Population Density and Image Texture: A Comparison Study

XiaoHang Liu, Keith Clarke, and Martin Herold

Abstract

The correlation between census population density and Ikonos image texture was explored. The spatial unit for the analysis was census blocks with homogenous land-use. Ikonos image texture was described using three methods: the gray-level co-occurrence matrix (GLCM), semi-variance, and spatial metrics. Linear regression was conducted to explore the correlation between image texture and population density. It was found that although correlation exists, its degree varies depending on the method used to describe image texture. The highest correlation is given by the spatial metrics method. This result suggests that the correlation between texture and population density is not strong enough to predict or forecast residential population. However, image texture does provide a base to refine census-reported population distribution using remote sensing. High-resolution satellite images therefore have the potential to support "smart interpolation" programs to estimate human population distribution in areas where detailed information is not available.

Introduction

Knowledge of the size and spatial distribution of human population in an urban area is essential for understanding social, economic, and environmental issues. Traditionally, census has been the primary source of information on population distribution and demographic characteristics. Due to the cost, frequency, and boundary designation problems associated with census (Openshaw, 1984), the utility of remote sensing for population estimation has been continuously explored since the 1950s. Various types of satellite imagery have been examined to study population distribution, including Landsat MSS (Iisaka and Hegedus, 1982), TM and ETM (Forster, 1985; Li and Weng, 2005), SPOT (Lo, 1995), and DMSP nighttime imagery (Dobson *et al.*, 2000; Sutton *et al.*, 2001). The successes of these studies vary, however one consensus reached is that for applications with small geographic extent, a spatial resolution of 0.5 to 5 m is necessary (Jensen and Cowen, 1999). Such a fine resolution was rarely available with previous satellite images.

The advent of very high spatial resolution satellite images such as Ikonos renewed the interest in urban remote sensing including using remote sensing to estimate human population count and its spatial distribution. It has been suggested that an objective assessment of the current status and future potential of urban remote sensing is necessary (Donnay *et al.*, 2001). For urban population studies, Ikonos provides urban specificities comparable to those derived from low-altitude aerial photographs. The panchromatic imagery of 1 m spatial resolution enables the counting of individual dwelling units, while the multispectral imagery of 4 m spatial resolution clearly reveals the differences between residential neighborhoods. With the recognized limitations of medium-resolution satellite sensors (e.g., MSS, TM, and SPOT) and excitement about the new generation of high spatial resolution images like Ikonos, an examination of the utility of the new sensors is necessary.

This paper intends to provide a discussion of the methodological challenges in conducting this line of research and present an examination of the correlation between Ikonos image texture and census population density. Remote sensing has long been used to estimate urban population and socio-economic parameters. Several surrogates derivable from remotely sensed imagery have been examined in the literature, such as the extent of an urbanized area, the spectral reflectance value, and the proportion of each land-use class. In this study, Ikonos image texture will be examined as a correlate of population density. The fundamental assumption is that neighborhoods with similar housing characteristics tend to have similar population density. Housing characteristics can be described by the size of the houses, the greenness, and other associated conditions. The interaction among these factors forms texture. Different housing conditions reveal different textures in remotely sensed imagery; consequently, if a relationship can be established between image texture and population density, texture can be use to inform studies of population count and its spatial distribution. Research in this paper will help to answer the following questions: Is there correlation between population density and Ikonos image texture?; Is the correlation strong enough to estimate population count?; Which method of texture description is most correlated with population density?. Answers to these questions will not only help to evaluate the utility of very high-resolution images for socioeconomic applications but also help to identify the challenges and future research needs in urban remote sensing.

Remote Sensing and Studies of Human Population

Research using remote sensing to study human population started in the 1950s and advanced in two interrelated

XiaoHang Liu is at San Francisco State University, Department of Geography & Human Studies, 1600 Holloway Avenue, HHS 279, San Francisco, CA 94132 (xhliu@sfsu.edu).

Keith Clarke and Martin Herold are at the University of California, Santa Barbara, CA 93106 (kclarke@geog.ucsb, martin@geog.ucsb.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 72, No. 2, February 2006, pp. 187–196.

0099-1112/06/7202–0187/$3.00/0
© 2006 American Society for Photogrammetry and Remote Sensing

directions. One direction is to use remote sensing to provide accurate estimation of population count (e.g., Wellar, 1969; Anderson and Anderson, 1973; Lo, 1995; Sutton *et al.*, 1997; Chen, 2002; Harvey, 2002), and the other is to use remote sensing as ancillary information to refine census-reported spatial distribution of population, i.e., the so-called dasymetric mapping with remote sensing (e.g., Donnay and Unwin, 2001; Mennis, 2003). Although their specific aims differ, both types of research are based on the assumption that correlation exists between remote sensing surrogates and population density. To establish the correlation, three questions have to be answered: (a) What is the spatial unit to be used?, (b) Which surrogates of remote sensing image are correlated with population density?, and (c) If a regression model is established between the remote sensing surrogate and population density, what form does the regression take?

There are many answers to these questions in the literature. Choices of spatial unit include land-use zone (Anderson and Anderson, 1973; Donnay and Unwin, 2002; Mennis, 2003), image pixel (Lo, 1995; Sutton *et al.*, 1997), and census-reporting unit (Chen, 2002). The selection of the spatial unit determines the difficulty of obtaining training samples and remote sensing surrogate. For example, if image pixel is used, obtaining the corresponding ground population can be quite challenging (Harvey, 2002). On the other hand, if census reporting unit is used, it is easy to collect the population information of the training samples, however, the land-use class and texture of a census-reporting unit may be heterogeneous, making the average digital number or texture not directly appropriate to study the correlation.

Once the spatial unit is determined, a remote sensing surrogate must be selected so as to examine its correlation with population density. The most widely used surrogate is categorical land-use information. The assumption behind this method is that each land-use class has a characteristic population density. By image interpretation, it is possible to obtain the area of each land-use class and hence the corresponding population count (e.g., Anderson and Anderson, 1973; Mennis, 2003). The problem with this approach is that although the difference between land-use classes is recognized, the differences within a land-use class are ignored, Not all residential areas have the same population density, as evidenced by the contrast between detached-housing and multiple-unit housing. Even within a detached-housing area, there is a difference between low-density, medium-density, and high-density. To incorporate such considerations, one can conduct a more detailed classification in residential area and associated different population densities with different residential categories (Donnay and Unwin, 2001). Although this method improves the estimation, the fundamental problem remains unsolved. Population density is a continuous variable, yet land-cover and land-use class is a discrete variable. To associate a land-use with a single population density is equivalent to modeling a continuous variable with a discrete variable and is bound to introduce errors unless remedial work is conducted. Another problem with the above method is that it fails to consider the impact of location on population density. Empirical studies have repeatedly shown that geographic location and neighborhood have an impact on population density. For example, residential areas close to urban center and transportation systems tend to have higher population density.

An alternative to the categorical land-use information is image spectral data or texture (Lo, 1995; Sutton *et al.*, 1997; Chen, 2002). Because digital value and texture are continuous, they are not subject to the problems associated with using categorical land-use information. The challenge in this method is to select an appropriate approach which can describe zone-average reflectance or texture. Between the two approaches, texture is believed to be more advantageous for high-resolution images such as Ikonos. This is because texture not only utilizes the spectral information but also takes into account the spatial configuration of pixels. In the real world, the arrangement of vegetation and dwelling units, the uniformity of dwelling units, and the distance between adjacent ones all contribute to explain the differences in population density.

This research examines the correlation between census population density and image texture. Three methods to describe image texture will be tested and compared. The result will provide guidance to future research on using high-resolution satellite image to improve population estimation.

Methods

Homogenous Urban Patches (HUP)

The spatial unit of this study is homogenous urban patch (HUP) as defined by Herold *et al.* (2002b). This concept is based on the concept of *photomorphic region* widely used in aerial photographic interpretation (Peplies, 1974). It has the following characteristics:

1. An HUP has a homogenous texture which is visibly different from that of the neighboring HUPs.
2. An HUP may have several land-cover types (e.g., built-up, vegetation, etc.) within it, but has only one land-use (e.g., commercial). The built-up areas in a residential HUP are usually similar in terms of size, density, and spatial arrangement.
3. Where possible, HUP boundaries follow streets and other relevant natural and anthropogenic features such that large built-up patches remain contiguous for their delineation.
4. An HUP should be sufficiently large. Very small homogenous areas are too small for urban land-use characterization.

The HUP concept is fundamental to subsequent examination of the relationship between image texture and population density. Compared to other spatial units, such as a quadratic filtering window (kernel) or pixel (Gong *et al.*, 1992) which are traditionally used in urban remote sensing, an HUP has the advantage of allowing the characterization of thematically defined, irregularly shaped areas. For raw Ikonos imagery, the boundaries of HUPs may be obtained using texture-based image segmentation which maximizes between-patch textural differences while minimizing within-patch differences. Since the focus of this study is the correlation between population density and image texture, the HUP boundaries were delineated by an experienced image analyst through visual interpretation. Field checking and peer evaluation confirmed that the boundaries of the HUPs are highly accurate, which minimizes the propagation of errors in HUP boundaries to further analysis. Figure 1 illustrates HUP boundaries. The land-use of each HUP is also coded according to a modified Anderson III classification scheme.

Study Area and Source Data

The study area of this research is coastal Santa Barbara County, California (Figure 2). It is located 170 km northwest of Los Angeles in the foothills of the California Coast Range. The area is characterized by different types of land-use including residential areas with different density and socioeconomic structure, commercial and industrial districts, and open spaces (e.g., farm land and wetland). In addition to Ikonos image and census data, rich datasets are available for this area, which significantly facilitates sampling, image interpretation, and accuracy assessment.

To examine the correlation between image texture and population density, population data on block level from Census 2000 and an Ikonos image were acquired. Errors and positional inaccuracy of census data was corrected with the use of parcel maps and building footprint maps which show the shape and position of the buildings in each parcel. The Ikonos imagery was processed by mosiacing seven individual multispectral Ikonos images acquired between March and July 2001. The resulting mosaic covers the entire study area. Since the images were acquired on different dates with varying atmospheric and illumination conditions, geometric and atmospheric corrections were conducted to create a geometrically rectified and normalized image mosaic. Details on the preprocessing of the Ikonos images can be found in Herold *et al.* (2002a).

Texture Measurement

To examine the correlation between image texture and population density, a method to describe the texture of an HUP must be identified first. Many methods exist to describe the texture of a remotely sensed image, for example, the standard deviation method (Arai, 1993), the contrast between neighboring pixels (Edwards *et al.*, 1988), local variance (Woodcock and Harward, 1992), and fractal dimensionality (Lam, 1990). Commonly used methods include the gray-level co-occurrence matrix (GLCM) (Haralick, 1973), semi-variance (Lark, 1996), and wavelet analysis (Zhu and Yang, 1998). These methods are typically applied with a rectangular window or kernel. In this study, the spatial unit for texture analysis is homogenous urban patch (HUP) whose shape can be irregular and non–uniform. The gray-level co-occurrence matrix and semi-variance are examined along with a third approach using spatial metrics. Note that GLCM and semi-variance work with gray-level pixel values whereas spatial metrics are applied to classified land-cover maps only. Below is the review of these three methods.

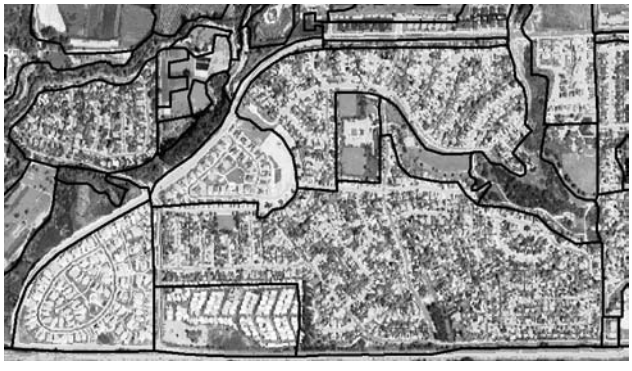

Figure 1. HUP boundaries in black lines overlaid on an Ikonos image. A color version is available at ASPRS website, URL: www.asprs.org.

Gray-level Co-occurrence Matrix (GLCM)

GLCM tabulates the frequency of one gray tone appearing in a specified spatial relationship with another gray tone within the area under investigation (Baraldi and Parmiggiani, 1995). Each element in GLCM is the estimated probability of going from gray-level i to gray-level j given the displacement vector which consists of a direction and distance. Detailed discussion on GLCM calculation can be found in Haralick *et al.* (1973). For this study, the GLCM analysis uses an isotropic displacement vector. This is because the purpose of texture analysis is to examine whether texture can be used to estimate residential population density. The population density of an HUP is determined by the number of buildings and the population count in each building. As long as these two numbers remain the same, whether the buildings are lined up horizontally, vertically, or diagonally should not affect the population density. Therefore, it is reasonable to assume that the direction factor in the displacement vector is isotropic. For the distance vector, since little theoretical guidance to its selection exists, an empirical experiment was conducted. Due to the intensive computational cost associated with GLCM, distance factors varying from one to nine pixels were examined. For an isotropic displacement vector with a distance of h, the frequency of going from pixel i to pixel j is counted as follows: for each pixel i in an HUP, find the square window of $(2h + 1) \times (2h + 1)$ centered on the pixel. The boundary of this window consists of $4 * (2h + 1)$ pixels. If any of the 8 h pixels does not belong to this HUP, pixel i is then excluded from GLCM calculation. Otherwise, for each gray-level j, find the number of pixels with intensity j in the window

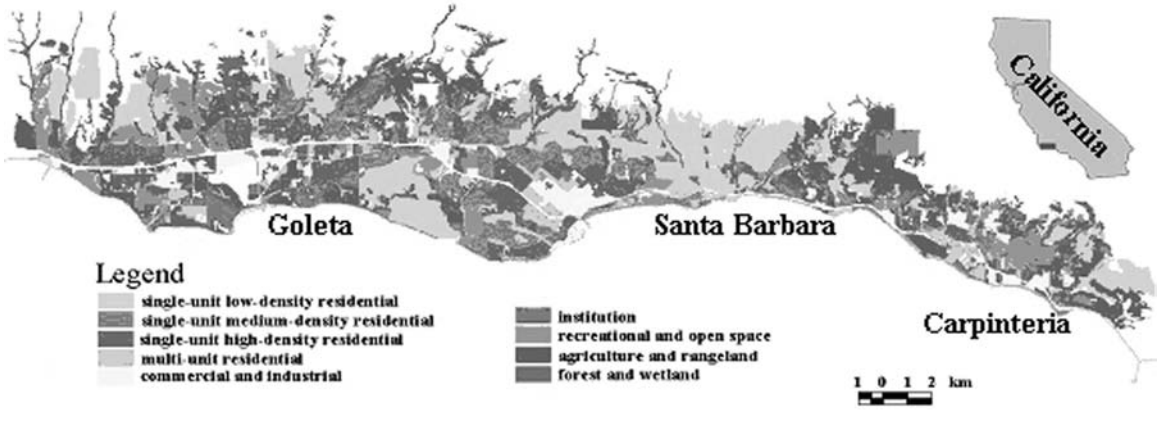

Figure 2. Map of the study area with land-use as the background (after Herold *et al.* (2003), used with permission). A color version is available at ASPRS website, URL: www.asprs.org.

boundary and update the (*i,j*) entry in GLCM. This process is applied to every pixel in the HUP. The result is a GLCM which is then normalized to obtain the normalized GLCM. The normalized GLCM is further compressed into six texture descriptors defined by Haralick (1973) and presented in Table 1.

GLCM requires the specification of the lag distance. Since no prior knowledge is available on the optimal lag distance, an isotropic displacement vector ranging from one to nine pixels was tested. Larger lag distance would result in very sparse GLCM matrixes and were not examined. For each HUP, the following GLCM-based texture descriptors are available:

$$T_i = [Energy_{1,\ldots,9}, Entropy_{1,\ldots,9}, Contrast_{1,\ldots,9}, Correlation_{1,\ldots,9}, Variance_{1,\ldots,9}, Homogeneity_{1,\ldots,9}]$$

where $Entropy_{1,\ldots,9}$ denotes $entropy_1$, $entropy_2$,..., $entropy_9$ which correspond to the entropy values calculated using isotropic displacement vector of one to nine pixels respectively. This notation applies to other variables also.

Texture Description Using Semi-variance

The pixel values within an HUP represent a regionalized variable whose spatial variation can be described by semi-variance (Issaks and Srivastava, 1989):

$$\hat{r}(h) = \frac{1}{2N_h}\sum_i (z_i - z_{i+h})^2 \quad (1)$$

where h is the lag distance along the specified direction, $\hat{r}(h)$ is the semi-variance at lag h, z_i and z_{i+h} are the values of a pair of pixels separated by a distance of h pixels, and N_h is the number of such pairs in the study area. Semi-variance can describe the spatial and directional dependence of the variance of pixel values at characteristic scales, as well as the spatial periodicity of this variation. It is therefore considered as a useful tool to describe texture (Woodcock *et al.*, 1988; Lark, 1996).

Like GLCM, the calculation of semi-variance requires the specification of a direction factor. This study examines isotropic semi-variance only. Isotropic semi-variances of each HUP were calculated using varying lag distances to produce an experimental semi-variogram. Typically, semi-variance rises with distance until a sill value is achieved. The distance where the sill is reached is referred to as *range*. Although it is possible to fit a theoretical semi-variogram model to the semi-variances, it is not done in this research for two reasons: (a) Accurate semi-variogram modeling usually requires manual curve fitting. To use a theoretical semi-variogram to describe the texture of an HUP, manual fitting would have to be conducted on HUPs one by one. Since there are more than 2000 HUPs in the study area, this process is not feasible. (b) Not every semi-variance can be easily fitted by a theoretical semi-variogram. For these reasons, empirical semi-variances corresponding to lag distances from 1 pixel to 20 pixels were calculated for each HUP. These 20 semi-variances make up one vector which has the form:

$$T_i = [\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_{20}] \quad (2)$$

where $\hat{\gamma}_j$ is the semi-variance calculated when lag distance is j pixels. This vector is used to describe the texture of a HUP. The calculation of the semi-variances was completed using UNIX-based GSLIB software (Deutsch and Journel, 1997).

Spatial Metrics

The goal of texture analysis in this study is to determine whether correlation exists between the population density and image texture of a residential HUP. Residential areas are characterized by diverse materials such as asphalt, metal, plastic, glass, shingles, water, vegetation, and bare soil. Although the digital number values of these materials are different, they can be grouped into a limited few categories such as vegetation, or built-up area. Ridd (1995) proposed a V-I-S model to categorize urban areas where V, I, and S stand for vegetation, impervious surface, and soil, respectively. For the purpose of studying population density in this research, his classes were modified into three new categories: vegetation, built-up area, and others. To characterize the spatial configuration of these materials, one could use their raw DN values as in GLCM or semi-variance, or use the categorical information as in the spatial metrics method.

Spatial metrics was developed in landscape ecology to quantify the environmental patterns of a natural landscape. Recently, it has been used to understand urban environments (Herold *et al.*, 2003). Many landscape metrics exist (McGarigal *et al.*, 1995). Unlike GLCM and semi-variance which uses the pixel values in a gray-level image, landscape metrics uses the patches in a categorical map as the spatial unit. In the context of this study, a patch is defined as a

TABLE 1. GLCM-BASED TEXTURE DESCRIPTORS

Texture Descriptor	Description
Energy $= \sum_{i,j=0}^{N_g-1} g^2(i,j)$	Measures texture uniformity, i.e., pixel pair repetitions. High energy occurs when the distribution of gray-level values is constant or period.
Entropy $= \sum_{i,j=0}^{N_g-1} g^2(i,j)\ln g(i,j)$	Highly correlated to energy. Measures the disorder of an image. Entropy is high when an image is not texturally uniform.
Contrast $= \sum_{i,j=0}^{N_g-1} (i-j)^2 g(i,j)$	Contrast measures the difference between the highest and lowest values of a contiguous set of pixels. A low-contrast image features low spatial frequencies.
Correlation (Corr) $= \sum_{i,j=0}^{N_g-1} (i-u)(j-u) g(i,j)/\sigma^2$	Measures the linear dependency in the image. High correlation values imply a linear relationship between the gray-levels of pixel pairs.
Variance (Var) $= \sum_{i,j=0}^{N_g-1} (i-u)^2 g(i,j)$	A measure of heterogeneity. Variance increases when the gray-level values differ from their mean.
Homogeneity (Homo) $= \sum_{i,j=0}^{N_g-1} \frac{1}{1+(i-j)^2} g(i,j)$	Measures image homogeneity. Sensitive to the presence of near-diagonal elements in a GLCM.

where N_g is the number of gray levels, $g(i,j)$ is the entry (i,j) in the gray-level co-occurrence matrix and $u = \sum_{i,j=0}^{N_g-1} i \cdot g(i,j)$ and $\sigma^2 = \sum_{i,j=0}^{N_g-1} (i-u)^2 g(i,j)$.

group of adjacent pixels classified into the same land-cover category. For example, a house occupying 20 × 20 meters on the ground corresponds to 5 × 5 pixels in the multispectral Ikonos image. Although these 5 × 5 pixels may have different DN values, they are all classified as built-up, and therefore form a single patch. Similarly, a contiguous vegetated area, such as a lawn, also forms a patch. There can be many patches (built-up, vegetation, and others) in an HUP. Their spatial configuration determines the texture of the HUP. To describe the characteristics of the texture of an HUP comprehensively, the following aspects were considered when selecting spatial metrics:

1. Composition describes the variety and abundance of patch types within an HUP without considering their spatial placement. This type of landscape metric is useful in differentiating residential types since the percentage of built-up area or vegetation in a HUP has a correlation with the population density. For example, a low-density residential area is usually characterized by a higher percentage of trees and grasses, while a high-density residential area has more built-up patches within it. The principal measures of composition in this study are Percentage of Landscape (PLAND) which measures the proportional abundance of built-up area and vegetation in an HUP, and Shannon's Diversity Index (SHDI) which is a composite measure of the number of patches present and the relative dominance of a patch.
2. Spatial configuration describes the spatial arrangement, position, or orientation of patches within an HUP. The principal aspects of spatial configuration are patch density, isolation/proximity, connectivity, and contagion. Patch density (PD) measures the number of patches per unit area. It can be used to compare the degree of fragmentation among HUPs. For example, houses in high-density, single-unit housing areas are smaller and are separated by vegetation and impervious surfaces, resulting in a higher degree of fragmentation. Isolation/proximity refers to the tendency for patches to be relatively isolated in space from other patches of the same or similar class. This study concerns the isolation of built-up patches only, and uses the Euclidean Nearest Neighbor Distance (ENN) to describe it. Both the Mean ENN (ENN_MN) and the Standard Deviation of ENN (ENN_SD) are calculated. ENN_MN measures the average distance between two adjacent houses while ENN_SD indicates the regularity of ENN. In a residential areas where houses display a high degree of orderliness, ENN_SD is expected to be low. Cohesion (COHESION) is computed from the information contained in patch area and perimeter. Contagion refers to the tendency of patch types to be spatially aggregated, that is, to occur in large, aggregated or "contagious" distributions. Contagion ignores patches *per se* and measures the extent to which cells of similar class are aggregated. The contagion index (CONTAG) used in this study is based on the probability of finding a cell of type *i* next to a cell of type *j*. This index increases in value as a landscape is dominated by a few large, contiguous patches, and decreases in value with increasing subdivision and interspersion of patch types. The index summarizes the aggregation of all classes and thereby provides a measure of overall "clumpiness" of the landscape. The spatial metrics used in this study are summarized in Table 2.

In this study, the landscape metrics-based texture analysis requires categorical land-cover information as the input. A land-cover map consisting of three categories: built-up area (built), vegetation (veg), and others were obtained using the method discussed in the next section. The texture of each HUP is thus described by the following vector:

$$T_i = [PLAND_{built}, PLAND_{veg}, PD_{built}, PD_{veg}, ENN\_MN_{built}, ENN\_SD_{built}, COHESION, CONTAG, SHDI]$$

Texture is an intrinsically complex feature. Although three texture- measuring methods are proposed, each has some shortcomings and may not capture all aspects. Nevertheless, they serve as a good starting point for exploratory analyses.

Data Preparation

Since GLCM and semi-variance analysis can only be conducted on gray-level images, a gray-level image must be created first. Two such images were prepared. The first uses the near infrared band (i.e., Band 4) of the Ikonos image. This band was selected because a standardized principal component analysis showed that the near infrared band was the most significant single band and accounted for 74.3 percent of the total variances in the multi-spectral image. The other gray-level image was based on the normalized difference vegetation index (NDVI) value of each pixel. NDVI is scaled to become an integer between 0 and 255. NDVI is very efficient to differentiate built-up areas from vegetation. Since population density correlates with the density of built-up area and vegetation, NDVI is used in this research. In the following sections, these two images will be referred to as the NIR image and the NDVI image, respectively.

The spatial metrics method runs on categorical maps only. The categorical map used in this study is the land-cover map obtained by applying object-oriented image analysis on the multi-spectral Ikonos imagery using eCognition®. The map has three classes: built-up, vegetation, and the others. Details on the pre-processing and image classification of Ikonos imagery can be found in (Herold *et al.*, 2002). By overlaying the land-cover map with HUP boundaries, it is possible to determine the patches within each HUP. The overall accuracy and Kappa coefficient of the land-cover classification is 82.4 percent and 71.4 percent, respectively.

As defined previously, each HUP may contain several land-cover patches, but has only one land-use. While delineating the HUP boundaries, the land-use of each HUP is also determined according to a modified Anderson III classification scheme. Residential areas were classified into low-density single-unit housing, medium density single-unit housing, high density single-unit housing, multiple-unit housing, and commercial-residential mixed land-use. Because the image analyst is very familiar with the study area, both the HUP boundaries and the land-use of each HUP are very accurate and serve as the reference for accuracy assessment. Figure 3 provides an illustration of the Ikonos composite, land-cover classification, and reference data showing the footprint of buildings and roads.

Training Samples

To examine the correlation between image texture and population density, some training samples must be collected. Image texture can be calculated using the three texture methods aforementioned. However, obtaining the population information for a selected spatial unit has been a challenge for some studies (Harvey, 2002). If HUP is used as the spatial unit, the corresponding population information is hard to determine since few HUPs have boundaries that coincide perfectly with that of census-reporting units. Conversely, although the population of each census unit is known, many of them have multiple land-uses, making the texture not uniform. Texture calculation on these units does not seem reasonable. Because the census population data is on block level, the final spatial unit utilized was those census blocks that fall completely into a single HUP. These census blocks are suitable as training data because their textures and population densities are known. Moreover, because HUP is defined as a spatial unit with homogenous image texture, a subset of an HUP should have the same image texture as the HUP. This suggests that the image

TABLE 2. DESCRIPTION OF SPATIAL METRICS

Spatial Metrics	Description
PLAND – percentage of landscape $\text{PLAND} = P_k = \frac{\sum_l a_{kl}}{A}(100)$ P_k = proportion of the landscape occupied by land-use class k. a_{kl} = area (m^2) of patch kl, *i.e.* patch l with land-use k. A = total landscape area (m^2)	PLAND quantifies the proportional abundance of each patch type in the landscape. For this research, it is used to describe the percentage of buildings, vegetation, and other land-cover types in a land-use polygon.
PD – patch density $\text{PD} = \frac{N}{A}(10{,}000)(100)$ N = total number of patches in the landscape A = total landscape area (m^2)	PD describes the number of patches on a per unit area basis that facilitates comparisons among land-use polygons of varying size. The higher the PD, the more fragmented the land-use polygon. High density single-unit housing is expected to have the highest PD value, while forest/rangeland may have the lowest.
ENN – Euclidean Nearest-Neighbor Distance $\text{ENN} = h_{kl}$ h_{kl} = distance (m) from patch kl to the nearest neighboring patch of the same class k, based on patch edge-to-edge distance, computed from cell center to cell center. ENN-MN: Euclidean mean nearest neighbor distance ENN-SD: Euclidean nearest neighbor distance standard deviation	Equals the distance (m) to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance. ENN is a simple measure of isolation. For buildings, ENN can help to describe whether the houses are spaced regularly. High density single-unit housing display highest orderliness and close to each other. Therefore, its ENN is expected to be small and has low standard deviation.
$\text{Cohesion} = \left[1 - \frac{\sum_j p_{kl}^*}{\sum_l p_{kl}^* \sqrt{a_{kl}^*}}\right] \cdot \left[1 - \frac{1}{\sqrt{z}}\right]^{-1} \cdot (100)$ p_{kl}^* = Perimeter of patch kl in terms of number of cell surfaces a_{kl}^* = area of patch kl in terms of number of cells Z = total number of cells in the landscape	Patch cohesion index measures the physical connectedness of the corresponding patch type. Cohesion approaches 0 as the proportion of the landscape comprised of the focal class decreases and becomes increasingly subdivided and less physically connected. Cohesion increases monotonically as the proportion of the landscape comprised of the focal class increases.
CONTAG – Contagion $\text{CONTAG} = \left[1 + \frac{\sum_{k=1}^{K}\sum_{n=1}^{K}\left[(P_k)\left[\frac{g_{kn}}{\sum_{n=1}^{m} g_{kn}}\right]\right] \cdot \left[\ln(P_k)\left[\frac{g_{kn}}{\sum_{n=1}^{m} g_{kn}}\right]\right]}{2\ln(K)}\right](100)$ K = the number of land-use classes g_{kn} = number of adjacencies between pixels of patch types k and n P_k = proportion of landscape occupied by patch type k	Contagion refers to the tendency of patch types to be spatially aggregated; that is, to occur in large, aggregated or "contagious" distributions. Contagion is an aspect of landscape texture. Contagion approaches 0 when the patch types are maximally disaggregated (i.e., every cell is a different patch type) and interspersed (equal proportions of all pairwise adjacencies). Contagion equals to 100 when all patch types are maximally aggregated, i.e., when the landscape consists of single patch.
SHDI – Shannon's Diversity Index $\text{SHDI} = -\sum_k (p_k^* \ln P_k)$ P_k = proportion of the landscape occupied by land-use i	SHDI equals to 0 when the landscape contains only 1 patch (i.e., no diversity). SHDI increases as the number of different patch types (i.e., patch richness, PR) increases and/or the proportional distribution of area among patch types become more equitable.

texture of a census block is similar to that of the HUP it falls in. On the other hand, the assumption behind using image texture to assist population estimation is that areas with similar image texture have similar population density. Since the census blocks have the same texture as the HUP it falls in, the population density of the HUP is similar to that of the census block. Therefore, the correlation between texture and population density on these training blocks can be used to infer the correlation on HUP level. There were 1,578 blocks used as training samples. Each of these blocks is located within a single HUP and has residential land-use. They display varying textures on Ikonos image and the land-cover map. Their populations were obtained from the census, and their textures were calculated using the three methods discussed previously.

Results and Discussion

Regression analysis was performed on the 1,578 training samples. The dependant variable was the average population

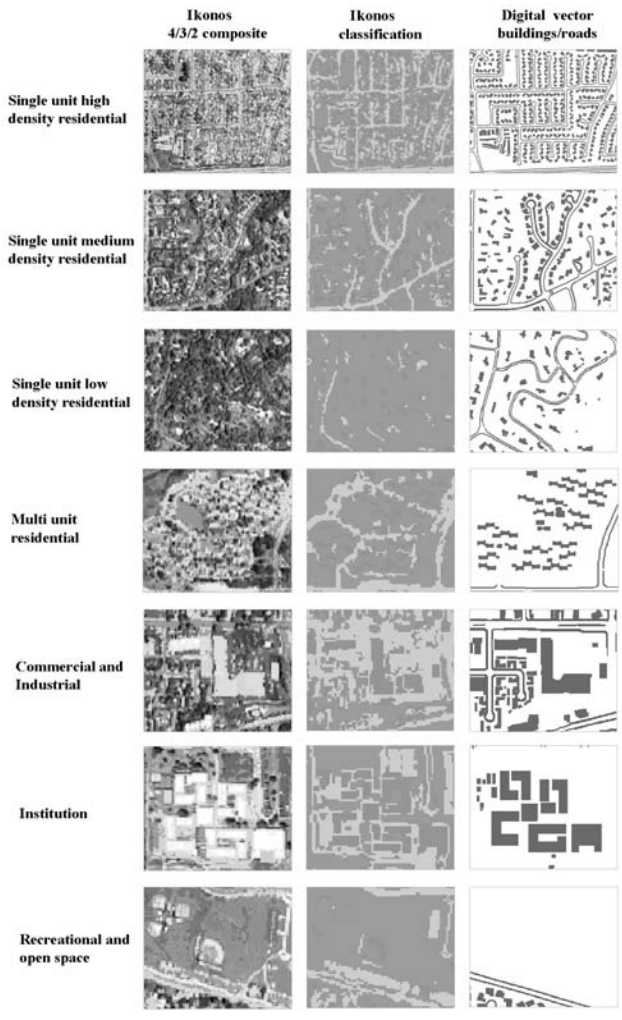

Figure 3. Examples of spatial land-cover configurations for major urban land-use categories shown as an Ikonos false color composite, an Ikonos classification result with buildings in red and vegetation in green. The digital vector building/roads data is included here to provide a ground reference of built-up area. (Modified after Herold *et al*., 2003) A color version is available at ASPRS website, URL: www.asprs.org.

density of each census block. The explanatory variables are texture descriptors. The three methods to measure texture (GLCM, semi-variances, and spatial metrics) were examined. Since in each texture method there are several texture descriptors which may not contribute equally to explain the variation in population density, stepwise regression was conducted in order to retain only the significant variables in the final regression model.

As an exploratory data analysis, the simplest linear regression model between population density and texture descriptors is examined:

$$p_i = \beta_0 + \sum_{j=1}^{n} \beta_j T_{ij} + \varepsilon_i \quad (3)$$

where p_i is the population density of a training sample, T_{ij} is texture descriptor j for sample i, n is the number of explanatory variables for indicators; and β_i are parameters to be estimated from the data. This form was reported by Chen (2002). However, the training data in this case study does not seem to support this model. In fact, few texture descriptors show a direct linear relationship with population density. Instead, it is the logarithmic transformation of population density which demonstrates a much stronger linear correlation with texture. The model was therefore modified to:

$$\ln p_i = \beta_0 + \sum_{j=1}^{n} \beta_j T_{ij} + \varepsilon_i \quad (4)$$

The logarithmic transformation has been reported by several other researchers, including Anderson and Anderson (1973), Sutton *et al*. (2001), Harvey (2002). The results of this study also seem to also support such a relationship.

To compare the correlation, regression coefficient (R^2) is used. Although there are imperfections and cautions associated with this measurement (Harvey, 2002), this measurement is popularly used in the literature. For the purpose of this research whose goal is to examine and compare the correlation between image texture and population density, R^2 is considered appropriate, but used with caution.

GLCM-based Regression

In GLCM-based analysis, the texture of each HUP is described using a vector containing six descriptors, each of which is calculated using lag distance from one to nine pixels. Each descriptor describes a different aspect of image texture; it is thus interesting to see which texture descriptor is mostly correlated with population density. The varying lag distances provide an opportunity to examine the impact of scale. Together, GLCM-based analysis will identify which texture descriptor at what scale is most significantly correlated with population density. On the basis of this study, a stepwise regression using all texture descriptors with all lag distances was conducted to obtain the best "fitting" model of GLCM. This model was subsequently compared with that obtained from semi-variances and spatial metrics based analysis.

The analysis was conducted on both NIR and NDVI images. The first set of experiments was conducted using each single GLCM texture descriptor. The correlation coefficients are listed in Table 3 where $entropy_{1,2}$ denotes that there are two independent variables, $entropy_1$ and $entropy_2$, which correspond to the entropy values calculated using isotropic displacement vector of one and two pixels, respectively. This notation applies to other variables also.

For the same texture descriptors, NDVI consistently demonstrates higher correlation than NIR as measured by R^2. Energy and variance did not seem to correlate well with population density, yielding an R^2 of nearly 0. In contrast, correlation and dissimilarity demonstrate higher correlation

TABLE 3. CORRELATION BETWEEN POPULATION DENSITY AND GLCM-BASED TEXTURE DESCRIPTORS

	R^2	
Texture Descriptor	NIR	NDVI
$entropy_{1,2}$	0.13	0.29
$energy_{1,2}$	0.03	0.16
$contrast_{1,2}$	0.02	0.24
$correlation_{1,2}$	0.06	0.29
$homogeneity_{1,2,3,4}$	0.02	0.19
$dissimilarity_{1,2}$	0.01	0.28
$variance_{1,2,3,4}$	0.01	0.02

TABLE 4. CORRELATION BETWEEN POPULATION DENSITY AND DIFFERENT TEXTURE DESCRIPTORS

		Independent Variables	R^2
GLCM	NIR	$entropy_{1,2}$, $energy_{1,2}$, $contrast_{1,2}$, $correlation_{1,2}$, $homogeneity_{1,2,3,4}$, $dissimilarity_{1,2}$, $variance_{1,2,3,4}$	0.35
	NDVI	$entropy_{1,2}$, $energy_{1,2}$, $contrast_{1,2}$, $correlation_{1,2}$, $homogeneity_{1,2,3,4}$, $dissimilarity_{1,2}$, $variance_{1,2,3,4}$	0.45
Semi-variance	NIR	var_i, $i = 1, \ldots, 20$; semi-variances with lags from 1 pixels to 20 pixels	0.20
	NDVI	var_i, $i = 1, \ldots, 20$; semi-variances with lags from 1 pixels to 20 pixels	0.20

with population density ($R^2 = 0.28 \sim 0.29$) when the NDVI image is used.

Once the individual texture descriptors are examined, they are combined to run a new iteration of regression analysis, i.e., using $entropy_{1,2}$, $energy_{1,2}$, $contrast_{1,2}$, $correlation_{1,2}$, hom$ogeneity_{1,2,3,4}$, and $dissimilarity_{1,2}$; the R^2 was raised to 0.45. The use of more displacement distances or texture measurements did not result in significant improvement in R^2 value. Therefore, an R^2 of 0.45 was considered as the highest correlation between image texture and population density.

Semi-variogram-based Analysis

For the semi-variogram method, semi-variances corresponding to lag distance between 1 to 20 pixels were examined on NDVI and NIR images. The R^2 obtained did not differ between the two images; both were 0.20. Including more semi-variance values calculated from more lag distances did not improve the R^2 significantly.

Spatial Metrics-based Analysis

Although nine spatial metrics were examined in land-use classification, only three of them showed significant correlation: percentage of built-up area, percentage of vegetation in the area, and the patch density of built-up area. The linear correlation has the following form:

$$\ln(d) = 8.819 + 1.772p_1 - 2.612p_2 + 0.0632p_3, R^2 = 0.55 \quad (5)$$

where d is the population density (people/km^2), p_1 is the percentage of built-up area ($PLAND_1$), p_2 is the percentage of vegetation ($PLAND_2$), and p_3 is the patch density of built-up area (PD_1).

The result of this analysis is interesting in the sense that the simple measurement of built-up and vegetation percentage together with patch density can explain a significant amount of variances in population density. Other landscape metrics, such as Euclidean distance and contagion, did not contribute as significantly as expected. The three straightforward descriptors appear to have much powerful explanatory power than the others.

The residuals of the linear regression are plotted in Figure 4a. A plot between texture-estimated population density and that of ground reference is shown in Figure 4b. It can be seen that the range of the population is quite wide, and there is no distinct threshold from low density to high density. The R^2 obtained is 0.55, which is significantly higher than that from the GLCM and the semi-variances method. This result suggests two things. One is that there is indeed some correlation between image texture and the natural logarithm of population density, although the correlation is not high enough to make reliable estimates of population. Moreover, the estimated

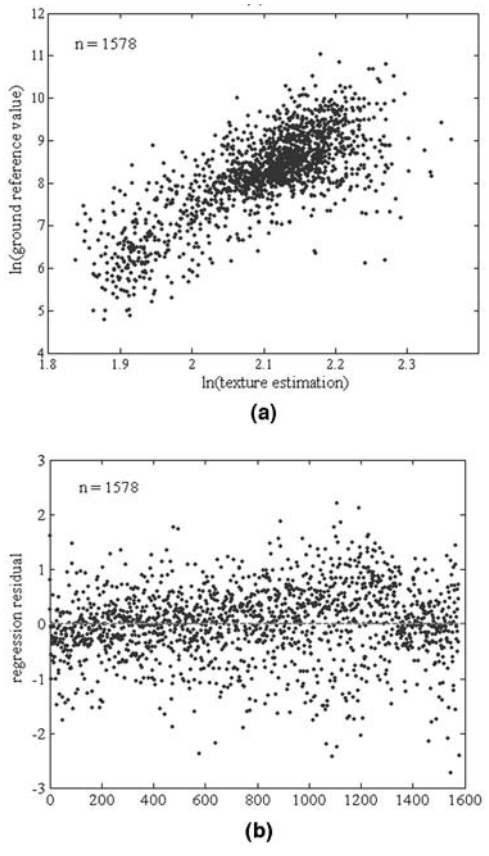

Figure 4. The correlation between spatial metrics and population density: (a) the comparison between texture-estimated population density and ground reference, and (b) plot of the residual of the regression.

population density is in logarithmic format, and exponential operations must be applied to obtain population density. During the process, small errors will be exponentialized, and thereby amplified. On the other hand, imperfect correlation indicates that significant residuals exist after the regression analysis. For studies on remote sensing surrogates such as texture or land-use category to improve census-reported population distribution, simple regression-based interpolation may not be sufficient. Residual modeling may be necessary especially if the total population count is to be preserved.

Summary

This study examined the correlation between image texture and residential population density. Homogenous urban patch (HUP) was used as the spatial unit for analysis. Three methods to describe image texture were examined: gray-level co-occurrence matrix (GLCM), semi-variances, and spatial metrics. Each method was examined using a various texture descriptors and lag distances, where appropriate. In all examinations, it was found that the natural logarithm of population density was linearly correlated with texture measurements. Among the three texture analysis methods, spatial metrics yielded the highest regression coefficient with population density. Three spatial metrics were found to be particularly useful: the percentages of built-up and vegetation in an area, and the patch density of built-up area.

The highest regression coefficient R^2 obtained was around 0.55. The significant residuals left from the linear regression analysis suggest two points: (a) To estimate population count through statistical regression using Ikonos imagery, texture alone is unlikely to be sufficient, and other interpretation keys or methods must be explored to achieve satisfactory accuracy; and (b) The correlation between texture and population density provides a basis to improve census estimation on the spatial distribution of population. The problem due to mixed land-use in a reporting unit can be potentially alleviated. However, residual modeling is necessary for accurate disaggregation, especially if the total population count is to be preserved. Few studies on population interpolation using remote sensing have discussed this point. Instead, the majority work focused on the correlation between remote sensing image feature and population density. Research on residual modeling and accuracy assessment are yet to be conducted.

The arrival of the new generation of high-resolution satellite imagery (e.g., Ikonos) has opened up the opportunity for detailed mapping and analysis of spatial characteristics within the urban environment. However, new methods are yet to be developed to best utilize this new opportunity. Many previous studies have examined the feasibility of using remote sensing to predict population or assist population interpolation. This research shows that remote sensing images can indeed help to estimate population density. However, the correlation may not be strong enough for empirical applications. In using remote sensing surrogates, the selection of image and method can all affect the final result. GLCM and semi-variances which have been used widely in the past on medium-resolution images may not be the best method to measure image texture for high-resolution satellite images. In contrast, spatial metrics method seems to be much more promising. However, the spatial metrics method requires information on HUP boundaries which is yet another challenge for the image interpretation community. In addition, texture alone does not seem to be powerful enough for this analysis. Shape, pattern, and association can be included in future studies. For high-resolution satellite images such as Ikonos, these factors are intuitively very important. Detailed analysis is yet to be conducted to examine quantitatively the utility of these elements.

Acknowledgments

The authors gratefully acknowledge the support by NSF Grant No. SBR-9817761. We would also like to thank Mellisa Kelly and Ryan Aubry for their contribution to this research.

References

Anderson, D.E., and P.N. Anderson, 1973. Population estimates by humans and machines, *Photogrammetric Engineering & Remote Sensing*, 39(1):147–154.

Arai, K., 1993. A classification method with a spatial-spectral variability, *International Journal of Remote Sensing*, 14: 699–709.

Baraldi, A., and F. Parmiggiani, 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters, *IEEE Transactions on Geoscience and Remote Sensing*, 33:293–304.

Chen, K., 2002. An approach to linking remotely sensed data and areal census data, *International Journal of Remote Sensing*, 23:37–48.

Deutsch, C., and A. Journel, 1997. *GSLIB: Geostastistical Software Library and User's Guide*, Second edition, Oxford University Press, New York, New York, 369 p.

Dobson, J.E., E.A. Bright, P.R. Coleman, R.C. Durfee, and B.A. Worley, 2000. LandScan: A global population database for estimating populations at risk, *Photogrammetric Engineering & Remote Sensing*, 66(7):49–857.

Donnay, J.-P., M.J. Barnsley, and P. Longley, 2001. Remote sensing and urban analysis, *Remote Sensing and Urban Analysis* (J.-P. Donnay, M.J. Barnsley, and P.A. Longley, editors), Taylor and Francis, New York, New York, pp. 3–18.

Donnay, J.-P., and D. Unwin, 2001. Modelling Geographical distributions in urban areas, *Remote Sensing and Urban Analysis* (J.-P. Donnay, M.J. Barnsley, and P.A. Longley, editors), Taylor and Francis, New York, New York, pp. 205–224.

Edwards, G., R. Landary, and K.P.B. Thomson, 1988. Texture analysis of forest regeneration sites in high-resolution SAR imagery, *Proceedings of the International Geosciences and Remote Sensing Symposium (IGARSS '88)*, 12–16 September, Edinburgh, Scotland (European Space Agency, ESA SP-284), pp. 1355–1360.

Forster, B., 1985. An examination of some problems and solutions in monitoring urban areas from satellite platforms, *International Journal of Remote Sensing*, 6, 139–151.

Gong, P., D.J. Marceau, and P.J. Howarth, 1992. A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data, *Remote Sensing of Environment*, 40:137–151.

Haralick, R., S.K., and I. Dinstein, 1973. Texture features for image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, 3:610–622.

Harvey, J.T., 2002. Estimating census district populations from satellite imagery: Some approaches and limitations, *International Journal of Remote Sensing*, 23:2071–2095.

Herold, M., A. Mueller, S. Guenter, and J. Scepan, 2002a. Object-oriented mapping and analysis of urban land use/cover using IKONOS data, *Proceedings of the 22nd EARSEL Symposium, Prague, Czech Republic*, 04–06 June, pp. 531–538.

Herold, M., J. Scepan, and K. Clarke, 2002b. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses, *Environment and Planning A*, (34):1443–1458.

Herold, M., X. Liu, and K. Clarke, 2003. Spatial metrics and local texture for mapping urban land-use, *Photogrammetric Engineering & Remote Sensing*, 69(9):991–1001.

Iisaka, J., and E. Hegedus, 1982. Population estimation from Landsat imagery, *Remote Sensing of Environment*, 12:259–272.

Isaaks, E., and M. Srivastava, 1989. *Applied Geostatistics*, Oxford University Press, New York, New York, 561 p.

Jensen, J.R., and D.C. Cowen, 1999. Remote sensing of urban/suburban infrastructure and socio-economic attributes, *Photogrammetric Engineering & Remote Sensing*, 65(5):611–622.

Lam, N.S.N., 1990. Description and measurement of Landsat TM images using fractals, *Photogrammetric Engineering & Remote Sensing*, 56(1):187–195.

Lark, R.M., 1996. Geostatistical description of texture on an aerial photograph for discriminating classes of land cover, *International Journal of Remote Sensing*, 17:2115–2133.

Li, G., and Q. Weng, 2005. Using Landsat ETM+ imagery to measure population density in Indianapolis, Indiana, U.S.A., *Photogrammetric Engineering & Remote Sensing*, 71(8):947–958.

Lo, C.P., 1995. Automated population and dwelling unit estimation from high-resolution satellite images: A GIS approach, *International Journal of Remote Sensing*, 16:17–34.

McGarigal, K., S.A. Cushman, M.C. Neel, and E. Ene, 2002. *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*, University of Massachusetts, Amherst, Massachusetts, URL: www.umass.edu/landeco/reesarch/fragstats/fragstats.html (last accessed: 14 November 2005).

Mennis, J., 2003. Generating surface models of population using dasymetric mapping, *The Professional Geographer*, 55(1):31–42.

Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data, *Environment and Planning A*, 16:17–31.

Peplies, R.W., 1974. Regional analysis and remote sensing: a methodological approach, *Remote Sensing: Techniques for Environmental Analysis* (J.E. Estes and L.W. Senger, editors), Hamilton Publishing Company, Santa Barbara, California, pp. 277–291.

Ridd, M.K., 1995. Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystems analysis through remote

sensing: comparative anatomy for cities, *International Journal of Remote Sensing*, 16:2165–2185.

Sutton, P., D. Roberts, C. Elvidge, and H. Meij, 1997. A comparison of nighttime satellite imagery and population density for the continental United States, *Photogrammetric Engineering & Remote Sensing*, 63(11):1303–1313.

Sutton, P., D. Roberts, C. Elvidge, and K. Baugh, 2001.Census from Heaven: An estimate of the global human population using nighttime satellite imagery, *International Journal of Remote Sensing*, 22(16):3061–3076.

Wellar, B, 1969. The role of space photography in urban and transportation data series, *Proceeding of Sixth International Symposium on Remote Sensing of Environment*, Volume II, 15 October, Ann Arbor, Michigan, pp. 831–854.

Woodcock, C.E., A.H. Strahler, and D.L.B. Jupp, 1988. The use of variogram in remote sensing: I. Scene models and simulated images, *Remote Sensing of Environment*, 25:323–348.

Woodcock, C., and V.J. Harward, 1992. Nested-hierarchical scene models and image segmentation, *Remote Sensing of Environment*, 25:323–348.

Zhu, C., and X. Yang, 1998. Study of remote sensing image texture analysis and classification using wavelet, *International Journal of Remote Sensing*, 13:3167–3187.

(Received 08 November 2004; accepted 11 January 2005; revised 28 January 2005)