Chen, K., 2003. Scale analysis for remotely sensed images and areal census data: Stressing objects. *Proceedings of the 30th International Symposium on Remote Sensing of Environment (CD-ROM) – Information for Risk Management and Sustainable Development*, 10-14 November 2003, Honolulu, Hawaii, 4 p.

Scale Analysis for Remotely Sensed Images and Areal Census Data: Stressing Objects

Keping Chen

Risk Frontiers – Natural Hazards Research Centre, Macquarie University
NSW 2109, Australia – kchen@els.mq.edu.au

Abstract – Spatial scale analysis for disparate geospatial data is facilitated by object-based reasoning and analysis. This paper introduces a wavelet transform-based approach for remotely sensed images and a variogram approach for areal census data; both stress objects as the core for performing scale analysis. By calculating a series of statistics from wavelet transform-based sub-images and energy signature images, it was found that with energy signature images, the approximate characteristic scale of scene variation can be suggested by the change rate of the standard deviation over two successive decomposition resolutions (ΔSD/ΔR). Four simulated fine spatial resolution images having different object specifications were analysed and in each case the method identified the approximate characteristic size of the salient objects in the image. An accurate result can be obtained more easily using the range of geostatistical variograms.

For areal census data, we correlate census dwelling data and image structural information. By calculating variograms at *individual areal units* rather than at *a whole scene level*, high correlations exist between variogram sills and dwelling densities of individual areal units. Thus, dwelling attributes across arbitrary areal units can be interpolated and scaled. This paper concludes that such scale analyses in a modern, digital spatial analysis environment should still conform to the classical principle of cartographic scale transform; that is, scale association with object representation. Using geostatistics for examining the cognate relations between scale and objects holds promise.

Keywords: Spatial scale; Objects; Remotely sensed images; Areal census data; Geostatistics; Wavelet transform.

1. INTRODUCTION

Exploring scale information inherent in geospatial data is critical to many multidisciplinary studies. This is particularly the case in the natural and social sciences given easy access to geospatial data such as remotely sensed images at different spatial resolutions and census data at hierarchical areal units.

The definition of scale is often ambiguous and this paper adopts contemporary usage that scale is equivalent to size (van Gardingen et al., 1997). Scale analysis often deals with pixel size changes of raster-based remotely sensed images and areal unit changes of vector-based census data. Unfortunately these abstract manipulations do not explicitly incorporate the topology of objects. This paper emphasises the importance of objects in scale analysis, and illustrates this using two examples. First, a wavelet transformed-based statistic is used to identify the characteristic scale of salient objects in images. Second, geostatistical variograms are calculated at areal unit zone levels (not at a whole scene level) to link variogram sills and underlying object counts.

2. REMOTELY SENSED IMAGES

Object-based methods are increasingly employed for image analysis, complementary to conventional pixel-based methods. Two conspicuous programs with emphasis on object-based image segmentation and classification are *eCognition* [www.definiens-imaging.com] and *Feature Analysis* [www.featureanalyst.com]. Scale analysis of remotely sensed images can benefit from object-based approaches, and two essential tasks include (1) identification of the characteristic scale of scene variations, and (2) scaling biophysical parameters.

2.1 Characteristic Scale of Scene Variation

Various techniques, including geostatistics, fractals (Quattrochi, and Goodchild, 1997; Tate and Atkinson, 2001), Scale-Space theory (Hay et al., 2002), and wavelet transform (Chen and Blong, 2003), have been employed to investigate the scale properties of images. Chen and Blong (2003) evaluated six statistics (mean, variance, standard deviation - SD, coefficient of variation, skewness and kurtosis) calculated from wavelet transform-based sub-images and energy signature images for identifying the characteristic scale of scene variation in both simulated and real world fine spatial resolution images. With energy signature images, the distinguishable change rate of standard deviation over spatial resolution ranges between two successive decomposition levels (ΔSD/ΔR) provided an estimate of the characteristic scale.

This paper further illustrates the utility of the statistic (ΔSD/ΔR) using images with the variable object attributes such as number, size, shape, intensity and background. A popular second-order Daubechies's wavelet was utilised. The four simulated images are shown in Fig. 1, each with the size of 1024 × 1024 pixels. Fig. 1(a) contains a simple scene with 100 black disc objects (diameter 33 pixels) on a white background; Fig. 1(b), 50 black disc objects at two sizes (diameters 25 and 101 pixels) on a grey background; Fig. 1(c), 50 rectangle objects (size 31 × 65 pixels) with random intensity levels on a white background; and Fig. 1(d), 100 disc objects (diameter 25 pixels) with different intensity levels on a noisy background. All objects are randomly distributed across the images. The relative position of each sub-figure in Figs. 2, 3 and 4 corresponds to the position of each image in Fig. 1. For example, Figs. 2(a), 3(a) and 4(a) all show the analysis results for Fig. 1(a).

The distinguishable change of ΔSD/ΔR with energy signature images (Fig. 2) indicates the characteristic scale for scene variations (corresponding to the size of salient objects), with consistent results. For example, in Fig. 2(a) a sharp change of ΔSD/ΔR at ΔR = 16 suggests that the size of objects is between ~16 and ~32 pixels. In Figs. 2(b), 2(c) and 2(d), the abrupt change of ΔSD/ΔR also closely corresponds with the size of the salient objets. By comparison, directional ΔSD/ΔR changes (Fig. 3) are more sensitive to object specifications.

By comparing the above wavelet-based results with the ranges of the geostatistical variograms, the geostatistical method correctly identifies the characteristic scale or size of objects (e.g., 25 pixels in Fig. 4(a)). This has advantages over many other methods, including the wavelet-based methodology shown above, because semivariances can be calculated at any lag and orientation. In contrast, standard wavelet transforms are only decomposed at limited spatial resolution and orientation levels.

The above analysis has significant implications. When designing statistics for revealing the characteristic scale of scene variations, both lag and orientation should be distinctly analysed. Variograms calculated by a simple mathematical formula provides an elegant example in this regard. The scale identified can be used to (1) quantify the scale threshold for best object representation, (2) segment image objects globally and locally, and (3) determine optimum size of texture filters used to improve image classification. Moreover, in landscape ecology studies, those diversity and contagion indices should be used with a great care as resolution changes; crucial ecological processes and functions can only be confidently inferred by valid patterns delineated.

2.2 Scaling Biophysical Parameters

The classification of *physically or ecologically important landscapes* as objects is instrumental for scaling biophysical parameters. A study area can be partitioned or segmented using nested-hierarchical scene models, such as those used in forest canopy reflectance models; vegetation can be classified into a hierarchy of distinctive classes (e.g., conifer, hardwood), stands, and trees (Woodcock and Harward, 1992; Collins and Woodcock, 2000). Recent studies show that biophysical parameters such as leaf area index can be better estimated using object-based contextures in images (e.g., Wicks et al., 2002). However, up-scaling parameters should be interpreted with caution if images are simply degraded (or generalised, resampled) from a fine to a coarser spatial resolution. This is because the value change of parameters is often nonlinear and scale-dependent; beyond the representative scale threshold, necessary and important information content relevant to objects may not be retained.

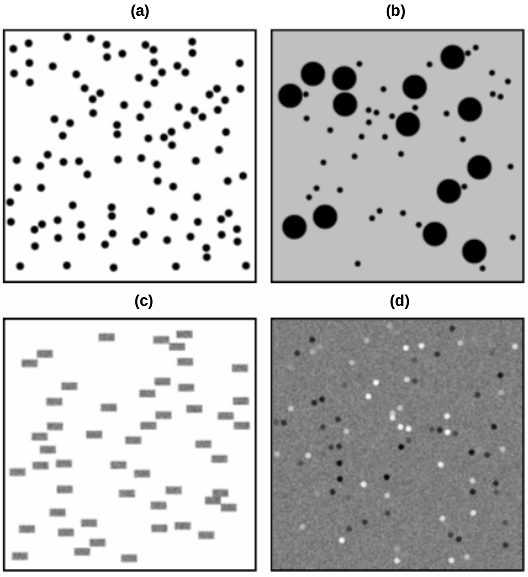

Figure 1. Four simulated images.

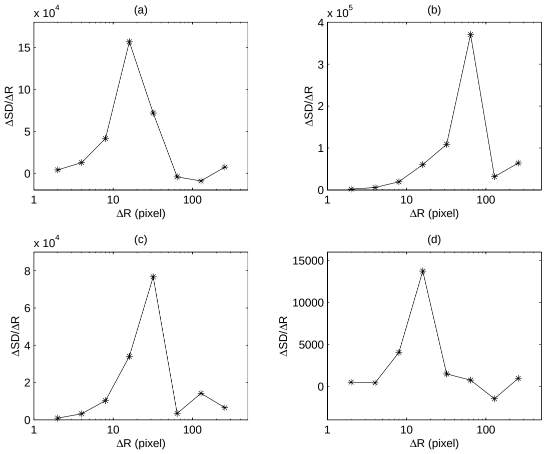

Figure 2. The change rate of standard deviation over two successive wavelet decomposition levels (ΔSD/ΔR) for energy signature images.

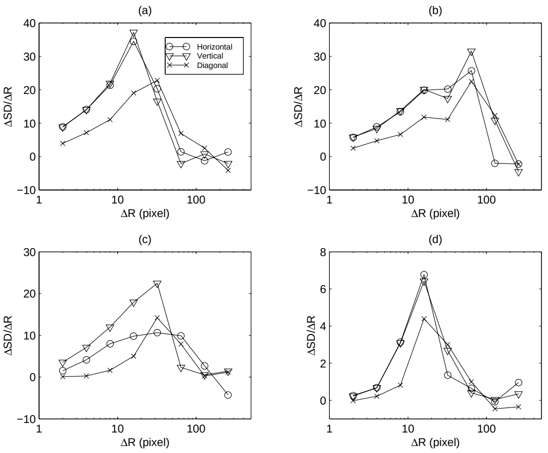

Figure 3. ΔSD/ΔR for horizontal, vertical and diagonal sub-images.

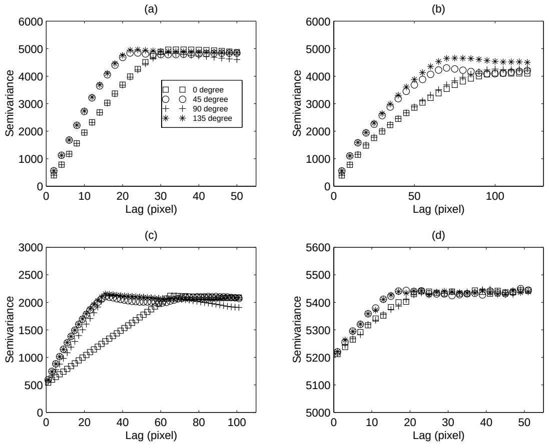

Figure 4. Directional variograms.

3. AREAL CENSUS DATA: AREAL INTERPOLATION WITH A VARIOGRAM APPROACH

Scale properties of census data at hierarchical areal units directly relate to the disaggregation of aggregate census data, areal interpolation and the modifiable areal unit problem (MAUP). While there is a suite of methods available, a fundamental conceptual solution to scale analysis lies in how more information from statistical inference (e.g., Cressie, 1996) and/or ancillary geospatial data (e.g., images, streets) can be wisely incorporated.

Remotely sensed images provide a rich tapestry of contextures relevant to objects. The delineation of *physically and socially-defined landscapes* (e.g., land covers and uses, separate dwellings) is important for scaling socioeconomic data across hierarchical areal units, for example, the use of categorical land covers or densities in dasymetric mapping approaches (Fisher and Langford, 1996; Chen, 2002; Mennis, 2003). This paper explores object structural information and spatial variations in a fine spatial resolution image for scaling areal census dwelling counts, using a variogram approach. Variogram serves as a descriptive tool for revealing the scale and pattern of spatial variability of an image (Curran and Atkinson, 1998). Semivariance is usually calculated for the *whole image scene*, but in this study semivariance is calculated at *different image zones* that are defined by areal census tracts. According to Woodcock et al. (1988), the range of variogram is related to the size of salient objects, and the sill of variogram is related to object density. If variograms are calculated for each separate zone, using indicative size of ground objects and sills, and the known area of each zone, the total number of ground objects of each zone can then be estimated. In this way, socioeconomic attributes (e.g., dwelling counts) can be estimated and interpolated across areal units.

A simulated image with four defined zones was used to illustrate this idea. Fig. 5 shows a simulated image with the size of 1024 × 1024 pixels, containing a simple scene with 50 black disc objects on a white background (horizontal or vertical diameter 7 pixels, size 37 pixels – see the enlarged object). All objects are randomly distributed. Semivariances of the four zones (A1, A2, A3 and A4, equivalent to census tracts here) were calculated (Fig. 6, Table 1), and sills have a pure linear correlation with their zone object densities ($R^2 = 1.00$) (Fig. 7). Based on this, one can predict object densities or the total number of objects (if the average size of objects is known) at *any* new user-defined zones where sample semivariances can still be calculated.

Fig. 8 defines another zoning system to test the methodology to predict the numbers of objects. With newly calculated sills of zones B1, B2, B3 and B4, and the regression equation in Fig. 7, densities of the four new zones were first predicted. Then by multiplying densities and the known areas of the new zones, the total area of objects for each zone was obtained. And finally, the total area of objects was divided by the size of each object (i.e., 37 pixels) to accurately predict the number of objects of each zone (Table 2).

Zone-based variograms provide a new and viable method for scaling areal census dwelling data and to reveal structural information hidden in a single band of a fine spatial resolution image. Further research on the use of real images with different land covers and fine spatial resolutions, and census data at different areal unit levels is under way. A recent SPOT-5 HRG panchromatic image (resolution 2.5 m) covering a predominant residential area east of Campbelltown, Sydney (Australia) was tested. For 36 census collection districts with an average number of dwellings of 240, high correlations exist between variogram sills and dwelling counts, with small absolute errors. Adjusted R^2 were 0.85 and 0.84 at identified lags of 10 m and 12.5 m (close to average dwelling size of 11 m).

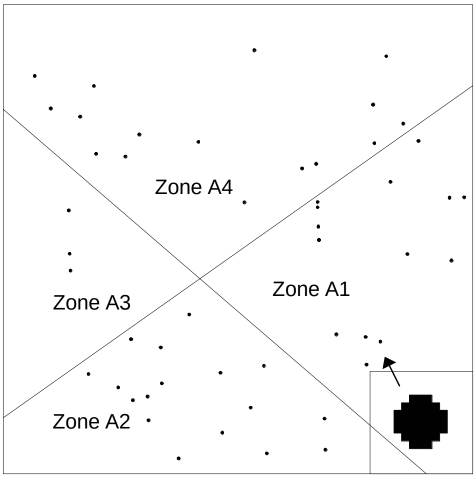

Figure 5. Simulated image with four zones A1, A2, A3 and A4, and an enlarged object shown at bottom-right corner.

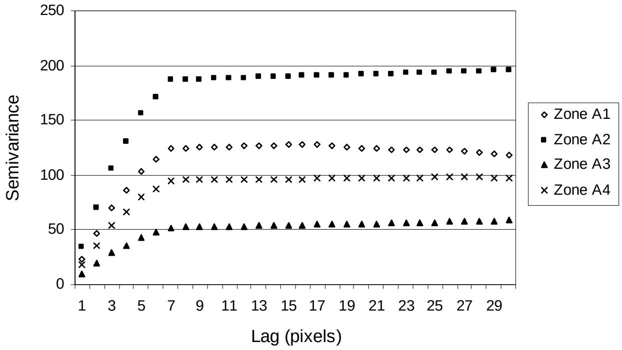

Figure 6. Variograms of four zones A1, A2, A3 and A4.

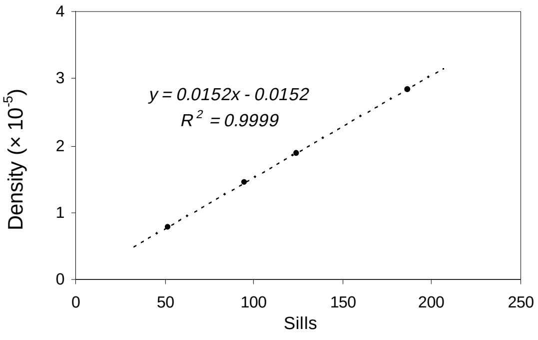

Figure 7. Linear correlation between densities and sills.

Table 1. Densities and Sills of Four Zones A1, A2, A3 and A4.

Zones	Zone areas (pixels)	Random discs		Densities (× 10^{-5})	Sills
		No.	Areas		
A1	276375	14	518	*1.87*	*124.55*
A2	221632	17	629	*2.84*	*187.05*
A3	142674	3	111	*0.78*	*52.29*
A4	407895	16	592	*1.45*	*95.35*

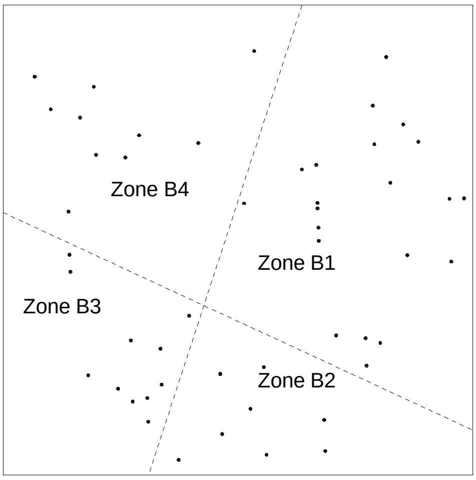

Figure 8. Simulated image and four new zones B1, B2, B3 and B4.

Table 2. Densities and Sills of Four Zones B1, B2, B3 and B4.

Zones	Sills	Predicted densities (× 10^{-5})	Discs		
			Areas (pixels)	Pred. No.	Known No.
B1	*128.70*	*1.95*	777.09	21.0	21
B2	*124.04*	*1.88*	295.93	8.0	8
B3	*149.60*	*2.27*	409.83	11.0	11
B4	*78.54*	*1.18*	367.62	9.9	10
Total	*115.51*	*1.75*	1831.96	49.5	50

4. CONCLUSIONS

Multidisciplinary studies use heterogeneous spatial data from geographical information systems, remote sensing systems, census systems and global positioning systems ("4S"). As more object-based spatial analysis approaches are applauded and enthusiastically assimilated by the general geospatial community, if both biophysical parameters and socioeconomic attributes can be explored and extracted using the same object-based scale analysis approach, our long-term aspiration for effective data integration and building truly integrative models across scales will be significantly boosted.

Stressing objects in scale analysis is one means of discovering geographical knowledge. In fact, our core spatial thinking has never strayed too far from the nature of objects, be they tiny or vast, definite or fuzzy, fleeting or lasting, bounded on the Earth or roaming in the universe. For an increasingly digitised kingdom of geospatial data and analyses, scale analysis can benefit tremendously by adhering to the very principle of cartographical scale transform: object representation associated with scale. The geostatistical approach capable of quantifying object-based spatial variations serves as an enlightening tool. It is anticipated that by investigating the profound and cognate relations between scale and objects, we can clear intriguing mysteries surrounding and promote best practice of scale analysis.

REFERENCES

Chen, K., 2002. An approach to linking remotely sensed data and areal census data. *International Journal of Remote Sensing*, 23(1), 37-48.

Chen, K., and Blong, R., 2003. Identifying the characteristic scale of scene variation in fine spatial resolution imagery with wavelet transform-based sub-image statistics. *International Journal of Remote Sensing*, 24(9), 1983-1989.

Collins, J.B., and Woodcock, C.E., 2000. Combining geostatistical methods and hierarchical scene models for analysis of multiscale variation in spatial data. *Geographical Analysis*, 32(1), 50-63.

Cressie, N., 1996. Change of support and the modifiable areal unit problem. *Geographical Systems*, 3, 159-180.

Curran, P.J., and Atkinson, P.M., 1998. Geostatistics and remote sensing. *Progress in Physical Geography*, 22(1), 61-78.

Fisher, P.F., and Langford, M., 1996. Modelling sensitivity to accuracy in classified imagery: A study of areal interpolation by dasymetric mapping. *Professional Geographer*, 48(3), 299-309.

Hay, G.J., Dubé, P., Bouchard, A., and Marceau, D.J., 2002. A scale-space primer for exploring and quantifying complex landscapes. *Ecological Modelling*, 153, 27-49.

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

Quattrochi, D.A., and Goodchild, M.F. (eds.), 1997. *Scale in Remote Sensing and GIS*. Lewis Publishers.

Tate, N.J., and Atkinson, P.M. (eds.), 2001. *Modelling Scale in Geographical Information Science*. John Wiley & Sons.

van Gardingen, P.R., Foody, G.M., and Curran P.J. (eds.), 1997. *Scaling-up: From Cell to Landscape*. Cambridge University Press.

Wicks, T.E., Smith, G.M., Curran, P.J., 2002. Polygon-based aggregation of remotely sensed data for regional ecological analyses. *International Journal of Applied Earth Observation and Geoinformation*, 4, 161-173.

Woodcock, C.E., and Harward, V.J., 1992. Nested-hierarchical scene models and image segmentation. *International Journal of Remote Sensing*, 12(16), 3167-3187.

Woodcock, C.E., Strahler, A.H., and Jupp, D.L.B., 1988. The use of variogram in remote sensing: II. real digital images. *Remote Sensing of Environment*, 25, 349-379.