International Journal of Remote Sensing
Vol. 31, No. 6, 20 March 2010, 1427–1439

Taylor & Francis
Taylor & Francis Group

Segmentation of multispectral high-resolution satellite imagery using log Gabor filters

PENGFENG XIAO*†, XUEZHI FENG†, RU AN‡ and SHUHE ZHAO†
†Department of Geographical Information Science, Nanjing University, Nanjing, 210093, PR China
‡College of Hydrology and Water Resources, Hohai University, Nanjing, 210098, PR China

Image segmentation has been recognized as a valuable approach that performs a region-based rather than a pixel-based analysis of high-resolution satellite imagery. A scheme for segmenting the multispectral IKONOS image based on frequency-domain filtering is presented. The frequency spectrum of typical landscape objects is analysed first. The spectrum curves are comparable in logarithmic coordinates rather than in Cartesian coordinates; therefore the Gabor filters are superseded by log Gabor filters to extract the multiscale texture features from panchromatic band. Edge features then are calculated from the pan-sharpened multispectral bands based on the vector field model. Finally, the texture-marked watershed segmentation algorithm is implemented and the segmentation accuracy is assessed. The experimental results show that the developed scheme generated an effective tool for automatic segmentation of multispectral high-resolution satellite imagery and suppressing the over-segmentation problem of watershed transform.

1. Introduction

With progresses in satellite sensor technology, higher spatial resolution increases the spectral within-field variability – in contrast to the integration effect of earlier sensors – and therefore may decrease the classification accuracy of traditional methods on a per-pixel basis. The object-based method has become one of the most commonly used strategies for processing high resolution imagery. Hence, image segmentation has already been recognized as a valuable approach that creates regions instead of pixels as carriers of features which are then introduced into the classification stage (Schiewe 2002).

Image segmentation techniques automatically group neighbouring pixels into meaningful regions based on homogeneity or heterogeneity criteria. They include the following strategies: pixel-oriented, edge-oriented, region-oriented and hybrid (Pal and Pal 1993). Even though image segmentation has been heavily studied in image processing and computer vision fields, and despite the early efforts that use spatial information for classification of remotely sensed imagery, segmentation algorithms have only recently started receiving emphasis in remote sensing image analysis. Examples of image segmentation in the remote sensing literature include region growing (Baatz and Schäpe 2000, Evans *et al.* 2002) and Markov random field models (Sarkar *et al.* 2002) for segmentation of natural scenes, region growing for object level change detection (Hazel 2001), and boundary delineation of agricultural fields

*Corresponding author. Email: xiaopf@gmail.com

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903475324

(Rydberg and Borgefors 2001). With the increased availability of very high-resolution imagery, image segmentation has become popular as a common variant of data interpretation (Pesaresi and Benediktsson 2001, Pekkarinen 2002). As a morphologic segmentation algorithm, watershed transform is used widely in remote sensing to the resulting closed and connected regions, e.g. Karantzalos and Argialas (2006) proposed a scheme for improving edge detection and watershed segmentation for detecting olive trees from IKONOS panchromatic image. In the studies of Meinel and Neubert (2004) and Neubert *et al.* (2006) the capabilities of available segmentation programmes for high resolution remote sensing data were assessed and compared.

Although there has been a lot of development in the segmentation of greyscale and colour images, there has been little progress in the segmentation of multispectral images. Moreover, remote sensing sensors are producing multispectral images so that in contrast to the most often used monochromatic image in the disciplines, not only the complexity but also the redundancy increases. Measuring all of the individual objects in the spatial domain would be very time-consuming and not particularly accurate, but the frequency-domain representation of the image shows the periodicity clearly.

A scheme for segmenting a multispectral IKONOS image based on frequency-domain filtering and watershed transform is presented in this paper. Firstly, a Fourier spectrum of typical landscape objects illustrates better distinguishable capability in logarithmic rather than Cartesian space, thus the log Gabor filters are used to extract the multiscale texture features from panchromatic band. Then, the edge features are calculated from the pan-sharpened multispectral bands based on a vector field model. Finally, the texture-marked watershed segmentation algorithm is implemented and the discrepancy is calculated to assess the accuracy of segmentation. The complete system used in the approach is outlined in figure 1 and consists of the following stages: frequency spectrum analysis, feature extraction, and feature segmentation.

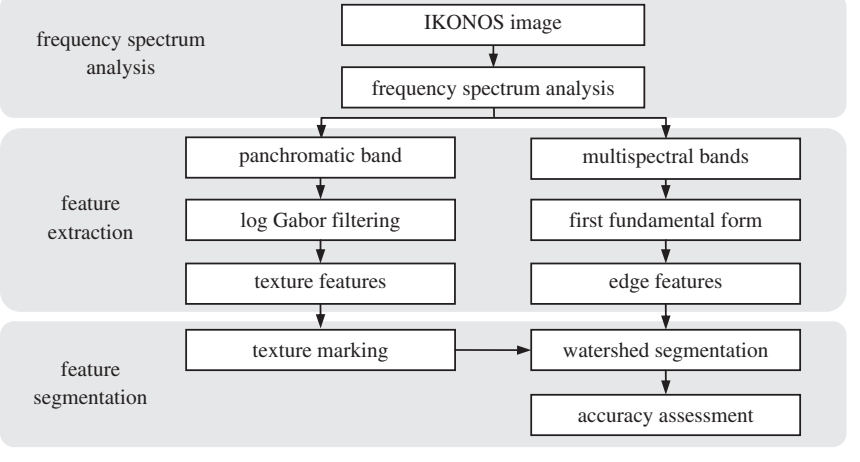


Figure 1. The scheme for segmentation of a multispectral IKONOS image based on frequency-domain filtering, which consists of three main stages: frequency spectrum analysis, feature extraction and feature segmentation.

2. Study area and data

Nanjing in China is a rapidly developing megacity and has been chosen as the study area. It is a central city (latitude 32° 03′ N and longitude 118° 46′ E) in the western part of the lower reaches of the Yangtze River Delta and has scenery with mountains, water and green trees. The downtown area of Nanjing is 2597 km^2, and the city has a population of more than six million people. Nanjing spans the southern and northern sides of the Yangtze River.

Very high-resolution IKONOS imagery is considered, which features a geometric and radiometric quality of one panchromatic band with 1-m spatial resolution and four multispectral bands with 4-m spatial resolution. The image was acquired on 15 September 2000. The size of the image is 9628 columns by 11 240 rows of pixels. To produce a high quality multispectral image, the low-resolution multispectral bands are first sharpened to 1-m resolution using a high-resolution panchromatic band based on the Pansharp method (Zhang 2002). The image primarily consists of water, paddy, vegetation, road, workshop and house objects. The false colour image is composed with infrared, red and green bands. In the false colour image, the red colour represents vegetation, black indicates water, and grey represents concrete structures, i.e. roads and houses.

3. Frequency spectrum analysis

Fourier transform is widely used in digital signal processing. It transforms energy into a frequency spectrum. The 2D discrete Fourier transform of an $M \times N$ image $I(x, y)$ can be written as

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y) e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}, \tag{1}$$

where j is the square root of −1. The transform function $F(u, v)$ is generally complex, the sum of a real part and an imaginary part. This has a polar form,

$$F(u,v) = |F(u,v)| e^{-j\varphi(u,v)}, \tag{2}$$

where $|F(u, v)|$ is magnitude and $\phi(u, v)$ is phase. The magnitude is commonly referred to as the frequency spectrum. The square of the magnitude $P(u, v) = |F(u, v)|^2$ is referred to as the power spectrum. By Parseval's theorem the variance of the image, i.e. energy, is equal to the integral of the power spectrum.

3.1 *Qualitative description*

Figure 2 shows six types of landscape object—water, paddy, woodland, road, workshop and house—in an IKONOS panchromatic image, along with their frequency spectrum, where the spectrum is shown with a logarithmic vertical scale for clarity. The origin is in the centre of the spectrum. Since different frequencies are represented at different distances from the origin, different directions represent different orientations in the original image, and the magnitude at each location shows how much of that frequency and orientation is present in the image. The low-frequency components near the origin of the plot provide the overall shape of the image, while the high-frequency components are needed to sharpen edges and provide fine detail. The spectrum character of the six types of landscape object can be described as follows.

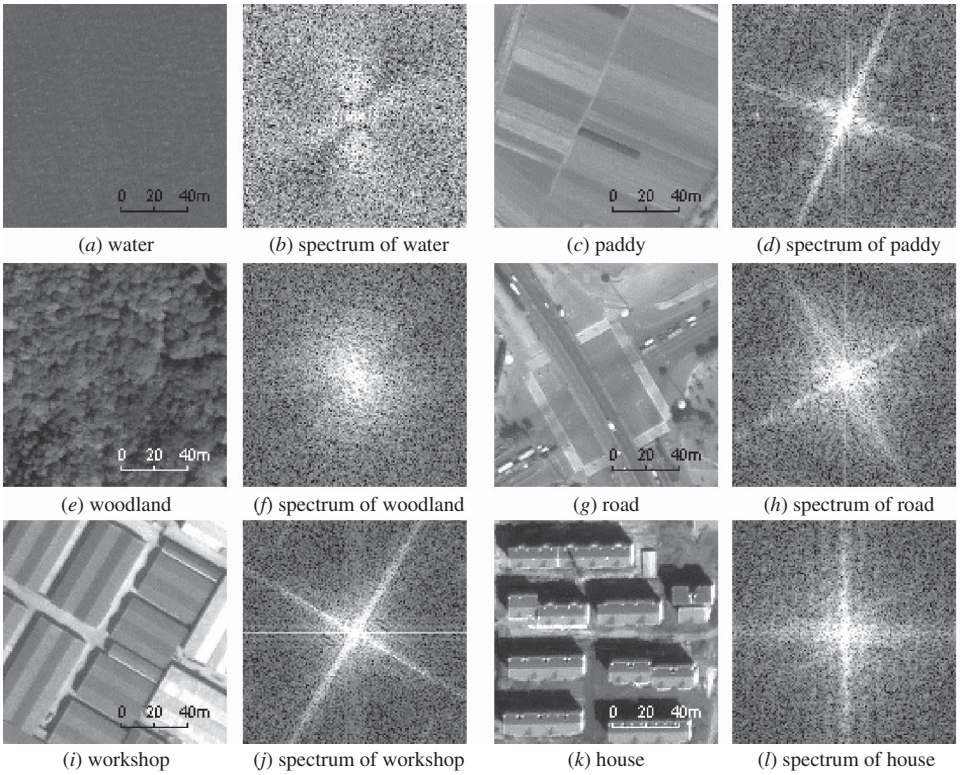


Figure 2. Six types of landscape object: water, paddy, woodland, road, workshop and house, in IKONOS panchromatic images, with their frequency spectrum, where the spectrum is shown with a logarithmic vertical scale for clarity.

1. Water: The direct centre (DC) component is not very distinctive, and the low and high-frequency components are also scattered. So it can be shown that the water image lacks texture features.
2. Paddy: The DC component is centralized. There are more low-frequency components than high-frequency components, shown that the large-area field blocks are more dominant than detailed structures in the image. Two spectrum lines are particularly noticeable: a 60° one indicating the orthogonal ridges of the fields, and a 150° one indicating the orthogonal roads between the fields in the image.
3. Woodland: There is not a distinct spectrum line. The low-frequency components are richer than the high-frequency components, shown that the canopy texture is dominant in contrast to the branch and leaf texture in the image.
4. Road: The concrete structure is highly reflective, thus it is rich in low- and high-frequency components. Spectrum lines are evident, indicating the dominant line structure of the road image.
5. Workshop: The average brightness of the image is high because there are many white roofs with high reflection, so three is prominent peak in the DC component. The rich low-frequency components primarily show the shape of the workshop. Two orthogonal spectrum lines indicate the boundary of the roofs.

6. House: This is rich in the low- and high-frequency components. The low-frequency components roughly indicate the flat grey of the ground, and the high-frequency components roughly indicate the house texture. Two orthogonal spectrum lines indicate the boundary of the houses and shadows.

The different character of landscape objects in the frequency spectrum is particularly useful for isolating periodic structures, which is difficult in the space domain.

3.2 *Quantitative analysis*

For quantitatively analysing the spectrum, we scan the spectrum in a different frequency to obtain the annular distribution first proposed by Conners and Harlow (1980):

$$M(r) = \sum_{\theta=0}^{\pi} |F(r,\theta)|, \tag{3}$$

where $|F(r, \theta)|$ is the polar form of the magnitude $|F(u, v)|$. The annular distribution of the magnitude can be used as a measure of texture coarseness and periodicity. As shown in figure 3(*a*), the house, workshop and road image has high magnitude, whereas the paddy and woodland image has low magnitude, and the water image has the lowest magnitude, which illustrates their differences in reflectivity: the house, workshop and road reflect high, the paddy and woodland reflect low, and the water reflects lowest. The magnitude of the house and workshop image has an evident peak in about four cycles/image and six cycles/image, respectively, indicating the shape structures in that frequency. However, they all show the greatest magnitude at low frequencies and decreasing magnitude as the frequency increases, thus the curves are staked together at high frequencies and are difficult to distinguish.

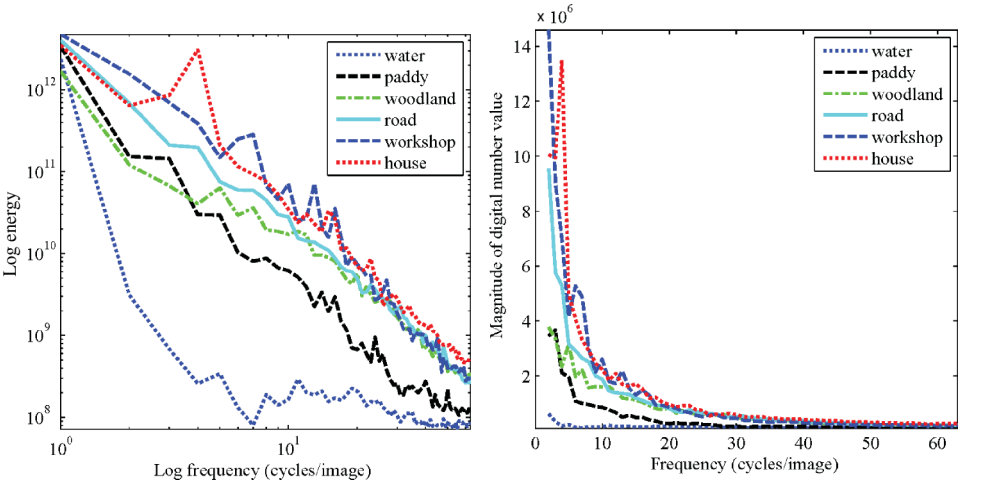


Figure 3. The annular distribution of magnitude (*b*) and logarithmic annular-distribution of power (*a*) for the six types of landscape object. The DC components are eliminated from the spectrum.

To make clear the difference at high frequency, which is important in filter design, we plot the annular distribution of the power spectrum on logarithmic coordinates,

$$\log[E(r)] = \log\left[\sum_{\theta=0}^{\pi} P(r,\theta)\right], \tag{4}$$

where $P(r, \theta)$ is the polar form of the power $P(u, v)$. The logarithmic annular-distribution of power can be used as a measure of energy character of the image. Figure 3(*b*) shows the logarithmic annular-distribution of power for the six landscape objects. On the logarithmic coordinates the power falls off by a factor of roughly $1/f^2$ (Field 1987). Although the results vary from image to image, these plots suggest that, in contrast to the annular distribution of magnitude, there is roughly comparable energy in any given octave, i.e. no particular frequency contains consistently more energy. These properties lead us to design frequency-domain filters in the logarithmic space.

4. Feature extraction

4.1 *Log Gabor filters and texture feature extraction*

In order to produce texture features we first need to characterize the texture content of the image at each pixel. One of the most popular techniques is the use of a set of differently scaled and orientated complex Gabor filters (e.g. Bovik *et al.* 1990, Jain and Farrokhnia 1991, Teuner *et al.* 1995, Arivazhagan *et al.* 2006). The Gabor filter has the capability of reaching the minimum bound for simultaneous localization in the space and frequency domains. The one dimensional Gabor function is defined as a complex sinusoidal plane wave of some frequency and orientation within a Gaussian envelope:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{x^2}{2\sigma^2} + ju_0x), \tag{5}$$

where σ is the space constants of the Gaussian envelope, and u_0 is the frequency of the sinusoidal plane wave.

However, the Gabor filter is mathematically pure in only the Cartesian coordinates where all the Gabor channels are the same size in frequency and hence have sensors that are all the same size in space. An objective of the filter design might be to obtain as broad as possible spectral information with maximal spatial localization. One cannot construct a Gabor function of arbitrarily wide bandwidth and still maintain a reasonably small DC component in the even-symmetric filter (Kovesi 1996).

Based on the annular distribution of energy in logarithmic coordinates, an alternative to the Gabor function is the log Gabor function, first proposed by Field (1987). On the linear frequency scale the log Gabor function has a transfer function of the form

$$G'(x) = \exp\left\{\frac{-[\log(x/u_0)]^2}{2[\log(\sigma/u_0)]^2}\right\}. \tag{6}$$

To obtain constant shape ratio filters the term σ/u_0 must also be held constant for varying u_0. There are two important characters for log Gabor function, as shown in figure 4. Firstly, log Gabor function, by definition, always has no DC component, and secondly, the transfer function of the log Gabor function has an extended tail at the high frequency end, which conquers the over-representation in low frequency components as discussed in §3.2.

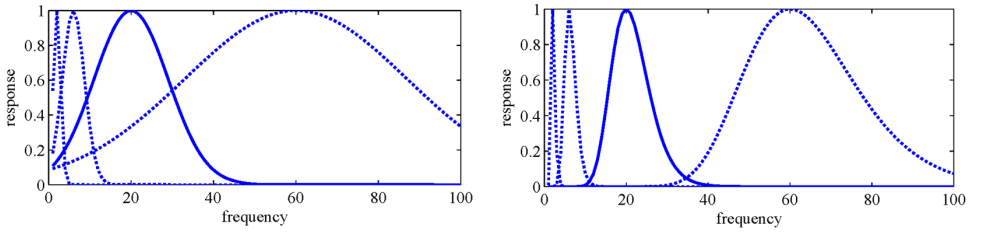


Figure 4. The transfer function of Gabor (*a*) and log Gabor (*b*) in Cartesian coordinates. The Gabor functions over-represent the low frequency, which is not the case for the log Gabor. The log Gabor functions always have no DC component, and have an extended tail at the high frequency end.

The panchromatic band is filtered separately using two-octave bandwidth log Gabor filters over six orientations and four frequencies. The wavelength of the highest frequency filter is three pixels, and the scaling between successive filters is 2. The filters are constructed directly in the frequency domain as polar-separable functions: a logarithmic Gaussian function in the radial direction and a Gaussian in the angular direction. The ratio between the angular spacing of the filters and angular standard deviation of the Gaussians is 1.2.

The magnitude responses of the log Gabor filters are regarded as the texture features in this study. The responses in four frequencies represent four scales of energy distribution. The first frequency texture is the highest frequency response, which indicates the periodical components of the most acute changes in the image, such as edge and noise. The second and third frequency textures indicate the transition of distribution of energy from edge to tone. The fourth frequency texture is the lowest frequency response, which mainly represents the tone features in the image. For clearly observing the detail of the processes and results, a sub-image of size 400 by 400 pixels is shown in the paper. Figure 5(*a*) shows the texture features of the

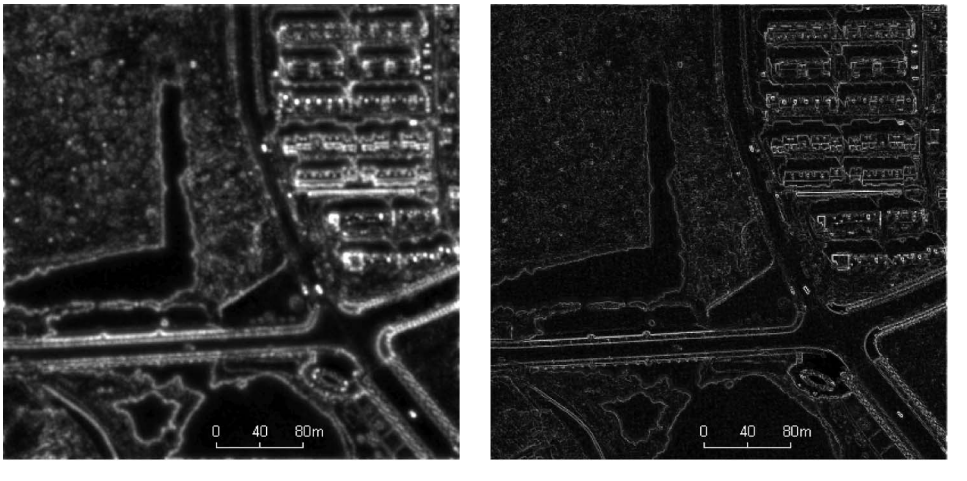


Figure 5. The texture features of the panchromatic band (*a*) and edge features of the pan-sharpened multispectral band (*b*). The texture features are the combination of (1 + 2 + 3) frequencies of the magnitude of log Gabor filtering. For clearly observing the detail of the processes and results, the sub-image of size 400 by 400 pixels is shown in the paper.

panchromatic band, which are the sum of the magnitude of the (1 + 2 + 3) frequencies.

4.2 *Multispectral edge feature detection*

Edge features provide most information in the image, thus it is very important to detect edge features for image segmentation. Remote sensors produce multispectral bands so that in contrast to the most frequently used monochromatic image, accurate edge detection is a serious challenge. Early approaches to detecting edge in multispectral images attempted to combine the response of single-valued edge detectors applied separately to each band. The responses for each band are generally combined heuristically, which has no theoretical basis. However, a principled way to look at gradient in multispectral images is the first fundamental form based on the vector field model (Sapiro and Ringach 1996).

Let $\mathbf{I}(x, y)$: $\mathbf{R}^2 \rightarrow \mathbf{R}^N$ be a multispectral image with bands $\mathbf{I}_i(x, y)$: $\mathbf{R}^2 \rightarrow \mathbf{R}$, $i = 1, \ldots, N$. The value of $\mathbf{I}$ at a given point (x_0, y_0) is a N-dimensional vector in $\mathbf{R}^N$, and the difference of image values at two points $P = (x_0, y_0)$ and $Q = (x_1, y_1)$ is given by $\Delta\mathbf{I} = \mathbf{I}(P) - \mathbf{I}(Q)$. When the Euclidean distance $d(P, Q)$ between P and Q tends to zero, the difference becomes the arc element

$$d\mathbf{I} = \frac{\partial \mathbf{I}}{\partial x} dx + \frac{\partial \mathbf{I}}{\partial y} dy \tag{7}$$

and its squared norm is given by

$$d\mathbf{I}^2 = \begin{pmatrix} dx \\ dy \end{pmatrix}^T \begin{pmatrix} \left(\frac{\partial \mathbf{I}}{\partial x}\right)^2 & \frac{\partial \mathbf{I}}{\partial x}\frac{\partial \mathbf{I}}{\partial y} \\ \frac{\partial \mathbf{I}}{\partial y}\frac{\partial \mathbf{I}}{\partial x} & \left(\frac{\partial \mathbf{I}}{\partial y}\right)^2 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} dx \\ dy \end{pmatrix}^T \begin{pmatrix} G_{xx} & G_{xy} \\ G_{yx} & G_{yy} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}. \tag{8}$$

This quadratic form is called the first fundamental form. It reflects the change in a multispectral image. The direction of maximal and minimal change is given by the eigenvectors of the 2×2 matrix. The corresponding eigenvalues denote the rates of change. Simple algebra shows that the eigenvalues are:

$$\lambda_{\pm} = \frac{G_{xx} + G_{yy} \pm \sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2}}{2} \tag{9}$$

Image discontinuities can be detected by defining a function $f = f(\lambda_+, \lambda_-)$ that measures the dissimilarity between λ_+ and λ_-. A possible choice is $f = f(\lambda_+ - \lambda_-)$, which has the convenient property of reducing to $d\mathbf{I}^2$ for the one-dimensional case.

It is shown in figure 5(*b*) that the edge detection results from the first fundamental form of the pan-sharpened multispectral band are uninterrupted, and important edges have been preserved. Furthermore, for multispectral image, a single-valued approach can be adopted by segmenting each band separately, or by first combining the bands into a single grey image. The concept of the first fundamental form however allows us to access gradient information from all bands simultaneously (Scheunders 2001).

5. Feature segmentation

5.1 *Texture-controlled watershed segmentation*

Texture features and edge features are all presented with gradient in this study. The watershed transform is a technique for segmenting a digital image; it uses a type of

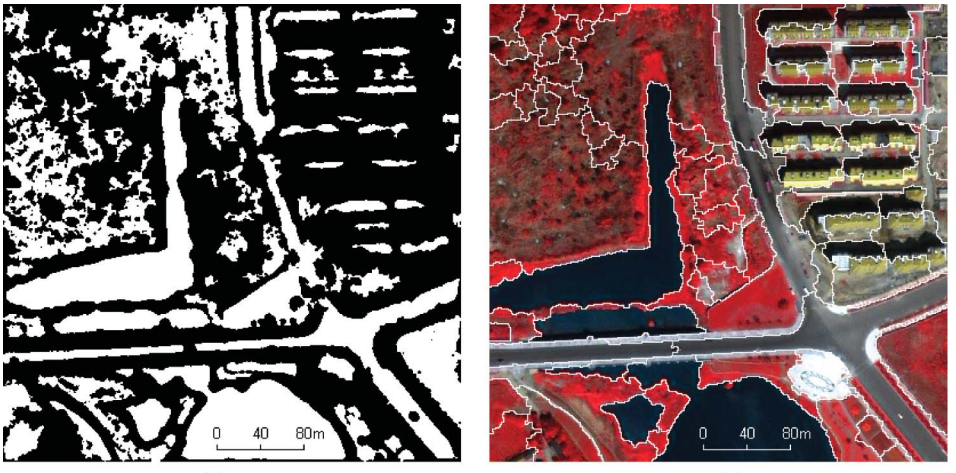
mathematical morphology method based on image gradient, and its resulting boundaries form closed and connected regions. It is thus one of the best choices for the segmentation of remotely sensed imagery, which needs all of the objects in the image to be recognized. The presented watershed algorithm is based on an immersion process analogy (Vincent and Soille 1991), in which the flooding of the water in the image is efficiently simulated using a queue of pixels. However, the standard approach in watershed segmentation causes severe over-segmentation. There are many solutions to the problem. A marker based solution is chosen in this study as basins are flooded from selected sources rather than minima.

To integrate the edge and texture information of the landscape objects in an image, watershed transform is implemented based on edge features, and the marker image is calculated from texture features. The texture features are segmented at first using a moving threshold algorithm developed from Hill *et al.* (2003). Algorithm 1 calculates the mean and standard deviation of the texture features. Then several binary images are produced at reasonably spaced thresholds using the mean and standard deviation. For each binary image, the number of closed and connected regions greater than the given minimum size is calculated. The threshold with the maximum number of connected regions is used as the output marker image.

```
Algorithm 1 MovingThreshold (minsz,G)
comment minsz: the minimum acceptable marker size
comment G: input gradient image
std = StandardDeviationOf(G)
mean = MeanOf(G)
threshs[11] = -1 to 0 step 0.1
for i = 1 to 11 {
    thresholdLevel = mean + threshs[i] × std
    thresholdImage = GTI (thresholdLevel,G)
    markerImage[i] = GCRGT (minsz)
    regionNumber[i] = NOR (markerImage[i])
    }
maxIndex = FindMaxValue(regionNumber)
return markerImage[maxIndex]
comment GTI(·): GetThresholdImage
comment GCRGT(·): GetConnectedRegionsGreaterThan
comment NOR(·): NumberOfRegions
```

With this algorithm, no a priori knowledge is required about the number of regions. One only needs to give the size of the minimum region, which is always constant in most images. The size of 50 pixels is used as the minimum region in this study. Figure 6(*a*) shows the marker image calculated from texture features in figure 5(*a*). The main landscape objects, including water, vegetation, road and house, are marked distinctly.

The edge features are then reconstructed with the marker image, that is, the basins are flooded from texture regions rather than minima. Then the reconstructed image is segmented based on the watershed transform. Moreover, the panchromatic band has produced four texture features after log Gabor filtering with four frequencies. Using different combinations of the texture features can produce segmentation results at different scales. Generally, combination of low-frequency texture features produces large-scale segmentation results, whereas a combination of high-frequency texture features produces small-scale segmentation results.



(*a*) (*b*)

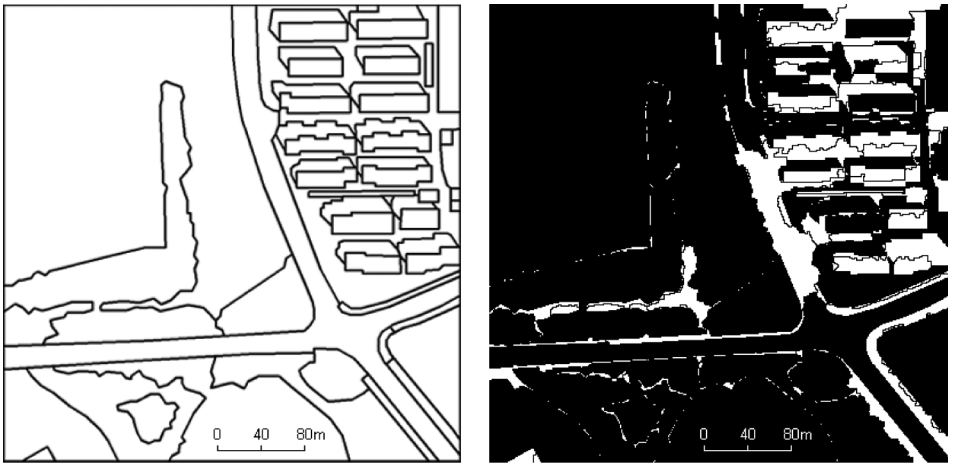
Figure 6. The marker image calculated from texture features (*a*) and the segmentation results overlaid on the original image (*b*). The false colour image is composed with infrared, red and green bands. The segment region (plotted with white boundaries) count is 80.

Figure 6(*b*) shows the image segmented with a combination of (1 + 2 + 3) texture features. It is shown that the region boundaries are congruent with most landscape objects. With texture marking, the over-segmentation problem is solved so that the count of segments is decreased to a meaningful number. But there is still a little over-segmentation in vegetation areas and under-segmentation in house areas, which may be further improved in future work.

5.2 *Accuracy assessment*

The decision about the best segmentation results usually relies on assessment. However, similar to the segmentation theory there is no established procedure for the assessment of the results. They are usually evaluated only visually, qualitatively or indirectly by the effectiveness of the segmentation on the subsequent processing steps. A general classification of assessment methods has been proposed by Zhang (1996), but only very few studies, e.g. Carleer *et al.* (2005), employ assessment on remote sensing data. Most existing approaches are supervised methods using discrepancy measures between a reference and the segmentation, which is also employed in this study.

The reference map is obtained based on human interpretation of the IKONOS false colour image, as shown in figure 7(*a*). These polygons are then converted to raster. To compare the discrepancies, the segmentation results are overlaid on the reference map. Observing from each segment region, the largest part of the reference map is regarded as the right segmentation of this region, and the other parts are regarded as mix-segmented pixels. Two discrepancy evaluations are calculated: percentage of right-segmented pixels in the whole image (PR), and ratio of region count in segmented image to reference map (RC). For an excellent segmentation, PR will be close to 100%, and RC will equal 1. If only PR is used to evaluate the segmentation, PR will be decreased with the fragmentation of the segmentation, and it will be equal to 100%



(*a*) (*b*)

Figure 7. The reference map (*a*) and corresponding mix-segmentation image (*b*). The reference regions count is 57. The mix-segmentation image is generated overlaying the segmentation results on the reference map. PR = 83.70%, RC = 1.40.

when the region size is equal to one pixel, which becomes a meaningless segmentation. So RC is an important indicator to restrict the segment region count close to the reference region count. It is shown in figure 7(*b*) that PR of the segmented sub-image is up to 83% when $1 < \mathrm{RC} < 2$. There are fewer mix-segmented pixels in simple structure areas, i.e. water and vegetation, and more in complicated structure areas, i.e. house.

The PR and RC of the segmentation results, marked with different combinations of frequency texture, are shown in table 1. It can be seen that the RC is increased with the transition to high-frequency marking. That is, the segmentation will become more fragmented if marked with texture of higher frequency. The sub-image is better segmented when marked with a combination of high-frequency texture, and reaches the best PR when mark with a combination of (1 + 2 + 3) frequency texture.

Table 1. Accuracy assessment for the segmentation of the sub-image.

Combination of marking frequencies	Segment region count	Ratio of region count in segmented image to reference map (*RC*)	Percentage of right-segmented pixels (*PR*)(%)
1	102	1.79	81.94
1+2	95	1.67	83.65
1+2+3	80	1.40	83.70
1+2+3+4	51	0.89	78.14
2+3+4	50	0.88	78.64
3+4	47	0.82	78.32
4	30	0.53	70.61

6. Conclusion

In conclusion, a scheme for segmenting multispectral high-resolution satellite imagery based on log Gabor filtering and marker-controlled watershed transform is proposed. The frequency spectra of typical landscape objects are analysed. They show greatest magnitude at low frequencies and decreasing magnitude as the frequency increases, thus the curves are staked together at high frequencys. But when the spectrum curves are transformed to logarithmic coordinates, there is roughly comparable energy in any given octave. Thus Gabor filters are superseded by log Gabor filters to extract the multiscale texture features from panchromatic band. Edge features are calculated from the pan-sharpened multispectral bands using the first fundamental form based on vector field model. Texture features and edge features are integrated in texture-marked watershed segmentation. The inclusion of texture features based on the actual frequency content of the image may ensure that differently textured regions are segmented effectively. Marking with a different combination of frequency texture may produce multiscale segmentation results.

In experiments, the proposed method demonstrates excellent performance even in complicated urban areas. In particular, the proposed approach gives a better solution for the segmentation of multispectral remotely sensed imagery. It also retains small but significant regions in image, and has an intrinsic hierarchy effect that reduces dramatically the problems of over-segmentation in the classical approach.

The drawback of the proposed method concerns the heavy computation of the multichannel filtering, which may prevent the approach being applied in real-time applications. Further study is needed on how to reduce the computation consumption of the filtering and to further improve the segmentation accuracy.

Acknowledgments

This work is supported by National High Technology Research and Development Program of China (Grant No. 2008AA12Z106) and by National Natural Science Foundation of China (Grant No. 40801166 and 40771137). The authors would like to express their thanks to Professor Fahui Wang of Northern Illinois University, USA for help for the paper. They also wish to thank four anonymous reviewers for valuable suggestions.

References

Arivazhagan, S., Ganesan, L. and Priyal, S.P., 2006, Texture classification using Gabor wavelets based rotation invariant features. *Pattern Recognition Letters*, **27**, pp. 1976–1982.

Baatz, M. and Schäpe, A., 2000, Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung* XII, J. Strobl, T. Blaschke and G. Griesebner (Eds), pp. 12–23 (Heidelberg: Wichmann-Verlag).

Bovik, A.C., Clark, M. and Geisler, W.S., 1990, Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, pp. 55–73.

Carleer, A.P., Debeir, O. and Wolff, E., 2005, Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 1285–1294.

Conners, R.W. and Harlow, C.A., 1980, A theoretical comparison of texture algorithms. *IEEE Transactions on Pattern Analyses and Machine Intelligence*, **2**, pp. 204–222.

Evans, C., Jones, R., Svalbe, I. and Berman, M., 2002, Segmenting multispectral Landsat TM images into field units. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, pp. 1054–1064.

Field, D.J., 1987, Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, **4**, pp. 2379–2394.

Hazel, G.G., 2001, Object-level change detection in spectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 553–561.

Hill, P.R., Canagarajah, C.N. and Bull, D.R., 2003, Image segmentation using a texture gradient based watershed transform. *IEEE Transactions on Image Processing*, **12**, pp.1618–1633.

Jain, A.K. and Farrokhnia, F., 1991, Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, **24**, pp. 1167–1186.

Karantzalos, K. and Argialas, D., 2006, Improving edge detection and watershed segmentation with anisotropic diffusion and morphological levellings. *International Journal of Remote Sensing*, **27**, pp. 5427–5434.

Kovesi, P., 1996, Invariant measures of image features from phase information. PhD dissertation, University of Western Australia, Nedlands, Western Australia, Australia.

Meinel, G. and Neubert, M., 2004, A comparison of segmentation programs for high resolution remote sensing data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XXXV**, Part B4, pp. 1097–1102.

Neubert, M., Herold, H. and Meinel, G., 2006, Evaluation of remote sensing image segmentation quality –further results and concepts. In *1st International Conference on Object-based Image Analysis*, 4–5 July 2006, Salzburg, Austria, **XXXVI 4/C42**, CD-ROM (Austria: ISPRS).

Pal, N.R. and Pal, S.K., 1993, A review on image segmentation techniques. *Pattern Recognition*, **26**, pp. 1277–1294.

Pekkarinen, A., 2002, A method for the segmentation of very high spatial resolution images of forested landscapes. *International Journal of Remote Sensing*, **23**, pp. 2817–2836.

Pesaresi, M. and Benediktsson, J.A., 2001, A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 309–320.

Rydberg, A. and Borgefors, G., 2001, Integrated method for boundary delineation of agricultural fields in multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 2514–2520.

Sapiro, G. and Ringach, D.L., 1996, Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, **5**, pp. 1582–1586.

Sarkar, A., Biswas, M.K., Kartikeyan, B., Kumar, V., Majumder, K.L. and Pal, D.K., 2002, A MRF model-based segmentation approach to classification for multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, pp. 1102–1113.

Scheunders, P., 2001, Multivalued image segmentation based on first fundamental form. In *Proceedings of 11th International Conference on Image Analysis and Processing*, 26–28 September 2001, Palermo, Italy (Los Alamitos, CA, USA: IEEE Computer Society), pp. 185–190.

Schiewe, J., 2002, Segmentation of high-resolution remotely sensed data—concepts, applications and problems. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **34**, 380–385.

Teuner, A., Picjler, O. and Hosticka, B.J., 1995, Unsupervised texture segmentation of images using tuned matched Gabor filters. *IEEE Transactions on Image Processing*, **4**, pp. 863–869.

Vincent, L. and Soille, P., 1991, Watershed in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, pp. 583–598.

Zhang, Y., 2002, Problems in the fusion of commercial high-resolution satellite images as well as Landsat 7 images and initial solutions. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **34**, part 4, pp. 236–242.

Zhang, Y.J., 1996, A survey on evaluation methods for image segmentation. *Pattern Recognition*, **29**, pp. 1335–1346.