2009 Urban Remote Sensing Joint Event

Super-resolution mapping of urban buildings with remotely sensed imagery based on prior shape information

Feng LING
Institute of Geodesy and Geophysics
Chinese Academy of Sciences
Wuhan, China
lingf@asch.whigg.ac.cn

Bo FU
School of Electrical and Electric Engineering
Hubei University of Technology
Wuhan, China
fubofanxx@yahoo.com.cn

***Abstract*—Urban building extraction is an important task of remotely sensed imagery analysis. Many automatic or semi-automatic algorithms have been proposed to address this problem using various remote sensing sources. Although very high spatial resolution remotely sensed imagery has been used in this field, the large number of mixed pixels that currently still exist in these images makes the extracted urban buildings inaccurate, especially with respect to determining building boundaries. Super-resolution mapping is a promising technology that will improve the spatial resolution of land cover mapping with remotely sensed imagery. This technology uses the fraction maps derived with soft classifications as input and converts them into high resolution land cover maps based on the land cover spatial pattern, which is often described with the maximum spatial dependence principle. Although previous research proved the effectiveness of this principle, it is still not suitable for some special land classes, especially those of man-made objects. In this study, we revised the normal spatial dependence principle to incorporate the prior shape information of urban buildings to make the super-resolution mapping technology more suitable for urban building extraction. The proposed algorithm was evaluated with several simulated images. Our results show that the proposed method can obtain more accurate maps than those produced by the standard super-resolution mapping method. Incorporating more specific prior information improves the performance of super-resolution mapping with remotely sensed imagery.**

I. INTRODUCTION

Urban building extraction from remotely sensed data has been an active research field for many years because of its use in various applications, including map updating, urban planning, and land use analyses. Many kinds of remotely sensed data sources, including multispectral imagery, synthetic aperture radar (SAR) imaging, and light detection and ranging (LIDAR) remote sensing, have been used to extract urban buildings using different automatic or semi-automatic algorithms [1-5]. With multispectral remotely sensed imagery, urban buildings are often extracted through image classification or segment technologies at the pixel scale. Although very high spatial resolution remotely sensed imagery has been used previously in this field, mixed pixels are still an important issue especially for those images with coarse spatial resolution. These mixed pixels make extracted urban buildings inaccurate with respect to building boundaries and must therefore be carefully resolved to improve the quality of urban building extraction.

Soft classification technology [6] provides a useful method that can be used to address the problem of mixed pixels in remotely sensed imagery. Unlike traditional hard classification technologies, soft classification methods do not assign mixed pixels as a single land cover class but instead predict the proportional cover of each land cover class within each mixed pixel. Many soft classification methods such as linear spectral mixture modeling and fuzzy c-means classifiers [7, 8] have been proposed. The fraction maps derived from the soft classification technology could provide more useful land cover information as compared with that derived from hard classification technologies. However, the spatial distribution of each class in these mixed pixels could not be determined. Only the building areas could be estimated more precisely using soft-classification technologies, whereas the boundaries of the urban buildings could not be determined.

Super-resolution mapping technology provides a promising method to predict the spatial land cover distribution within mixed pixels [9]. In general, super-resolution mapping can be considered as the post procession of soft classification, which is a technique that transforms the fraction maps derived from soft classification into a finer scaled hard classification map, i.e., soft classification yields fraction values, whereas super-resolution mapping techniques use these fraction values as input to retrieve an appropriate spatial location for these land cover fractions [10].

The purpose of this research is to investigate the performance of super-resolution mapping technologies for urban building extraction. Unlike normal land cover classes, urban buildings always have unique shape features, which provide useful prior information input for super-resolution mapping. In this paper, we propose a novel super-resolution mapping algorithm that may resolve the most suitable building boundary using urban building shape information to improve the quality of urban building extraction.

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

Section II of this paper provides a brief literature review on the theory and existing algorithms for super-resolution mapping of remotely sensed imagery. In Section III, we introduce our methods for urban building mapping incorporating its prior shape information. We present the results that we obtained from several simulated images in Section IV and conclude with a discussion of the limitations and open problems of the proposed schemes in Section V.

II. BACKGROUND

A. Super-resolution mapping

The purpose of super-resolution mapping is to obtain the most suitable locations for the different class fractions within a pixel. According to a pre-defined zoom factor, each coarse resolution image pixel was first divided into a fixed number of sub-pixels. Every land cover class within each coarse resolution pixel was then assigned a number of sub-pixels based on the fraction values of the soft classification used as input. Finally, all possible spatial patterns of these sub-pixels were evaluated using some principle, and the most suitable configuration was thought to be the super-resolution mapping result [10].

Apparently, the key issue in super-resolution mapping is how to define the principle that describes the spatial land cover pattern. Because more information about the sub-pixel land cover pattern within the coarse resolution pixel could not be provided, super-resolution mapping is generally often accomplished assuming spatial dependence, i.e., the tendency for spatially proximate observations of a given property to be more similar than that of more distant observations, which is a simple method to describe the land cover spatial distribution pattern.

Super-resolution mapping thus can be formulated as an inverse problem, which reconstructs a fine spatial resolution map of land cover class labels from a set of class fractions that the low resolution image provides. Many methods for solving the super-resolution mapping problem have been proposed, such as Hopfield neural networks [11], the genetic algorithm [12], and the pixel-swapping algorithm [13]. The pixel-swapping method, which is briefly introduced in the following section, was used in this research because of its simplicity and efficiency.

B. Pixel-swapping algorithm

With the maximal spatial dependence principle mentioned above, the super-resolution mapping problem can be described as follows:

$$\text{Maximize} \quad S = \sum_{k=1}^{NLC} \sum_{i=1}^{NP} x_{ik} * PLC_{ik} \tag{1}$$

which is subject to

$$\sum_{k=1}^{NLC} x_{ik} = 1 \tag{2}$$

and

$$\sum_{i=1}^{NP} x_{ik} = NPLC_k \tag{3}$$

where NLC is the land cover classes; NP is the number of sub-pixels in each coarse-resolution pixel; $NPLC_k$ is the number of sub-pixels assigned to land cover class k, which is derived from the fraction images; and PLC_{ik} is the spatial dependence calculated for land cover class k and for each sub-pixel i.

Within each iteration, the spatial dependence PLC_{ik} of each sub-pixel i for a particular class k is predicted as a distance-weighted function of its neighbors, as follows:

$$PLC_{ik} = \sum_{j=1}^{n} \lambda_{ij} Z_k(x_j) \tag{4}$$

where n is the number of neighbors, $Z_k(x_j)$ is the value of the class k (i.e., now constrained to be either 0 or 1) at the jth pixel location x_j, and λ_{ij} is a distance-dependent weight predicted as

$$\lambda_{ij} = \exp(\frac{-h_{ij}}{a}) \tag{5}$$

where h_{ij} is the distance between pixel x_i (i.e., for which the attractiveness is desired) and the neighbor x_j, and a is the non-linear parameter of the exponential distance decay model.

The algorithm ranks the attractiveness scores on a pixel-by-pixel basis. In each pixel, the locations of the least attractive target sub-pixel (e.g., a 1 surrounded mainly by 0s) and the most attractive background sub-pixel (e.g., a 0 surrounded mainly by 1s) are stored. If swapping the two sub-pixels increases the spatial correlation of the sub-pixels (i.e., the attractiveness of the least attractive target location is less than that of the most attractive background location), the sub-pixels are swapped; otherwise, no swap is made. One sub-pixel pair per pixel is swapped per iteration and the algorithm runs for either a specified number of iterations or until no further swaps are made [14].

III. METHODS

A. The prior information

Although the spatial dependence principle has been widely used for geosciences research and has been proven to be acceptable for super-resolution mapping in most situations, the principle is not always suitable to some specific research fields, such as those involving man-made objects. Incorporating additional site-specific land cover information, such as vector boundaries [15], the features of linear land cover classes [14], and digital elevation models [16], into the super-resolution mapping procedure can improve the accuracy of the resulting maps.

With respect to urban buildings, many buildings have mutually perpendicular directions. If the main axis of the whole building is resolved, all building boundaries can then be determined to be either "horizontal" or "vertical" to the main-

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

axis direction. This prior information has been widely used for urban-building extraction or for boundary regularization [3, 4]. In this situation, the urban building does not have maximal spatial dependence, and the resulting map may not be accurate if the maximal spatial dependence is still used as the principle of super-resolution mapping technology. Replacing the standard spatial dependence principle with this prior shape information of the urban building can improve the performance of super-resolution mapping technologies.

Therefore, the urban building super-resolution mapping problem can be expressed as follows:

(1) The fraction maps and the main axis of the building are either known or have been calculated by pre-procession methods.

(2) The number of building sub-pixels in each pixel is then decided with the fraction maps as input.

(3) The exact locations of urban building sub-pixels are assigned to make the final building boundaries either “horizontal” or “vertical” to the main axis of the building.

B. The proposed algorithm

In this study, we further developed the original pixel-swapping algorithm to incorporate the prior urban building shape information into the super-resolution mapping technology.

Note that the spatial dependence of building sub-pixels can be quite different in respect to the locations of the sub-pixels. The spatial dependence of those sub-pixels inside the building is the same in all directions, whereas the spatial dependence for those sub-pixels in the area of building boundaries reaches a maximum in the same direction of the building boundary. To represent this special feature, we revised the spatial dependence of each sub-pixel in the standard pixel swapping algorithm by creating a unique anisotropic exponential distance-decay window for every sub-pixel based on the main building axis. Because the boundary sub-pixels may belong to lines at directions either “horizontal” or “vertical” to the main axis, we considered both directions as possible solutions and therefore calculated two spatial dependences at both directions and set the maximal one to be the spatial dependence of this sub-pixel.

Fig. 1 illustrates the revised spatial dependence principle. For each sub-pixel, we applied a moving window with the center pixel as the “target” to calculate the revised spatial dependence. Four directions (i.e., at 0°, 22.5°, 45°, and 67.5°) and their vertical directions (i.e., at 90°, 112.5°, 135°, and 157.5°) are shown in Fig. 1 within a 7 x 7 window. The defined directions within the moving window were then used to modify the exponential distance-decay model. Fig. 2(a) shows the “standard” isotropic exponential distance-decay model for one sub-pixel, and Fig. 2(b), Fig. 2(c), and Fig. 2(d) show the anisotropic exponential distance-decay models with directions at 0°, 22.5°, and 45°, respectively.

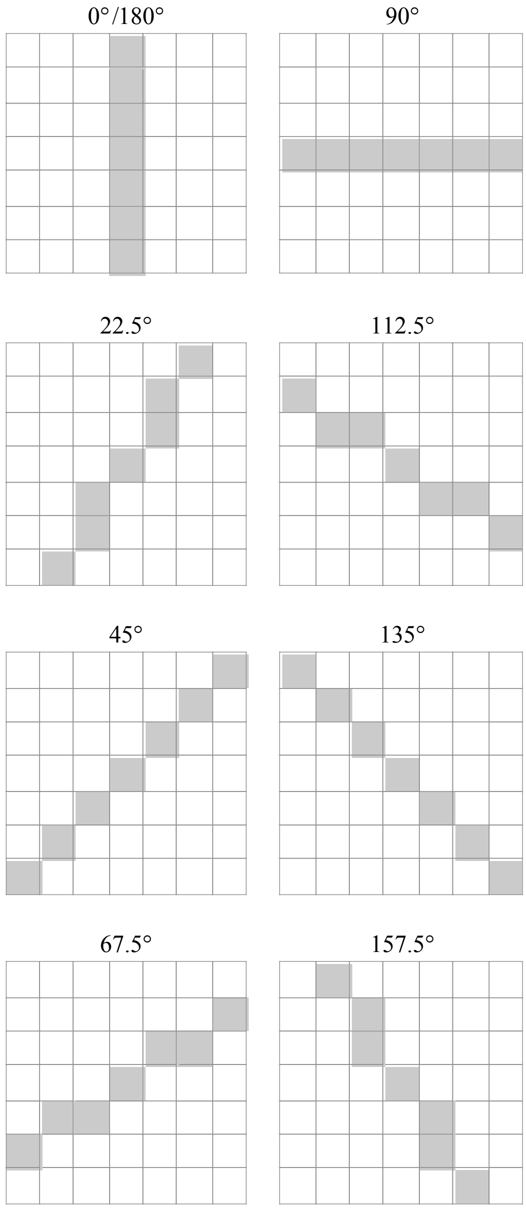

Fig. 1. Four directions and their vertical directions in a 7 x 7 window.

Once the revised spatial dependence principle is determined, the super-resolution mapping procedure can be performed as mentioned above. To avoid getting into a local maximum, we ran the pixel-swapping algorithm based on the anneal simulation procedure, whereby the possible swapped sub-pixels were not the least attractive target sub-pixel and the most attractive background sub-pixel but instead were chosen randomly.

IV. Application and Discussion

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

3.5	4.5	5	4.5	3.5
4.5	7	10	7	4.5
5	10	X_i	10	5
4.5	7	10	7	4.5
3.5	4.5	5	4.5	3.5

a

0	0	5	0	0
0	1	10	1	0
0	1	X_i	1	0
0	1	10	1	0
0	0	5	0	0

b

0	0	0	4.5	0
0	0	10	7	0
0	1	Xi	1	0
0	7	10	0	0
0	4.5	0	0	0

c

0	0	0	0	5
0	0	1	10	0
0	1	X_i	1	0
0	10	1	0	0
5	0	0	0	0

d

Fig. 2. Fig. 2(a) is the isotropic exponential distance-decay model. Figs. 2(b), (c), and (d) are anisotropic exponential distance-decay models at directions of 0°, 22.5°, and 45°.

To illustrate and validate the performance of the proposed algorithm, we applied it to three simulated artificial images representing different shapes of urban buildings, as shown in Fig. 3(a1), Fig. 3(b1), and Fig. 3(c1). All images comprised 70 x 70 pixels, and two classes representing the background (white) and the building (black) were considered. The main-axis directions of these simulated urban buildings are at 0°, 22.5°, and 45°, respectively.

To avoid introducing error from the uncertainty of the soft classification result, we obtained the simulated data by degrading these original high-resolution images. We then calculated the area proportions of both classes in each low-resolution pixel in a window size according to the zoom factor. The corresponding fraction images were considered as the soft classification result and were then used as the inputs for the subsequent super-resolution mapping algorithm. The zoom factor of 10 was set in this study, thus the simulated low resolution images contained 7 x 7 pixels, and each low-resolution pixel contained 10 x 10 pixels of the original high resolution image. Because previous research shows that super-resolution mapping is more accurate than the hard classification method, we only compared the urban-building extraction using the proposed method with that of the standard super-resolution mapping technology.

Visual comparison of the results showed that the revised super-resolution mapping algorithm was more effective than the standard algorithm in resolving urban buildings. For the first shape shown in Fig. 3(a1), the resulting maps of both super-resolution mapping technologies are shown in Fig. 3(a2) and Fig. 3(a3), respectively. The resulting map produced by the proposed method was more similar to the reference than that of the standard method. The right angle corners in this shape were good examples illustrating the difference in performance of both technologies. The traditional method did not present all corners well; the resulting corners are more circular in shape. However, a noticeable improvement can be seen with the proposed method as it preserved the exact shape of all corners. For the second and third shapes shown in Fig. 3(b1) and Fig. 3(c1), although the improvement was different, the resulting maps of the proposed technology were all more similar to the reference map than that of the traditional method. Fig. 3(b3) and Fig. 3(c3) show that the proposed method reconstructed both original shapes effectively.

The error-mapping-pixels (EMP) statistical index was used to validate the performance of the proposed method quantitatively, and the results are shown in Table 1. The EMPs of three shapes were significantly different because the spatial pattern of the target shape had an effect on the performance. However, compared with the standard technology, the resulting map of the proposed method always had a considerably lower EMP index, which equates to an increase in mapping accuracy. For the first and the third shapes, EMPs of the proposed method, even when equal to 0, meant that the resulting maps were exactly the same as the original image. However, the EMPs were 80 and 70, respectively, with the standard super-resolution mapping technology. For the second shape, the EMP of the standard super-resolution mapping technology was 112, and the EMP of the proposed technology was only 20. Although the proposed method did not reconstruct the original shape exactly, mainly because the direction at 22.5° cannot be precisely expressed with the moving window shown in Fig. 1, this significant improvement still showed the efficiency of the proposed algorithm.

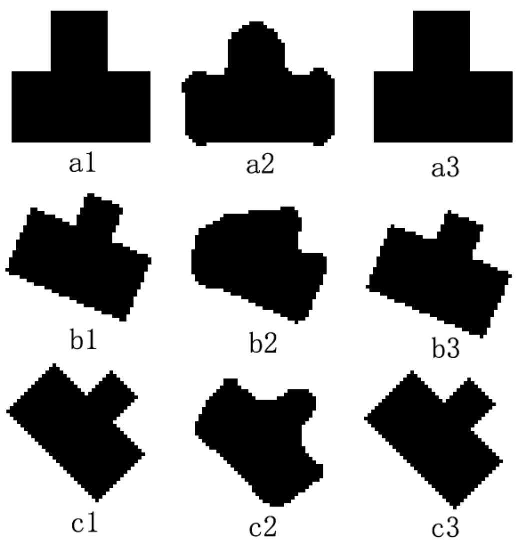

Fig. 3. The simulated artificial imagery and super-resolution mapping results. Figs. 3(a1), (b1), and (c1) are three simulated artificial images (70 x 70 pixels); Figs. 3(a2), (b2), and (c2) are the resulting maps of the standard algorithm; and Figs. 3(a3), (b3) and (c3) are the resulting maps of the proposed algorithm.

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

Table 1. Results of accuracy analysis for simulated images.

EMP	0°	22.5°	45°
Standard method	70	112	80
Proposed method	0	20	0

V. CONCLUSIONS

In this paper, we proposed a novel super-resolution mapping technology for urban building extraction. In the proposed algorithm, prior urban building shape information was incorporated into the super-resolution mapping procedure to improve the accuracy of the resulting maps. To address this issue, we revised the pixel-swapping algorithm in the proposed procedure by replacing the standard isotropic exponential distance-decay model for the pixel-swapping algorithm with the directional anisotropic exponential distance-decay model, which sets the maximal spatial dependencies at a direction either “horizontal” or “vertical” to the main axis of urban buildings as the final spatial dependence of this pixel.

The proposed algorithm was evaluated with three simulated artificial images. The results show that the proposed method can obtain more accurate maps than the standard super-resolution mapping method both visually and quantitatively. Incorporating more specific prior information is a promising method to improve the performance of super-resolution mapping technologies. Although the proposed technology showed good performance, there are still several issues that need to be studied further. As the simulated results in this research were obtained under ideal conditions without any additional errors, the uncertainty caused by soft classification needs to be investigated in more depth. Special spatial pattern descriptions for other objects must also be studied further to increase the performance of the super-resolution technology for practical applications.

ACKNOWLEDGMENT

This work was funded by the National Natural Science Foundation of China (No.40801186, No.60702079) and the Hubei Provincial Natural Science Foundation of China (No. 2008CDZ080).

REFERENCES

[1] L. C. Chen, T. A. Teo, C. Y. Kuo, and J. Y. Rau, "Shaping polyhedral buildings by the fusion of vector maps and lidar point clouds," Photogrammetric Engineering and Remote Sensing, vol. 74, pp. 1147-1157, Sep 2008.

[2] D. S. Lee, J. Shan, and J. S. Bethel, "Class-guided building extraction from Ikonos imagery," Photogrammetric Engineering and Remote Sensing, vol. 69, pp. 143-150, Feb 2003.

[3] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne lidar point clouds," Photogrammetric Engineering and Remote Sensing, vol. 73, pp. 805-812, Jul 2007.

[4] G. Sohn and I. Dowman, "Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction," Isprs Journal of Photogrammetry and Remote Sensing, vol. 62, pp. 43-63, May 2007.

[5] A. Thiele, E. Cadario, K. Schulz, U. Thonnessen, and U. Soergel, "Building recognition from multi-aspect high-resolution InSAR data in urban areas," in 6th European Conference on Synthetic Aperture Radar, Dresden, GERMANY, 2006, pp. 3583-3593.

[6] G. M. Foody, "Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data," International Journal of Remote Sensing, vol. 17, pp. 1317-1340, May 1996.

[7] P. M. Atkinson, M. E. J. Cutler, and H. Lewis, "Mapping sub-pixel proportional land cover with AVHRR imagery," International Journal of Remote Sensing, vol. 18, pp. 917-935, Mar 1997.

[8] L. Bastin, "Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels," International Journal of Remote Sensing, vol. 18, pp. 3629-3648, Nov 1997.

[9] P. M. Atkinson, "Mapping Sub-pixel Boundaries from Remotely Sensed Images," in Innovations in GIS 4 London, UK: Taylor and Francis, 1997, pp. 166-180.

[10] J. Verhoeye and R. De Wulf, "Land cover mapping at sub-pixel scales using linear optimization techniques," Remote Sensing of Environment, vol. 79, pp. 96-104, Jan 2002.

[11] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Super-resolution target identification from remotely sensed images using a Hopfield neural network," Ieee Transactions on Geoscience and Remote Sensing, vol. 39, pp. 781-796, Apr 2001.

[12] K. C. Mertens, L. P. C. Verbeke, E. I. Ducheyne, and R. R. De Wulf, "Using genetic algorithms in sub-pixel mapping," International Journal of Remote Sensing, vol. 24, pp. 4241-4247, Nov 2003.

[13] P. M. Atkinson, "Sub-pixel target mapping from soft-classified, remotely sensed imagery," Photogrammetric Engineering and Remote Sensing, vol. 71, pp. 839-846, Jul 2005.

[14] M. W. Thornton, P. M. Atkinson, and D. A. Holland, "A linearised pixel-swapping method for mapping rural linear land cover features from fine spatial resolution remotely sensed imagery," Computers & Geosciences, vol. 33, pp. 1261-1272, Oct 2007.

[15] P. Aplin and P. M. Atkinson, "Sub-pixel land cover mapping for per-field classification," International Journal of Remote Sensing, vol. 22, pp. 2853-2858, Sep 2001.

[16] F. Ling, F. Xiao, Y. Du, H. P. Xue, and X. Y. Ren, "Waterline mapping at the subpixel scale from remote sensing imagery with high-resolution digital elevation models," International Journal of Remote Sensing, vol. 29, pp. 1809-1815, 2008.

978-1-4244-3461-9/09/$25.00 ©2009 IEEE