# A rapid matching approach for large-scale urban imagery

*Xiucheng Yang, Xuebin Qin, Jianhua Wang, Xin Ye, Jun Wang, Qiming Qin*

*Institute of Remote Sensing and Geographic Information System, Peking University, Beijing, 100871, China*

*Corresponding author: Qiming Qin. Tel: +86-10-62754674. E-mail address: qmqinpku@163.com.*

**KEY WORDS** – regional feature matching, topological descriptor, perceptual organization, unsupervised classification

**ABSTRACT** Image matching is a common technique in series of remote sensing imagery applications. The proposed method conducts image matching rapidly in complex urban based on widely existing rectangular buildings. The optimized framework combines the external contour and internal structure to construct regional feature representing building roof, in which perceptual grouping and classification is introduced simultaneously. Then a newly proposed Euler Number matrix, combined with conventional area index and histogram, describes the regional feature, and relevant matching metrics measure the similarity. Specific experiments performed on urban areas with repetitive patterns and poor texture justify the effectiveness and robustness of the proposed image matching approach. The method can require the demand of image retrieve, image mosaic, and serve as primary match of geometric measurement applications.

## I. INREODUCTION

Image matching is the process of overlapping two or more images of the same scene. Image matching is an inevitable problem arising in many applications, such as object recognition, image retrieve, image mosaic, and geometric measure, in special fields of urban planning, digital mapping, disaster monitoring and et al. using the aerial remote sensing imagery of urban areas. During the last three decades, many researchers have intensely studied this problem and presented many algorithms, which have been reviewed and classified in detail (Atallah et al., 1999; Mamic & Bennamoun, 2000; Shokoufandeh et al., 2012). And the current image matching techniques can be most widely divided into two categories, i.e. area-based match and feature-based match.

Area-based matching techniques are the oldest and simplest of the matching algorithms. Each image point to be matched is in fact the center of a small window of points in the reference image, and this window is statically compared with similarity sized windows of points in the target image. The measure is either a difference metric that is minimized, such as RMS difference, or more commonly a correlation measure that is maximized, such as mean- and variance-normalized cross-correlation (Hannah, 1988). Correlation works well most of the time, but encounters difficulties when the two images are of a scene that does not contain adequate visual texture, with many depth discontinuities, or with markedly repetitive patterns. Feature-based matching are in contract more robust, and they have been the topic of much recent research. The goal of feature match is to establish the correspondence between the features detected in the reference and target image. Based on information content and complexity, it is divided into local low-level features (e.g. point primitives, line primitives), and regional high-level features such as regular shaped structures directly grouped by local features describing specific object containing semantic information. Local features have high locating accuracy and simple description, whereas they are characterized of large number and little information, which makes it problematic in cases of poor-texture and repetitive patterns despite of introducing constraint criteria or matching strategies. Regional features contain sufficient image information and include few elements so as to realize fast matching easily. Yet the extraction and description is away from matured, and the location accuracy is comparatively low in 1-2 pixel.

Large-scale urban aerial remote sensing imagery is mostly man-made objects based instead of series terrain. The image features depth discontinuities, occlusion, shadow, low-texture and repetitive patterns, which brings more challenging problems in the match process, especially the area-based match is invalid and local feature match is problematic. Yet if building roof can be detected fundamentally served as high-level feature, the regional feature match is easier to realize, the results of which can require the demand of object recognition and image retrieve, and act as the coarse match in hierarchical matching aiming at survey applications.

With the development of object recognition, various kinds of methods for building extraction from very high remote sensing imagery have been proposed (Mayer, 1999; Ünsalan & Boyer, 2011). Many studies focus on the rectangular roof generally existing in urban areas, and methods based of perceptual organization, arranging the line segments of building's edges to reconstruct the structure of the roof, reach good extraction effect (Benedek el al., 2010; Izadi & Saeedi, 2012).

The paper takes rectangular buildings, the dominant regional features in urban images, as matching object, to realize the fast matching among very high resolution remote sensing images in urban areas. We adopted a fast automatic rectangular building extraction approach (Wang el al., 1013) to segment the regional features, and unsupervised classification to increase heterogeneous spatial distribution. Once detecting the regional feature, area index, topology containment matrix and gray histogram describes the feature in aspect of contour, internal structure and spectrum respectively. Finally, similar metrics and confidence rule succeeds to find correspondence efficiently.

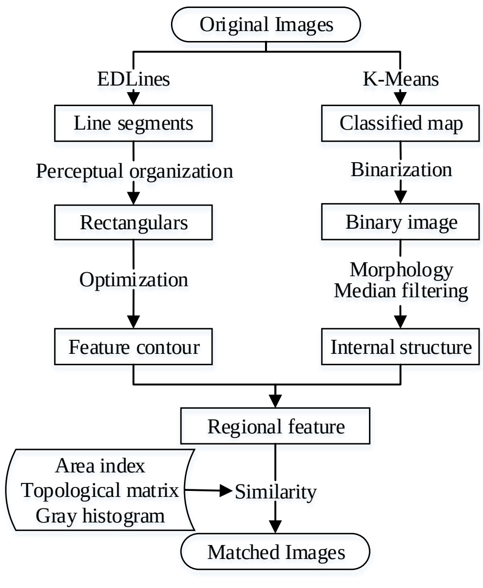


Figure 1. Workflow of the fast urban image matching approach

## II. METHODOLOGY

A framework combing external contour extraction and internal classification is adopted to construct regional-feature representing building roof. Based on Marr's vision theory, the contour extraction strategy starts from low-level image primitives (e.g., linear structures) to high-level geographical semantic polygon sets (e.g., building roofs). Yet the contour feature still cannot refrain from the ambiguity caused by urban repetitive structures and roof poor texture. Given that there may exist shadow areas, sub-objects such as chimney and calorifier, or ununiformity of gray distribution, classification based approach is adopted to obtain the distinctively internal spatial structure. Then not only the conventional histogram and area index measures the similarity, but also an Euler Number matrix, representing the heterogeneous spatial distribution in roof regions, is introduced to optimize the matching metrics. The overall methodology is presented schematically in Figure 1.

### 2.1 Regional feature detecting

Feature-based matching relies on the correctness and distinguishability of detected features. We segment the interests of roof and classify the heterogeneous regions at the same time, constructing a regional feature representing the building. Thus the feature contains a wealth of information, contour, internal structure, and spectrum.

#### A. Contour extraction

An approach for rapid automatic building extraction from very high resolution remote sensing imagery is introduced. The method conducts automatic and accurate building extraction combining low-level image primitive detection with real-time line segment detection and middle-level perceptual grouping. The approach consists of three stages: Firstly, the imagery is preprocessed by histogram equalization and bilateral filter to enhance image contrast and decrease noises simultaneously. Secondly, a line segment detector called EDLines is introduced for real-time accurate extraction of building line segments. Finally, a graph search-based perceptual grouping approach is introduced to hierarchically group previously detected line segments into candidate rectangular buildings. The recursive process was improved through efficient examination of geometrical information with line linking and closed contours search. And it has been proved that the approach can remarkably improve building extraction efficiency, which is valid and feasible for fast image matching.

#### B. Internal structure generating

K-Means is a classical unsupervised classification algorithm, which only an initial number of distinct categories need to be set, and classification map can be obtained rapidly. Binary image is then transformed by the classified. Afterwards, morphology processing and median filtering is performed on the binary image to obtain regular patches with clear boundary and homogeneous regions.

### 2.2 Feature description and similarity metric

To measure the feature similarity, compact and stable descriptors are necessary. We take area index, gray histogram and topological matrix describing contour feature, original spectrum characteristics and internal heterogeneous structure respectively. Then the regional features are compared one by one between reference image and target image, according to the three similarity metrics, and the most similar one within certain threshold as well is considered as a matching pair. Instead of complex constraint and strategy, the simplest traversal pattern can accomplish fast and accurate image matching, benefit from the small number of high-level features.

#### A. Area index

Area of figure measures the distribution range, which is a simplest descriptor to represent contour for rapid matching judgment. Perspective distortion and grouping deviation makes the same roof differ between the matched images, yet the difference would be a relatively small value unless the perceptual organization leads to wrong or inexact result. Although the area index is not certain to judge whether match or not, the similarity can serve as an optimum technological process to eliminate the absolute mismatches estimated as following:

$$\frac{\mathrm{area}\left(\left|a_i - b_j\right|\right)}{\min(a_i, b_j)} < \delta \qquad i = 1, \cdots, m;\ j = 1, \cdots, n;$$

#### B. Gray histogram

The spectrum of the original image can also make a difference. Although the poor texture makes the pixel-to-pixel comparisons problematic, and the perspective or illumination changes may affect the gray values, region histogram considers the all pixels and many proposed measurement can keep from the local and global problem. In the paper, histogram intersection (*d*) helps to select correspondence, in which the smaller the value is, the higher the similarity is.

$$d_i(\mathrm{H}_1,\mathrm{H}_2)=1-\sum_i \min(\mathrm{H}_1(\mathrm{i}),\mathrm{H}_2(\mathrm{i}))$$

**C. Topological matrix**

Topological relationship has the invariance of translation, rotation and scaling, and the satisfied effects are acquired in image matching. Building contour topologically contains internal heterogeneous blocks, so the relationship can construct the other descriptor to examine the similarity. To overcome the limitation to represent spatial distribution of conventional topological descriptors (e.g., adjacent graph, Euler Number), we propose a Euler Number matrix, which divides the rectangular roof into four quadrants, calculates the Euler Number in each subdomain, and construct a 2×2 matrix. The similarity distance is calculated as the sum of absolute difference between corresponding quadrants. The Euler Number matrix describes internal spatial structure so that it gives more proofs of distinguishability than Euler Number, and the low computational complexity is suitable for fast image matching.

Given a reference image and a target image, once the regional roof is detected, area index firstly chooses the candidate matches, and then histogram similarity and Euler Number matrix distance decides the best to be the match pair.

## III. EXPERIMENTS

The proposed method were carried out over an area in Yangjiang, China. The images are captured by SWDC-5 looking nadir with 60% overlapping, which supplies favorable data for image matching. The ground resolution of images is close to 10cm when the flying height is about 850m, which is clear enough to detect building roof. To test the robustness and efficiency in urban remote sensing image matching of the proposed approach, we take an experimental area with similar buildings as shown in Figure 2.

Figure 2. The experimental images taken by different views

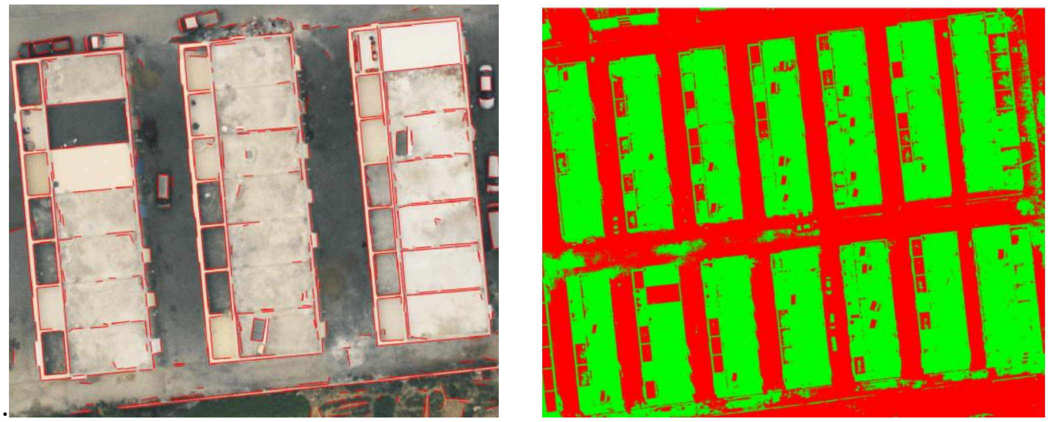

Figure 3. The line segments extracted by EDLines (left) and the classified map processed by K-Means (right)

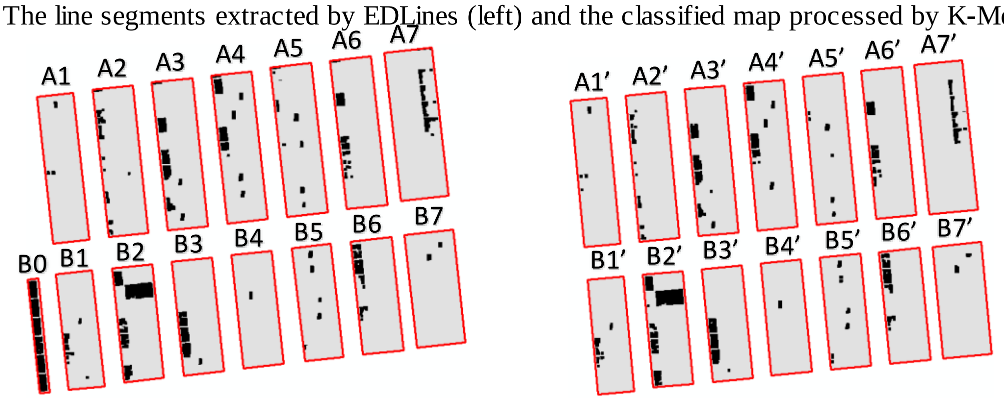


Figure 4. Regional features constructed of internal structure and contour

For each image, pre-processing of external contour and internal structure is fundamentally conducted as shown in Figure 3. As the results show, line segments are extracted soundly by EDLines detector. Although the fragmentation is remarkable in roof contour, the edge feature is fully detected so that it is feasible to restore the straight line by re-organization. Meanwhile, the classified map has properties of simple structure, quick response, and robustness.

Figure 4 represents the regional feature constructed combining contour and internal blocks. The results reveal that the adopted building extraction algorithm can remarkably restore the contour (completely accurate in the experimental area) and the classification based approach can robustly obtain simple heterogeneous distribution, which provides the stable and distinctive feature for image matching.

Once extracted the regional feature, image match can be conducted based on feature description and similarity measurement. Finally, the repetitive and poor-textured areas are matched well.

## CONCLUSION

A fast and automatic image matching method is presented for urban aerial image. First interests of roofs are produced by perceptual grouping and unsupervised classification, then three descriptors characterize and match the regional feature effectively. The experimental results reveal that introduced building roof extraction approach can remarkably improve detection efficiency and accuracy, and the proposed heterogeneous classification framework can divide internal structure validly and robustly. Meanwhile, through the constructed similarity measurements fully describing the regional feature, we have obtained rapid and promising image matching results overcoming the complex urban scenes. The experimental area is filled with repetitive patterns and poor texture, the results are almost completely correct. Thus, the proposed approach can significantly realize urban image matching rapidly, which can be applied to image retrieve, image mosaic, and geometric measure in complex urban scenes.

## REFERENCES


Atallah, M.J., Génin, Y., Szpankowski, W., 1999. Pattern matching image compression: Algorithmic and empirical results. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(7), pp.614-627.

Benedek, C., Descombes, X., et al., 2012. Building development monitoring in multi-temporal remotely sensed image pairs with stochastic birth-death dynamics. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(1), pp.33-50.

Hannah, M.J., 1988. Digital stereo image matching techniques. In: International Archives of Photogrammetry and Remote Sensing, 27(B3), pp.280-293.

Izadi, M. and Saeedi, P., 2012. Three-dimensional polygonal building model estimation from single satellite images. IEEE Transactions on Geoscience and Remote Sensing, 50(6), pp.2254-2272.

Mamic, G.J., Bennamoun, M., 2000. Review of 3D object representation techniques for automatic object recognition. In: Visual Communications and Image Processing 2000. International Society for Optics and Photonics, pp.1185-1197.

Mayer, H., 1999. Automatic object extraction from aerial imagery—A survey focusing on buildings. Computer Vision and Image Understanding, 74(2), pp.138-149.

Shokoufandeh, A., Keselman, Y., Demirci, M.F., et al, 2012. Many-to-many feature matching in object recognition: a review of three approaches. Computer Vision, 6(6), pp.500-513.

Ünsalan, C. and Boyer, K. L., 2011. Review on building and road detection. Multispectral Satellite Image Understanding. Springer London, pp.139-144.

Wang, J., Qin, Q., Chen, Li., et al., 2013.Automatic building extraction from very high resolution satellite imagery using line segment detector. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp.212-215.