

Isotropic surround suppression and Hough transform based target recognition from aerial images



Hafiz Suliman Munawar *, Adnan Maqsood, Zartasha Mustansar

Research Centre for Modeling and Simulation, National University of Sciences and Technology, Islamabad, Pakistan

ARTICLE INFO

Article history:

Received 18 January 2017

Received in revised form

7 June 2017

Accepted 23 June 2017

Keywords:

Isotropic surround suppression

Target recognition

Optical images

Bridge recognition

Runway recognition

Road recognition

ABSTRACT

In this paper a procedure for target recognition of linear shaped landmarks (bridges and runways) from optical imagery is proposed. This study has done to propose a method for efficient target recognition as there is a need to have surveillance measures for defense, trade and disaster management. Keeping this in view, in this study image segmentation is performed using canny edge detection technique. Edge maps with more texture environs are suppressed by means of introducing a way that can combat isotropic surround. The inhibitor should apply lesser weights on the edges in case of texture environs as compared to the edges having defined boundaries. An efficacious method of computation for calculations of the isotropic surround suppression is used that accelerates the proposed algorithm. Subsequently, Hough transform is used on edge image along with its suppressed energy to extract all possible lines. Finally, the target regions are segmented from surrounding regions by examining geometric information and the resemblance between target and its surroundings. A number of the existing methods have been viewed and explored; the final proposal is refined to match the current trends and needs. The results indicate that the new method is efficient and effective for extracting target in optical images acquired by Unmanned Air Vehicles and it improves target detection significantly.

© 2017 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Airports and bridges play a central role and serve as the most convenient route for all intents in terms of a country's economy, travel and trade. Owing to such major roles, runways and bridges hold huge importance worldwide and also hold a big share in any country's economy, directly capturing the financial conditions of a nation. Bridge and runway detection from aerial images have vital importance in military, surveillance, disaster management and relief missions. Automatic detection of such landmarks from multispectral aerial / satellite images is an emerging area of research with articulate demands in military, security and civil sector. Specifically, during floods/earthquakes, relief operations are carried out by Unmanned Air Vehicles (UAVs). They can provide the damage assessment of bridges/runways for community relief, rescue and help identifying alternative transport networks.

Therefore it is critical to have these locations under surveillance. There is a need to enhance technology and make automatic detection of such strategic locations. These aspects form the motivation of this paper. Over the last few years, extensive research has been carried out on detection of bridges from infrared, Synthetic Aperture Radar (SAR) and visible images. Specifically, several Knowledge-Based Methods (KBMs) (Wenguang et al., 2009) are developed that work on a specific set of rules. These KBMs utilize the difference between the object and background and spatial relationship. For-instance, in SAR images (Grigorescu et al., 2004):

- A bridge is always over water and utilizing this contextual information, the search area for bridge detection can be reduced.
- A bridge stands out from the other elements in the image, as it has strong backscattering in SAR images due to its metallic structure.
- On the other hand, water has smaller backscattering because the echo reflected by water is weak.
- A bridge is approximately perpendicular to the river and its length is related to the width of the river.

* Corresponding Author.

Email Address: hafizsuliman91@gmail.com (H. S. Munawar)

<https://doi.org/10.21833/ijaas.2017.08.006>

2313-626X/© 2017 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Similarly in infra-red and optical images, the methods to detect bridge assume the following a priori knowledge (Mumford-Shah Model (Min et al., 2006):

- (a) Grey levels of a bridge and the land are higher as compared to water.
- (b) Area encompassed by river is larger than bridge's area.
- (c) The edges of a bridge are considered to be two near parallel lines.
- (d) The river is divided into two homogeneous regions by the bridge situated across it.
- (e) The length of a bridge is greater than its width.
- (f) When a bridge is close to the image acquisition system, it can be clearly seen as the divider between the two sides of a river. However, when seen from afar, the bridge appears as a thick dark line.

However, the background complexity is enhanced in terms of piers and linked cables. In addition, in low detailed image the area ratio of bridge or runway to image is very small. It becomes hard to recognize target perceived by information driven procedure.

Abu-Jbara et al. (2015) presented an approach for runway detection. The potential application of this method is for vision-based navigation. Specifically, a Region of Interest (ROI) in the image based on a runway model that is two line segments, one for each edge of the runway. The typical runways are flat enough to be considered as two parallel straight lines, which may appear non-parallel due to perspective projection in the image. Although runways may consist of multiple intersecting runways, in such a case the assumption of two line segments may fail.

Similarly, earlier studies by Gupta and Agrawal (2007) for runway detection comprise of two main stages. First, the binary classification based on textural properties and later on the analysis of these specified regions is based on their shape. Subsequently, the shape detection algorithm that discovers long parallel lines segments on the possible runway regions is applied. However, this technique is limited only to orthogonal images where parallel lines exist.

Recent method, proposed by Abuthahir et al. (2014), for runway detection work on textural properties. It uses different mathematical parameters that include mean, standard deviation of image intensity, Zernike moments, Circular-Mellin, Haralick, Fourier power spectrum, wavelets, and Gabor filters. Finally, Adaboost classification method is used for runway detection. This approach takes about 5 seconds to classify all the blocks in a test image.

Although bridge extraction from aerial images is not extensively studied yet, but still a few methods are there which have been analyzed and have also been implemented so far. Abraham and Sasikumar (2014) used fuzzy threshold segmentation for pre-

processing and remove the small undesired objects and fill the bridge gaps by a sequence of morphological erosion and dilation. The technique used is classical and uses simple thresholding which fails with the intensity change (Liu et al., 2013) proposed bridge detection in forward looking infrared images that uses Gabor filter in fixed orientation. His proposed method considers horizontal bridges with piers. Bridges have no standard structure that can be considered it vary from each other in many aspects whether it have piers, cables etc.

Chaudhuri and Samal (2008) used knowledge-based method by defining a set of rules for initial classification of the multispectral images of bridges into eight land-cover types (Chaudhuri and Samal, 2008). The technique commonly known as 'supervised classification' is done at three levels. However, this technique has a drawback that it is likely to overlook a narrow bridge if that happens to be present alongside other bridges of varying width in any particular image. These techniques cannot wipe out the interference of either the shoals or the water clutter.

Another group proposed method uses fractal theory in conjunction with contextual information to identify the ROIs for bridge detection (Yan et al., 2007). The grey image is converted into a three-dimensional fractal surface to assess the complexity and irregularity of the textures of image elements. Dynamic threshold segmentation and fractal threshold segmentation is applied to detect rivers. Numerous other algorithms (Yuan et al., 2003; Trias-Sanz and Loménie, 2003; Ando, 2000; Ballard, 1981; Zai-hua and Shu-qian, 1998) recognize bridges by knowledge or context-based information. These method mostly uses Otsu threshold, fuzzy threshold, parallel lines detection scheme, and by line detection using standard Hough Transform.

Grayscale differentiation between bridges/runways and their surroundings add to the complexity of the problem. Under cluttered background in infrared images, the gray level difference between bridge and water gets low or even bridge or dam body looks darker than water (Liu et al., 2013). This difference exists clearly between the runway and its surroundings. Recognition of the bridge using simple thresholding method gives positive results on shoals. It is obligatory to differentiate between bridge and shoal. On the other hand, it is difficult to extricate the parallel lines as there exist mostly a single-straight line.

In our approach presented here, we have used simply determinable features that best describe the textural changes near crosswise region of the target. Textural features that include mean and standard deviation of the gradient of the image and the image intensity are used. Further sections of this paper describe the following: Section 2 is Region of Interest (ROI) extraction; Section 3 is devoted to target recognition in light of fractal hypothesis; and Section

4 provides experimental results followed by conclusion.

2. Roi extraction

Region of Interest (ROI) is extracted using line detection. For this firstly, the Canny (1986) edge detection is used to extract edges on RGB/synthetic images as shown in Figs. 1a and 1b. Canny edge detection method uses a multi-stage algorithm to detect a wide range of edges in images. The Canny algorithm works on multiple adjustable parameters that affect the computation time and overall effectiveness of the algorithm. To maximize the optimization of canny operator, threshold parameters can be determined by Otsu (1979) a method of threshold segmentation that achieves automatic optimum threshold selection through utmost inter-class variance as shown in Fig. 1c.

Based on the gray features of images, classification into background and target can be achieved on images. The Otsu method has an advantage of easy operation and fast processing, this ensures that the method is widely used in the threshold segmentation. Further this threshold can be utilized as a part of the canny calculation to identify the object's edge. Also size of the Gaussian (smoothing) filter used in the canny algorithm directly affects the results. Using a larger sized filter causes more blurring. If we spread out the value of a given pixel over a larger area of the image, this scheme become more useful for detecting larger and smoother edges as shown in Fig. 1d.

Once the edge image is formed, its magnitude is used to formulate edge energy by isotropic surround suppression at any given pixel (x, y) . Lower weights are carried by the edges with strong surround suppression (Guo et al., 2009) resulting in the formation of peaks. The edges in texture regions and complex backgrounds have votes which are of less implication and thus lowered or even demolished by surround suppression. However, the edges that has peaks gathered from clear boundaries between dissimilar entities, like sky (background) and buildings or vegetation and roads slightly get affected by surround suppression.

In this paper the surround suppression used is preferred owing to its efficacy and an intuitive method of pointing out the confusion about edge pixels that are supposed to be eradicated. Results of Isotropic surround suppression are shown in Fig. 3. Let $I(r)$ be an input gray level image, with $\mathbf{r} = (x, y) \in \mathbb{R}^2$, and let $\nabla\sigma I(r)$ be its Gaussian gradient defined as the convolution between $I(r)$ and the gradient of a Gaussian function $g_\sigma(r)$ (Eq. 1)

$$\nabla\sigma I(r) \triangleq (I * \nabla g_\sigma)(r), g_\sigma(x, y) \triangleq \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

The inhibition term is computed as a weighted local average of the Gaussian gradient magnitude $|\nabla\sigma I(r)|$ over a ring around each pixel. While weighting function $w_\sigma(x, y)$, defined as a normalized difference of Gaussians (Eq. 2),

$$w_\sigma(x, y) = \frac{1}{A} \left| \frac{1}{2\pi k^2 \sigma^2} e^{-\frac{x^2+y^2}{2k^2\sigma^2}} - \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \right| \quad (2)$$

where A is a normalization factor defined such that (Eq. 3)

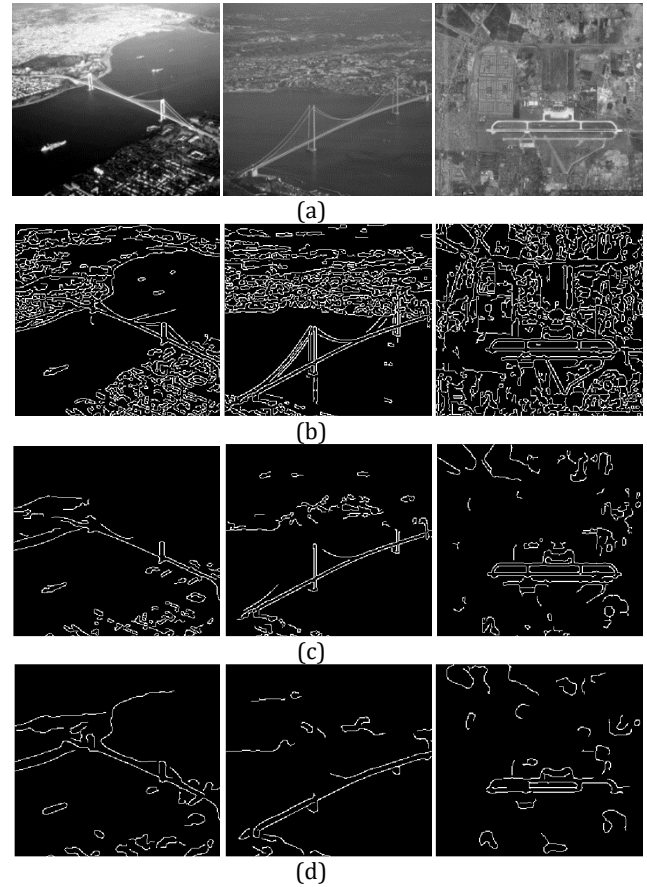


Fig. 1: a) Input images b) Canny edge output c) Canny edge output with Otsu Thresholding Method d) Canny Edge output with Otsu Thresholding and Larger Gaussian Mask

$$\iint w_\sigma(x, y) dx dy = 1 \quad (3)$$

and k is the ratio between the scale parameters of the two Gaussians, whose ideal value has been found to be 4 and the symbol $| \cdot |^+$ is defined as (Eq. 4)

$$|u|^+ \triangleq \begin{cases} u, & u > 0 \\ 0, & u \leq 0 \end{cases} \quad (4)$$

The weighting function w_σ with $\sigma = 1$ is shown in Fig. 2. The inhibition term $T(r)$ is computed as the convolution of the Gaussian gradient magnitude $|\nabla\sigma(r)|$ with the inhibition filter $w_\sigma(r)$: In practice, the influence of points having larger distances ($>8\sigma$) from origin are negligible, and in such a case, the weighting function shows a weighted ring-shaped neighborhood centered at the pixel being considered. When such a region. i.e. the 8σ -radius region is used, the entire region size is $(2 \times 8\sigma + 1) \times (2 \times 8\sigma + 1)$ (Eq. 5)

$$T(r) \triangleq \{ |\nabla\sigma I| * w_\sigma \}(r) \quad (5)$$

Further ROI is obtained by means of straight line extraction using Hough transform. Extracted features from the input image are used for

generating voted for the parameter sets by mapping the image on parameter space. Features with a high number of votes are identified by looking for significant local maxima in the accumulator array.

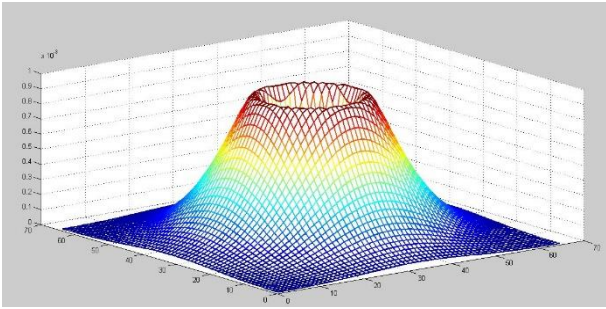


Fig. 2: Weighting function w_σ

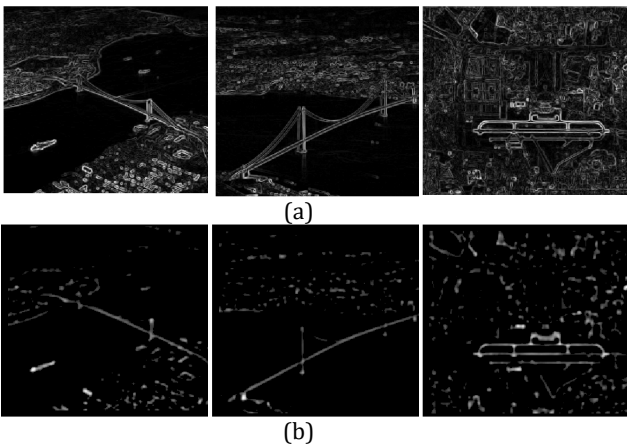


Fig. 3: Edge magnitude as energy (a) without isotropic surround suppression (b) with isotropic surround suppression

Hough transform and phase grouping are the basis of many classical methods for line grouping. In terms of real world images, the Standard Hough transform and its derived methods with standard voting scheme give low detection rates when used on real world images. One major reason for this decline is the non-linearity of edges. Further, any false peak with high votes in the Hough space results in suppression of a nearby true peak, which then leads to a missing line. Observations show that, in general, real-world images, having complex backgrounds or texture regions in the images causes noise edges comprising a substantial portion, which degrades the result quality of line detection. These detailed edges in texture regions or complex background typically do not possess high perceptual importance (Knierim and Van, 1992) when segmentation of target body from background or from each other needs to be extracted such as locating the region of interest among the surrounding artificial objects. Also, some texture regions may not be having real lines, likewise, lines produced by grills of the bridge. These lines do not describe distinct features defining the object. Instead they are regarded merely as the components of the texture. Suppressing the influence of edge pixels is reasonable in case of the pixels formulated by complex texture regions and backgrounds so that

quality of peaks detected in Hough space may be lifted.

Jones et al. (2001) proposed that from neurophysiology, this point is also provided that, the presence of a complex surrounding decreases the perceptual importance of the point under concern in human visual system. We have used an approach that works by giving a weight to each of the edge according to its position while being voted into Hough space, depending on strength of the surrounding suppression in this position. Fig. 4 shows the comparative results of image Fig. 1a of both techniques using standard Hough transform Fig. 4a and an improved Hough transform voting scheme utilizing surround suppression in Fig. 4b, which shows smooth form of Hough space.

By bringing forward a method of isotropic surround suppression, edge pixels are treated differently. By giving minor weights to texture edges and larger weights to strong and perceptual, instead using edge intensity idea is to gather votes in Hough space. This results in reduction of false peaks that are formed by texture edges. These edges are suppressed in order to enhance the overall quality for the detection results.

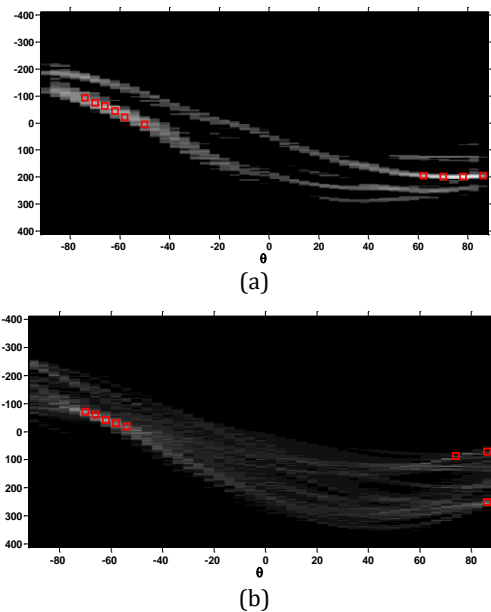


Fig. 4: a) Peaks shown on Standard Hough space b) Peaks shown on Hough space using voting scheme

3. Target body extraction

This section describes the detection of linear targets using the resulting line(s) which are ROI positions. Proposed method doesn't require the whole water area segmentation. Instead variance map is used to measure non homogeneity around the target. Variance map of an image is computed by taking a disk-shaped window of a set size around a center pixel as shown in Fig. 5. The variance within the window can be computed using following Eq. 6.

$$Var(X) = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N-1} \quad (6)$$



Fig. 5: Variance map of an image

Research so far illustrates that for urban areas (area with more features), as compared to that of rural areas, rule based verification helps in finding linear targets. The rule based verification sequence which helps to reduce the false alarm in detection of the bridge is as follows.

Considered linear targets commonly are brighter than their surroundings and have a uniform gray level. Hence, the intensity levels and the variation can be defined by the means and the variances of intensity, and the gradient of intensity inside the image blocks, respectively.

Finally, the target is extracted by crosswise area analysis by examining the following:

1. The line(s) which are less than threshold length L_0 are eliminated.
2. Grayscale difference on both side of the line. This should be minimal in case of bridge and should be greater in runway.
3. Variance score on both sides should be minimal.
4. Mean value for bridge should be greater than its crosswise mean value. If bridge pixels are darker than its surroundings it may be bipolar issue. In that case both side pixels mean value should be greater than bridge or dam mean value.
5. Length to width ratio should be greater.

Simple thresholding methods often miss the true segmentation of the objects. For this, pixel traversing is used to measure orthogonal intensities alongside target where minimum grayscale values correspond to high value area.

Non homogeneity of surrounding exists in those areas where greater extent of terrain changes or man-made structures is constructed. Even in that case saliency remains high. Proposed algorithm locates the ROI sub images by means of first showing single straight line extraction and then evaluating for qualification as shown in Fig. 6.



Fig. 6: Linear target detection results using proposed method

4. Computational time

The effectiveness of proposed algorithm is shown in this section. To validate the efficacy and the extent of the proposed algorithm, 4 images are tested in

this paper Experiments are performed using MATLAB on a machine with 2 GHz processor and memory of 4 GB. It takes less than 1s to process one image. On image size of 576 x 768 following are the computation time (Table 1).

Table 1: Computational time

| Methods | Image 1 | Image 2 | Image 3 |
|----------------------|----------|----------|----------|
| Canny | 0.057848 | 0.051521 | 0.046558 |
| Surround Suppression | 0.477023 | 0.467194 | 0.500323 |
| Hough | 0.356396 | 0.315913 | 0.282036 |
| Total Time (s) | 0.8913 | 0.8346 | 0.8289 |

5. Conclusion

In this research, a multi-stage technique to detect bridges and runways from low and high oblique aerial images instead of straight down (orthogonal views) is proposed. The camera and environmental parameters are assumed to be known apriori. During the first stage, corner and edge detection is carried out by using gradient covariance. Edges with more texture environs are suppressed by introducing a measure of isotropic environs and allocating large weights to the clear boundaries and the strong edges. For the second stage, an enhanced Hough transform voting scheme is used on edge image along with its energy to extract all possible lines as per the relevant description of the target orientation and length.

The target is identified by evaluating crosswise feature examination that seeks self-resemblance across the target. An efficient method of computation for the calculations of the isotropic surround suppression is used that accelerates the proposed algorithm. Result shows that this approach consists of basic and effective method with as low as 0.9 s to recognize the target bridge from the image.

The focus for future work will be mainly in two directions. Firstly, the current work is designed in order to train the object designer in such a way that it identifies objects pertaining to one particular category. It will further be extended for the detection of different objects pertaining to multiple categories. Secondly, this current work focused on using the spatial information but the future aims include combining it with rich spatial information in order to achieve better and more precise object detection.

References

- Abraham L and Sasikumar M (2014). Detection of bridges using different types of high resolution satellite images. *International Journal on Recent Trends in Engineering and Technology*, 10(2): 107-116.
- Abu-Jbara K, Alheadary W, Sundaramorthi G, and Claudel C (2015). A robust vision-based runway detection and tracking algorithm for automatic UAV landing. In the *International Conference on Unmanned Aircraft Systems*, IEEE, Denver, USA: 1148-1157. <https://doi.org/10.1109/ICUAS.2015.7152407>
- Abuthahir A, Arasi MM, Kumari KGE, Nagarajan P (2014). Airport runway detection based on adaboost algorithm. *International Journal of Advanced Research in Electronics and Communication Engineering*, 3(11): 1481-1485

- Ando S (2000). Image field categorization and edge/corner detection from gradient covariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(2): 179-190.
- Ballard DH (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2): 111-122.
- Canny J (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, 8(6): 679-698.
- Chaudhuri D and Samal A (2008). An automatic bridge detection technique for multispectral images. *IEEE transactions on geoscience and remote sensing*, 46(9): 2720-2727.
- Grigorescu C, Petkov N, and Westenberg MA (2004). Contour and boundary detection improved by surround suppression of texture edges. *Image and Vision Computing*, 22(8): 609-622.
- Guo S, Pridmore T, Kong Y, and Zhang X (2009). An improved Hough transform voting scheme utilizing surround suppression. *Pattern Recognition Letters*, 30(13): 1241-1252.
- Gupta P and Agrawal A (2007). Airport detection in high-resolution panchromatic satellite images. *Journal of Institution of Engineers (India)*, 88(5): 3-9.
- Jones HE, Grieve KL, Wang W, and Sillito AM (2001). Surround suppression in primate V1. *Journal of Neurophysiology*, 86(4): 2011-2028.
- Knierim JJ and Van Essen DC (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67(4): 961-980.
- Liu S, Sun G, Niu Z, and Chen Z (2013). Bridge recognition based on Gabor filter in forward-looking infrared images. In the 8th International Conference on Multispectral Image Processing and Pattern Recognition, International Society for Optics and Photonics, Wuhan, China: 891803-891803. <https://doi.org/10.1117/12.2031233>
- Min L, Tang YD, and Shi ZL (2006). Segmentation and recognition of bridge over water based on Mumford-Shah model. *Infrared and Laser Engineering*, 35(4): 499-504.
- Otsu N (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1): 62-66.
- Trias-Sanz R and Loménie N (2003). Automatic bridge detection in high-resolution satellite images. In the International Conference on Computer Vision Systems, Springer Berlin Heidelberg, Heidelberg, Germany: 172-181. http://doi.org/10.1007/3-540-36592-3_17
- Wenguang W, Jinping S, Ru H, and Shiy M (2009). Knowledge-based bridge detection from SAR images. *Journal of Systems Engineering and Electronics*, 20(5): 929-936.
- Yan X, Shi C, Zhao BJ, and He PK (2007). Recognition of bridges over water in infrared image based on fractal theory and background knowledge mining. *Systems Engineering and Electronics*, 29(7): 1031-1033.
- Yuan XH, Jin LZ, Li JX, and Xia LZ (2003). Recognition of bridges over water through detecting and analyzing regions of interest. *Journal of Infrared and Millimeter Waves*, 22(5): 331-336.
- Yuan Zai-hua and Yang Shu-qian (1998). Automatic recognition and tracking techniques for infrared bridge image. *Infrared and Laser Engineering*, 27(1): 4-8.