An Improved Approach to Automatic Recognition of Civil Infrastructure Objects

Fengliang Xu, Xutong Niu, and Rongxing Li

Department of Civil and Environmental Engineering and Geodetic Science
The Ohio State University, Columbus, OH 43210

Abstract

Civil infrastructure objects are important elements in GIS applications. Due to the wide variety of object types, automatic recognition of infrastructure objects from imagery has been a challenging issue for the last two decades. Different approaches have been developed for recognition of buildings, one kind of infrastructure object most frequently dealt with in GIS, by defining and employing individual criteria. A Hopfield neural network can effectively combine different criteria in an overall network structure for a global optimization in finding object. In this paper we develop a holistic feature extraction approach including edge extraction, noise edge elimination by Gabor filters, contour extraction based on morphological operations, polygon simplification by local Hough transform, and building roof candidate selection using central contour sequence moment. In addition, shadows associated with buildings are extracted. This improved feature extraction approach greatly enhances the quality of recognition of objects, such as peaked-roofed and flat-roofed buildings, by a Hopfield neural network that accommodates similarity measures using the extracted features in a structured way. The achieved results demonstrate a promising approach for building recognition and can be extended to other infrastructure objects.

I. INTRODUCTION

Civil infrastructure objects, such as buildings, roads, traffic lights, and light poles, are important elements in GIS (Geographic Information System) applications. For some objects with less complex shapes, such as light poles (Figure 1a), a three-dimensional (3-D) model can easily be built. The light poles can then be extracted and recognized by using a Hopfield neural network [8] that compares similarities between the back-projected linear features of the 3-D model of the light poles and those of images [12]. More complicated objects, for example trucks on roads or parking lots (Figure 1b), can be recognized from aerial images by incorporating additional information such as the truck-shadow relationship in a 3-D truck model. A two-layer Hopfield neural network is capable of integrating linear or area features to recognize the trucks by comparing the aerial images with the 3-D truck model and the shadow information [19]. A framework of spatial feature recognition using a Markov random field model is demonstrated by an example of traffic light extraction and recognition [20].

Among civil infrastructure objects, buildings are the most frequently encountered objects, and they are difficult to model and recognize due to their complex variations in building shape, size, and other factors. Various methods have been developed for building extraction from remotely sensed data including monocular images, stereo images, and LIDAR data. Efforts based on monocular images for building extraction include Venkateswar and Chellappa [21] in which they constructed polygons by analyzing edges and right-angled corners and used shadows to verify roofs. Huertas and Nevatia [10] used shadow corners to initialize a process of rectangle construction. Lin [14] extended Huertas' work by considering scenes from a nonvertical pointing angle where parallelograms must be handled instead of rectangles. To compensate for this shape deviation, they added a wall verification in addition to a shadow verification. Irvin and McKeown [11] started by using shadows to detect the existence of buildings and then applied criteria such as shadow width and shadow consistency for verification. McGlone [17] used vanishing points to find vertical lines (walls), applied a voting space to find the two major horizontal line directions of building roofs, and then combined corners and vertical lines to detect buildings with flat and peaked roofs. Liow and Pavlidis [16] used morphological operations and quad-tree segmentation to extract shadows and identify building candidates.

Weidner and Forstner [22] calculated a high-resolution DEM from stereo images and then detected building candidates from the DEM using morphological filters. Eckstein and Munkelt [6] used texture filters to differentiate trees and buildings from DEM. Berthod et al. [2] used hierarchical parallaxes from high-resolution image serials to segment buildings from the ground. Collins [5] extracted features from single images, used feature-based matching to generate a hierarchical DEM, and then extracted buildings. Brunn and Weidner [3] used a Bayesian network to detect regions of interest (building candidates) in a hierarchical way and then extracted corners, edges, and faces from original images to reconstruct buildings. Haala [7] used stereo images to generate a DEM that is then used to match line segments for forming 3-D lines that are subsequently compared to building models. Baillard et al. [1] used feature-based matching to generate digital surface models (DSM) and

1082-4006/03/0901~2-78 $5.00

(a) (b)

Figure 1. Examples of civil infrastructure objects that can be recognized by Hopfield neural networks: (a) light poles extracted from a mobile mapping image and (b) trucks from an aerial image.

segmented and grouped 3D points in the DSM into an elevation connection graph. They then used a Markov random network to decrease the overall energy for extraction of the aboveground elements. Afterwards these elements were classified according to entropy differences into buildings and natural features such as trees and others. Xu and Li [23] proposed a fast building extraction method by merging image data and laser ranging data from an airborne scanner.

This paper presents the improved results of our continuous research efforts in civil object recognition using Hopfield neural networks [8]. Particular contributions were made in more robust feature extraction and in house recognition by comparing a number of feature similarities between building models and building candidates. The advantage of this method is that it applies individual feature similarities in an interconnected way in the neural network and does not depend on just one specific feature, which is especially effective when dealing with complex objects like buildings. The network structure is now flexible by allowing insertion of additional feature similarities, including topological ones.

II. A HOLISTIC APPROACH TO FEATURE EXTRACTION

The quality of feature extraction directly affects the results of object recognition. Robust recognition methods may be able to handle more noisy input. However, it is recommended that great efforts be made to extract high quality features so that a higher success rate can be achieved by the object recognition method. We propose a holistic approach to feature extraction that prepares feature candidates for building recognition based on Hopfield neural networks. This approach employs a set of progressive processing tools to obtain the feature candidates, including polygons from building roofs and polygons from building shadows. We use an aerial image throughout the paper to explain the progressive results of feature extraction and object recognition.

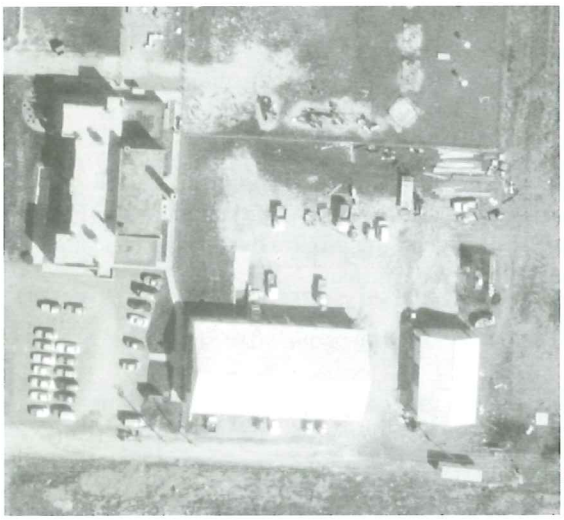
Histogram adjustment and edge extraction

Eleven images were taken using an aerial camera with a focal length of 152.7mm at an elevation of around 1,570 meters (see one of them in Figure 2a). Each of these images may have different illumination and contrast levels. To achieve a consistent result from the images, a histogram adjustment is applied for an overall brightness equalization (Figures 2b and 2d). The dark features in the original image (Figure 2a) become darker in the brightness equalized image (Figure 2b) and details in the dark areas are less identifiable. Therefore, another histogram adjustment is used to expand the lower part of the histogram and to improve the dark features (Figures 2c and 2e).

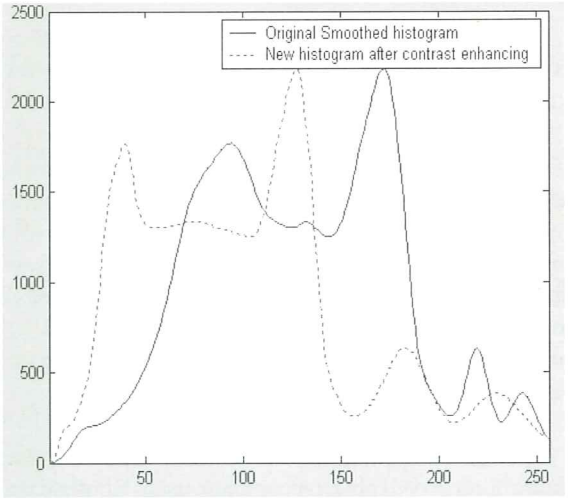
In implementation, a histogram curve is first smoothed. Significant peaks and valleys along the histogram curve are then extracted. To equalize the brightness, the extracted curve peaks are placed evenly within the entire brightness range. This is especially helpful for enhancing the contrast and making the roofs more distinguishable. Similarly, to enhance dark features, the lower part of the histogram is shifted toward the higher end. Edges are then extracted from the histogram-adjusted images (Figures 2b and 2c) using a Canny filter [4]. In Figure 3a edges are consistent in the entire image. In Figure 3b there are more details in the dark areas, which are helpful for distinguishing different dark features.

(a)

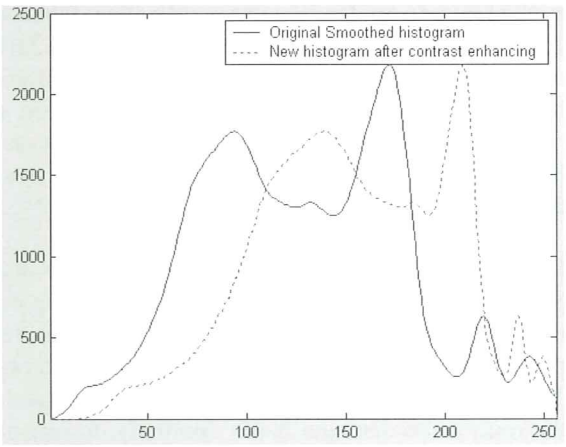
(b)



(c)



(d)



(e)

Figure 2. Two histogram enhancement processes: (a) Original image; (b) Result of a brightness equalization on (a); (c) Result of enhancement of dark features; (d) Histogram curves before (solid line) and after (dotted line) the brightness equalization; and (e) Histogram curves before (solid line) and after (dotted line) the dark feature enhancement.

Noise edge elimination

The long linear edges may contain the candidates for man-made civil objects. On the other hand, short and nonlinear edges may mainly represent noise. The linear edges of civil objects are mostly anisotropic, while the nonlinear noise edges are basically isotropic. Gabor filters, which are directional texture filters, can be used for elimination of such noise edges [24]. A Gabor filter emphasizes the fluctuation of 2-D signals in one direction and smoothes changes in the opposite direction (Figure 4a). It is expressed by the following equation:

$$Gabor_{\theta}(x, y, v_0; \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}\right)\cos\left(2\pi v_0(-\sin\theta_x+\cos\theta_y)\right) \quad (1)$$

where q is the filter direction with q_x and q_y as its components in x and y directions; v_0 is the spatial frequency for fluctuation

(a) (b)

Figure 3. Edges extracted by a Canny filter: (a) from the brightness equalized image (Figure 2b) and (b) from the dark feature enhanced image (Figure 2c)

usage; and (s_x,s_y) are parameters of the Gaussian kernel.

We perform Gabor filtering systematically in different directions to extract directional features. For example, horizontal edges receive strongest responses in a horizontal filtering output (Figure 4b) and weakest responses in a vertical filtering output (Figure 4c). For most noise edges, filtering in different directions produces similar results. Therefore, differences of Gabor filtering outputs in different directions can be used to evaluate the degree of isotropy, and the noise edges with high isotropy values are eliminated. The remaining edges are either long anisotropic edges or directional short edges (Figure 4d).

To preserve the long directional edges, the remaining short edges are then filtered out. The resulting long directional edges are subsequently used as seeds to trace long edges in the Canny filtering outputs, which represent civil objects more naturally (Figure 5a).

Region segmentation and contour generation

Regions are generated from the long edges in Figure 5a using morphological operations [18]. First, the edges are treated as background and blank areas as regions. There may be small gaps along the edges resulting in regions being connected through these gaps. Second, an erosion operation is applied using a 5x5 structuring element to separate different regions by sealing any gaps in the edges. The result is regions that are completely segmented from each other (Figure 5b). Third, a dilation operation is employed to restore the region shapes in Figure 5c with inner boundaries differentiating adjacent boundaries, and outer boundaries are overlaid on the original image (Figure 5d).

The next step is to select contours with regular shapes representing man-made civil objects. We measure the degree of regularity of a contour by a Normalized Central Contour Sequence Moment (NCCSM) defined in [18]. An r^{th}-order NCCSM is calculated as follows:

$$\bar{\mu}_r = \frac{\mu_r}{(\mu_2)^{r/2}} = \frac{\frac{1}{N}\sum_{i=1}^{N}[z(i)-m_1]^r}{(\frac{1}{N}\sum_{i=1}^{N}[z(i)-m_1]^2)^{r/2}} \qquad (2)$$

where i is the index of a point along a contour and N is the number of points of the contour; $z(i)$ is the Euclidean distance between the centroid of the contour and point i; and m_1 is the mean value of $z(i)$, called first-order contour sequence moment. A contour with an irregular shape usually has a smooth curve representing the 1st to 10th order NCCSM (Figure 6a). On the other hand, a contour with a regular shape should have a zigzag curve (Figure 6b). Therefore, roof candidates are now selected by discarding contours with smooth NCCSM curves or of small sizes (Figure 6c).

Contour generalization

The contours extracted thus far are still complicated and do not consist of simple components such as straight lines and sharp corners that express characteristics of buildings. We generalize the contours by two methods, namely, a polygon approximation and a local Hough transform.

In the polygon approximation method, a curve is recursively divided until each curve segment can be approximated by a straight line segment and the segment lengths are within a threshold. The connected line segments form a polygon (Figure 7a). This method is computationally efficient. However, it does not preserve corner information very well and may have multiple edge directions on one polygon side. The local Hough transform retains only edges in the major directions and is

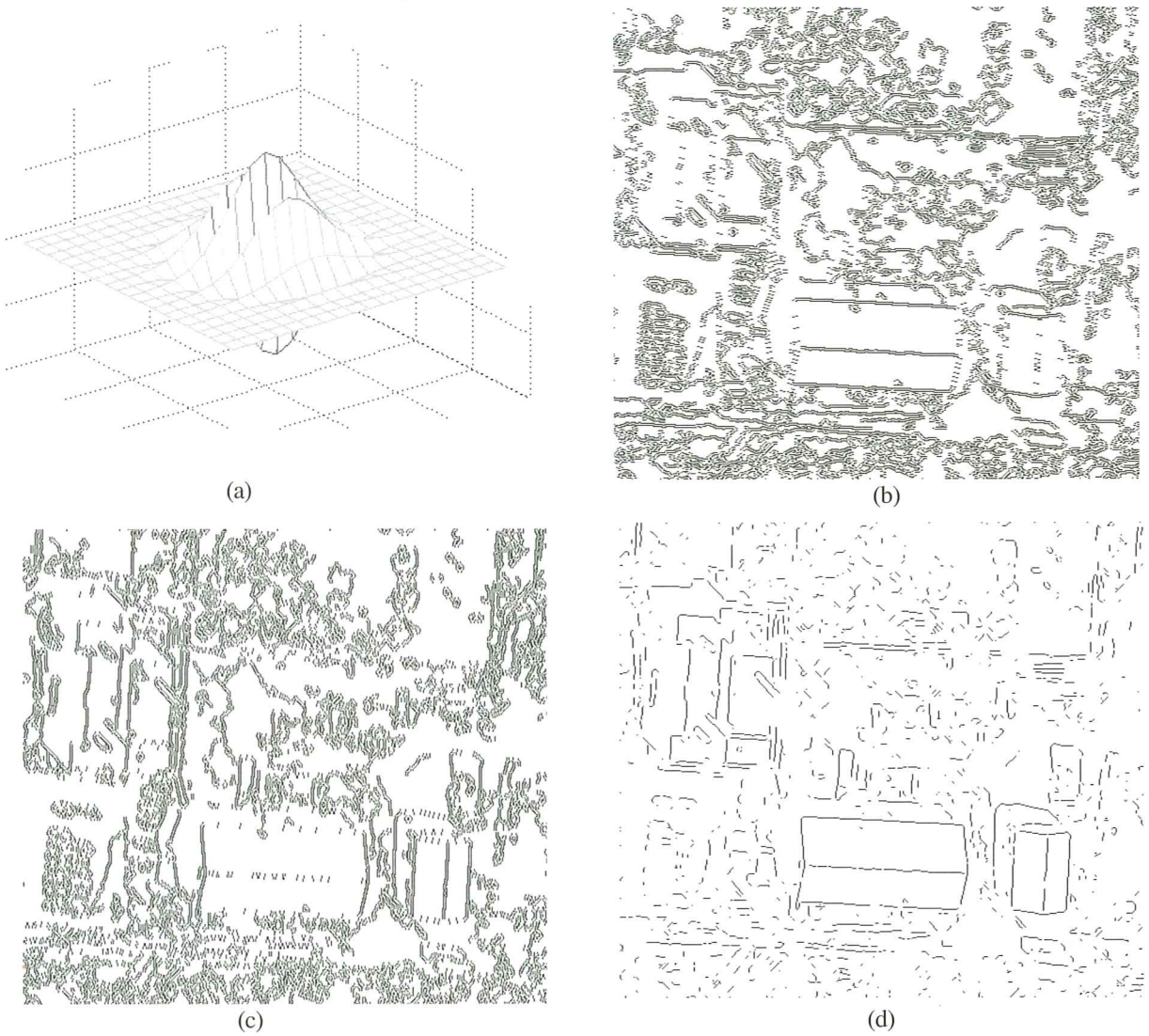


Figure 4. Noise edge elimination using a Gabor filter: (a) Shape of a Gabor filter, (b) Result of Gabor filtering in horizontal direction, (c) Result of Gabor filtering in vertical direction, and (d) Edges after noise edge elimination

more robust in preserving corners.

Hough transform [9] is well known and has a two-parameter ($\tilde{n}$ - range and $\grave{e}$- azimuth in a polar coordinate system) voting space derived from a 2-D binary edge image. A hotspot (high count) in the voting space indicates the possible existence of a straight line. For a complicated scene with many curves, the voting space may have too many local hotspots and noises, making it too complicated and time-consuming to analyze. However, in our case, we have only one contour to analyze at a time. We simplify the general Hough transform into a local Hough transform by reducing the 2-D voting space (Figure 8a) into two 1-D voting spaces to speed up the processing. The two 1-D voting spaces are the projections of the 2-D voting space in the direction of $\tilde{n}$ (Figure 8c) and $\grave{e}$ (Figure 8d), respectively. Thus, the order of computational intensity is reduced from $O(n^2)$ to $O(n)$ with n denoting the dimension of the voting space.

If a contour represents a part of a building, we assume that most line segments of the contour are in two dominant directions, which can be extracted based on the peaks of voting curves and verified in the edge images. In Figure 8b the contour describes one of the building roofs that is more complicated in shape. The straight lines show the dominant directions computed from the contour using the local Hough transform. If necessary, line segments in minor directions can be obtained by removing the contour pixels in the dominant directions and repeating the above voting process. The line segments are intersected to produce corners and to form a polygon. Figure 7b shows the resulting polygons using the local Hough transform.

Shadow extraction

Shadow extraction is an integral part of this Hopfield neural network based building recognition system. To distinguish

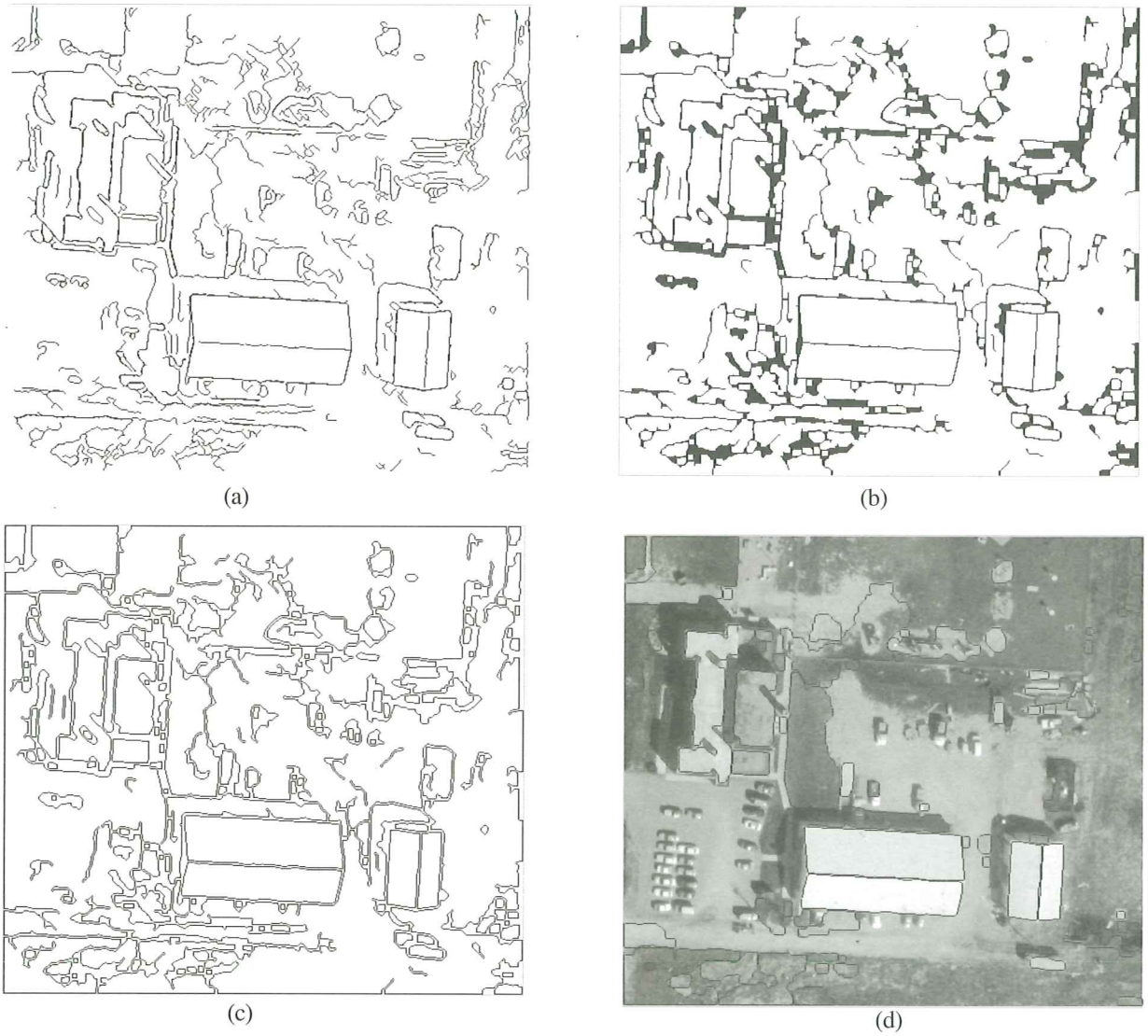


Figure 5. Roof candidates extracted by morphological operations: (a) Extracted long edges, (b) Segmented regions, (c) Regions with inner and outer boundaries, and (d) Outer boundaries overlaid on the original image

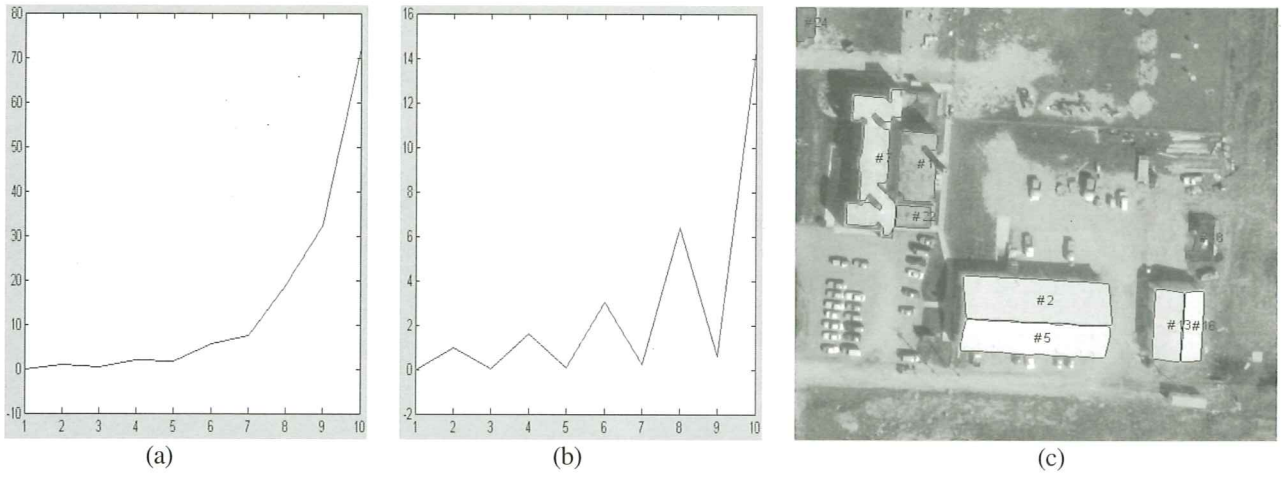
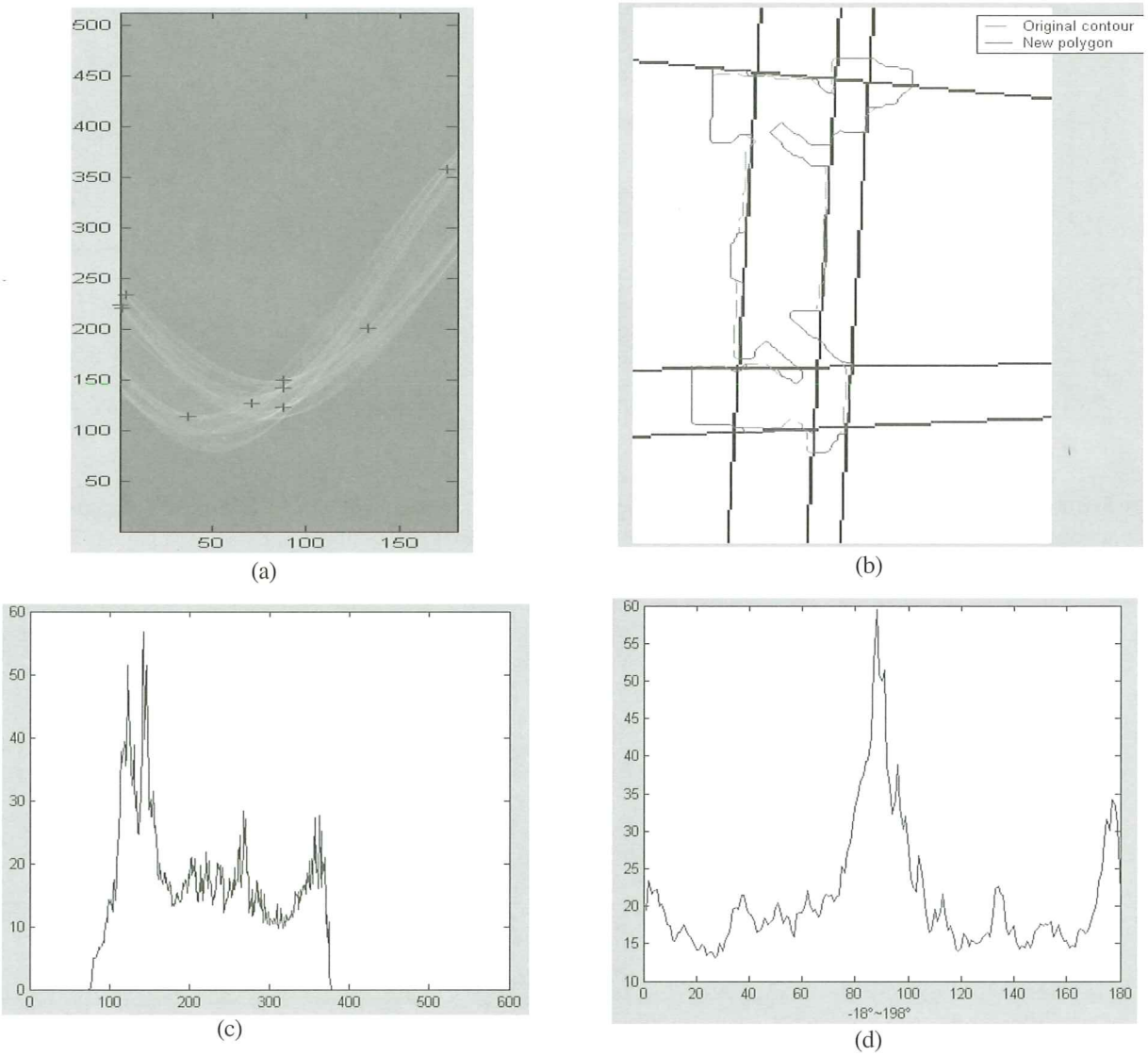


Figure 6. (a) NCCSM curve of an irregularly shaped contour, (b) NCCSM curve of a regularly shaped contour, and (c) Selected regular shaped contours

(a) (b)

Figure 7. Contour generalization: (a) Result of a polygon approximation and (b) Result of a local Hough transform



(a) (b)

(c) (d)

Figure 8. Local Hough transform for detection of straight line segments of a contour: (a) 2-D voting space, (b) Line segments in two dominant directions, (c) 1-D voting space in ? direction, and (d) 1-D voting space in ? direction

possible shadows, we first separate the original image into shadow and nonshadow areas (Figure 9a) using a histogram segmentation. Edges of the shadow areas are detected by a Canny filter (Figure 9b). These shadow areas are obviously overly segmented. For example, some grassland has similar gray scales as the shadows. Next, given a sun incident direction, line segments of the above extracted building polygons are examined. Those line segments that have the potential to produce shadows in the sun incident direction, which are not obstructed by other building polygons, are selected and called shadowing edges. Then, each shadowing edge sweeps in the sun incident direction to create an artificial shadow area. The final extracted shadow areas are obtained where matches between the shadow candidate areas in Figure 9b and the artificial shadow areas exist. Finally, additional template information such as the gray scale difference on both sides of a shadow edge (darker inside and brighter outside) is also used to maximize the success rate of shadow extraction (Figure 9c). This method performs well for flat-roofed buildings where roof edges are projected as parallel edges in the shadow polygons. For peak-roofed buildings, however, it is sometimes difficult to extract correct shadows.

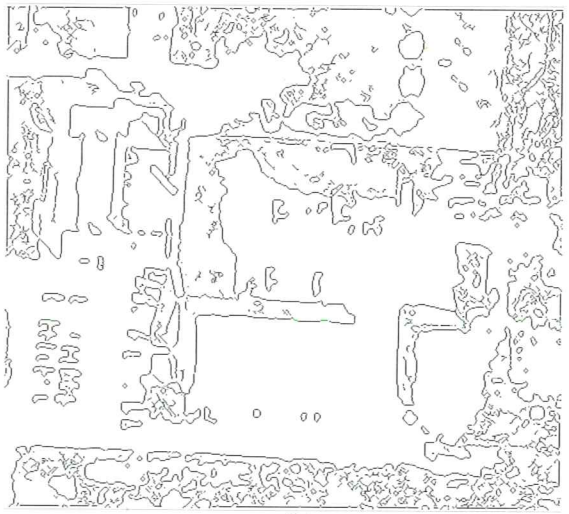
III. BUILDING RECOGNITION USING HOPFIELD NEURAL NETWORKS

The hopfield neural network

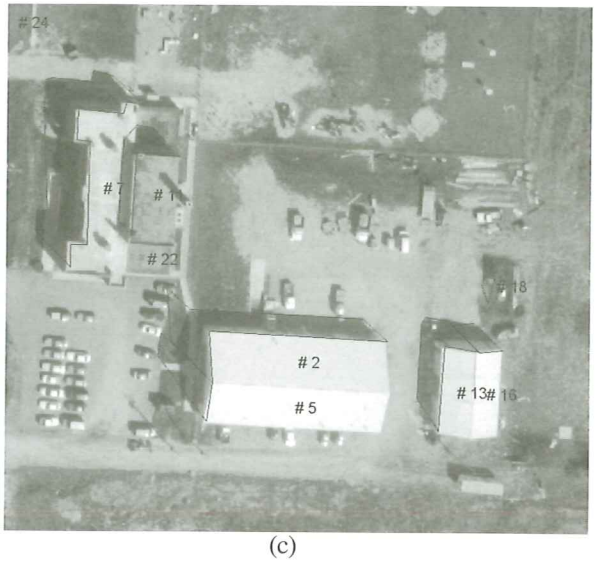
A Hopfield neural network [8] has a two-dimensional N×M grid structure with N elements of candidates and M elements of the model of the object to be recognized. Each grid node (i, k) is a neuron and all neurons are interconnected [13,15,12]. The object can be recognized by matches between the model and candidate elements through a global minimization of the following energy function:

$$E = -A\sum_i\sum_k\sum_j\sum_l C_{ikjl}V_{ik}V_{jl} + B\sum_i(1-\sum_k V_{ik})^2 + C\sum_i\sum_k\sum_{l\neq k}V_{ik}\times V_{il} + D\sum_k(1-\sum_i V_{ik})^2 + E\sum_k\sum_i\sum_{j\neq i}V_{ik}\times V_{jk} \quad (3)$$

(a)



(b)



(c)

Figure 9. Shadow extraction: (a) Segmented shadow and nonshadow areas, (b) Edges of shadow candidate areas, and (c) Extracted shadow polygons

where V_{ik} and V_{jl} are neuron states for the candidate-model pairs (i, k) and (j, l), respectively. C_{ikjl} denotes the similarity between the candidate image pair (i, j) and model feature pair (k, l) and serves as a connection between the neurons. It is expressed as:

$$C_{ikjl} = C^1_{ik} + C^1_{jl} + C^2_{ikjl} \cdot$$

$$C^1_{ik} = \sum_{n=1}^{N_1} w^1{}_n f_n{}^1(x_{in}, y_{kn})$$

$$C^2_{ikjl} = \sum_{n=1}^{N_2} w^2{}_n f_n{}^2(x_{ijn}, y_{kln}) \qquad (4)$$

where $C^1{}_{ik}$ and $C^2{}_{ikjl}$ denote unary and binary similarity measures between the candidate and model elements. $C^1{}_{ik}$ is calculated by N_1 unary similarities (e.g. N_1=5 in a peak-roofed case and 4 in a flat-roofed case). $C^2{}_{ikjl}$ is calculated by N_2 binary similarities (e.g. N_2=5 in a peak-roofed case and 2 in a flat-roofed case). W_n is the weight of the nth similarity measure, with the sum of all weights equal to 1. $f(x, y)$ is a sigmoid function where a difference of the similarity measure between the candidate x and the model y is evaluated against a threshold. $C^1{}_{jl}$ has a similar equation as $C^1{}_{ik}$. Discussions on a more detailed implementation of the network for mapping object recognition can be found in [12] and a later section of this paper. By iteratively updating the states of the interconnected neurons until the neural network is stabilized, the neuron state V_{ik} converges to 1.0 if the candidate image element i matches the model feature element k perfectly; otherwise it is equal or close to 0.

Implementation for building extraction

In this study, two types of buildings are to be recognized using Hopfield neural networks: peak-roofed buildings and flat-roofed buildings. For a typical peak-roofed house, a house model generally has two slant rectangular faces (Figure 10a). For a vertical aerial image, the base part of the house does not appear significantly and may not be included in the model.

To recognize a peak-roofed house using a Hopfield neural network, we project the 3-D house model onto the image using known image position and orientation parameters. The rectangular roof faces in 3-D space become parallelograms after the projection and thus are similar to the extracted roof polygons in the aerial image. Also, the shadow of the house model can be generated by the projection using the sun incident angle.

The recognition process is then to measure similarities between the roof candidates and the projected house model in the image space through the neural network. Both unary and binary similarity measures in Equation 4 are implemented (Figure 11). The unary similarity measures include

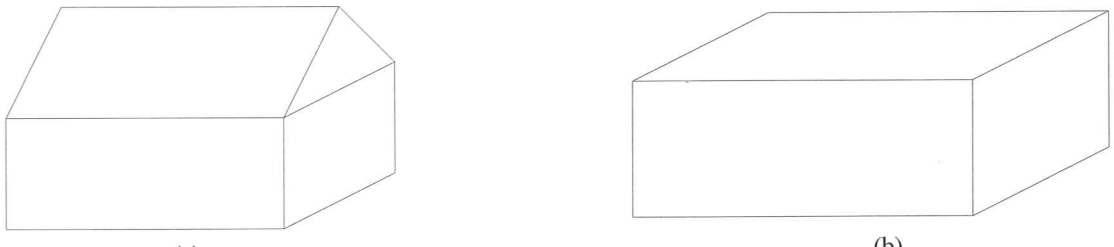
- Area and perimeter of each face (Figure 11a),
- Fourier descriptor of the face boundary that provides shape information (Figure 11b),
- Ratio between the long side and short side of the face (Figure 11c), and
- Angle between the long side and short side (Figure 11c).

Furthermore, the binary similarity measures specify relationships between the two faces:

- Gray scale difference between two roof faces caused by different normal vector directions (Figure 11d),
- Distance between them (Figure 11e),
- Ratio between the two long sides (Figure 11f),
- Ratio between the two short sides (Figure 11g), and
- Angle between the joint sides of the two faces and the short side, each from one parallelogram (Figure 11h).

The above unary and binary similarity measures between candidates and the model are evaluated to produce C coefficients in Equation 4. The threshold values were obtained based on experimental data. The similarity measure differences are amplified through the sigmoid function, and then weighted and summed to serve as the linkage between neurons in the Hopfield neural network.

For flat-roofed buildings, the 3-D building model is generalized to a box (Figure 10b). Deviation from this generalization is usually reflected in the roof shapes and would often be too complicated to represent by a generic building. In practice the relationship between the roof and the shadow plays an important role to compensate for this drawback and to achieve a high success rate.



(a) (b)

Figure 10. Building models: (a) Peak-roofed model (e.g. base height = 4.5m, length = 19.5m, width = 14.1m, roof height = 1.57m, and ratio of two sides of the parallelogram = 6.2/7.9m), and (b) Flat-roofed model (e.g. height = 4.5m, length = 15.1m, and width = 12.5m)

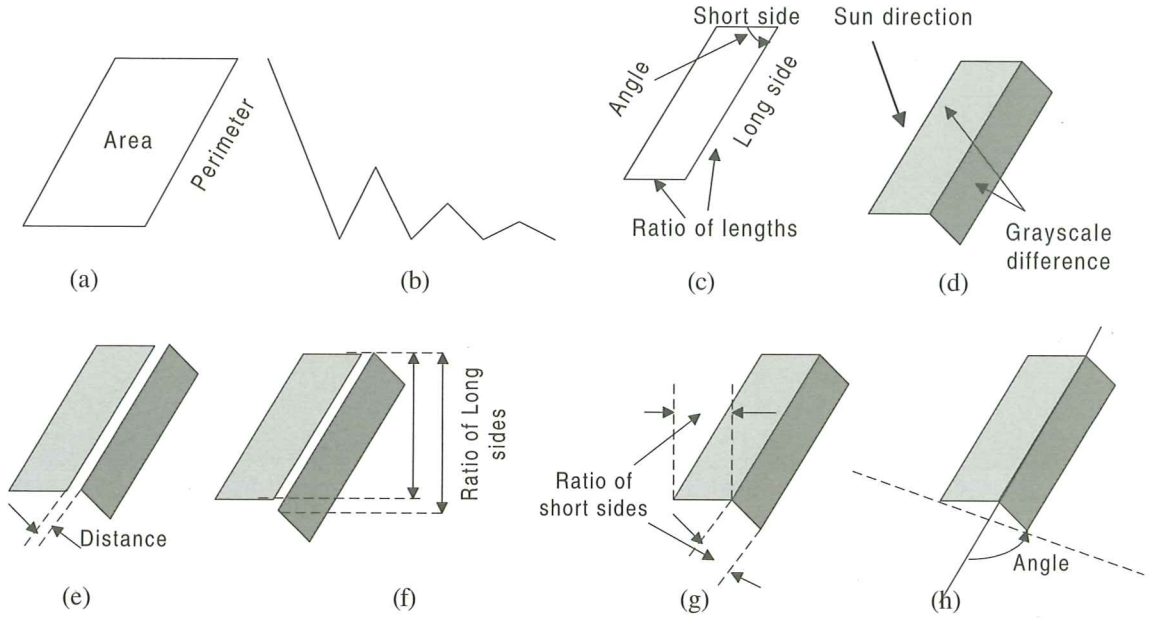


Figure 11. Unary and binary similarity measures for a peak-roof with two faces (see text)

The roof is represented by a rectangle in the image. The unary similarity measures for a flat-roofed building include

- Area and perimeter of the roof polygon,
- Mean right angle of the roof polygon (Figure 12a),
- Shadow width consistency (Figure 12b), and
- Ratio between a shadow edge (in the sun direction) and the support edge (Figure 12c).

The only binary similarity measure used for flat-roofed buildings is the distance between the building polygon and the shadow polygon (Figure 12d).

IV. DISCUSSION AND CONCLUDING REMARKS

The recognition results for the peak-roofed and flat-roofed buildings using the Hopfield neural network as well as parameters employed in the neural network computation are shown in Figures 13 and 14.

In the feature extraction stage, 79 contours were extracted after brightness equalization. Among them, only nine contours were selected by applying the NCCSM criterion. These nine contours were then improved to produce candidate polygons using a local Hough transform. From the shadow enhanced image, seven shadow polygons were also extracted. Consequently, these 16 candidate polygons were the input to the Hopfield neural network computation.

Two peak-roofed buildings were recognized correctly (Figure 13a). Furthermore, two additional polygons that passed the feature extraction process were correctly rejected. The complex flat-roofed building has two levels with shadows. It is represented by three extracted roof polygons and two shadow polygons. In the process of flat-roofed building recognition, the roof polygons were simplified as rectangles and considered as individual building candidates (Figure 14a). One of the three roof polygons of the flat-roofed building complex is not recognized as a flat-roofed building because no shadow is associated with it, while two other roof polygons are recognized as individual buildings.

Overall, the improved holistic feature extraction approach demonstrated significant effectiveness in noise suppression, roof candidate generation, and shadow extraction. The high quality feature extraction results greatly enhanced the success

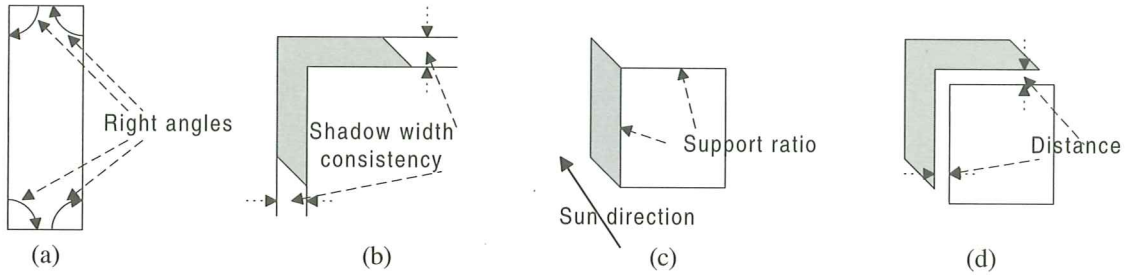
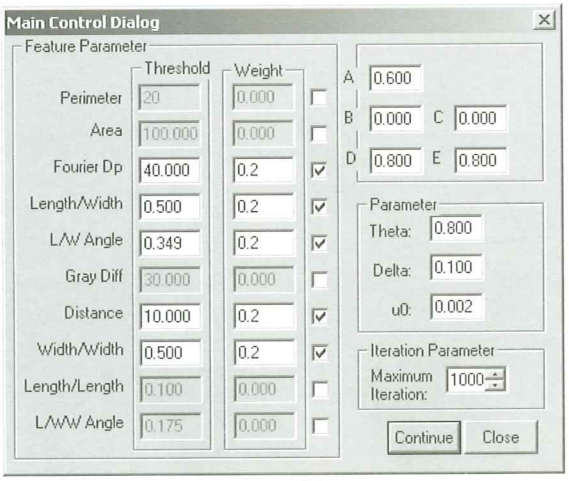


Figure 12. Unary and binary similarity measures for a flat-roofed building (see text)

(a)



(b)

Figure 13. Recognition result of peak-roofed buildings: (a) Buildings recognized in yellow, red polygons of candidates determined as not desired objects, and blue polygons as the back-projected building model; and (b) Parameters used in the network computation

rate of the building recognition process. Experiment results also show that the requirements for selection of the parameters, including thresholds, weights, and model parameters, used in the Hopfield neural network can be relaxed because of the high quality of the input data.

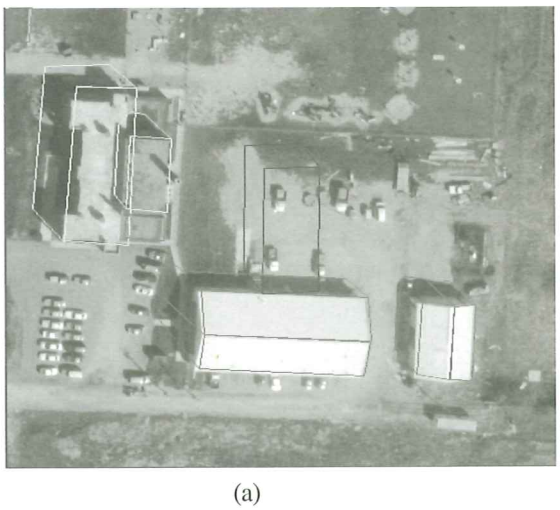
The above results demonstrated that the introduced holistic feature extraction approach performs well on aerial images for generation of quality high-level civil infrastructure object candidates. The Hopfield neural network is capable of accepting supplied object candidates from aerial images and desired 3-D object models to successfully recognize civil infrastructure objects such as peak-roofed and flat-roofed buildings. Currently, complex building structures are extracted as separate roof polygons and recognized as individual buildings. More efforts should be made to incorporate 3-D complex building models and to recognize the complex building as one object.

ACKNOWLEDGEMENTS

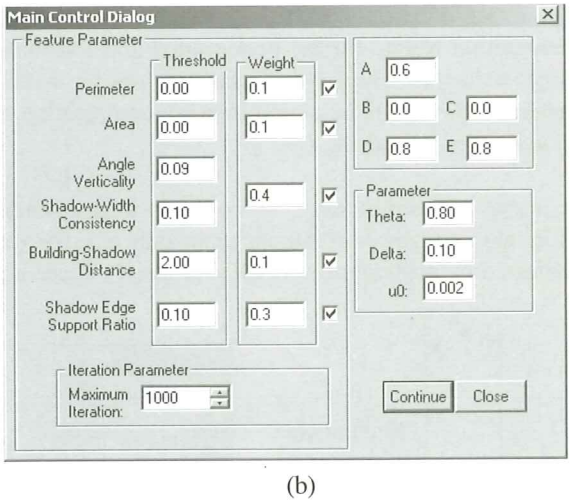
This research was supported by National Science Foundation Grant CMS-9812783.

REFERENCES

[1] Baillard, C., O. Dissard, O. Jamet, and H. Maitre, 1998, Extraction and textural characterization of aboveground areas from



(a)



(b)

Figure 14. Recognition result of flat-roofed buildings: (a) Recognized roofs and shadows in yellow, red polygons of candidates determined as not desired objects, and blue polygons as the flat-projected building model; and (b) Parameters used in the network computation

aerial stereo pairs: a quality assessment, *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(2): 130-141.

[2] Berthod, M., L. Gabet, G. Giraudon, and J. L. Lotti, 1995, High-resolution stereo for the detection of buildings, *Automatic Extraction of Man-Made Objects From Aerial and Space Images*, Grun.A, Kubler.O, and Agouris.P, Birkhauser Verlag, 1995, Basel, Switzerland, 135-144.

[3] Brunn, A. and U. Weidner, 1998, Hierarchical Bayesian nets for building extraction using dense digital surface models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53:296-307.

[4] Canny, J. F., 1986. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679-698.

[5] Collins, R. T., A. R. Hanson, E. M. Riseman, and H. Schultz, 1995, Automatic extraction of buildings and terrain from aerial images, *Automatic Extraction of Man-Made Objects From Aerial and Space Images*, Grun.A, Kubler.O, and Agouris.P, Birkhauser Verlag, 1995, Basel, Switzerland, 169-178.

[6] Eckstein, W. and O. Munkelt, 1995, Extracting objects from digital terrain models, *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, Proceedings of SPIE, the International Society for Optical Engineering*, Schenk.T, 1995, 9-10 July, San Diego, California, 2572:43-51.

[7] Haala, N. and M. Hahn, 1995, Data fusion for the detection and reconstruction of buildings, *Automatic Extraction of Man-Made Objects From Aerial and Space Images*, Grun.A, Kubler.O, and Agouris.P, Birkhauser Verlag, Basel, Switzerland, 211-220.

[8] Hopfield, J. J. and D. W. Tank, 1985, Neural computation of decisions in optimization problems, *Biological Cybernetics*, 52:141-152.

[9] Hough, P. V. C., 1962, A method and means for recognizing complex patterns, *U.S. Patent*, 3,069,654

[10] Huertas, A. and R. Nevatia, 1988, Detecting buildings in aerial images, *Computer Vision, Graphics, and Image Processing*, 41:131-152.

[11] Irvin, R. B. and D. McKeown, 1989, Methods for exploiting the relationship between building and their shadows in aerial imagery, *IEEE Transactions on Systems, Man and Cybernetics*, 19(6): 1564-1575.

[12] Li, R., W. Wang, and H. Z. Tseng, 1999, Detection and location of objects from mobile image sequences by Hopfield neural networks, *Photogrammetric Engineering & Remote Sensing*, 65(10): 1199-1205.

[13] Li, W. and N. M. Nasrabadi, 1989, Object recognition based on graph matching implemented by a Hopfield-style neural network, *IJCNN International Joint Conference on Neural Networks*, 1989, 18-22 June, Washington, DC, vol. 2, 287-290.

[14] Lin, C., A. Huertas, and R. Nevatia, 1995, Detection of buildings from monocular images, *Automatic Extraction of Man-Made Objects From Aerial and Space Images*, Grun.A, Kubler.O, and Agouris.P, Birkhauser Verlag, 1995, Basel, Switzerland, 125-134.

[15] Lin, W. C., F. Y. Liao, C. K. Tsao, and T. Lingutla, 1991. A hierarchical multiple-view approach to three-dimensional object recognition, *IEEE Transactions on Neural Networks*, 2(1), 84-92.

[16] Liow, Y. and T. Pavlidis, 1990, Use of shadows for extracting buildings in aerial images, *Computer Vision, Graphics, and Image Processing*, 242-277.

[17] McGlone, J. and J. Shufelt, 1994, Projective and object space geometry for monocular building extraction, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1994, 21-23 June, Seattle, Washington, 54-61.

[18] Sonka, M., V Hlavac, and R Boyle, 1998, *Image Processing, Analysis, and Machine Vision, Second Edition*, PWS Publishing, Pacific Grove, California, 770 p.

[19] Tu, Z. and R. Li, 1999, A multi-layer Hopfield neural network for 3-D object recognition, *Proceedings of the International Workshop on Mobile Mapping Technology* , 1999, 21-23 April, Bangkok, Thailand, vol. 7A-3, 1-6.

[20] Tu, Z. and R. Li, 2002. A framework for automatic recognition of spatial features from mobile mapping imagery, *Photogrammetric Engineering & Remote Sensing*, 68(3), 267-276.

[21] Venkateswar, V. and R. Chellappa, 1990, A framework for interpretation of aerial images, *Proceedings of the IEEE Computer Society 10th International Conference on Computer Vision and Pattern Recognition*, 1990, 16-21 June, Atlantic City, New Jersey, vol. 1, 204-206.

[22] Weidner, U. and W. Forstner, 1995, Towards automatic building extraction from high-resolution digital elevation models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), 38-49.

[23] Xu, F. and S. Li, 2001, Automatic building extraction of ALIMS, *The 3rd International Symposium on Mobile Mapping Technology*, 2001, 3-5 January, Cairo, Egypt, 7(4) (CDROM).

[24] Yamamoto, S., Y. Nakajima, S. Tamura, Y. Sato, and S. Harino, 1999, Extraction of fluorescent dot traces from a scanning laser ophthalmoscope image sequence by spatio-temporal image analysis Gabor filter and radon transform filtering, *IEEE Transactions on Biomedical Engineering*, 46(11): 1357-1363.