require different amounts of work. Among these schemas, the cluster schema is the most convenient to conduct, with much less labor than the other schemas. In this study, the cluster schema was applied to the study area by sampling the cluster test data over the satellite data with the aid of visual interpretation of the satellite data, digital orthophoto data, and in situ data. Moreover, very-high-resolution data from Google Earth is also used to help visually interpret the satellite data for the result validation in this study.

3.6 Conclusion

This chapter presented the methodology used in this book. A general framework of methodology was first introduced to explain the logical relationship of land cover diversity, remote sensing responses, and ISE. Second, the per-pixel modeling of ISE was presented as a basic strategy to estimate impervious surfaces in this research, followed by an introduction of the approaches to investigate the seasonal effects of ISE in tropical and subtropical regions. Third, the feature extraction methods were presented in detail. Following an introduction to the conventional feature extraction approach based on GLCM, a novel approach based on the SAN was presented with technical details. Fourth, a methodological framework of fusing the optical and SAR
data was presented with methods of image coregistration, investigation of the advantages and disadvantages of optical and SAR data, comparison of different fusion levels, and the fusion procedure with supervised classifiers. Finally, the sampling methods for training and testing datasets and the accuracy assessment approach were presented.