ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

ISPRS Journal of Photogrammetry & Remote Sensing 1253 (2003) 1–19

PHOTOGRAMMETRY & REMOTE SENSING

www.elsevier.com/locate/isprsjprs

A comparison of three image-object methods for the multiscale analysis of landscape structure

Geoffrey J. Hay[a,*], Thomas Blaschke[b], Danielle J. Marceau[a], André Bouchard[c]

[a] Geocomputing Laboratory, Département de Géographie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montreal, Quebec, Canada H3C 3J

[b] Department of Geography and Geoinformation, University of Salzburg, Austria, Hellbrunner Str. 34, A-5020 Salzburg, Austria

[c] IRBV, Université de Montréal, Jardin Botanique de Montréal, 4101 Sherbrooke Est, Montreal, Quebec, Canada H1X 2B2

Received 3 March 2002; accepted 12 July 2002

Abstract

Within the conceptual framework of *Complex Systems*, we discuss the importance and challenges in extracting and linking multiscale objects from high-resolution remote sensing imagery to improve the monitoring, modeling and management of complex landscapes. In particular, we emphasize that remote sensing data are a particular case of the modifiable areal unit problem (MAUP) and describe how *image-objects* provide a way to reduce this problem. We then hypothesize that multiscale analysis should be guided by the intrinsic scale of the dominant landscape objects composing a scene and describe three different multiscale image-processing techniques with the potential to achieve this. Each of these techniques, i.e., *Fractal Net Evolution Approach (FNEA)*, *Linear Scale-Space and Blob-Feature Detection (SS)*, and *Multiscale Object-Specific Analysis (MOSA)*, facilitates the multiscale pattern analysis, exploration and hierarchical linking of image-objects based on methods that derive spatially explicit multiscale contextual information from a single resolution of remote sensing imagery. We then outline the weaknesses and strengths of each technique and provide strategies for their improvement.

Keywords: complex systems theory; fractal net evolution approach; image-objects; multiscale object-specific analysis

1. Introduction

Landscapes are complex systems composed of a large number of heterogeneous components that interact in a non-linear way and exhibit adaptive properties through space and time. In addition, complex systems exhibit characteristics of emergent properties, multiscale hierarchical interactions, unexpected behavior and self-organization (Forman, 1995; Wu and Marceau, 2002), all of which produce characteristic patterns that (appear to) change depending on their scale of observation (Allen and Starr, 1982). Thus, the roles of the *observer* and of *scale* are fundamental in recognizing these patterns, which in turn are necessary for understanding the processes that generated them.

Conceptually, scale corresponds to a 'window of perception'. More practically, scale represents a measuring tool composed of two distinct components: grain and extent. *Grain* refers to the smallest intervals

* Corresponding author. Tel.: +1-514-343-8073; fax: +1-514-343-8008.

E-mail address: ghay@sympatico.ca (G.J. Hay).

0924-2716/03/$ - see front matter © 2003 Published by Elsevier Science B.V.
doi:10.1016/S0924-2716(02)00162-4

in an observation set, while *extent* refers to the range over which observations at a particular grain are made (O'Neill and King, 1998). From a remote sensing perspective, grain is equivalent to the spatial, spectral and temporal resolution of the pixels composing an image, while extent represents the total area, combined bandwidths and temporal duration covered within the scene (Hay et al., 2001). In addition, remote sensing platforms are the primary data source from which landscape patterns can be assessed. Therefore, to fully understand, monitor, model and manage our interaction within landscapes, three components are required: (1) remote sensing data with a fine enough grain and broad enough extent to define multiscale landscape patterns; (2) methods and theory capable of identifying pattern components, i.e., real-world objects at their respective scales of expression; and (3) the ability to link and query these objects within appropriate hierarchical structures. In this paper, *small* or *fine scale* refers to a small measure, i.e., pixel size or area, while *large* or *coarse scale* refers to a large measure.

Multiscale analysis is composed of two fundamental components: (1) the generation of a multiscale representation and (2) information extraction. To achieve the innate pattern recognition abilities of humans, a number of image-processing techniques have been developed that incorporate concepts and theory from computer vision and machine learning. These include edge detectors (Canny, 1986), mathematical morphology (Haralick et al., 1987), texture analysis (Hay et al., 1996), spectral unmixing (Settle and Drake, 1993), neural nets (Foody, 1999), Bayesian networks (Robert, 2001), fuzzy logic (Wang, 1990) and multiscale techniques such as pyramids (Jähne, 1999), wavelets (Salari and Ling, 1995) and fractals (Chaudhuri and Sarkar, 1995). However, results from these methods often fall short when compared with those of human vision. This is in part because the majority of these techniques do not generate explicit object topology, or even incorporate the concept of *object* within their analysis. Yet this is innate to humans (Biederman, 1987). Furthermore, when these techniques are applied to remote sensing data, their output are typically used only as additional information channels in per-pixel classification techniques of multidimensional feature space, rather than in object delineation (Skidmore, 1999). While many of these techniques provide interesting and useful results over a single scale, or narrow range of scales, the ability to apply these methods for the automatic analysis of multiscale landscape patterns and the hierarchical linking of their components through a scale continuum is not well defined (Hay et al., 1997).

Remote sensing images are composed of pixels, not objects and there are no explicit scaling laws that define where to scale to and from within an image, the number of scales to assess, or the appropriate upscaling method(s) to use (Hay et al., 2001). To overcome these limitations, we hypothesize that the analysis of multiscale landscape structure should be guided by the intrinsic scale of the varying sized image-objects that compose a scene. To facilitate this, we provide a brief background on the modifiable areal unit problem (MAUP), image-objects and hierarchy (Section 2). We then describe three different multiscale techniques: the *Fractal Net Evolution Approach (FNEA)*, *Linear Scale-Space and Blob-Feature Detection (SS)*, and *Multiscale Object-Specific Analysis (MOSA)* (Section 3). Each of these techniques facilitates the multiscale pattern analysis, exploration and hierarchical linking of image-objects based on methods that derive spatially explicit, multiscale contextual information from a single remote sensing image. Finally, we outline the strengths and weaknesses of each technique and provide strategies for their improvement (Section 4).

2. Background: MAUP, image-objects and hierarchy

2.1. Remote sensing and the modifiable areal unit problem

While remote sensing data are often visually impressive, they also correspond to an arbitrary spatial sampling of the landscape and thus represent a particular case of the MAUP (Marceau et al., 1994). The MAUP originates from the use of arbitrarily defined spatial units for data acquisition and analysis. The consequence is that data and results achieved from them are dependent upon the spatial units, i.e., pixels, used to collect them (Openshaw, 1981, 1984; Marceau, 1999). Though recognized in the Social and Natural Sciences for several decades (Openshaw and

Taylor, 1979), we suggest that few understand the real challenges this poses, especially when multiscale analysis is applied to remotely sensed data (for an in depth review of MAUP, see Marceau, 1999; Marceau and Hay, 1999). Fortunately, several solutions to the MAUP have been proposed. In particular, the use of objects represents the clearest way out of MAUP, as an analyst works with spatially discrete entities rather than arbitrarily defined areal units (Fotheringham and Wong, 1991; Hay et al., 2001). However, a remote sensing image is not composed of spatially discrete real-world entities that contain explicit object topology. Instead, its fundamental primitive is typically a square pixel that only exhibits simple topological adjacency.

2.2. Image-objects

Despite this topological limitation, humans can cognitively group similar toned and spatially arranged pixels into meaningful *image-objects* that correspond to real-world entities within the geographic extent of the scene being assessed. The term *image-objects* (Hay and Niemann, 1994; Hay et al., 1997, 2001) refers to individually resolvable entities located within a digital image that are perceptually generated from high-resolution pixel groups. According to Woodcock and Strahler (1987), High-resolution (H-res) corresponds to the situation where a single real-world object is visually modeled by many individual pixels; whereas low-resolution (L-res) implies that a single pixel represents the integrated signal of many (smaller) real-world objects. In a remote sensing image, both H- and L-res situations occur simultaneously. For example, in a 1.0-m-resolution image of a forest canopy, where each tree crown exhibits a 10-m diameter, each crown image-object will be composed of many pixels. In this situation, each 1.0 m pixel is *part of* an individual crown; thus, it is H-res in relation to the crown-object it models. However, each 1.0 m pixel will also be *composed of* the integrated reflectance from many needles/leaves and branches; thus, it will be L-res in relation to these individual crown components. As a result, an image-object tends to be composed of spatially clustered pixels that exhibit high spectral autocorrelation because they are all part of the same object. Consequently, they have similar gray values. These characteristics correspond to Tobler's first law of Geography where 'objects are related to all other objects, but proximal objects are more likely to be related to each other' (Tobler, 1970). In an image-object, this relationship is both spatial and spectral.

2.3. Hierarchy

Similar to Tobler's first law, Ecologists have long recognized that in nature, many processes produce clusters of entities that are typically generated by a small set of self-organizing principles (Allen and Starr, 1982). These entities emerge at specific scales and result in visually distinct spatial patterns. Therefore, one way to understand, explain and forecast the effects of *natural processes* is to examine these *natural patterns* at their corresponding *natural scales* of emergence (Wessman, 1992; Levin, 1999). To assist in this task, the conceptual framework of Hierarchy theory has been developed that builds upon this idea of *natural scales*. Conceptually, a hierarchically organized system can be seen as a nested system in which levels exhibiting progressively slower behavior are at the top (Level +1), while those reflecting successively faster behavior are seen as a lower level in the hierarchy (Level −1). The level of interest is referred to as the *focal level* (Level 0) and it resides between the other two. From a landscape ecology perspective, Hierarchy theory states that complex ecological systems, such as landscapes, are composed of loosely coupled levels (*scale domains*), where each level operates at distinct time and space scales. *Scale thresholds* separate scale domains and represent relatively sharp transitions, i.e., critical locations where a shift occurs in the relative importance of variables influencing a process (Wiens, 1989). Thus, interactions tend to be stronger and more frequent *within* a level of the hierarchy than *among* levels (Allen and Starr, 1982). This important fact enables the perception and description of complex systems by decomposing them into their fundamental parts and interpreting their interactions (Simon, 1962).

However, to achieve this, *objects*, i.e., fundamental parts, need to be clearly defined. Rowe (1961) distinguishes between two fundamental object types: *integrated objects* and *aggregate objects*. Integrated objects contain structurally organized parts, while aggregate objects occupy a common area, but have no structural organization. Furthermore, integrated objects have *intrinsic scale*, whereas aggregates do not. From a remote sensing perspective, image-objects are *integrated objects* that exhibit an intrinsic scale and are composed of structurally connected parts, i.e., H-res pixels. To understand how image-objects interact within and across scale domains, we need techniques to automatically define them in remote sensing data and the ability to link them within appropriate hierarchical structures—thus reducing MAUP. The primary unknowns to achieve this are:

- What are the 'optimal' scales to evaluate the varying sized, shaped and spatially distributed image-objects within a scene?
- At what scales should hierarchies be established?

We suggest that there is no single 'optimal' scale for analysis. Rather there are *many optimal scales* that are specific to the image-objects that exist/emerge within a scene (Hay and Niemann, 1994; Hay et al., 1997, 2001). Therefore, we hypothesize that multiscale analysis should be guided by the intrinsic scale of the dominant landscape objects, i.e., image-objects, composing a scene.

3. Materials and methods

In this section, we introduce the study site and data set used. We then briefly describe three different image-processing approaches, each of which facilitates the multiscale pattern analysis, exploration and hierarchical linking of image-objects, from a single resolution of remote sensing imagery. They are referred to as *Fractal Net Evolution Approach (FNEA)*, *Linear Scale-Space and Blob-Feature Detection (SS)*, and *Multiscale Object-Specific Analysis (MOSA)*.

3.1. Study site and data set

The data used throughout this paper is a 500 × 500 pixel sub-image of an IKONOS-2 (Geo) scene acquired in August 2001 (Fig. 1a). Geographically, this area represents a portion of the highly fragmented agro-forested landscape typical of the Haut Saint-Laurent region of southwest Quebec, Canada (Fig. 1b). IKONOS-2 provides 11-bit multispectral data in the red, green, blue and near-infrared (NIR) channels at 4.0 m spatial resolution and an 11-bit panchromatic (PAN) channel at 1.0 m resolution. Due to the computational demands required by SS processing (Section 3.3), all data were linearly scaled to

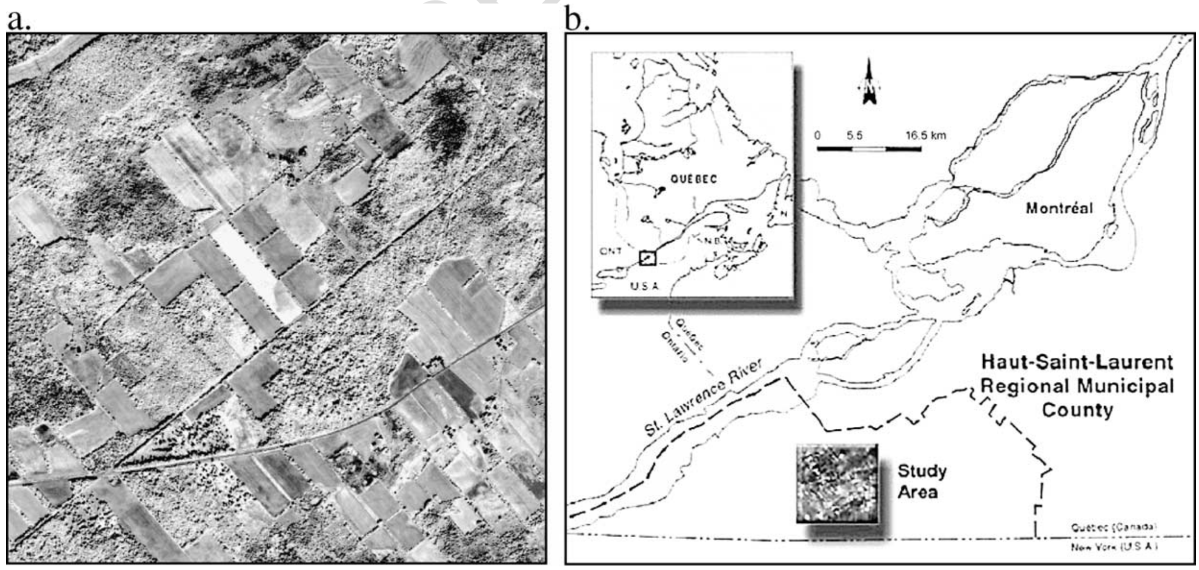

Fig. 1. IKONOS-2 sub-image and study site map: (a) a 500 × 500 pixel IKONOS-2 image of the study site; (b) map location of the image.

8-bit. Since the PAN channel covers a significant portion of the wavelengths represented by the four multispectral channels, a geographically corresponding portion of the 1.0-m PAN image was selected and resampled to 4.0 m using Object-Specific Upscaling (Section 3.4), which is considered a robust upscaling technique (Hay et al., 1997). During SS analysis, only the single PAN image was assessed. During FNEA and MOSA analysis, all five channels were evaluated.

3.2. Fractal net evolution approach

The fractal net evolution approach (FNEA) is embedded in a commercial software environment (Definiens, 2002). It utilizes fuzzy set theory to extract the objects of interest, at the scale of interest, segmenting images simultaneously at both fine and coarse scales. By operating on the relationships between networked, i.e., linked objects, it is possible to use local contextual information, which, in addition to the images' spectral information, can be combined with image-object form and texture features to improve classifications.

From a FNEA perspective, image information are considered *fractal* in nature. Here the term 'fractal' refers to the same degree of non-regularity at all scales, or self-similarity across scales. That is, structures typically appear at different scales in an image simultaneously and exhibit commonalities (Bobick and Bolles, 1992). To extract meaningful image regions, the user has to take into account the scale of the problem to be solved and the type of image data available. Consequently, users are required to focus on different scale levels because almost all attributes of image structure such as color, texture, or shape, are highly scale dependent. This is different from approaches that do not require user-defined parameters, i.e., most region growing and watershed algorithms, multi-fractal-based segmentation and Markov random fields. In cases where input parameters may be required, i.e., user-defined cliques in Markov random fields, or seeds in watershed segmentation, the user does not explicitly focus on a specific 'level' or 'scale', rather they focus on certain heterogeneity criteria for the resulting image-objects. In FNEA, defining a specific level of analysis leads to defining objects at a unique scale.

FNEA starts with a single pixel and a pairwise comparison of its neighbors with the aim of minimizing the resulting summed heterogeneity. The common solution for this pairwise cluster problem is *global mutual best fitting*. Global mutual best fitting is the strongest constraint for the optimization problem and it reduces heterogeneity following a pure quantitative criterion. However, there is a significant disadvantage to global mutual best fitting. It does not use the distributed treatment order and—in connection with a heterogeneity definition for color—builds initial segments in regions with a low spectral variance. This leads to an uneven growth of image-objects over a scene and to an imbalance between regions of high and low spectral variance. To overcome this, FNEA incorporates *local mutual best fitting*, which always performs the most homogeneous merge in the local vicinity following the gradient of best fitting. That is: each initial pixel or group of pixels grows together with its respective neighboring pixel(s), which results in the lowest heterogeneity of the respective join. To achieve this, an iterative heuristic optimization procedure aims to get the lowest possible overall heterogeneity across an image. The basis for this is the degree of difference between two regions. As this difference decreases, the fit of the two regions appears to be closer. These differences are optimized in a heuristic process by comparing the attributes of the regions (Baatz and Schäpe, 2000). Thus, given a certain feature space, two image-objects f_1 and f_2 are considered similar when they are near to each other in this feature space. For an n-dimensional feature space (f_{nd}), the heterogeneity h is described as:

$$h = \sqrt{\sum_d (f_{1d} - f_{2d})^2} \quad (1)$$

Examples of object features include mean spectral values, or texture features, such as spectral variance. These distances can be further normalized by the standard deviation of the feature in each dimension (σ_{fd}) using Eq. (2).

$$h = \sqrt{\sum_d \left(\frac{f_{1d} - f_{2d}}{\sigma_{fd}}\right)^2} \quad (2)$$

Eq. (3) describes the difference in heterogeneity (h) of two regions (h1 and h2) before and after a virtual merge[1] (hm). Given an appropriate definition of heterogeneity for a single region, the growth in heterogeneity of a merge should be minimized. There are several possibilities for describing the heterogeneity change (hdiff) before and after a virtual merge—but they are beyond the scope of this paper. For more information, see Baatz and Schäpe (2000).

$$\mathrm{hdiff} = \mathrm{hm} - (\mathrm{h1} + \mathrm{h2})/2 \quad (3)$$

hdiff allows us to distinguish between two types of objects with similar mean reflectance values but different 'within-patch heterogeneity'. An application based on this type of heterogeneity was described by Blaschke et al. (2001) where they used the *mean spectral difference between all sub-objects* as one example of heterogeneity applied to pastures and conservation changes in a cultural heritage landscape in central Germany. This involved a multiscale delineation of images and image semantics that incorporated image-structure and image-texture characteristics. It was found that they could distinguish three levels of delineation appropriate for three different key species, which resulted in the construction of a hierarchical network of image-objects and semantic rules between these levels.

Since its recent introduction by Baatz and Schäpe (2000), FNEA has been applied to various research projects in Europe (Blaschke et al., 2000; Blaschke and Strobl, 2001; Schiewe et al., 2001), many of which have demonstrated the potential of this multiscale segmentation approach. In particular, the 'realistic' appearance (Fig. 2) of the resulting segmented patches of forests, pastures, fields and built-up areas have motivated several European agencies to seriously evaluate this approach.

3.3. Scale-space

The following overview represents a multiscale approach as described by Lindeberg (1994) that is composed of two principal components: *Linear Scale-Space and Blob-Feature Detection*. For a more detailed non-mathematical description of both, see Hay et al. (2002a). *Linear Scale-Space* (SS) is an uncommitted framework[2] for early visual operations that was developed by the computer vision community to automatically analyze real-world structures at multiple scales—specifically, when there is no a priori information about these structures, or the appropriate scale(s) for their analysis. When scale information is unknown within a scene, the only reasonable approach for an uncommitted vision system is to analyze the input data at (all) multiple scales. Thus, a SS multiscale representation of a signal (such as a remote sensing image of a landscape) is an ordered set of derived signals showing structures at coarser scales that constitute simplifications, i.e., smoothing, of corresponding structures at finer scales. Smoothed layers are created by convolving the original image with a Gaussian function, where the scale of each derived signal is defined by selecting a different standard deviation of the Gaussian function. This results in a scale-space cube or 'stack' of progressively 'smoothed' image layers, where each new layer represents convolution at an increased scale. Each hierarchical layer in a stack represents convolution at a fixed scale, with the finest scale at the bottom and the coarsest at the top (Fig. 3).

The use of Gaussian filters is essential to linear SS theory as they satisfy necessary conditions or axioms for an uncommitted framework (Weickert et al., 1997). These include (among others) linearity (no knowledge, no model, no memory), spatial shift invariance (no preferred location), isotropy (no preferred orientation) and scale invariance (no preferred size or scale). In addition, a Gaussian kernel satisfies the linear diffusion equation; thus, Gaussian smoothing is considered as the diffusion of gray-level intensity over scale (t), instead of time.

The second SS component we use is referred to as *Blob-Feature Detection* (Lindeberg, 1994). The primary objective of this non-linear approach is to link structures at different scales in scale-space, to higher

[1] Different possibilities for merging neighboring regions and their respective heterogeneity are calculated in computer memory (i.e., virtually), and then the candidate with the lowest resulting heterogeneity is actually merged.

[2] The term *uncommitted framework* refers to observations made by a *front-end vision system*, i.e., an initial-stage measuring device, such as the retina or a camera that involves 'no knowledge' and 'no preference' for anything.

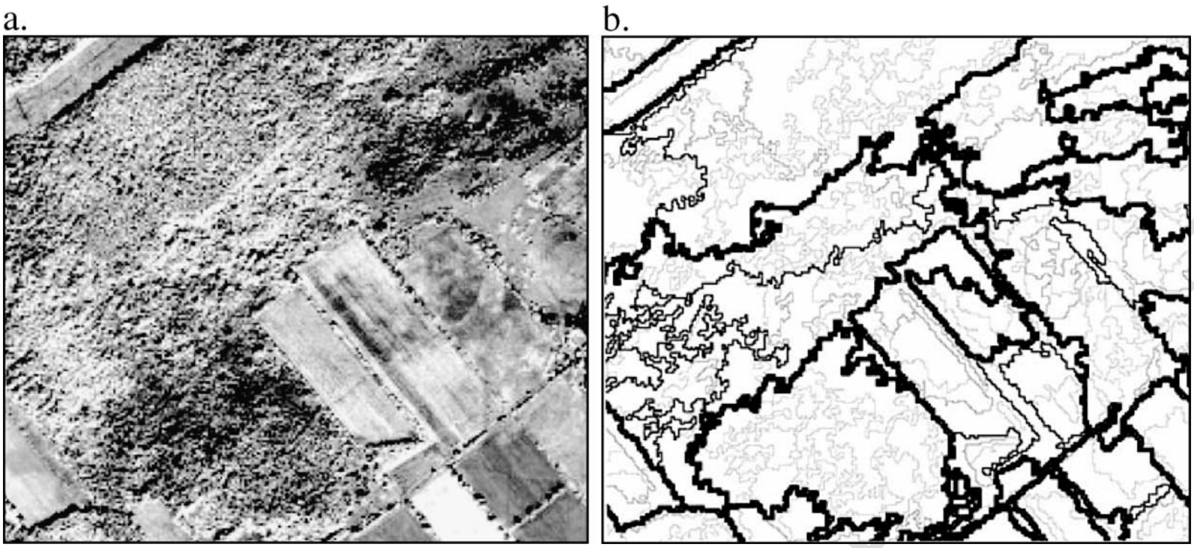

Fig. 2. Three different levels of FNEA segmentation: (a) panchromatic IKONOS-2 sub-image (245 × 210 pixels) extracted from the top right corner of Fig. 1a. (b) A close up of typical FNEA results using different segmentation levels. These levels roughly correspond to the smallest units of interest, e.g., single groups of trees/bushes illustrated by bright gray lines. Medium-sized black outlines represent a medium segmentation level, which corresponds best to 'forest stands'. Bold black lines indicate the coarsest level of segmentation where semantically different landscape objects have merged, but can still be exploited as 'super-objects'.

order objects called 'scale-space blobs' and to extract significant features based on their appearance and persistence over scales. The main features of interest at each scale within a stack are smooth regions, which are brighter or darker than the background and which stand out from their surrounding. These regions are referred to as 'gray-level blobs'. When blobs are evaluated as a volumetric structure within a stack, it becomes apparent that some structures persist through scale, while others disappear (Fig. 4). Therefore, an important premise of SS is that structures which persist in scale-space are likely candidates to corre-

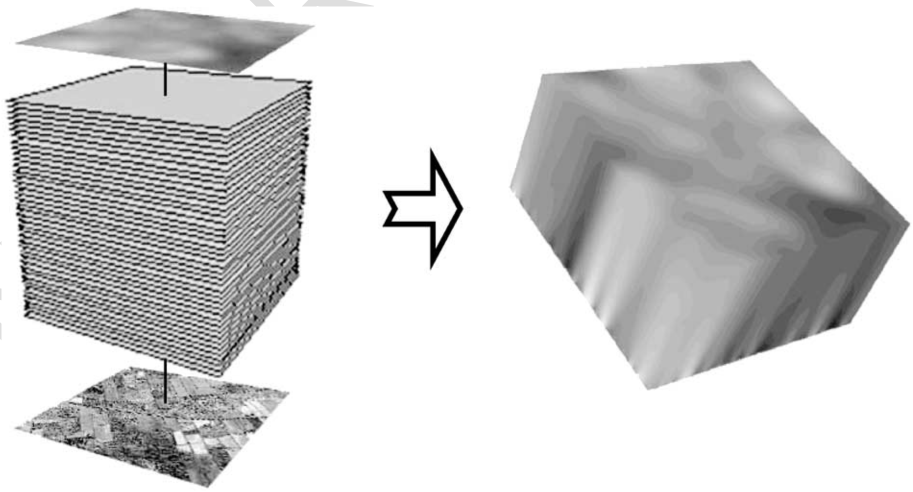

Fig. 3. Linear scale-space 'stack'. The finest scale is on the bottom and the coarsest scale, i.e., the most smoothed, is on the top. At the margins of the right figure, the diffusive pattern of scale-space objects through scale can be seen.

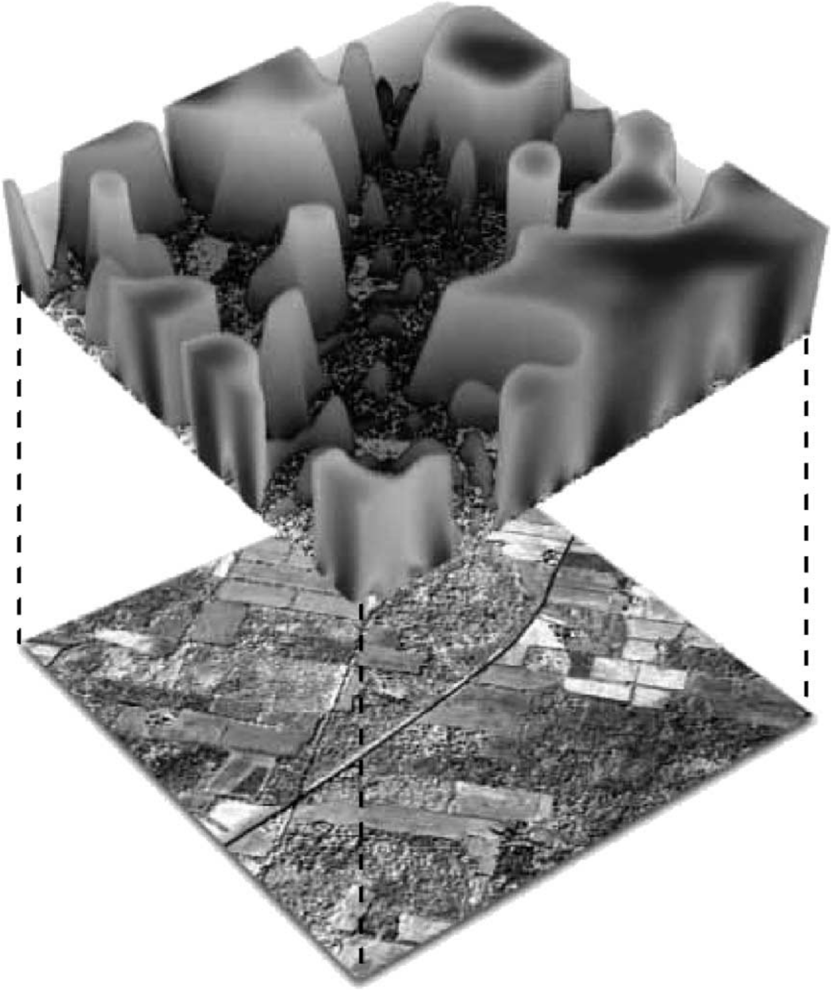

Fig. 4. Grey-level stack, with opacity filters and pseudo-shading applied to illustrate the persistence of blob structures through scale (for visualization purposes only).

spond to significant structures in the image and thus in the landscape.

In simple terms, gray-level blobs at each scale in the stack are objects with extent both in 2D space (x, y) and in gray-level (z-axis)—thus in 3D. Grey-level blob delineation may be visualized as follows: at each scale in the stack, the image function is virtually flooded. As the water level gradually sinks, peaks of the blobs appear. As this continues, different peaks will eventually become connected. The corresponding ‘connected’ contour delimits the 2D spatial extent, or ‘region of support’ of each blob, which is defined as a binary blob (Fig. 5a). 2D binary blobs are then linked over scale to create 3D hyper-blobs (Fig. 5b).

Within a single hyper-blob four primary types of ‘bifurcation events’ may exist: *annihilations* (A), *merges* (M), *splits* (S) and *creations* (C). These SS-events represent critical component of SS analysis, as scales between bifurcations are linked together forming the lifetime (Lt_n) and topological structure of individual SS-blobs (Fig. 5c). Next, the integrated normalized 4D volume (x, y, z, t) of each individual SS-blobs is defined. The blob behavior is strongly dependent upon image structure and the volume depends on the scale. To normalize for this dependency, statistics are extracted from a large number of stacks generated from random images. These statistics describe how random noise blobs behave in scale-space and are used to generate a normalized 4D SS volume for each SS-blob.

Normalized volumes are then ranked and a number of *significant* SS-blobs are defined, from which the scale (t) representing the maximum 3D gray-level blob volume (x, y, z) of each hyper-blob is extracted. From these, the 2D spatial support, i.e., binary blob, is identified. Thus, based on the underlying initial premise, 4D scale-space blobs are simplified to 3D gray-level blobs, which are further simplified to their 2D support.

3.4. Multiscale object-specific analysis

Multiscale object-specific analysis (MOSA) is composed of three primary components: *Object-Specific Analysis* (OSA), *Object-Specific Upscaling* (OSU) and *Marker-Controlled Watershed Segmentation* (MCS). OSA is a multiscale approach that automatically defines unique spatial measures specific to the individual image-objects composing a remote sensing scene (Hay et al., 1997, 2001). These object-specific measures are then used in a weighting function to automatically upscale (OSU) an image to a coarser resolution. Then MCS is applied to these multiscale data to automatically segment them into topologically discrete image-objects that strongly correspond to image-objects as though defined by visual analysis.

An underlying premise of OSA/OSU is that all pixels within an image are exclusively considered H-res samples of the scene-objects they model, even though, as previously discussed, both H- and L-res exist. Thus, we use pixels—the fundamental image primitive—to define the spatial extent of the larger image-objects they are a *part of*. Hay et al. (1997) noted that when plotting the digital variance of pixel gray-values located with increasingly larger kernels,

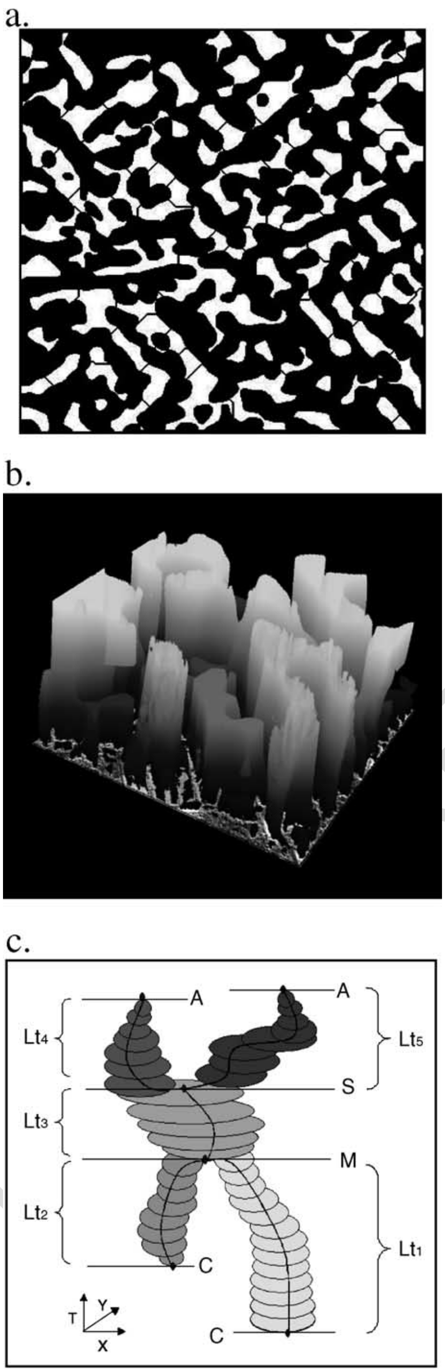

Fig. 5. (a) Binary blobs at scale 20 (i.e., t 20). (b) A hyper-blob stack composed of 2D binary blobs. For illustration only, each binary layer has been assigned a value equal to its scale. Thus, dark values are on the bottom, while bright values are near the top. (c) Idealized hyper-blob illustrating four different SS-events: annihilation (A), creation (C), merge (M) and split (S). The number of scales between SS-events represents the lifetime (Lt_n) of a SS-blob. Five different Lt_n are illustrated.

while centered on an image-object of known size, the resulting plot tended to produce curves with distinct breaks, or thresholds in variance as the analyzing kernel contacted the image-object's edges. When this 'threshold curve' was reached, the corresponding mean and variance values were also recorded for the pixel under analysis within the defined window size. This process was then locally applied to all remaining pixels within the original image, resulting in corresponding Variance (V_I), Area (A_I) and Mean (M_I) images. This form of processing is referred to as *object-specific analysis*. It is important to note that

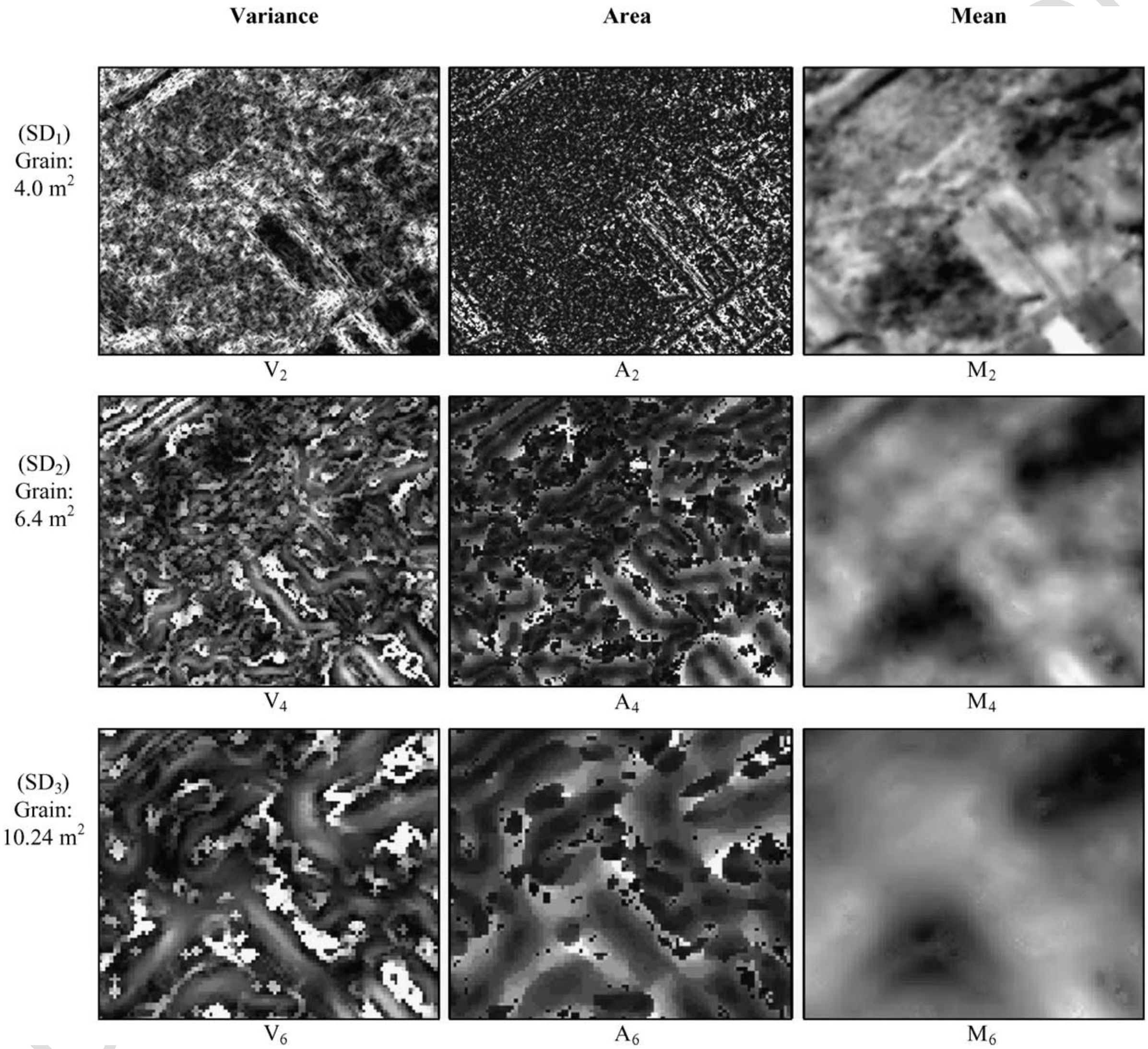

Fig. 6. Variance (V_I), Area (A_I) and Mean (M_I) from the first three scale domains SD_{1-3}. These data correspond to the sub-image illustrated in Fig. 2a. In the variance images, dark tones represent low variance, i.e., pixel groups that are more 'object-like', while bright tones represent high-variance edges between two or more image-objects. Area images define the spatial extent of individual objects at a particular scale. Dark tones represent small spatial extents, i.e., closer pixels are more 'object-like', while bright values represent large spatial extents as they are less 'object-like'. Mean images represent the average of the H-res pixels that constitute part of individual objects assessed within each object-specific threshold window.

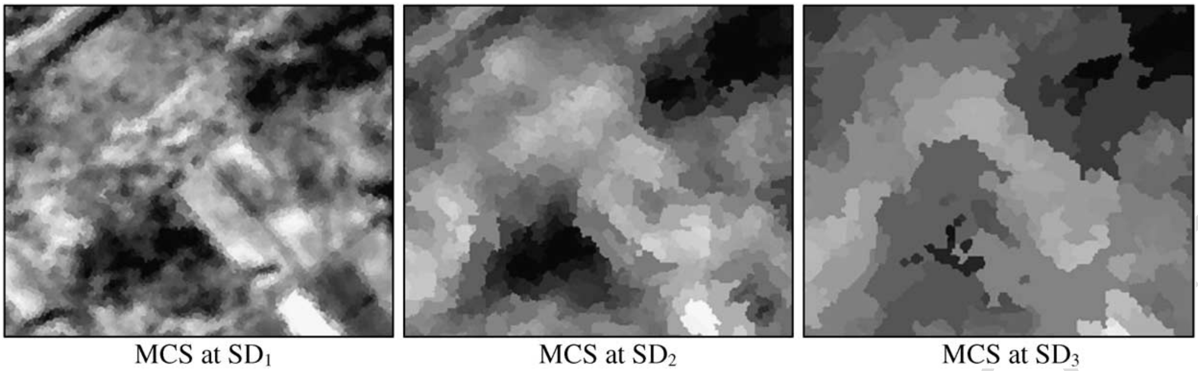

Fig. 7. Marker-controlled watershed segmentation (MCS) results from the first three mean images illustrated in Fig. 6. Each gray-tone represents a topologically distinct 'watershed' image-object.

the specific kernel size defined at each variance threshold was strongly related to the known size of the differently sized image-objects that each pixel was a part of (thus, object-specific). The area values associated with each pixel can then be used as part of a weighting scheme to upscale an image to a coarser resolution. This is referred to as *object-specific upscaling*. The resampling criterion for OSU is premised on the relationship between pixel size, image-objects and the point-spread function of modern sensors[3] and is further discussed in Hay et al. (2001).

Based on promising results from early research, Hay et al. (1997) recognized that the application of OSA/OSU rules reveal patterns that accurately correspond to the spatial extent of objects at the next coarser scale. This led to the hypothesis (Hay and Marceau, 1998) that by continuously applying object-specific rules to the M_I generated at each OSA iteration, new spatial patterns will emerge that represent dominant landscape objects and that these patterns will correspond to real-world objects through a wide range of scales.

To test this hypothesis, Hay et al. (2001) developed an iterative multiscale framework that represents a nested hierarchy of image-sets (IS_t) consisting of two versions of V_I, A_I and M_I, each of which possesses membership in a unique scale domain (SD_n). They recognized that there is often a range of scales between the end point of identifiable scale domains where certain image-objects exist and the point where new image-objects emerge at their next scale of expression. To exploit this information, the initial framework was modified as follows: at the first OSA iteration, every pixel is locally assessed within progressively larger kernels until a local *maximum* variance threshold is reached. When applied to the entire image, this process generates the first image-set (i.e., V_1, A_1, M_1)—as previously described. In the second iteration, each pixel in the newly generated M_1 is locally assessed until a *minimum* variance threshold is reached. The resulting images become the second image-set (i.e., V_2, A_2, M_2) where they represent the beginning scale of all newly emergent image-objects (Fig. 6). Recall that minimum variance indicates that pixels are very similar; thus, the corresponding image structures are most 'object-like'. Consequently, odd-numbered OSA iterations define scales representing the spatial extent or 'end' of objects, while even-numbered OSA iterations define the beginning scale of the next emergent object that each pixel is a part of. Therefore, the even-numbered OSA iteration (IS_2) is selected for upscaling (OSU) because it contains the new image-objects we are interested in defining. This entire OSA/OSU process is then repeated (iterated) on the newly upscaled Mean images until the numbers of pixels composing them are too small for further processing.

The result of this iterative approach is a nested hierarchy of image-sets (IS_t), each composed of two

[3] The point spread function defines the 'spread' of an infinitely small point of light resulting from lens aberrations as well as light interference at the aperture.

V_I, A_I and M_I that have membership in a unique scale domain (SD_n). Within a single SD_n, each image shares the same grain and extent and represents the result of multiscale analysis specific to the image-objects composing it. However, each SD_n has a coarser grain than the previous SD_{n-1} (due to upscaling), though it shares the same extent through all image-sets. OSU is applied to ensure that the original image-object heuristics maintain the same conditions for which they were originally designed and to reduce unnecessary computation and data generation/storage.

Once OSA/OSU processing has been completed, MCS is applied to the resulting Mean images (Fig. 7). MCS is a watershed transformation technique that detects regional similarities as opposed to an edge-based technique that detects local changes (Beucher and Lantuéjoul, 1979; Meyer and Beucher, 1990). The key characteristic of this technique is the ability to reduce over-segmentation by placing markers or 'seeds' in user-specified areas. The elegance of integrating MCS within MOSA is that it requires data inputs that are automatically and explicitly met by V_I, A_I and M_I. Specifically, the V_I represents the edges or 'dams' that define individual watersheds in an image. These watersheds contain regional minima, which are naturally represented by the A_I values due to a low internal variance inherent to image-objects. Once each watershed-object perimeter is automatically defined, it is labeled with the average of the pixels located with the corresponding M_I, resulting in topologically discrete image-objects.

4. Discussion

In this section, we outline the principal strengths and limitations of each technique. Based on this, we suggest strategies for their improvement by integrating appropriate characteristics from the other techniques.

4.1. Strengths of FNEA, SS and OSA/OSU

FNEA software was developed to simultaneously identify and extract objects of interest at different scales within textured imagery, i.e., radar and H-res satellite or airborne data, through multi-resolution segmentation. A commercial software product is available that can be integrated within commonly used image-processing packages. This has aided in the development of a growing user base and novel applications that range from Landscape Ecology to Proteomics. Additional strengths of FNEA include the following.

- The FNEA region-based approach involves generating a hierarchical segmentation at various scales that yield satisfying results with respect to the desired geometrical accuracy of image-object outlines/boundaries and their unique class membership within a single region. In addition, several studies illustrate that this type of classification improves land-use classification results—rather than land cover (for an overview, see Blaschke and Strobl, 2001; Schiewe et al., 2001).
- Another aspect beyond a simple improvement of image classification is the potential to differentiate 'object-classes' within the same image 'on-demand' with different levels of detail for different applications. For example, contrary to the static view of a map, all forest areas in an image could be treated as relatively homogeneous (although in reality they are not) and grassland-objects could be explored in greater detail or vice versa (Blaschke et al., 2001).
- FNEA offers the possibility to reproduce image-object boundaries across different data sets, e.g., medium and H-res imagery, regional to local, and it allows for a transparent inspection of results.
- The definition of heterogeneity in FNEA allows for both the description of an object by means of its absolute heterogeneity, as well as the comparison of objects according to the difference between adjacent regions. By this means, the user can compare different segmentation results and their corresponding optimization procedures.
- Image-object texture is expressed more explicitly in FNEA than when using moving window texture filters. Furthermore, a comparison of results between FNEA and traditional region growing segmentation algorithms indicates that FNEA results in less over-segmentation within an image and overall, produces far more recognizable image-objects within the scene (Blaschke et al., 2000).
- FNEA incorporates semantics based on user knowledge. For example, if an object spectrally

corresponds to *woodland* and some of its direct neighbors or objects within a certain distance are classified as *urban*, then the object could be classified as *urban park*.

Due to its non-committed framework, linear SS in combination with non-linear blob-feature detection represents a powerful, predominantly automated, multiscale object-detecting framework, the strengths of which include the following.

- SS requires no a priori scene information.
- SS is based upon well-established mathematical formulation and theory, i.e., group invariance, differential geometry and tensor analysis.
- The use of Gaussian kernels satisfies the linear diffusion equation. Therefore, the diffusion, i.e., smoothing, of the gray-level intensity over scale is well modeled and understood. Furthermore, Gaussian kernels meet axioms required by an uncommitted vision system and they exhibit similarity with biological vision. In particular, SS operators closely resemble receptive field profiles registered in mammalian retina and the visual cortex (Koenderink, 1984).
- SS produces visually impressive and intuitive results. The spatially explicit nature of these results, i.e., binary-blobs that result from blob feature detection, can be converted from raster to vector topology for use in a GIS, in spatial models and or by spatial statistical packages to evaluate landscape structures and their associated metrics.
- The hierarchical nature of the scale-space primal sketch lends itself well to multiprocessing and distributed-network computing solutions.

MOSA represents a non-linear framework for generating a multiscale representation of a scene that allows dominant image-objects to emerge at their respective scales. This framework exhibits the following properties.

- OSA/OSU was statistically proven to produce better upscaling results than cubic convolution, bilinear interpolation, nearest neighbor and non-overlapping averaging (Hay et al., 1997).
- OSU incorporates object-specific weights, thus minimizing the effects of MAUP.
- It is based upon concepts related to how humans perceive texture (Hay and Niemann, 1994; Hay et al., 1997) and it incorporates 'generic' point spread function characteristics in relation to object size for determining an appropriate upscale resolution for the next iteration of processing (Hay et al., 2001).
- OSA/OSU allows for upscaling between objects and within an image hierarchy. The underlying ideas and heuristics are conceptually simple, are based upon strong empirical evidence, and follow many concepts of Complex Systems and Hierarchy theory.
- For iterative OSA/OSU, no a priori scene information is required. Essentially, computation proceeds until there is not enough information remaining to upscale.
- OSU takes into account the relationship between the pixel size and the image-objects from which the original OSA heuristics were developed. Thus, at fine scales, results visually model known image-objects very well. Therefore, a precedent exists over results at coarser unverifiable image-scales.
- Land-cover classifications improve with the addition of object-specific data sets as additional logic channels.
- MCS is well documented in the literature, and watershed segmentation algorithms are commonly available in popular image-processing packages.

4.2. Limitations of FNEA, SS and OSA/OSU

Although FNEA is already embedded in a commercial software environment, it cannot be fully exploited as long as a theoretical framework remains undefined and users have to find useful segmentation levels by 'trial and error'. In particular:

- FNEA requires that the user must know the scale of the objects of interest to select appropriate segmentation heuristics. We suggest that this is not reasonable when working at scales beyond common spatial and or spectral experience, or when conducting baseline analysis in areas where no a priori information exists.
- FNEA requires a significant amount of exploratory work to define the appropriate segmentation levels. In addition, a unique solution for a single image is

not fully operational and transferable across scenes without major corrections.
- There is no continuous upscaling mechanism for scaling between hierarchical levels or image-objects.
- There is no sound ecological theory presented for linking/defining structures through scale.

The mathematical formulation of SS is extremely rigorous. Unfortunately, this makes it nontrivial for laypersons to understand. Furthermore, to the best of our knowledge, no commercial software exists;[4] thus, image processing and topological tools must be developed in-house, which limits its widespread utility. In this paper, all SS and MOSA programming have been performed using IDL (IDL, 2002). Other limitations include the following.

- Pixel size does not change through scale; thus, each stack represents large amounts of redundant information. This poses a real challenge when using large-scale remote sensing data sets. In our work, hundreds of gigabytes of statistical (white noise) processing were required before generating normalized 4D volumes for a stack representing 500×500 pixels $\times$ 200 channels. However, once generated, these statistics can be stored in a database and used for any other data set with the same grain, extent, scale increment, and number of scales.
- Within a stack, high-contrast features tend to persist in scale, regardless of whether or not such features have ecological meaning or meaning for the application.
- Values for optimal scale generation, i.e., number of scales in a stack, the selected scale increment, and the number of ranked 'significant blobs' to evaluate are all arbitrarily defined. We suggest that these represent the most fundamental weaknesses of this framework, because these values are critical 'scale' components from which corresponding entities emerge. However, in most cases, reasonable assumptions can be made regarding the number of scales to assess and their scale increment. But determining the number of ranked blobs is not trivial. We allowed 20% of the blobs to be ranked. This resulted in 2537 individual blobs, many of which appeared to overlay each other (Fig. 8), making evaluation difficult.
- SS uses discrete data, i.e., individual pixels at specific scales, to represent what is essentially a continuous process, i.e., an object's persistence through a range of scales. Thus, SS events, conceptualized as points in space, are actually modeled as single blobs. This means that a decision has to be made whether a blob is a member of the SS-blob below it, above it, or if it is a new blob. This in turn affects the 4D-volumetric measure and ultimately the ranking of significant 2D-blobs. Consequently, this is not a trivial problem to solve, though 'work-a-rounds' exist (Hay et al., 2002b).

One of the greatest limitations of the MOSA is that it has not yet been tested over a large number of different landscapes, or by a significant number of researchers. However, further testing and validation are underway. In addition, no commercial software is

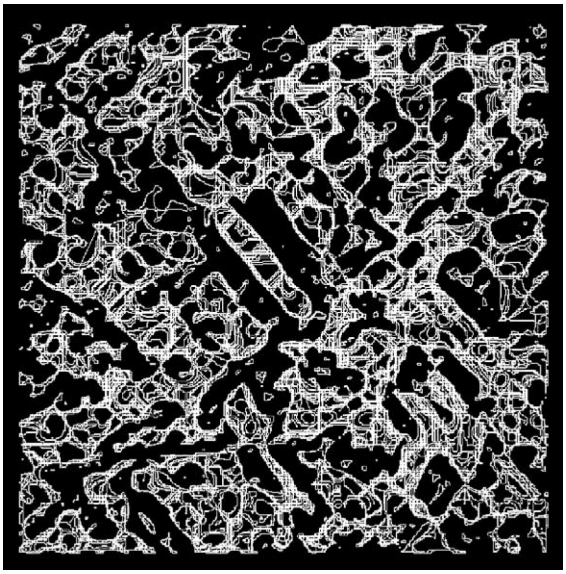

Fig. 8. Vectorized ranked blobs. Note how polygons overlay each other making analysis nontrivial (cf. with Fig. 1a).

[4] ter Haar Romeny and Florack (2000) present a scale-space workbook using the computer algebra package *Mathematica* (http://www.wolfram.com). Code for edge, ridge and corner detection is provided, but blob-detection is not described.

available and like FNEA, its object modeling is done empirically. Thus, the results of multiscale analysis require validation against field data, which becomes difficult if not impossible as scales become coarser. The addition of MCS to the iterative OSA/OSU framework allows for automatically segmented image-objects, but there is no presented topological solution for hierarchically linking them through scale.

4.3. Strategies for improving results in FNEA, SS and OSA/OSU

As discussed in the preceding section, each of the three techniques, while novel and powerful in the area of remote sensing, also exhibit limitations that make them less than ideal for the automatic detection, extraction, and hierarchical linking of ecologically meaningful multiscale image-objects within remotely sensed data. In this section, we describe a number of strategies that draw upon the strengths of each individual method and discuss how they may be applied to enhance the capabilities of the other respective techniques.

FNEA requires that a user possesses a priori information regarding the scene. To overcome this limitation, particularly when conducting baseline mapping, where no ‘ground-truth’ data are available, we hypothesize that the scale of expression and location of significant scale-space blobs may be used as early visual operators to automatically define and or refine the aggregation semantics of the FNEA. Unfortunately, based on preliminary results (Blaschke and Hay, 2001), it does not appear possible to incorporate SS-results for the a priori determination of the most relevant FNEA segmentation levels, because SS-blobs cannot be associated to a single level of segmentation. However, it may be possible to combine both approaches during the classification and/or interpretation processes, respectively. Because FNEA produces several levels of objects and the classification process utilizes this multi-level information explicitly, the analyst has to determine how many levels will be incorporated for a certain class definition. In this case, the SS rank-id and the domain/lifetime attributes provide important and potentially useful information.

In a SS-cube, a significant amount of redundant data results in large stack sizes, which in our research range from 200 to 980 MB each. To reduce the memory requirements when defining SS-blob topology, we have integrated a three-tier approach drawn from Hierarchy theory with the capability of IDL to ‘parallel-process’ multidimensional array structures (Hay et al., 2002b). Thus, instead of loading the entire stack into memory, we only need to load three scales of a SS-cube into memory at a time. From a Hierarchy theory perspective, we evaluate the blob locations at the ‘focal’ scale and establish links with blobs in the scale above and with those in the scale below. We then shift up an additional scale in the cube, while dropping the bottom scale, always keeping only three scales in memory at once. We then repeat this procedure until the last scale has been processed.

To overcome evaluation problems resulting from the large number of ranked SS blobs that visually obscure each other when overlaid on the study area, we suggest that SS-events represent critical thresholds within a hyper-blob, where fundamentally different geometric structures exist both in scale and the landscape. Thus, from an ecological perspective, the lifetime of a SS-blob may be considered as levels within a specific scale-domain. To define this domain, each hyper-blob is topologically registered as a unique entity and its corresponding SS-events are isolated. That is, the first SS-event of each hyper-blob is geometrically defined regardless of where and what scale they exist within the stack, i.e., x, y, t. Then the second, third and nth-events of each hyper-blob are isolated until the last possible event is defined. These event values are then considered as ‘scale domain attributes’ and are assigned to their corresponding ranked blobs. This domain attribute provides a unique way to query, partition and evaluate the resulting multiscale ‘domain’ surface structures, as many blobs can and do exist within a single domain, but no more than one blob can exist within the same ‘x, y, z domain’ space. Thus, the overlapping/obscuring problem is resolved and it allows us to evaluate the resulting multiscale surface structures in terms of *critical scale-specific thresholds*. In addition, by integrating these hierarchical concepts with geostatistics and 3D visualization techniques, domains can be visually modeled as ‘scale-domain manifolds’ (Hay et al., 2002b), which we suggest correspond to the ‘scaling ladder’ as described by Wu (1999) in his outline of the Hierarchical Patch Dynamics Paradigm.

Experience and knowledge gained from SS related to the importance of Gaussian filters and the axioms they satisfy as an uncommitted vision system have also been applied to OSA/OSU. In particular, to reduce the diagonal bias introduced by the square kernel originally used in OSA, we have incorporated the use of a round filter similar to that used in SS. Though not truly Gaussian, it is a pixelated approximation of a round kernel[5] that results in a more isotropic filter. To implement this change, the variance threshold heuristics have been modified and tested accordingly. The most important result of this implementation is that when analysis is conducted over large window sizes, diagonal artifacts (due to a square kernel) are significantly visually reduced within the image. Furthermore, to increase computational efficiency when using this filter, a 'bank' of varying sized round filters could be generated once and called as needed; and convolution in the Fourier domain (as done for SS processing) can be used to reduce the need to apply a moving window routine.

To provide a topological solution for hierarchically linking image-objects through scale, the topological tools developed to assess multiscale SS-blob structure can be used to establish hierarchical links with individual image-objects through all MCS data sets. Consequently, each watershed-object and its associated spatial attributes can be explicitly modeled and analyzed within a GIS and/or used as an additional logic channel for improved land-cover classification results. Thus, the ability exists to create a true object-oriented topology like FNEA, but with a number of the SS advantages inherent to an uncommitted framework. In addition, no user interaction is required and the system and its results are fully decomposable, i.e., tractable, through scale.

5. Conclusions

Complex systems are hierarchically structured, scale dependent and composed of many interacting components. These components are of two fundamental types: integrated objects and aggregate objects. From a remote sensing perspective, image-objects are integrated objects that exhibit an intrinsic scale and are composed of structurally connected parts, i.e., H-res pixels. In this paper, we hypothesize that multiscale analysis should be guided by the intrinsic scale of the image-objects composing the scene. Thus, image-objects should play a key role in the multiscale analysis, exploration and hierarchical linking of remote sensing data. As a basis for this, we describe and compare the limitations, strengths and results of three technically and theoretically different multiscale approaches, each with a common theme: their focus on intrinsic pattern and their multiscale exploration of image-objects within a single image.

FNEA automatically isolates image-objects of different sizes and shape dependent upon scales predetermined by the analyst. The resulting image-objects correspond strongly to differently sized landscape components as an experienced image interpreter would delineate them. However, a significant amount of exploratory work is required to define appropriate segmentation levels, and a unique segmentation solution is not fully operational and transferable to another image without major corrections.

Linear SS is an uncommitted vision framework; thus, it requires very little user interaction, or a priori scene information. However, a range of scales and their scale increment must be defined to generate a multiscale representation. Unlike FNEA, SS combined with blob-feature detection does not provide explicit object delineation, but rather provides a more generalized representation that can support or guide later stage visual processing.

MOSA represents a hierarchical framework by which the *spatially dominant* components of an image will emerge at coarser scales because analysis and segmentation are specific to the differently sized, shaped, and spatially distributed image-objects that compose the scene. These image-objects are visually meaningful, hierarchically tractable, reduce the effects of MAUP and require no a priori scene information for image-object delineation to occur. MOSA also provides an object-specific mechanism for upscaling, though we note that this framework is relatively recent and requires further evaluation.

Because each of the described techniques have evolved beyond individual pixel analysis to analyzing

[5] The method used is that of Michener's, modified to take into account the fact that IDL plots arrays faster than single points. See Foley and van Dam (1982, p. 445) for the algorithm.

the explicit contextual structure of image-objects, they have significant potential for ecological applications, for example:

- At fine scales, each technique could be used for individual tree crown, forest-object and landscape patch recognition, though we note that only SS and MOSA offer an 'unsupervised' approach for object delineation. In addition, FNEA output can be used to improve *land-use* classifications and MOSA output can automatically be generated to improve *land-cover* classifications.
- The explicit delineation of image-objects defined by FNEA can be used for baseline mapping and/or for updating existing geo-information. Additionally, FNEA has the ability to differentiate different 'object-classes' within the same image 'on-demand' for different applications. This could be especially useful for defining different habitat maps within the same scene based upon different habitat scale requirements.
- Animation of binary blobs within a stack could be used to visually assess fragmentation and connectivity of dominant forest ecosystem components through scale.
- Blob-events could be used to find critical scale-dependent landscape thresholds.
- Within the MOSA framework, individual ecosystem components (i.e., trees, vegetation components, etc.) and edge effects become real objects that evolve and are measurable through scale. This could have important implications in reserve and habitat planning. Model development and data type selection could also be guided by SS and MOSA scale-domains and landscape threshold patterns.

In this work, we examine the potential of three multiscale methodologies for the modeling of landscape structure in H-res imagery and have provided a comparison of each approach. While the methodological comparison is technically interesting, the overall goal of this study is to contribute to a more coherent understanding of landscape structures, their representation in remote sensing images, and mechanisms for their linking through scale. The authors independently started from paradigms that most remote sensing and GIS methodologies do not readily support: Geographic entities are represented at a variety of scales and levels of abstraction, within a single image. All three approaches incorporate a *bottom-up* approach designed to convert lower level observational data into higher level geographic entities. Thus, armed with a visual perspective of the patterns generated at different scales, and methods to decompose them into their constituent multiscale image-objects, we suggest that the ability to understand the processes behind multiscale landscape patterns is significantly enhanced.

Acknowledgements

This work has been supported by a team grant from FCAR, Gouvernement du Québec, awarded to Dr. André Bouchard and Dr. Danielle Marceau and by a series of scholarships awarded to Dr. Geoffrey Hay which include a Marc Bourgie Foundation PhD Award of Excellence, a Biology graduate scholarship of excellence from the University of Montréal, and a GREFi PhD scholarship (*Groupe de recherche en écologie forestière interuniversitaire*). Dr. Thomas Blaschke is thankful to Definiens Imaging, Munich, Germany for close cooperation with the University of Salzburg. We also thank Mr. Patrick Dubé for his assistance with scale-space coding and express our appreciation to three anonymous referees for their constructive comments.

References

Allen, T.F.H., Starr, T.B., 1982. Hierarchy Perspective for Ecological Complexity. University of Chicago Press, Chicago. 310 pp.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multiscale image segmentation. In: Strobl, J., Blaschke, T. (Eds.), Angewandte Geogr. Informationsverarbeitung, vol. XII. Wichmann, Heidelberg, pp. 12–23.

Beucher, S., Lantuéjoul, C., 1979. Use of watersheds in contour detection. Int. Workshop on Image Processing, Real-time Edge and Motion Detection/Estimation, Rennes, France, 17–21 September. CCETT/IRISA (Centre commun d'études de télédiffusion et télécommunications /Institut de recherche en informatique et systèmes aléatoires). Report No. 132, 2.1–2.12.

Biederman, I., 1987. Recognition-by-components: a theory of human image understanding. Psychological Review 94 (2), 115–147.

Blaschke, T., Hay, G.J., 2001. Object-oriented image analysis and scale-space: theory and methods for modeling and evaluating multiscale landscape structure. International Archives of Photogrammetry and Remote Sensing 34 (Part 4/W5), 22–29.

Blaschke, T., Strobl, J., 2001. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS-Zeitschrift für Geoinformationssysteme (6), 12–17.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: Cremers, A., Greve, K. (Eds.), Environmental Information for Planning, Politics and the Public, vol. 2. Metropolis Verlag, Marburg, pp. 555–570.

Blaschke, T., Conradi, M., Lang, S., 2001. Multi-scale image analysis for ecological monitoring of heterogeneous, small structured landscapes. Proceedings of SPIE 4545, 35–44.

Bobick, A., Bolles, R., 1992. The representation space paradigm of concurrent evolving object descriptions. IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (2), 146–156.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (6), 679–697.

Chaudhuri, B., Sarkar, N., 1995. Texture segmentation using fractal dimension. IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1), 72–77.

Definiens, 2002. www.definiens-imaging.com. Accessed October 30, 2002.

Foley, J.D., van Dam, A., 1982. Fundamentals of Interactive Computer Graphics. Addison-Wesley, Reading, MA. 664 pp.

Foody, G., 1999. Image classification with a neural network: from completely-crisp to fully-fuzzy situations. In: Atkinson, P.M., Tate, N.J. (Eds.), Advances in Remote Sensing and GIS Analysis. Wiley, Chichester, pp. 17–37.

Forman, R.T.T., 1995. Land Mosaics: The Ecology of Landscapes and Regions. Cambridge Univ. Press, Cambridge. 632 pp.

Fotheringham, A.S., Wong, D.W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. Environment and Planning A 23, 1025–1044.

Haralick, R.M., Sternberg, S.R., Zhuang, X., 1987. Image analysis using mathematical morphology. IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (4), 532–550.

Hay, G.J., Niemann, K.O., 1994. Visualizing 3-D texture: a three-dimensional structural approach to model forest texture. Canadian Journal of Remote Sensing 20 (2), 90–101.

Hay, G.J., Marceau, D.J., 1998. Are image-objects the key for upscaling remotely sensed data? Proceedings of Modelling of Complex Systems, July 12–17, New Orleans, USA, pp. 88–92.

Hay, G.J., Niemann, K.O., McLean, G., 1996. An object-specific image-texture analysis of H-resolution forest imagery. Remote Sensing of Environment 55 (2), 108–122.

Hay, G.J., Niemann, K.O., Goodenough, D.G., 1997. Spatial thresholds, image-objects and upscaling: a multi-scale evaluation. Remote Sensing of Environment 62 (1), 1–19.

Hay, G.J., Marceau, D.J., Dubé, P., Bouchard, A., 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16 (6), 471–490.

Hay, G.J., Dubé, P., Bouchard, A., Marceau, D.J., 2002a. A scale-space primer for exploring and quantifying complex landscapes. Ecological Modelling 153 (1–2), 27–49.

Hay, G.J., Marceau, D.J., Bouchard, A., 2002b. Modelling multiscale landscape structure within a hierarchical scale-space framework. International Archives of Photogrammetry and Remote Sensing 34 (Part 4), 532–536.

IDL (Interactive Data Language), 2002. www.rsinc.com. Accessed October 30, 2002.

Jähne, B., 1999. A multiresolution signal representation. In: Jähne, B., Haußecker, H., Geißler, P. (Eds.), Handbook on Computer Vision and Applications. Academic Press, Boston, USA, pp. 67–90.

Koenderink, J.J., 1984. The structure of images. Biological Cybernetics 50 (5), 363–370.

Levin, S.A., 1999. Fragile Dominions: Complexity and the Commons. Perseus Books, Reading. 250 pp.

Lindeberg, T., 1994. Scale-Space Theory in Computer Vision. Kluwer Academic Publishing, Dordrecht, The Netherlands. 423 pp.

Marceau, D.J., 1999. The scale issue in the social and natural sciences. Canadian Journal of Remote Sensing 25 (4), 347–356.

Marceau, D.J., Hay, G.J., 1999. Contributions of remote sensing to the scale issues. Canadian Journal of Remote Sensing 25 (4), 357–366.

Marceau, D.J., Howarth, P.J., Gratton, D.J., 1994. Remote sensing and the measurement of geographical entities in a forested environment: Part 1. The scale and spatial aggregation problem. Remote Sensing of Environment 49 (2), 93–104.

Meyer, F., Beucher, S., 1990. Morphological segmentation. Journal of Visual Communication and Image Representation 1 (1), 21–46.

O'Neill, R.V., King, A.W., 1998. Homage to St. Michael; or, why are there so many books on scale? Ecological Scale Theory and Applications. Columbia Univ. Press, pp. 3–15.

Openshaw, S., 1981. The modifiable areal unit problem. In: Wrigley, N., Bennett, R.J. (Eds.), Quantitative Geography: A British View. Routledge and Kegan Paul, London, pp. 60–69.

Openshaw, S., 1984. The modifiable areal unit problem. Concepts and Techniques in Modern Geography (CATMOG), vol. 38. Geo Books, Norwich. 40 pp.

Openshaw, S., Taylor, P., 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley, N. (Ed.), Statistical Applications in the Spatial Sciences. Pion, London, pp. 127–144.

Robert, C.P., 2001. The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation. Springer Verlag, New York. 624 pp.

Rowe, J.S., 1961. The level-of-integration concept and ecology. Ecology 42 (2), 420–427.

Salari, E., Ling, Z., 1995. Texture segmentation using hierarchical wavelet decomposition. Pattern Recognition 28 (12), 1819–1824.

Schiewe, J., Tufte, L., Ehlers, M., 2001. Potential and problems of multiscale segmentation methods in remote sensing. GIS-Zeitschrift für Geoinformationssysteme (6), 34–39.

Settle, J.J., Drake, N.A., 1993. Linear mixing and the estimation of ground cover proportions. International Journal of Remote Sensing 14 (6), 1159–1177.

Simon, H.A., 1962. The architecture of complexity. Proceedings of the American Philosophical Society 106, 467–482.

Skidmore, A., 1999. The role of remote sensing in natural resource management. International Archives of Photogrammetry and Remote Sensing 32 (Part 7C2), 5–11.

ter Haar Romeny, B.M., Florack, L.M.J., 2000. Front-end vision, a multiscale geometry engine. Proc. First IEEE International Workshop on Biologically Motivated Computer Vision (BMCV2000), May 15–17, Seoul, Korea. Lecture Notes in Computer Science, vol. 1811. Springer Verlag, Heidelberg, pp. 297–307.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46 (2), 234–240.

Wang, F., 1990. Improving remote sensing image analysis through fuzzy information representation. Photogrammetric Engineering and Remote Sensing 56 (8), 1163–1169.

Weickert, J., Ishikawa, S., Imiya, A., 1997. On the history of Gaussian scale-space axiomatics. In: Sporring, J., Nielsen, M., Florack, L., Johansen, P. (Eds.), Gaussian Scale-Space Theory. Kluwer Academic Publishing, Dordrecht, The Netherlands, pp. 45–59.

Wessman, C.A., 1992. Spatial scales and global change: bridging the gap from plots to GCM grid cells. Annual Review of Ecological Systems 23, 175–200.

Wiens, J.A., 1989. Spatial scaling in ecology. Functional Ecology 3 (4), 385–397.

Woodcock, C.E., Strahler, A.H., 1987. The factor of scale in remote sensing. Remote Sensing of Environment 21 (3), 311–332.

Wu, J., 1999. Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing 25 (4), 367–380.

Wu, J., Marceau, D., 2002. Modelling complex ecological systems: an introduction. Ecological Modelling 153 (1–2), 1–6.