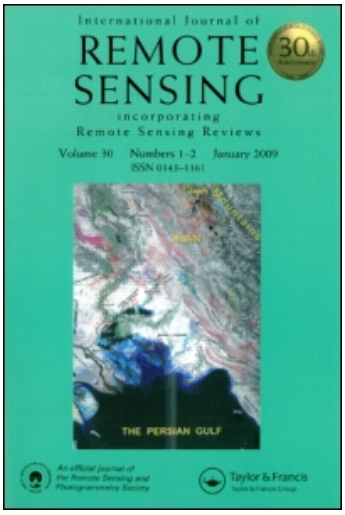

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Settlement extraction in the North China Plain using Landsat and Beijing-1 multispectral data with an improved watershed segmentation algorithm

Lei Wang [a]; Peng Gong [ab]; Qing Ying [a]; Zhenzhong Yang [c]; Xiao Cheng [d];Qiong Ran [e]
[a] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing, PR, China [b] Division of Ecosystem Sciences, University of California, Berkeley, CA, USA [c] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications of Chinese Academy of Sciences, Beijing, PR China [d] College of Global Change and Earth System Science, Beijing Normal University, Beijing, PR China [e] Beijing Landview Mapping Information Technology Co., Beijing, PR China

Online publication date: 30 March 2010

To cite this Article Wang, Lei , Gong, Peng , Ying, Qing , Yang, Zhenzhong , Cheng, Xiao andRan, Qiong(2010) 'Settlement extraction in the North China Plain using Landsat and Beijing-1 multispectral data with an improved watershed segmentation algorithm', International Journal of Remote Sensing, 31: 6, 1411 — 1426
To link to this Article: DOI: 10.1080/01431160903475332
URL: http://dx.doi.org/10.1080/01431160903475332

International Journal of Remote Sensing
Vol. 31, No. 6, 20 March 2010, 1411–1426

Taylor & Francis
Taylor & Francis Group

Settlement extraction in the North China Plain using Landsat and Beijing-1 multispectral data with an improved watershed segmentation algorithm

LEI WANG†, PENG GONG*†‡, QING YING†, ZHENZHONG YANG§, XIAO CHENG¶ and QIONG RAN||

†State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, PR China

‡Division of Ecosystem Sciences, University of California, Berkeley, CA 94720-3114, USA

§State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications of Chinese Academy of Sciences, Beijing 100875, PR China

¶College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, PR China

||Beijing Landview Mapping Information Technology Co., Beijing 100096, PR China

In this paper we present an improved watershed segmentation algorithm for settlement mapping from medium resolution satellite data over plain areas in China. The algorithm can increase the computational efficiency of the fastest reported watershed segmentation algorithm by 30–40%. We apply this method to a selected study area in southern Hebei Province, China. We acquired a Landsat Enhanced Thematic Mapper Plus (ETM+) image over this area in May 2000, two Landsat Thematic Mapper (TM) images in August 2004 and April 2005, and two Beijing-1 satellite images in May 2006 and May 2007. The three types of images have three similar spectral bands (green, red and near-infrared) with similar spatial resolution (30–32 m). Only the red and near-infrared bands were used in image segmentation for settlement area extraction. The extracted settlement results are compared with manual interpretation results by two people. We assumed the human interpretation results are of higher accuracy than the segmentation results. Our results indicated that our settlement area extraction method is effective. With high quality images, the overall accuracies are nearly 94%, the kappa coefficient can be greater than 0.85.

1. Introduction

Knowledge about the size, distribution and change of human settlement areas is a critical component in studies of land use and land use change. Areas developed into settlements are surveyed every year by Statistics Bureaus at the state, provincial and county levels in China. However, this source of data is not generally considered accurate as there is a large portion of developed land not reported for various reasons (e.g. Seto *et al.* 2000). It is desirable to use an independent source of data to obtain spatial temporal information about the dynamics of human settlement areas in China in support of policy and decision makers as well as planners and environmental scientists.

*Corresponding author. Email: gong@irsa.ac.cn

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160903475332

Remote sensing has long been used to meet such purposes (Ridd 2006). However, the application of remote sensing to study human settlements has mostly been devoted to the mapping of various land use types in urban areas (e.g. Jensen and Toll 1982, Forster 1985, Gong and Howarth 1990). In many parts of the world, settlements are scattered in rural areas, and are therefore difficult to map with medium spatial resolution satellite data such as Landsat Thematic Mapper (TM) imagery. In China, the rural population tends to be grouped together in villages of plain areas, which makes it possible to map rural settlements using satellite data with medium spatial resolution.

During the past 30 years, settlement development in China has grown rapidly due to strong economic growth, higher demand for living space and population increase. Urban areas and also rural areas have been expanding in China. This trend has been most observable in cities, townships and villages that are located in relatively flat areas in plains and basins where the primary land use is agriculture. While definitions of rural and urban settlements are mostly determined by their size of population and level of complexity of human activities, in China urban areas and townships are determined administratively with social, economic and political considerations. Thus, rural settlement areas in China are usually seen as having relatively small populations, and occupying a small area in space. The primary activity of rural residents is neither industrial/commercial nor service/administrative. Most urban residents do not undertake agricultural work (agricultural activities in China, in a broader sense, include crop and vegetable growing, grazing and fishing; however, most foresters are treated as non-agricultural workers and included in the urban population). Spatially, rural settlement areas usually have smaller constructions and fewer and narrower streets than urban areas, and are therefore harder to map. More than two thirds of the 1.3 billion people in China live in rural areas, and there are hundreds of thousands of villages that have not been mapped. It is important to monitor the changes of such a huge number of rural settlements in a timely and economically efficient manner. Although it has been possible to acquire images from satellites at the metre/sub-metre level of spatial resolution, it is not feasible to use such imagery to map large areas due to the lack of spatial coverage and funding limitations. Therefore, it is worthwhile to investigate the quality of rural settlement mapping with relatively less expensive satellite data such as the 30-m resolution TM and Enhanced Thematic Mapper Plus (ETM+), or similar data products. TM images have been widely used to map built-up areas in various parts of the world and to monitor changes in those areas (e.g. Seto and Fragkias 2005).

There are a number of ways to map settlement areas. These include field surveying, digitizing settlement boundaries from topographic maps, visual interpretation of aerial and satellite images, and image classification. However, topographic maps are usually out of date in rapidly changing areas. This is true for both China and countries in the developed world. Digitizing maps and visual interpretation of images are time consuming for large areas. Therefore, it is necessary to develop automatic or semi-automatic techniques for settlement mapping (Stuckens *et al.* 2000, Zha *et al.* 2003). Many automatic/semi-automatic methods have been developed based on image classification (Gong 2006). Most existing classification algorithms employ per-pixel classifiers whose classification decision is based either solely on data taken at each particular pixel location or on data describing the neighbourhood of the pixel to be classified (Gong and Howarth 1992, Gong *et al.* 1992). Recently, image segmentation techniques have been widely used in image classification (Benz *et al.* 2004, Yu *et al.* 2006). With this type of algorithm, image segmentation is the first step: high resolution images (1–2 m resolution) are segmented to form homogeneous regions, often called objects.

Classification is then made of each object as a whole based on various features extracted from each object. Spectral, textural, size and shape data are usually used as features in object classification. This type of object-based image classification has also been used in mapping urban areas using very high-resolution imagery (Shackelford and Davis 2003, Al-Khudhairy *et al.* 2005, Walker and Briggs 2007). They used IKONOS images or aerial photographic images. In fact, it is not necessary or efficient to map settlement areas in rural areas with very high spatial resolution imagery (0.5–4 m level) partly due to some of the reasons discussed earlier and partly because settlement areas are highly complicated in pattern and covered by impervious road surfaces, water bodies, roof tops, lawns and trees. All these cover types can be segmented into many individual objects in high spatial resolution images, and then effort is needed to re-group these small objects. With medium-resolution imagery, this re-grouping step could be largely saved because the variability of distinct cover types would have been smoothed out when the spatial resolution is coarser than 30 m in China. Therefore, it is reasonable to evaluate object-based image classification algorithms to extract settlement areas with medium resolution images.

The purpose of this paper is two-fold. Firstly, we present an improved watershed method and evaluate the effectiveness of a watershed segmentation algorithm for settlement mapping. Secondly, we compare the performance of Landsat TM and Beijing-1 imagery in settlement mapping using the watershed segmentation algorithm. Watershed segmentation algorithms have not been widely used in classification problems with remotely sensed data. Effective tools isolate or detect targets that are surrounded by relatively homogeneous backgrounds (Wang *et al.* 2004). We choose this technique and further improved its implementation because we believe this type of algorithm would work well in extractions of settlements with medium resolution images when the background has high contrast with settlements, as is the case in plain and basin areas of China. In the rest of this paper we first describe the watershed segmentation algorithm implemented here, and then we illustrate its application to map settlement areas with different satellite images acquired over an area in the North China Plain. This is followed by a presentation of some experimental results with some discussions of their implications. Finally we draw some conclusions.

2. Settlement extraction method

Watershed segmentation is a method widely used as a preprocessing step in image processing, for subsequent detailed analysis of images. This kind of algorithm has a number of advantages: it has a linear running time; its process is easy to understand; and it produces closed edges, which is a great advantage over most edge detection methods. This makes it possible to extract settlement areas that are usually distinctive from their surroundings. The drawback of watershed segmentation is its sensitivity to noise and small changes in grey levels. This property often leads to over-segmentation, i.e. too many partitions are produced by the watershed segmentation approach. To overcome the over-segmentation problem, one could either carry out some preprocessing before running the watershed segmentation approach, do some post-processing after the watershed segmentation approach, or use both techniques (Vincent and Soille 1991, Shafarenko *et al.* 1997, Haris *et al.* 1998, Bieniek and Moga 2000, Malpica *et al.* 2003, Oapos-Callaghan and Bull 2005, Sun *et al.* 2005, Osma-Ruiz *et al.* 2007). In order to keep edges at their right positions, one usually uses the gradient image instead of the original image (Sun *et al.* 2005).

In our implementation of the watershed segmentation algorithm for settlement mapping, we divide the processing into four successive steps: Normalized Difference Vegetation Index (NDVI) image generation, morphological gradient calculation, watershed segmentation and segment merging. In the following sections we explain each of these steps.

2.1 *Normalized Difference Vegetation Index image generation*

Settlement areas in the North China Plain are mostly surrounded by crop fields, and they usually have large distinctions with built-up lands during the growing season. There are many indices for obtaining grey-level images, such as NDVI and the normalized difference building index (NDBI) (Zha *et al.* 2003). NDVI uses the red (RED) and near-infrared (NIR) bands, which are commonly available in most remotely sensed data. We chose NDVI to enhance the contrast between vegetation and built-up areas. In this study, we simply applied NDVI = (NIR − RED)/(NIR + RED + 0.01) to the red and near-infrared bands of the TM, ETM+ and Beijing-1 images. Since we do not compare NDVI changes over time, the original data were not pre-processed to convert digital numbers into radiance. Because the subsequent watershed segmentation approach is applied to each individual NDVI image, we consider this raw calculation of NDVI appropriate.

2.2 *Morphological gradient calculation*

Since the contrast between vegetation and built-up areas is high in NDVI images, we simply choose the morphological gradient operator. The morphological operator calculates the gradient image according to the following equation:

$$G(f) = (f \oplus B) - (f \ominus B), \tag{1}$$

where f is the image to be calculated, the size of B is 3×3, and $\oplus$ and $\ominus$ are dilation and erosion operators.

In order to reduce local minima, we use a reconstruction method. A closing is applied to $G(f)$ with a 3×3 structural element, in order to remove small local minima, denoted by $C(G(f))$. After the closing operation of $G(f)$, a reconstruction operator is applied by eroding $C(G(f))$ from $C(G(f)) + h$, in order to remove the local minima with contrasts less than h (see Wang (1997) for more detail). The constant h is used to control the number of segmentation regions. As h increases, the number of regions produced decreases. The reconstruction by erosion fills all of the local minima where the contrast is lower than h, irrespective of their absolute values.

2.3 *Improved watershed segmentation*

There are two major kinds of watershed segmentation methods. The first kind is based on immersion simulation (Vincent and Soille 1991, Gao *et al*. 2006); however, this kind of algorithm does not partition the watershed pixels (Sun *et al*. 2005). The second kind does not have that problem. These kinds of algorithms are developed with all pixels in the image transformed to a unique basin, and no boundary pixels are produced. Most of these algorithms are based on rainfall simulation (Bieniek and Moga 2000, Sun *et al.* 2005, Osma-Ruiz *et al.* 2007).

The watershed segmentation method used in our experiments is based on the second kind of algorithms. Unlike the method described in Osma-Ruiz *et al.* (2007),

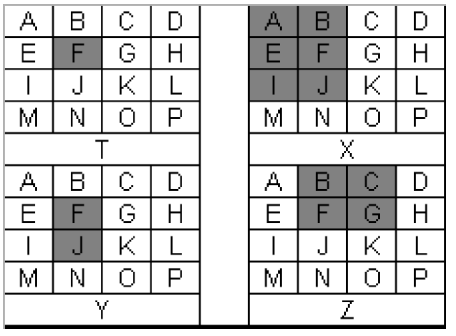

Figure 1. An illustration of the obtainable information obtained before processing the pixel. T refers to the pixel being transformed; and X, Y, Z refer to the smallest pixel of those in shade, respectively.

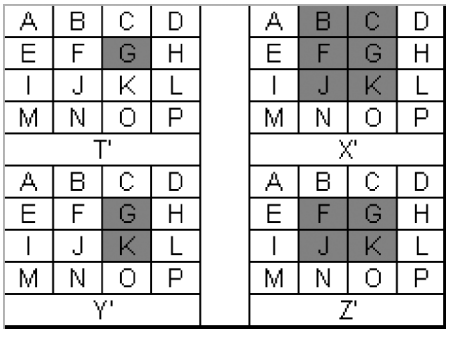

Figure 2. An illustration of the obtainable information after processing the pixel. T′ refers to the pixel being transformed; X′, Y′, Z′ refer to the smallest pixel of those in shade, respectively.

we improve the speed of the algorithm by taking advantage of the pixels that have already been formerly processed/scanned. This requires scanning the image in order, from left to right and from top to bottom.

The first scanning is as described in Bieniek and Moga (2000) (BM), and solves pixels with a lower neighbour. We have made some improvements over the BM algorithm. We describe them in detail by using a 4×4 matrix, as shown in figure 1.

We transformed these pixels in the matrix in a pre-defined order (top to bottom, left to right). In figure 1, X, Y and Z can be obtained from a previous transformation. We can get X and Y when dealing with pixel E. We can also get Z when dealing with pixel B. Dealing with F requires X, Y and Z be updated. The new X, Y and Z are shown in figure 2, where an array with the size of the image width is required.

We use X′, Y′ and Z′ to store the information contained in X, Y and Z. First, we compare pixel G with pixel K where G and K are pixels in the 4×4 matrix in figure 1, and the smaller one is stored in Y. Second, we compare Y′ with Y, and store the smaller one in Z. Third, we compare Z′ with Z, and store the smaller one in X. Finally, we compare X′ with X, and the smaller value is the smallest value associated with pixel F and its surrounding pixels in the 4×4 matrix. If the smallest one is not pixel F, let pixel F point to the smallest one. Otherwise, assign pixel F an unsolved mask. Here, it should be noted that if pixel F has neighbours with the same grey level, and none of the

remaining neighbours of pixel F has a smaller grey level, pixel F belongs to a plateau. Assign pixel F an unsolved mask.

In our method, each pixel is transformed with at most four comparisons in the 4×4 matrix. Pixels in the right-most column and the bottom row need no more than one comparison instead of neighbour search, which needs eight comparisons. The above two are superior compared to the BM algorithm.

The second scanning solves all the pixels which have no neighbours with lower values. The process is similar to other rainfall-based watershed method (see Bieniek and Moga (2000) for more detail).

The third scanning assigns all pixels a label following the direction they point to. All pixels are then transformed into different basins.

2.4 *Segment merging and settlement area extraction*

The use of reconstruction in the gradient calculation reduces the local minima. As a result it reduces the number of partitions in the watershed segmentation. However, the image is still over-segmented. Here, we use a merging operator described in Malpica *et al.* (2003), illustrated in figure 3 to merge over-segmented objects.

The flowchart in figure 3 is self-explanatory. T and $n_{\min}$ are defined by the user. After the merging step, the partitions of the image are merged into a meaningful one.

3. Study site and data

The North China Plain, with an area of approximately 300 000 km^2, is the second largest plain in China. Approximately 437 million people lived in the North China Plain in 2000, accounting for 34.8% of China's total population. Our study site (37.69° N, 114.80° E) is located to the south of Shijiazhuang City and to the north of Xingtai City, in the southern part of Hebei Province.

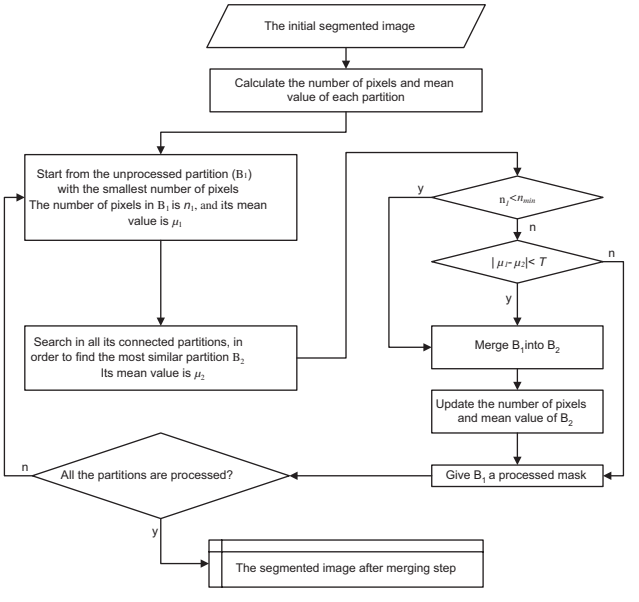

Figure 3. Flowchart of the segment merging process.

Table 1. Some of the characteristics of the Beijing-1 satellite data.

	High resolution sensor	Medium resolution sensor
Spatial resolution	4 m	32 m
Spectral range	500–800 nm	Green 520–620 nm Red 630–690 nm NIR 760–900 nm
Swath	24 km	600 km
Revisit capability	5–7 days	3–5 days

In China, a small satellite named Beijing-1 was launched in October 2005, carrying onboard two CCD pushbroom multispectral cameras and one panchromatic camera (table 1). It is one of the Disaster Monitoring Constellation satellites manufactured in the UK (Surrey Satellite Technology Co., Ltd, Surrey, UK). The spatial resolution of its multispectral cameras and the three bands in green, red, and near-infrared (NIR) are similar to the three corresponding Landsat TM bands. It is useful to compare the performance of the two types of sensors in mapping both urban and rural settlement areas.

As an experiment, the study site is set to a size of 1000 × 1000 pixels. Since the site is flat, no radiometric correction is performed. For the purpose of comparison over different times in the same area, a geometric correction is applied to all images according to topographic maps. The 32 m Beijing-1 data were resampled to 30 m. Registration errors are less than one pixel.

The data we use are: one ETM+ image acquired on 7 May 2000; two TM images acquired on 30 August 2004 and 27 April 2005; and two Beijing-1 images acquired on 14 May 2006 and 16 May 2007. The Beijing-1 image in 2006 is blurry and has the poorest quality due to the fact that it was acquired while the receiving station was still in testing mode.

Subimages with a size of 1000 × 1000 pixels were clipped from each image (figure 4). The selected area has two county headquarters of Ningjin Xian (located on the upper-left of the clipped image) and Zhao Xian (located on the lower-right of the clipped image). It also contains several other towns and villages of four counties (Ningjin Xian, Zhao Xian, Gaoyi Xian and Baixiang Xian). The majority of the settlements in the subimages are villages. In this study area, crop fields are covered by green vegetations in May and in August. Therefore, settlement areas have a strong contrast to the vegetation fields making them suitable for applications of the watershed segmentation technique.

From figure 4 it can be seen that although the nominal spatial resolution for Beijing-1 imagery is only two metres poorer than that of the ETM+ image, the Beijing-1 image appears to be quite blurred. We experimented with post-processing based on the modulation transfer function (MTF). It can substantially improve the clarity of the Beijing-1 images but the final visual effect is still inferior to TM and ETM+ images. However, Beijing 1 images in figure 4 were not enhanced with MTF.

4. Results

We experimented with different parameters in the watershed segmentation algorithm in order to obtain most satisfactory results. Figure 5 shows the NDVI calculation of the different images. In this figure, the settlement areas are in lower grey levels, while the vegetation is in higher grey levels.

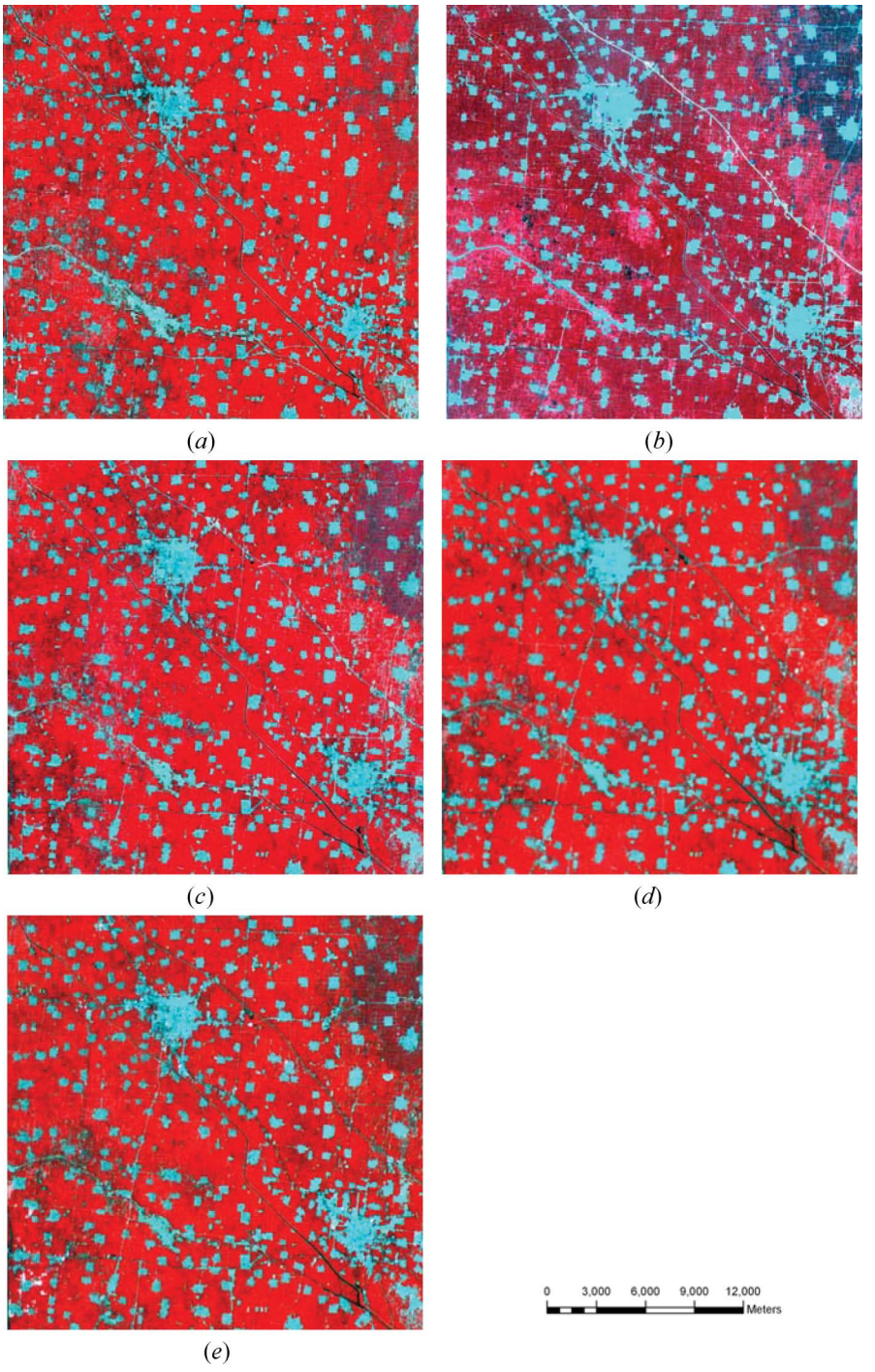

Figure 4. Clipped images from satellite data obtained at different times. NIR, red and green bands assigned to red, green and blue colour guns. Images are clipped from (*a*) ETM+ false-colour image on 7 May 2000; (*b*) TM false-colour image on 30 August 2004; (*c*) TM false-colour image on 27 April 2005; (*d*) Beijing-1 false-colour image on 14 May 2006; (*e*) Beijing-1 false-colour image on 16 May 2007.

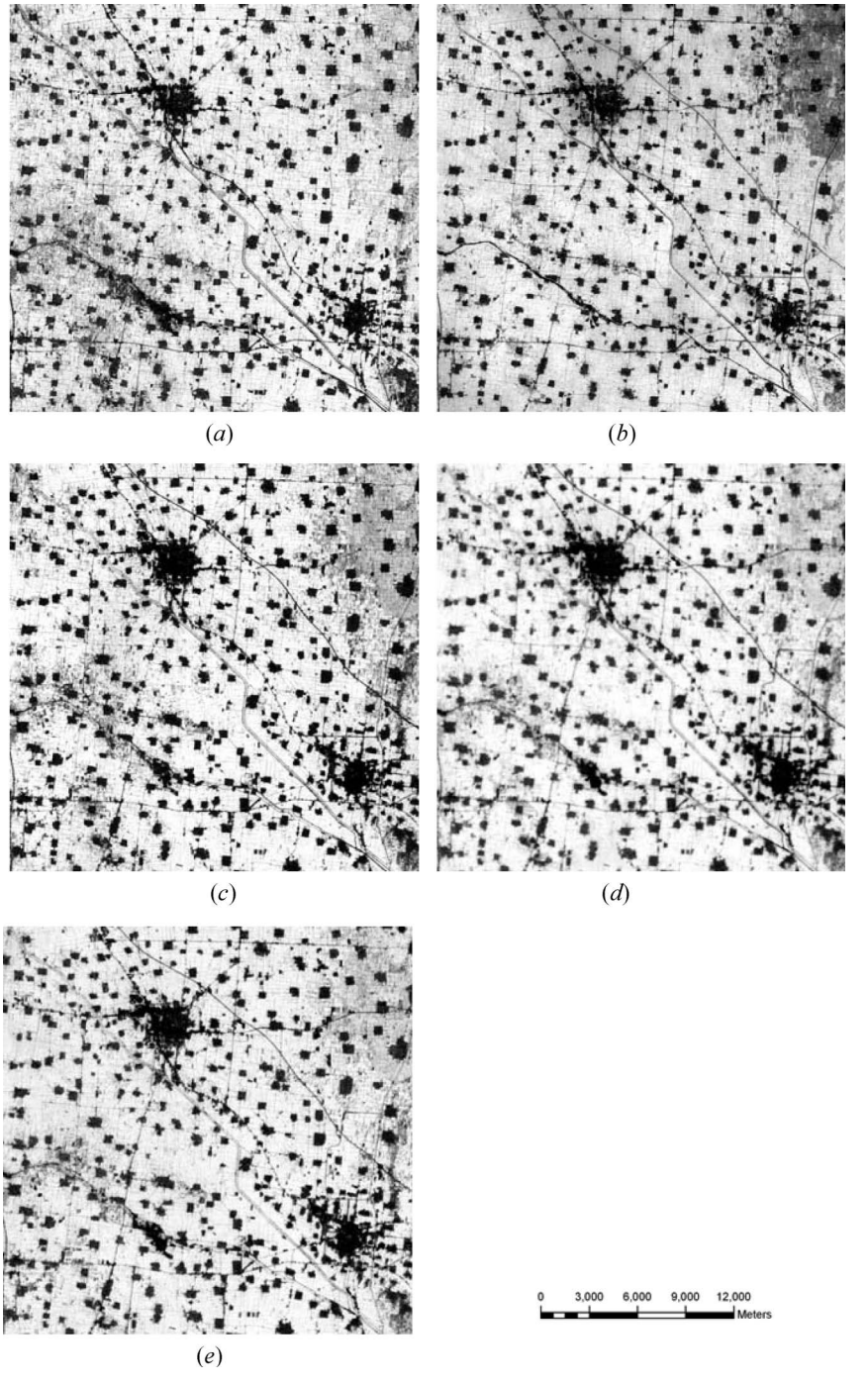

Figure 5. NDVI images correlated to the clipped images. (*a*), (*b*), (*c*), (*d*) and (*e*) are responding to (*a*), (*b*), (*c*), (*d*) and (*e*) in figure 4, respectively.

Table 2. Comparison of two watershed segmentation methods.

Test images		Running time			Catchment basins segmented	
Date	Type	Algorithm OGSG (ms)	New algorithm (ms)	Improvement (%)	OGSG algorithm	New algorithm
7 May 2000	ETM+	371	232	37.5	59145	59145
30 August 2004	TM	412	241	41.5	41295	41295
27 April 2005	TM	422	243	42.4	37857	37857
16 May 2006	BJ-1	435	286	34.3	15575	15575
14 May 2007	BJ-1	417	251	39.8	32101	32101

In calculating the gradient image step, the local minima with a contrast smaller than 8 are filled up. During the watershed segmentation step, we compared the new watershed segmentation method with the existing fastest algorithm (OGSG method; Osma-Ruiz *et al.* 2007) with the five gradient images. The comparison results are shown in table 2.

At the segment merging step, we tested the results with different values of T from 30 to 90, and chose 70 as the optimal threshold, except the image acquired on 30 August 2004 (its optimal threshold was 40).

In order to evaluate the performance of the watershed segmentation algorithm, we undertook a visual image interpretation by manually tracing the boundaries of settlement areas contained in the area of interest on each subimage by two interpreters. Figure 6 shows the automatic extraction results overlaid with one set of the manually interpreted results.

We then compared the results between automatic extraction results and manually interpreted results. We used the areas of interest (white areas on figure 6(*f*)) to carry out a pixel by pixel comparison. Tables 3 and 4 are the comparison results.

Confusion matrices, overall accuracies and kappa values are presented in tables 3 and 4. From these tables, it can be seen that the agreements between segmentation results and interpretation results are all greater than 90%. Specifically, for the first interpreter, the overall accuracies are 94.0%, 94.3%, 93.5%, 93.6% and 94.1% respectively for 2000, 2004, 2005, 2006 and 2007 images. For the second interpreter, the overall accuracies are 92.9%, 93.1%, 94.1%, 93.1% and 93.7% respectively for those corresponding 5 years. Table 5 shows some comparisons between the two interpreters. The overall accuracies are 93.7%, 94.8%, 94.9%, 94.4% and 91.1%. From above, the results obtained by the segmentation algorithm agree well with the interpretation results.

5. Discussion

In table 2, both methods produced the same number of basins with the images. This means that the two methods achieved the same accuracy in function. However, the method proposed in this research is faster than the OGSG method. With five different images with the size of 1000 × 1000, the improvements are 37.5%, 41.5%, 42.4%, 34.3% and 39.8%, respectively. This is helpful in practice when datasets are getting bigger.

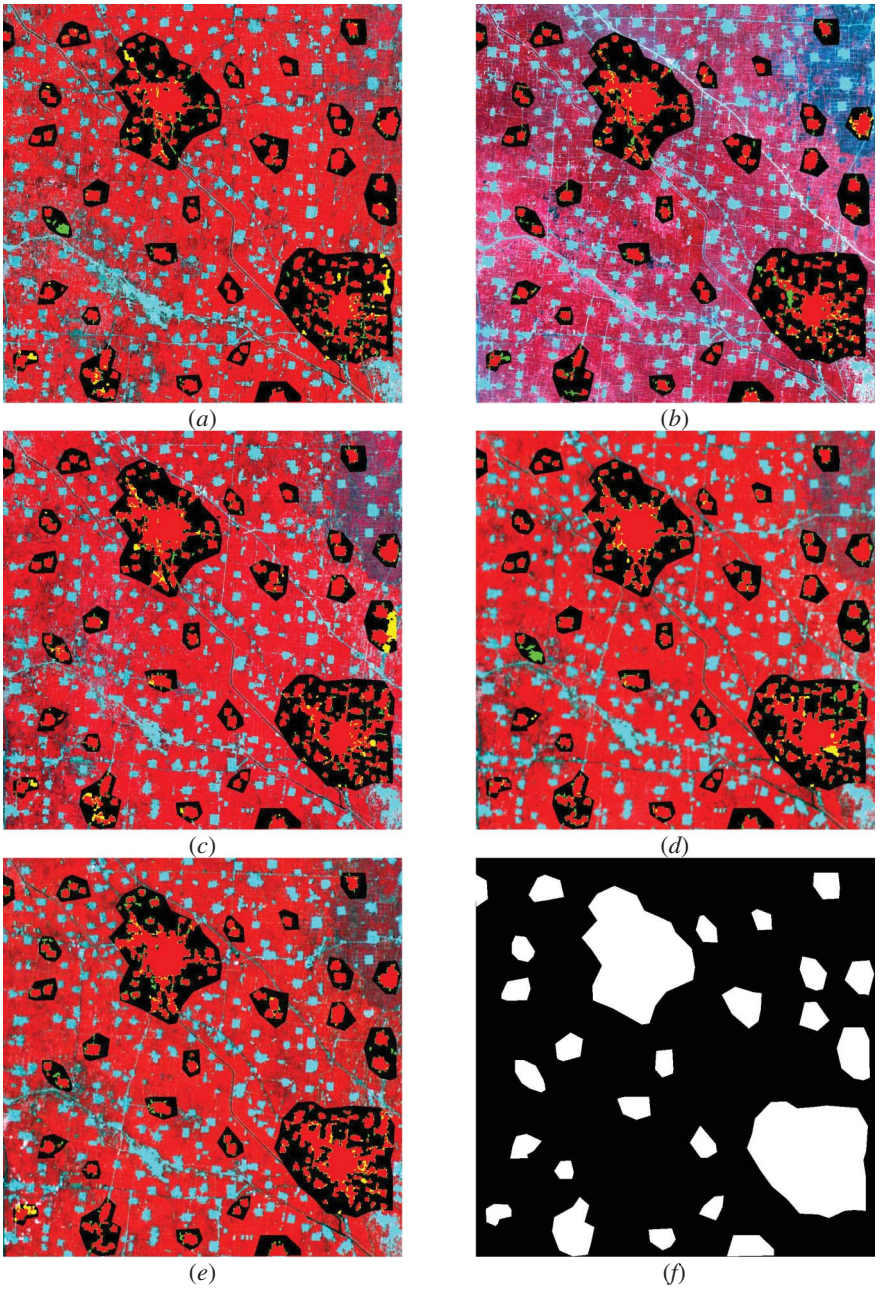

Figure 6. Final settlement extraction results and the area of interest. (*f*) the white areas are manually interpreted sites, while (*a*), (*b*), (*c*), (*d*) and (*e*) are the comparison results between automatic extraction results and manually interpreted results on the area of interest overlaid on the images of figure 4, respectively. There are four colors in the area of interest: the red pixel (automatic extraction result: settlement area; manually interpreted result: settlement area); the yellow pixel (automatic extraction result: settlement area; manually interpreted result: not settlement area); the black pixel (automatic extraction result: not settlement area; manually interpreted result: settlement area); the black pixel (automatic extraction result: not settlement area; manually interpreted result: not settlement area); (*a*), (*b*), (*c*), (*d*) and (*e*) are responding to (*a*), (*b*), (*c*), (*d*) and (*e*) in figure 4, respectively.

Table 3. Segmentation results versus interpretation results by the first interpreter (Person 1).

Image	Segmentation results	Interpreted by Person 1			Overall (%)	Kappa
		Settlement	Non settlement	Total		
7 May 2000 ETM+	Settlement	60539	7098	67637	94	0.858
	Non settlement	6168	148282	154450		
	Total	66707	155380	222087		
30 August 2004 TM	Settlement	56816	7239	64055	94.3	0.86
	Non settlement	5377	152655	158032		
	Total	62193	159894	222087		
27 April 2005 TM	Settlement	68952	6265	75217	93.5	0.856
	Non settlement	8125	138755	146880		
	Total	77077	145020	222097		
16 May 2006 BJ-1	Settlement	69338	9894	79232	93.6	0.858
	Non settlement	4349	138506	142855		
	Total	73687	148400	222087		
14 May 2007 BJ-1	Settlement	69149	7430	76579	94.1	0.869
	Non settlement	5679	139829	145508		
	Total	74828	147259	222087		

Table 4. Segmentation results versus interpretation results by the second interpreter (Person 2).

Image	Segmentation results	Interpreted by Person 2			Overall (%)	Kappa
		Settlement	Non settlement	Total		
7 May 2000 ETM+	Settlement	59100	8224	59100	92.9	0.831
	Non settlement	7607	147156	7607		
	Total	66707	155380	66707		
30 August 2004 TM	Settlement	58770	11910	70680	93.1	0.836
	Non settlement	3423	147984	151407		
	Total	62193	159894	222087		
27 April 2005 TM	Settlement	69801	5922	75723	94.1	0.868
	Non settlement	7276	139088	146364		
	Total	77077	145010	222087		
16 May 2006 BJ-1	Settlement	69893	11550	81443	93.1	0.848
	Non settlement	3794	136850	140644		
	Total	73687	148400	222087		
14 May 2007 BJ-1	Settlement	70541	9694	80235	93.7	0.862
	Non settlement	4287	137565	141852		
	Total	74828	147259	222087		

In table 3, the manual results were carefully interpreted by the first interpreter. The interpreted settlement areas are 67637, 64055, 75207, 79232 and 76579 pixels respectively for the five images in figure 4. Since the five images are ordered by time, there are two decreases in B and E. The image B was acquired on 30 August, while all others are acquired in May. In the North China Plain, vegetation grows better in late August

Table 5. Interpretation results by the first interpreter (Person 1) versus interpretation results by the second interpreter (Person 2).

Image	Interpreted by Person 1	Interpreted by Person 2			Overall (%)	Kappa
		Settlement	Non settlement	Total		
7 May 2000 ETM+	Settlement	60437	6887	67324	93.7	0.85
	Non settlement	7200	147563	154763		
	Total	67637	154450	222087		
30 August 2004 TM	Settlement	61625	9055	70680	94.8	0.878
	Non settlement	2430	148977	151407		
	Total	64055	158032	222087		
27 April 2005 TM	Settlement	69808	5915	75723	94.9	0.886
	Non settlement	5399	140965	146364		
	Total	75207	146880	222087		
16 May 2006 BJ-1	Settlement	74153	7290	81443	94.4	0.879
	Non settlement	5079	135565	140644		
	Total	79232	142855	222087		
14 May 2007 BJ-1	Settlement	71888	8347	80235	94.1	0.872
	Non settlement	4691	137161	141852		
	Total	76579	145508	222087		

than in May. The interpretation is based on the assumption that the land was covered by vegetation except roads and settlement areas. In image B, settlement areas should be more distinctive from non-settlement areas. The images acquired in May might contain open fields near settlement areas that can be mistakenly included into settlement areas. This partly contributed to the lower number of pixels in image B than in image A as interpreted by the first interpreter. However, the smaller number of pixels for the settlement in 2007 in image E cannot be explained with similar reasons. This could be caused by the poorer blurry image quality in image D, leading to an over--estimation of settlement areas. Nevertheless, the large discrepancy between the results obtained from the 2004 and 2000 images could not be totally explained by the seasonal differences. In addition, the discrepancy between the results obtained from the 2006 image and 2007 image could also not be totally explained by the image quality. Manual interpretation error/inconsistencies might have played a role as well.

In table 4, the settlement area results interpreted manually by the second interpreter are 67324, 70680, 75723, 81443 and 80235 pixels respectively for the five images in figure 4. Interestingly, this interpreter obtained more settlement areas for 2004, but still obtained a lower number of pixels for settlement areas in 2007 than 2006. A close examination between the two years shows relatively little change in settlement growth. The decrease in 2007 compared to 2006 might be caused by error of interpretation. The difference is only approximately 1.5%.

When we compare the interpretation results between the two interpreters, their levels of overall agreement range from 93.7% for 2000 to 94.9% for 2005. If we assume interpretation errors are random, these levels of agreement indicate a less than 4% inconsistency between the two interpreters. The image segmentation results are close to either set of the image interpretation results. If they are within 6–7% in error, the segmentation results should contain less than 10% of overall mapping errors.

In this study, we are more interested in how settlement areas are extracted from different satellite images with the watershed segmentation technique. In table 3(4):

- 60539(59100) pixels out of 67637(67324) pixels have been classified as settlement areas, leading to a correct classification rate of 89.5%(87.8%) on image A in figure 4;
- 56816(58770) pixels out of 64055(70680) pixels have been classified as settlement areas, leading to a correct classification rate of 88.7%(83.1%) on image B in figure 4;
- 68952(69801) pixels out of 75207(75723) pixels have been classified as settlement areas, leading to a correct classification rate of 91.7%(92.2%) on image C in figure 4.
- 69338(69893) pixels out of 79232 (81443) pixels have been classified as settlement areas, leading to a correct classification rate of 87.5%(85.8%) on image D in figure 4;
- 69149(70541) pixels out of 76579(80231) pixels have been classified as settlement areas, leading to a correct classification rate of 90.3%(87.9%) on image E in figure 4.

The segmented settlement represents over 88% of the total settlement area on average for the five images with two interpreters. For the first interpreter, the lowest agreement is with image D (a heavily blurred image). For the second interpreter, the lowest agreement is with image B. The single class accuracy inconsistency for settlement extraction by the first interpreter is less than 5%. For the second interpreter it is less than 9%. Both of them are acceptable. The high rate of image E shows that with a high quality of data, we can use Beijing-1 satellite data to extract settlement areas. The average accuracy difference between Beijing-1 and Landsat images is less than 1%. Both datasets are acceptable for mapping settlement areas with the watershed image segmentation algorithm.

The level of error may also come from human interpretation. While the blurred image might be a preferable one for some automatic algorithms that are sensitive to local variation and noise, it is certainly not preferred by human interpreters. Therefore, the error level involved in manual interpretation of the Beijing-1 image might also be larger. Therefore, determining whether or not the segmented results for the Beijing-1 image are poorer than other images requires more assessment with more accurate sources of information, perhaps from IKONOS or Quickbird images.

One of the advantages of Landsat TM or ETM+ sensors is their long history. Launched in 1982, Landsat TM could be a valuable source of information for us to study settlement changes in China because major settlement development started in the early 1980s. Unfortunately, due to the malfunctioning of the scan line corrector (SLC) of Landsat 7 beginning on 31 May 2003, the usable swath range of Landsat ETM+ is limited to the central portion of the image. The 600-km wide swath in the multispectral cameras of the Beijing-1 satellite has the advantage of obtaining more frequent global coverage compared to the 185 km swath of Landsat TM (and narrower than 185 km after the SLC is turned off). Beijing-1 satellite data can be considered as a complement to Landsat ETM+ data. The wider swath of Beijing-1 is also more helpful in areas where clouds are a major obstacle for clear-day image acquisition.

One of the advantages of the watershed segmentation algorithm is that no training is required and the separation of settlement and crop field over the study area is quite easy provided the images are acquired in the season of high vegetation coverage. For the North China Plain, April to May and August to September are the most suitable times.

Further work is needed to select more areas over the North China Plain to test the performances of the algorithm and the medium resolution data such as TM and Beijing-1 multispectral imagery. Other settlement extraction methods can be compared with the method used here. The two-band requirement for the watershed

segmentation algorithm is both an advantage and a disadvantage. Other classification algorithms can take advantage of information in more spectral bands. Future work would include further assessment of algorithms for selecting the best band combinations for improved algorithm performance when contrast between settlements and their surroundings are not large enough. It would be useful to acquire some high spatial resolution satellite data (e.g. IKONOS and Quickbird, or aerial photography data) as references for compassion with the mapping results obtained from the medium resolution images.

6. Conclusions

We have assumed that manual interpretation of settlements from medium resolution satellite data has a higher accuracy than automatic methods, so we used manual interpretation results to assess the performance of an improved watershed segmentation algorithm in settlement extraction with TM, ETM+ and Beijing-1 data. The watershed image segmentation algorithm can on average reach better than 88% accuracies for these types of images in settlement area extraction. This is rather encouraging and acceptable to map settlement areas with these types of data. Although there are some quality issues with a Beijing-1 image (due to image blurring) used in this study, because it was acquired during the test and adjustment period of the satellite system, the result is acceptable. Beijing-1 images can be used as substitute or as a continuation data source to Landsat TM and ETM+ for settlement mapping in the future.

Acknowledgement
This research was partially supported by a major project grant from the National Natural Science Foundation of China (30590370), an NSFC grant (40671125), two grants from the Ministry of Science and Technology (number: 2006AA12Z112; 2006BAJ01B02) and a hundred people grant from the Chinese Academy of Sciences.

References

Al-Khudhairy, D.H.A., Caravaggi, I. and Glada, S., 2005, Structural damage assessments from Ikonos data using change detection, object-oriented segmentation, and classification techniques. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 825–837.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M., 2004, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp. 239–258.

Bieniek, A. and Moga, A., 2000, An efficient watershed algorithm based on connected components. *Pattern Recognition*, **33**, pp. 907–916.

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

Gao, H., Lin, W., Xue, P. and Siu, W.C., 2006, Marker-based image segmentation relying on disjoint set union. *Signal Processing: Image Communication*, **21**, pp. 100–112.

Gong, P., 2006, Information extraction. In *Manual of Remote Sensing: Remote Sensing of Human Settlements*, 3rd edn, M. Ridd and A. Renz (Eds), pp. 275–334 (Bethesda, MD: American Society for Photogrammetry and Remote Sensing).

Gong, P. and Howarth, P.J., 1990, The use of structural information for improving land-cover classification accuracies at the rural-urban fringe. *Photogrammetric Engineering and Remote Sensing*, **56**, pp. 67–73

Gong, P. and Howarth, P.J., 1992, Frequency-based contextual classification and grey-level vector reduction for land-use identification. *Photogrammetric Engineering and Remote Sensing*, **58**, pp. 423–437.

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

Haris, K., Efstratiadis, S.N., Maglaveras, N. and Katsaggelos, A.K., 1998, Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, **7**, pp. 1684–1699.

Jensen, J.R. and Toll, D.L., 1982, Detecting residential land-use development at the urban fringe. *Photogrammetric Engineering and Remote Sensing*, **48**, pp. 629–643.

Malpica, N., Ortuno, J.E., and Santos, A., 2003, A multichannel watershed-based algorithm for supervised texture segmentation. *Pattern Recognition Letters*, **24**, pp. 1545–1554.

Oapos-Callaghan, R.J. and Bull, D.R., 2005, Combined morphological-spectral unsupervised image segmentation. *IEEE Transactions on Image Processing*, **14**, pp. 49–62.

Osma-Ruiz, V., Godino-Llorente, J.I., Sanchez-Lechón, N. and Gómez-Vilda, P., 2007, An improved watershed algorithm based on efficient computation of shortest paths. *Pattern Recognition*, **40**, pp. 1078–1090.

Ridd, M. (Ed.), 2006, *Manual of Remote Sensing: Remote Sensing of Human Settlements*, 3rd edn, A. Renz (Ed.) (Bethesda, MD: American Society for Photogrammetry and Remote Sensing).

Seto, K.C. and Fragkias, M., 2005, Quantifying spatiotemporal patterns of urban land use changes with four cities of China with time series landscape metrics. *Landscape Ecology*, **20**, pp. 871–888.

Seto, K.C., Kaufimann, R.K. and Woodcock, C.E., 2000, Landsat reveals China's farmland reserves, but they're vanishing fast. *Nature*, **406**, pp. 121.

Shackelford, A.K. and Davis, C.H., 2003, A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 2354–2363.

Shafarenko, L., Petrou, M., and Kittler, J., 1997, Automatic watershed segmentation of randomly textured color images. *IEEE transactions on Image Processing*, **6**, pp. 1530–1544.

Stuckens, J., Coppin, P.R. and Bauer, M.E., 2000, Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment*, **71**, pp. 282–296.

Sun, H., Yang, J. and Ren, M., 2005, A fast watershed algorithm based on chain code and its application in image segmentation. *Pattern Recognition Letters*, **26**, pp. 1266–1274.

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

Walker, J.S. and Briggs, J.M., 2007, An object-oriented approach to urban forest mapping in Phoenix. *Photogrammetric Engineering and Remote Sensing*, **73**, pp. 577–583.

Wang, D., 1997, A multiscale gradient algorithm for image segmentation using watershelds. *Pattern Recognition*, **30**, pp. 2043–2052.

Wang, L., Gong, P. and Biging, G.S., 2004, Individual tree crown delineation and treetop detection in high spatial resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 351–367.

Yu, Q., Gong, P., Clinton, N., Biging, G. and Schirokauer, S., 2006, Object-based detailed vegetation mapping using high spatial resolution imagery. *Photogrammetric Engineering and Remote Sensing*, **72**, pp. 799–811.

Zha, Y., Gao, J. and Ni, S., 2003, Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, **24**, pp. 583–594.