COMBINE MARKOV RANDOM FIELDS AND MARKED POINT PROCESSES TO EXTRACT BUILDING FROM REMOTELY SENSED IMAGES

Dengfeng Chai[a], Wolfgang Förstner[b], Michael Ying Yang[c]

[a]Institute of Spatial Information Technique, Department of Earth Science, Zhejiang University
No. 38, Zheda Road, Hangzhou, Zhejiang, 310027, China, chaidf@cad.zju.edu.cn
[b]Department of Photogrammetry, Institute of Geodesy and Geoinformation, University of Bonn
Nussallee 15, 53115 Bonn, Germany, wf@ipb.uni-bonn.de, http://www.ipb.uni-bonn.de
[c]Institute for Information Processing (TNT), Leibniz University Hannover
Appelstr. 9A, 30167 Hannover, Germany, yang@tnt.uni-hannover.de

Commission ICWG III/VII

KEY WORDS: Building, extraction, segmentation, combination

ABSTRACT:

Automatic building extraction from remotely sensed images is a research topic much more significant than ever. One of the key issues is object and image representation. Markov random fields usually referring to the pixel level can not represent high-level knowledge well. On the contrary, marked point processes can not represent low-level information well even though they are a powerful model at object level. We propose to combine Markov random fields and marked point processes to represent both low-level information and high-level knowledge, and present a combined framework of modelling and estimation for building extraction from single remotely sensed image. At high level, rectangles are used to represent buildings, and a marked point process is constructed to represent the buildings on ground scene. Interactions between buildings are introduced into the the model to represent their relationships. At the low level, a MRF is used to represent the statistics of the image appearance. Histograms of colours are adopted to represent the building's appearance. The high-level model and the low-level model are combined by establishing correspondences between marked points and nodes of the MRF. We adopt reversible jump Markov Chain Monte Carlo (RJMCMC) techniques to explore the configuration space at the high level, and adopt a Graph Cut algorithm to optimize configuration at the low level. We propose a top-down schema to use results from high level to guide the optimization at low level, and propose a bottom-up schema to use results from low level to drive the sampling at high level. Experimental results demonstrate that better results can be achieved by adopting such hybrid representation.

1 INTRODUCTION

With the progresses in image capture techniques, more and more remotely sensed images with high spatial, spectral, temporal and radiometric resolutions are available. With the popularity of tools such as Google Map, more and more up-to-date geoinformation are demanded by people. As a traditional research topic in photogrammetry and remote sensing, building extraction from remotely sensed images is a topic much more significant than ever.

In spite of the research efforts of the past decades fully automatic extraction is still a challenging task. The key issue is representation of objects and images (Mayer, 1999, Sowmya and Trinder, 2000, Baltsavias, 2004). Statistical approaches provide a strong framework of modelling and estimation. Markov random fields and marked point processes represent context-dependent entities well (Winkler, 1995, Li, 2009, Baddeley and van Lieshout, 1993). Based on Markov random fields (MRF), low-level information referring to the single image pixels and interaction between neighbouring pixels are represented concisely. However, high-level knowledge, such as free semantic structures and variable topology, can not be represented by MRFs conveniently. Based on spatial point process, high-level knowledge can be introduced via marks attached to the points and the relationships between neighbouring points. While specific shapes can be represented by geometric marks, general shape can not be determined based on image content. This problem results from the weakness of representing low-level information.

Motivated by the complementary characteristics of Markov random fields and marked point processes, we combine them to represent both low-level information and high-level knowledge.

Based on this representation, we propose an automatic approach for extracting buildings from single remotely sensed image.

2 PREVIOUS WORKS

2.1 Markov random fields based representation

Markov random fields provide a natural representation of context-dependent entities (Besag, 1974, Geman and Geman, 1984). A set of sites are used to represent pixels or primitives, and a set of labels attached to each site are used to denote events that may happen at the site. Furthermore, a neighbourhood system is used to describe the interrelationships between sites. Benefiting from the equivalence between MRF's and Gibbs distribution (Hammersley and Clifford, 1971), MRF can be described by local characteristics. Moreover, MRF-MAP estimation can be rigorously achieved in case of binary classification by Graph Cut algorithm (Boykov et al., 2001) and approximately e. g. by belief propagation (Felzenszwalb and Huttenlocher, 2004).

Although there are some high-level MRF models (Li, 1994), their structures are fixed in modelling and estimation. They are not flexible enough to represent random structures. Successes are demonstrated in low level, especially for regular lattices corresponding to image grids. For example, image restoration (Felzenszwalb and Huttenlocher, 2004), stereo matching (Tappen and Freeman, 2003), image segmentation (Rother et al., 2004) or clustering (Zabih and Kolmogorov, 2004) are formulated as pixel labelling problems with labels denoting pixel intensity, disparity or object category respectively.

Researchers introduced strong priori information into MRFs to improve its performance of representation. (Kumar et al., 2010) combined pictorial structures and MRFs and proposed an object category specific MRFs model for detecting and segmenting instances of a particular object category. The object-specific shape prior is represented using pictorial structures, and it relies on a large library of exemplars. (Winn and Shotton, 2006) adopted a part labelling which densely covers the object and proposed a Layout Consistent Random Field (LayoutCRF) model to impose asymmetric local spatial constraints on these labels to ensure the consistent layout of parts. It can detect and segment partially occluded objects of a known category. They introduced shape priori for objects and layout of object parts. Therefore, single object or a small number of objects can be segmented cleanly.

Since the number of objects presented in remotely sensed image simultaneously is large, object shapes and scene topology are too complex to be modelled using above mentioned approaches. On the other hand, it is important to utilize such information to segment object precisely.

2.2 Marked point processes based representation

Marked point processes provide a useful representation of spatially distributed objects. A set of points randomly distributed are used to represent objects. The number of points, their positions, and their interactions are random. Furthermore, marks are attached to each point to represent high-level knowledge such as category or geometric shape. Marked point processes are flexible enough to model the scene at the object level. Given a proposed model, reversible jump Markov Chain Monte-Carlo (Green, 1995) can be adopted to explore the configuration space and annealing schema can be adopted to simulate the objective distribution to find the optimal solution.

The Ariana Research Group CNRS/INRIA/UNSA introduced marked point process into remotely sensed image analysis. They demonstrated that marked point processes have a great potential in object extraction from remotely sensed image. (Ortner et al., 2007, Ortner et al., 2008) adopted rectangle to represent building footprint, and proposed an approach for building footprint extraction from DSM. (Lafarge et al., 2008) adopted a 3D model to represented buildings, and proposed an approach for building reconstruction from DSM based on a library of 3D models. (Tournaire et al., 2010) adopted above framework and formulated the energy in an efficient way, easy to parameterization and fast to compute.

The devised marks, however, can only represent specific shapes. Due to computational limitation, it is impossible to adopt a huge number of marks to represent general shapes. And, general shapes can not be determined based on image content adaptively. Moreover, images are linked with the model via a data term computed using hypothesis testing schema, which can not make full use of low-level information.

3 BAYESIAN FRAMEWORK FOR BUILDING EXTRACTION

As a whole, we represent buildings as foreground X and the rest as background $\overline{X}$, and formulate building extraction as foreground/background segmentation in a Bayesian manner.

Given the observed image I, the buildings X can be estimated by maximizing the posterior:

$$X = \arg\max_X P(X, \overline{X}|I; \Theta) \qquad (1)$$

$$= \arg\max_X P(I|X, \overline{X}; \Theta) P(X, \overline{X}|\Theta) \qquad (2)$$
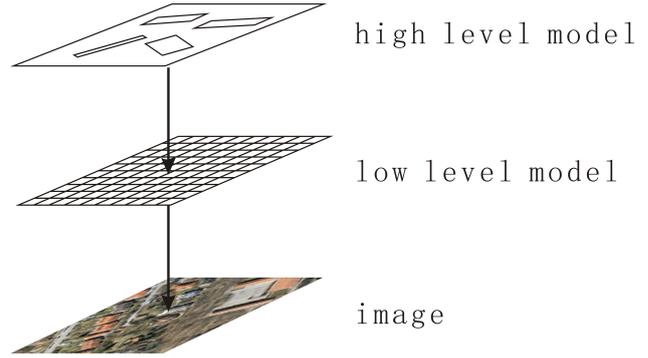
Figure 1: Hybrid representation.

where, Θ is a set of models and parameters, $P(X, \overline{X}|\Theta)$ is the priori probability of a specific configuration $X, \overline{X}$ conditioned on Θ, and $P(I|X, \overline{X}; \Theta)$ is the likelihood of observing an image I given the configuration $X, \overline{X}$ conditioned on Θ.

We will address issues of modelling and estimation in section 4 and section 5 respectively.

4 MODELLING

4.1 Hybrid representation

We combine marked point processes and Markov random fields and propose a hybrid representation for building extraction. As illustrated in Fig.1, marked point process is adopted to represent the high-level knowledge, i.e. the buildings and their distribution; Markov random field is adopted to represent the low-level information, i.e. the properties of all pixels.

4.2 High-level model

At high level, only buildings are represented explicitly, the rests are represented implicitly. The buildings are modelled as a marked point process.

4.2.1 Marked point process A rigorous definition of spatial point process involves measure theory which is difficult for the readers who have not studied the subject. Instead, we present only a simple description in this section.

Let $S \subset R^2$ be a compact set, and Ω_n be the set of configurations $X = \{x_1, \ldots, x_n\}$ that consist of n unordered points of S, the probability density of a specific configuration X is defined as follows:

$$h(X) = \alpha \beta^{n(X)} \prod_{x_i \in X} \phi^{(1)}(x_i) \cdots \prod_{\{x_{i_1}, \ldots, x_{i_k}\} \in X} \phi^{(k)}(x_{i_1}, \ldots, x_{i_k})$$

$$(3)$$

where, α is the normalizing constant, $n(X)$ is the number of points, β is a positive constant, and $\phi^{(k)}(x_{i_1}, \ldots, x_{i_k})$ reflects the interactions among k-tuplets neighbouring points.

By attaching a geometric mark $m_i = (l_i, w_i, dir_i)$ to each point $x_i \in X$, we can augment a spatial point process to be a marked point process, where, the triplet denotes the length, width and main direction of a rectangle. The marked point is denoted as $\hat{x_i}$. All buildings are represented by a rectangle-marked point process.

Return to Bayesian formulation in section 3, $h(X)$ is computed from a prior term and a data term:

$$h(X) \propto h_p(X) h_d(X). \qquad (4)$$

where, the prior term $h_p(X)$ measures the priori probability of different scenes, the data term $h_d(X)$ measures the coherence between the scene and the image, which is identified with the likelihood term. $h_p(X)$ and $h_d(X)$ correspond to $P(X, \overline{X}|\Theta)$ and $P(I|X, \overline{X}; \Theta)$ in Eq.(2) respectively.

4.2.2 Priori Most existing approaches rely on very specific priori (Ortner et al., 2007, Lafarge et al., 2008). (Lafarge et al., 2010, Benedek et al., 2010) adopted the relation overlap as a general interaction and developed a concise priori model.

In this paper, we argue that buildings can not overlap with each other and use this condition as a hard constraint. It is realized by defining the following density:

$$h_p(X) \propto \exp(\theta_0 o(X)) \quad (5)$$

where, $\theta_0 = -\infty$ prevents overlapped rectangles to appear in configuration, $o(X)$ is the number of pairs of overlapped rectangles:

$$o(X) = \sum_{x_i, x_j \in X} \mathsf{overlap}(\hat{x_i}, \hat{x_j}) \quad (6)$$

where,

$$\mathsf{overlap}(\hat{x_i}, \hat{x_j}) = \begin{cases} 1 & \hat{x_i} \text{ overlaps with } \hat{x_j} \\ 0 & \hat{x_i} \text{ does not overlap with } \hat{x_j} \end{cases} \quad (7)$$

To reflect the sparse distribution of buildings, small distances between neighbouring buildings are penalized. We augment above model by defining the density of valid configuration as unconditional Strauss process (Strauss, 1975):

$$h_p(X) \propto \exp(\theta_1 n(X) + \theta_2 p(X, r)) \quad (8)$$

where θ_1 and θ_2 are two parameters of this model, $n(X)$ is the number of points in X, $p(X, r)$ is the number of pairs of points that are nearer than r:

$$p(X, r) = \sum_{x_i, x_j \in X} d(x_i, x_j) \leq r \quad (9)$$

where, $d(x_i, x_j)$ is the distance between x_i and x_j.

To reflect parallel of buildings, large difference of directions between neighbouring buildings are penalized. We augment above model by adding a direction term:

$$h_p(X) \propto \exp(\theta_3 q(X, r)) \quad (10)$$

where, $q(X, r)$ is sum of square of direction differences between neighbouring buildings no far than r:

$$q(X, r) = \sum_{x_i, x_j \in X, d(x_i, x_j) \leq r} \alpha^2(\hat{x_i}, \hat{x_j}) \quad (11)$$

where, $\alpha(\hat{x_i}, \hat{x_j})$ is the difference of the directions of x_i and x_j:

$$\alpha(\hat{x_i}, \hat{x_j}) = dir(\hat{x_i}) - dir(\hat{x_j}) \quad (12)$$

where, $dir(\hat{x_i})$ is the main direction of the rectangle attaching to x_i. Furthermore, $\alpha(\hat{x_i}, \hat{x_j})$ is compared with $0, \pi/2, \cdots, 2\pi$ and the minimal difference is adopted as the result. This measure can reflect both parallel and orthogonal of buildings.

Finally, we combine above models and get a full distribution as follows:

$$h_p(X) \propto \exp(\theta_0 o(X) + \theta_1 n(X) + \theta_2 p(X, r) + \theta_3 q(X, r)) \quad (13)$$

where the first term assures that neighbouring buildings do not overlap with each other, the second and third terms assure that all building distribute sparsely, the last term assures that neighbouring buildings align with each other. $\theta_0 \in \Theta, \theta_1 \in \Theta, \theta_2 \in \Theta, \theta_3 \in \Theta, r \in \Theta$ are all parameters of the model.

4.2.3 Data term The data term measures the coherence between the scene and the image, i.e. the likelihood of presenting an image given the scene:

$$h_d(X) = P(I|X, \overline{X}; \Theta) \quad (14)$$

Since the low-level model links both high-level model and image, it serves as an intermediate model between marked point process and image, the data term is calculated based on low-level model. Details will be presented in section 4.4.2.

4.3 Low-level model

At the low level, both foreground and background are modelled explicitly and together as a Markov Random Field.

4.3.1 Markov random field Since foreground and background can be denoted by two labels, we model them as an Ising model (Li, 2009). The probability of presenting a specific configuration f is computed as follows:

$$P(f) = \frac{1}{Z} \exp(-U(f)) \quad (15)$$
$$\propto \exp(-(U_p(f) + U_d(f))) \quad (16)$$

where, Z is a normalizing constant called the partition function, which is common to all configurations and can be ignored in computation. $U_p(f)$ and $U_d(f)$ correspond to $P(X, \overline{X}|\Theta)$ and $P(I|X, \overline{X}; \Theta)$ in Eq.(2) respectively.

4.3.2 Priori The priori energy $U_p(f)$ for Ising model is calculated as follows:

$$U_p(f) = \sum_{i,j} |f_i - f_j| \quad (17)$$

where, i and j are horizontally or vertically neighbouring pixels, all neighbouring pixels with different labels contribute to the total energy.

Above priori term is based on the assumption that the random field varies smoothly everywhere. Every pair of pixels with different labels will increase the priori energy and decrease the probability. In fact, neighbouring pixels in foreground or background regions should have the same labels. neighbouring pixels across the region boundaries should have different labels. In other words, label field should be allowed to change at region boundaries without increasing the priori energy.

To reflect such priori knowledge, neighbouring pixels across regions should contribute nothing to the priori energy, while as neighbouring pixels within regions should contribute to the priori energy as traditional way. We augment above priori term to be:

$$U_p(f) = \sum_{i,j} \beta_{i,j} |f_i - f_j| \quad (18)$$

where, if i and j across foreground and background regions, $\beta_{i,j}$ should be 0; otherwise, it should be 1.

Such discontinuity preserving constraint is more reasonable than the simplest constraint making configuration varies smoothly everywhere. However, It is difficult to express the discontinuity preserving constraints because nothing is known in advance about

the regions and their boundaries. In the proposed hybrid representation, the high-level model provides a guess of such knowledge, it can be utilized to calculate $\beta_{i,j}$. Details will be presented in section 4.4.1.

4.3.3 Data term The data term $U_d(f)$ corresponds to the likelihood is defined as follows:

$$U_d(f) = U_d(I|f) = \sum_i U(I_i|f) \tag{19}$$

where, the energy $U(I_i|f)$ is calculated as follows:

$$U_d(I_i|f) = \begin{cases} -\log(P(I_i|H_b)) & f_i = 0 \\ -\log(P(I_i|H_f)) & f_i = 1 \end{cases} \tag{20}$$

where, $H_b \in \Theta$ and $H_f \in \Theta$, are two normalized histograms for background and foreground respectively. $P(I_i|H_b)$ and $P(I_i|H_f)$ measures the likelihood of the pixel with colour I_i belonging to background and foreground respectively.

4.4 Linking high-level and low-level models

Each marked point at the high level denotes one building, and it corresponds to one rectangular region in the Markov random field at the low level. The high-level model and the low-level model are combined together by establishing correspondences of marked points at the high level and regions (each one consists of a set of pixels) at the low level. High-level knowledge is introduced as a priori term in the MRF and low-level information is introduced into data term in the marked point process. In this way, a flexible and robust representation is achieved.

4.4.1 Priori As pointed out in section 4.3.2, high-level knowledge can be utilized to construct a discontinuity preserving a priori term for the low-level model.

Suppose that there are a set of marked points at the high level, we can get a set of rectangular regions at low level by projecting the marked points on to the grid of Markov random field. Each pixel i has one label f_i which denotes its class, i. e. foreground or background. Without loss of generality, suppose that i and j are neighbouring pixels. The discontinuity preserving priori term is constructed by defining $\beta_{i,j}$ as follows:

$$\beta_{i,j} = \begin{cases} 0 & f_i \neq f_j \\ 1 & f_i = f_j \end{cases}, \tag{21}$$

i. e. the label configurations of the Markov random field are not evaluated at the borders induced by the marked point process.

4.4.2 Data term Suppose that there is a set of marked points at the high level. We can obtain a set of rectangular regions at the low level by projecting the marked points on to the grid of Markov random field. Each pixel i has one label f_i which denotes its class, i. e. foreground or background. Using Eq.(19), we can calculate the data term for the high level model.

More specifically, summing the data terms over one rectangular region, we get the data term corresponding to one marked point at the high level.

5 OPTIMIZATION

We adopt simulated annealing to simulate the posterior distribution so that an optimal configuration can be achieved as the temperature gradually approaches zero. It iteratively simulates the distribution $h^{\frac{1}{T}}(X)$ with T gradually decreasing to 0.

Furthermore, we use reversible jump Markov random Monte Carlo (rjMCMC) techniques to explore the configuration space at the high level, an uniform birth and death kernel and a translation kernel are developed to generate new states according the current state. The former randomly generates a point together with a rectangle (length, width and direction) in the image region, while the latter randomly selects an existing rectangle and adjusts its parameters randomly. The new state is adopted with an accept rate to keep the detailed balance.

What distinguishes our approach from existing ones is that a top-down schema and a bottom-up schema are proposed for random sampling. In each sample, it first generates a new state at high level, then uses it to guide the optimization at low level, then uses the optimization results to adjust state at high level.

5.1 Top down schema

Since there are a large number of buildings presented in remotely sensed image simultaneously, previous MRF based approaches need seed pixels provided manually. Otherwise, neighbouring buildings can not be distinguished well without knowledge about their spatial distribution. We, however, use the results at the high level to provide the information about the spatial distribution and approximate shapes of buildings.

Low-level optimization is conducted only when a new marked point is birthed. Given a new states, i. e. a new rectangle-marked point. we use it to guide the optimization at low level:

1. Project the rectangle into image to get projected regions;

2. Construct discontinuity preserving priori term for low-level model;

3. Adopt Graph Cut algorithm (Greig et al., 1989) to optimize the proposed object function with discontinuity preserving priori term, the optimization is conducted in the projected regions and it results in some building regions.

5.2 Bottom up schema

Motivated by the data driven MCMC (Tu and Zhu, 2002), we use results at low level to drive the sampling, i. e. compute new state according results of optimization at low level. In their data driven MCMC, the low-level results are computed by edge detection or segmentation and are fixed in the process of MCMC sampling. In our approach, the low-level results are computed from low level optimization and vary in the process of MCMC sampling.

Given the optimization results achieved at low level, i. e. some building regions at low level, we use them to adjust the new state at high level:

1. Select the largest one and find a minimal rectangle enclosing the selected region;

2. Adjust the new marked point to be the rectangular region found above;

3. Calculate the data term based on the optimization result instead of the original marked point;

4. Calculate the accept rate using the data term calculated above;

5. Move the current state to the (adjusted) new state with the accept rate.
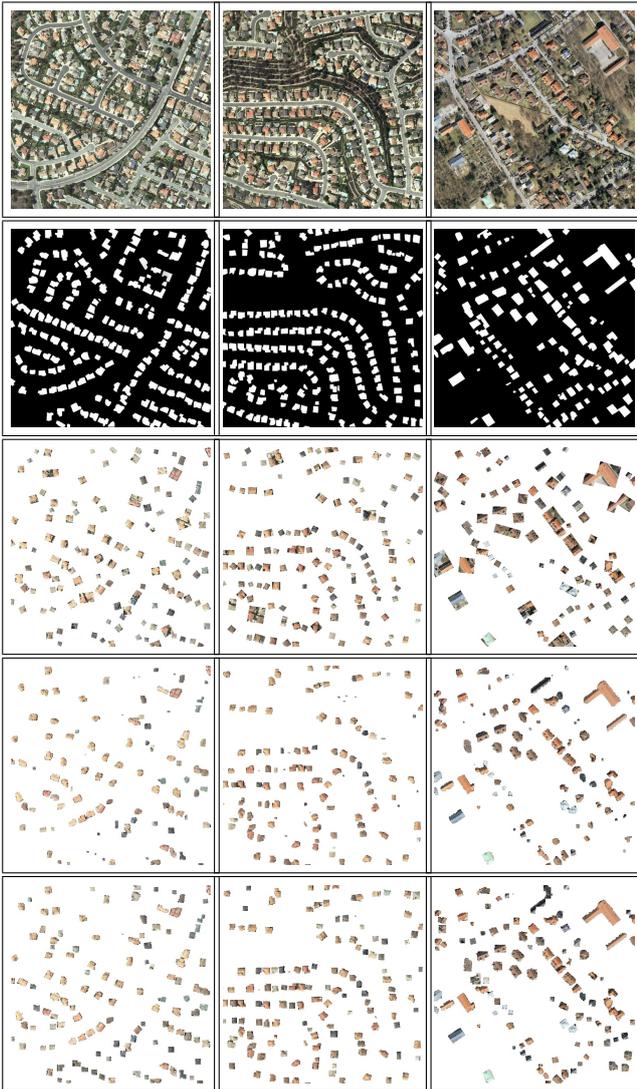
Figure 2: Extraction results: the first row illustrates the original images, the second row illustrates the reference extraction results, which are annotated manually, the rest rows illustrate extraction results based on Marked point processes, Markov random fields, and hybrid representation.

Benefited from optimization at low level, more precise building region can be found in each random sampling, this may improve the definition of both shape and data term. As a result, this schema may drive rjMCMC sampling to achieve better results. Theoretical foundation is guaranteed by simulated annealing and rjMCMC technique, which is our computation framework.

6 EXPERIMENTS

6.1 Results and comparison

We apply our approach to extract buildings from three satellite images of developed urban or suburban areas. For comparison, we also apply marked point processes based approach and Markov random fields based approach to extract buildings.

In this experiment, the histograms H_b and H_f (i.e. three dimensional arrays) are learned from real images by manually annotating images. The rest parameters are set as follows: $\theta_0 = -\infty, \theta_1 = 1, \theta_2 = -0.01, \theta_3 = -0.01, r = 10 pixels$.

The original images, manually annotated reference images, and extraction results are presented in Fig.(2). As illustrated, there are some clear errors in the third row achieved by marked point processes based approach. Since the data term for a rectangle is calculated as a whole but not pixelwise, some regions that contain buildings may be recognized as a building region by mistake. Theoretically, infinite sampling can remove such cases, however, it can not be achieved in practice.

As illustrated, there are many buildings missed in the fourth row achieved by Markov random fields based approach. Since the priori term (smooth term) drive the results as smooth as possible, some regions with low density of being buildings are segmented as background by mistake. On the contrary, some regions between neighboring buildings are segmented as foreground.

As illustrated, the last row achieved by our hybrid representation is better than above rows. Benefited from the hybrid representation, both high-level knowledge and low-level information are well-represented and utilized in the process of building extraction. The point distribution at high level provides a topology structure of the scene. Based on the topology, the optimization at the low level is expected to achieve robust results. The detailed data terms computed at low level improve foreground and background distinguishing at high level.

We also recorded the number of building pixels extracted as building pixels, building pixels extracted as non-building pixels, and non-building pixels extracted as building pixels. They are divided by the total number of building pixels and presented in Tab.(1, 2 and 3) respectively. As illustrated, the results achieved by hybrid representation are much better than those of Markov random fields based approach, they are also better than those of marked point processes based approach, especially indicated by False Negative, which means the number of non-building pixels extracted as building pixels.

Table 1: Quantitative evaluation on first image

	True	False Positive	False Negative
MPP-based	0.53	0.47	0.14
MRF-based	0.41	0.59	0.04
Hybrid-based	0.52	0.48	0.09

Table 2: Quantitative evaluation on second image

	True	False Positive	False Negative
MPP-based	0.56	0.44	0.15
MRF-based	0.44	0.56	0.04
Hybrid-based	0.57	0.43	0.08

Table 3: Quantitative evaluation on third image

	True	False Positive	False Negative
MPP-based	0.74	0.26	0.34
MRF-based	0.69	0.31	0.07
Hybrid-based	0.78	0.22	0.13

7 CONCLUSION

This paper presents a hybrid representation for buildings in remotely sensed image and an approach for building extraction from single remotely sensed image. First, it formulates building extraction in a Bayesian framework. Then, it addresses modelling issue and optimization issue respectively. Buildings are modelled at two levels. At the high level, marked point processes are used to represent such high-level knowledge as topology structure of

a scene. At the low level, a Markov random field is used to represent pixel colour and interaction. After establishing a link between high-level model and low-level model, it proposes a top-down schema and a bottom-up schema optimizing an objective function. Benefited from the hybrid representation and optimization schema, good extraction results are achieved as demonstrated by experiments presented in this paper.

To our knowledge, it is the first work on the combination of marked point process and Markov random fields. Therefore, there are many issues to be investigated in the near future. First, the optimization schema can be improved greatly since the interactions between high-level model and low-level model are not fully utilized. Second, much more information from the image data need to be explored to improve the extraction quality since histograms of colours do not fully represent information contained in the images, this can be seen in the density images calculated using histograms.

ACKNOWLEDGEMENTS

REFERENCES

Baddeley, A. and van Lieshout, M., 1993. Stochastic geometry models in high-level vision. Statistics and Images 1, pp. 231–256.

Baltsavias, E., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. ISPRS Journal of Photogrammetry and Remote Sensing 58(3-4), pp. 129–151.

Benedek, C., Descombes, X. and Zerubia, J., 2010. Building detection in a single remotely sensed image with a point process of rectangles. In: 20th International Conference on Pattern Recognition (ICPR), 2010, pp. 1417 –1420.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society. Series B (Methodological) 36(2), pp. pp. 192–236.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Transaction on Pattern Analysis and Machine Intelligence 23(11), pp. 1222–1239.

Felzenszwalb, P. and Huttenlocher, D., 2004. Efficient belief propagation for early vision. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 261–268.

Geman, S. and Geman, D., 1984. Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. IEEE Transaction on Pattern Analysis and Machine Intelligence 6(6), pp. 721–741.

Green, P. J., 1995. Reversible jump markov chain monte carlo computation and bayesian model determination. Biometrika 82(4), pp. pp. 711–732.

Greig, D., Porteous, B. and Seheult, A., 1989. Exact maximum a posteriori estimation for binary images. J. Royal Statistical Soc., Series B 51(2), pp. 271–279.

Hammersley, J. M. and Clifford, P., 1971. Markov field on finite graphs and lattices. unpublished.

Kumar, M. P., Torr, P. and Zisserman, A., 2010. Objcut: Efficient segmentation using top-down and bottom-up cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, pp. 530–545.

Lafarge, F., Descombes, X., Zerubia, J. and Pierrot-Deseilligny, M., 2008. Automatic building extraction from dems using an object approach and application to the 3d-city modeling. ISPRS Journal of Photogrammetry and Remote Sensing 63(3), pp. 365–381.

Lafarge, F., Gimelfarb, G. and Descombes, X., 2010. Geometric feature extraction by a multi-marked point process. IEEE Transaction on Pattern Analysis and Machine Intelligence 32(9), pp. 1597–1609.

Li, S., 1994. A markov random field model for object matching under contextual constraints. In: Proceedings of Computer Vision and Pattern Recognition, 1994, pp. 866 –869.

Li, S., 2009. Markov Random Field Modeling in Image Analysis. Springer.

Mayer, H., 1999. Automatic object extraction from aerial imagery–a survey focusing on building. Computer Vision and Image Understanding 74(2), pp. 138–149.

Ortner, M., Descombes, X. and Zerubia, J., 2007. Building outline extraction from digital elevation models using marked point processes. International Journal of Computer Vision 72(2), pp. 107–132.

Ortner, M., Descombes, X. and Zerubia, J., 2008. A marked point process of rectangles and segments for automatic analysis of digital elevation models. IEEE Transaction on Pattern Analysis and Machine Intelligence 30(1), pp. 105–119.

Rother, C., Kolmogorov, V. and Blake, A., 2004. Grabcut: Interactive foreground extraction using iterated graph cuts. ACM Transaction on Graphics 23(3), pp. 309–314.

Sowmya, A. and Trinder, J., 2000. Modelling and representation issues in automated feature extraction from aerial and satellite images. ISPRS Journal of Photogrammetry and Remote Sensing 55(1), pp. 34–47.

Strauss, D. J., 1975. A model for clustering. Biometrika 62(2), pp. 467–475.

Tappen, M. and Freeman, W., 2003. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In: Proc. of International Conference on Computer Vision, pp. 900–907.

Tournaire, O., Brdif, M., Boldo, D. and Durupt, M., 2010. An efficient stochastic approach for building footprint extraction from digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing 65(4), pp. 317–327.

Tu, Z. and Zhu, S., 2002. Image segmentation by data-driven markov chain monte carlo. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, pp. 657–673.

Winkler, G., 1995. Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction. Springer.

Winn, J. and Shotton, J., 2006. The layout consistent random field for recognizing and segmenting partially occluded objects. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, Vol. 1, pp. 37 – 44.

Zabih, R. and Kolmogorov, V., 2004. Spatially coherent clustering using graph cuts. In: Proceedings of Computer Vision and Pattern Recognition, 2004, Vol.2, pp. 437 –444.