The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

A Method of Target Recognition from Remote Sensing Images

Yili Fu, *Member, IEEE*, Kun Xing, Xianwei Han, Shuguo Wang

Abstract—**According to the characteristics of airfield and harbor from remote sensing images, a method of large target recognition based on the combination of target region and shape features is presented. First, edge detection and improved Hough transform are used to select line segments, the region including regular-array line segments in image is considered as region of interesting (ROI). ROI detection is the base for recognition. Target geometry shape is extracted from ROI using optimum threshold segmentation, which removes location effect and improves efficiency. As calculating shape principal orientations, all shapes are rotated to the same horizontally right to avoid rotation effect. The features extracted from shape implement multi-levels representation with moment features, normalized moment of inertia, length-width ratio and compact ration. Finally, feature vectors are normalized to measure similarity between target and template. Experiments show that target regions can be located accurately using ROI detection and it is effective for target recognition. Besides, the extracted features have good invariability with respect to rotation, translation and scaling, and they comprise local and overall consistency of the target, therefore, the recognition results meet expectations well.**

I. Introduction

Target recognition is a way of detecting, locating and classing using target features extracted from images. Remote sensing image is of great importance in military reconnaissance, precision attack and civil activities, but the targets are difficult to identify because of their great information, susceptible to interference and multi-source. In many cases, objects can be distinguished by their shapes, which is different from color and texture [1]. Good shape descriptors should be accurate, simple and easy-to-use. It also should have good invariability with respect to rotation, translation and scaling [2]. There are two ways to describe shapes: edge-based and region-based. Edge-based methods use edge information, like edge curves, edge direction histogram, corner point and point of interesting, etc [3,4]. While region-based shape detection uses gray distribution features inside the region, such as invariable moments and wavelet transform [5]. Region-based methods are much easier to obtain semantic information. Most image retrieval methods were operated like this [6-8]: segment image and extract features, then match vectors to indicate the most similar muster. However, there are still disadvantages: First, for remote sensing image, segmentation technology cannot extract target accurately at present, many regions in image are interference; second, the extracted features are not satisfied with similarity measure.

The central contribution of this paper is an efficient method for recognition of large artificial target in panchromatic remote sensing images. This approach is based on the combination of target region and shape features, and avoids such disadvantages mentioned above. A new voting procedure is developed for extraction of line segments using the Hough transform. The region including regular-array line segments is considered as ROI. We also construct geometrical invariant features to measure similarity for target retrieval. It allows a software implementation to perform while being robust to the target recognition and navigation. Data are from one high-resolution satellite.

The remaining of the paper is organized as follows: Section II describes the presented algorithm, whose results, advantages, and limitations are discussed in Section III. Section IV summarizes the paper and points some directions for future exploration.

II. Algorithm Description

A. ROI Detection

The topography and landform beside targets in remote sensing images are usually complex. For artificial targets, there are obvious linear features, such as edge or contour of airfield runway and harbor breakwater. So we identify ROI mainly based on linear features: the region with regular-array line segments is probably the region with target.

There are three steps of line segment extraction: edge detection, preliminary line extraction and Hough transform (HT). First, Canny operator is used to detect edge points. Canny operator has advantages of smoothing image, binarizing image and suppressing the big noise caused by rough surface. Second, use chain code tracing to extract preliminary lines, which are formed by some collinear feature points. Preliminary lines are defined as $L(x,y)=\{L_m(x,y)\,|\,m=1,2,3,\cdots\}$ stored with a set of collinear feature points (x,y) according to scanning priority. m is the symbol of the lines. Preliminary line extraction can remove no value short lines by the threshold number of

Manuscript received February 28, 2009. This work was supported the National Natural Science Foundation of China (No.60575016).

Yili Fu is with the State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin, 150001, China. (E-mail: meylfu@ hit.edu.cn).

Kun Xing is with the State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin, 150001, China. (Corresponding author to provide phone: 086-13836147540; e-mail: xingkun@ hit.edu.cn).

Xianwei Han is with the State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin, 150001, China. (e-mail: xianweihan@ 163.com).

collinear points and reduce calculated amount for next HT, however the line is not the strict straight line and the information is not completely obtained.

HT can conveniently extract any analytic curve [9]. But it has proven had some limitations for extraction of image straight lines, as better functions were developed and the power assist of HT became effective [10-12]. Compared with improved HT mentioned above, we develop a new voting process that votes are cast using many-to-one mapping scheme. The new process aims to avoid loss of spatial information as well as to eliminate spurious peaks and reduce discretization errors. A merging procedure is applied to fit accumulator data and to check the endpoints dynamically.

In order to reduce spatial complexity, Our HT is used respectively for each extracted line $L_m(x,y)$. The improved process can be described as follows:

1) Two feature points $L_m(x_j,y_j)$ and $L_m(x_k,y_k)$, extracted from line space $L_m(x,y)=\{L_m(x_i,y_i)\,|\,i=1,2,3,\cdots\}$ and $L_m(x,y)=\{L_m(x_i,y_i)\,|\,i=j+1,j+2,j+3,\cdots\}$ according to sequence, constitute point pairs. In order to reduce calculated amount, the distance between two points must be greater than a threshold T_1.

2) For each point pairs, calculated

$$\theta_{jk}=\arctan(\frac{x_j-x_k}{y_k-y_j}) \tag{1}$$

$$\theta_{jk}=\begin{cases}\theta,\theta\geq 0\\ \pi+\theta,\theta<0\end{cases} \tag{2}$$

$$R_{jk}=[y_j-\frac{y_k-y_j}{x_k-x_j}\cdot x_j]\cdot\sin\theta_{jk} \tag{3}$$

3) On the basis of HT principle, every feature point $L_m(x_i,y_i)$ and θ_{jk} can determine one straight line in image space.

$$\rho=x_i\cos\theta_{jk}+y_i\sin\theta_{jk} \tag{4}$$

$S_{R_{jk}\theta_{jk}}(x,y)$ is defined as a set of voting points received by

$$S_{R_{jk}\theta_{jk}}(x,y)=\left\{\begin{array}{l}(x,y)\,|\,\left|R_{jk}-x\cos\theta_{jk}-y\sin\theta_{jk}\right|\leq\delta\\ (x,y)\in L_m(x,y)\end{array}\right\} \tag{5}$$

An accumulator $A(R_{jk},\theta_{jk})$ stores the number of voting points. During the voting phase, collinear edge points increase the same accumulator $A(R,\theta)$ and are assigned to the $S_{R\theta}(x,y)$. Such a line representes also a common reference system with origin in the foot point (R,θ) of the normal to the line. After all the points in the line are processed by Eq. (5), point pairs need re-select to calculate by Eq. (3). At last we get several accumulators, $A_{\max}$ is the maximal number in all accumulators, if $A_{\max}$ is greater than the threshold T_2, the corresponding line is the real line in image space. Voting process for one line is end.

Unlike other approaches, the voting operation is restricted to the single ρ value, and vote spreading is allowed. However, as all maxima detection, a merging procedure is applied to recover the spatial information spread in neighbor buckets. For all $S_{R\theta}(x,y)$, (R_i,θ_i) and (R_j,θ_j) are compared in order to check whether they are the same straight line, the following fusion rules are applied:

$$\begin{cases}\left|\theta_i-\theta_j\right|\leq 2\Delta\theta\\ \left|R_i-R_j\right|\leq 2\Delta R\end{cases} \tag{6}$$

Where $\Delta\theta$ and ΔR are parameter space quantization steps. In this way, all collinear points are set in one group according to sequence. A distance threshold is needed to classify and cluster collinear points into different line segments. Each line segment can be uniquely identified by two points, P_{beg} and P_{end}. Whenever a segment receives a point from the group, the following step are performed: the endpoints of the segment are updated, or a new segment is created, depending on whether or not the distance of the current point to the next point is greater than the threshold T_3. The n segment endpoints (P^n_{beg},P^n_{end}) are automatically updated by means of the relations:

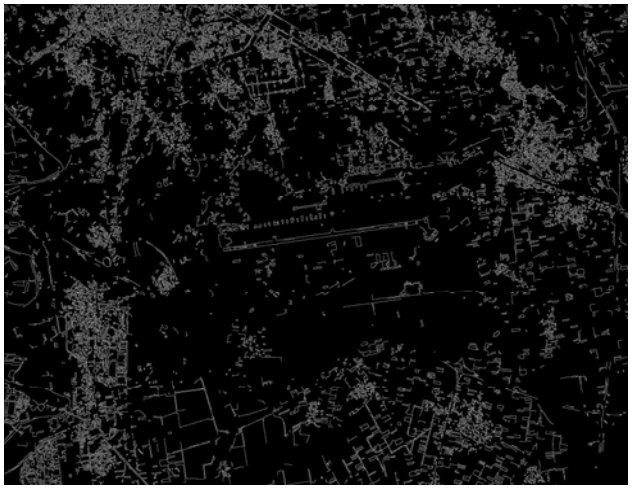
$$P^n_{beg}=\left\{\begin{array}{l}(x_i,y_i)\,\big|\,d[(x_i,y_i),(x_{i+1},y_{i+1})]\leq T_3\\ (x_i,y_i),(x_{i+1},y_{i+1})\in S_{R\theta}(x,y)\end{array}\right\} \tag{7}$$

$$P^n_{end}=\left\{\begin{array}{l}(x_i,y_i)\,\big|\,d[(x_i,y_i),(x_{i+1},y_{i+1})]>T_3\\ (x_i,y_i),(x_{i+1},y_{i+1})\in S_{R\theta}(x,y)\end{array}\right\} \tag{8}$$

Where $d[(x_i,y_i),(x_{i+1},y_{i+1})]$ is Euclidean distance between two points. Other situations cannot occur, as there cannot exist two separate segments that are close to point.

Fig.1 shows how to detect ROI in a large remote sensing image (2336×1800 pixels, 5-meter resolution). The results of line segments extraction are shown in Fig.1 (d) with $T_1=200$, $T_2=100$ and $T_3=200$. According to two parallel straight lines as shown in Fig.1 (d) and their location, the ROI with airfield can be determined.

(a) Original remote sensing image



(b) Edge detection

(c) Preliminary line extraction

(d) Line segments extraction by improved HT

Fig.1. Running example of airfield detection

The background of target region is very simple after ROI detection. Then, target geometry shape is extracted from ROI using optimum threshold segmentation. The threshold should be the gray value nearest to the smallest probability between two maximums of normal distribution. After segmentation, there are still some discrete noises with dozens of pixels or more, which are far smaller than targets. We use connectivity sign to eliminate them.

B. Target Shape Rotation

Target shape principal orientation is determined by shape features. It is in the straight line with the inclination α through the center of gravity of shape. All shapes are rotated to horizontally right by principal orientation before feature extraction to avoid the effects of orientation. According to Kahunen-Loeve (KL) transform [13]:

$$\alpha = \arctan\left[\frac{\mu_{02}-\mu_{20}+\sqrt{(\mu_{02}-\mu_{20})^2+4\mu_{11}^2}}{2\mu_{11}}\right] \tag{9}$$

Where μ_{pq} is $p+q$ order central moment of the region rounded by shape.

$$\mu_{pq}=\sum_x\sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \tag{10}$$

$\bar{x}$ and $\bar{y}$ are barycentric coordinates of target. Target region is defined as $f(x,y)$. As calculating α, the principal orientation φ is defined as:

$$\varphi=\begin{cases}\alpha+\pi & \mu_{30}>0\\ \alpha & \mu_{30}\le 0\end{cases} \tag{11}$$

Fig.2 shows the example of shape rotation. Fig.2 (a) is the ROI with airfield detected from original image. Fig.2 (b) and (c) are target shapes extracted from ROI. The shapes are same but the principal orientation φ are different. φ is 45° in Fig.2 (b), and -90° in Fig.2 (c). They are rotated around their barycentric by $-\varphi$, and the common result is shown in Fig.2 (d). The shapes are rotated to the same horizontally right.

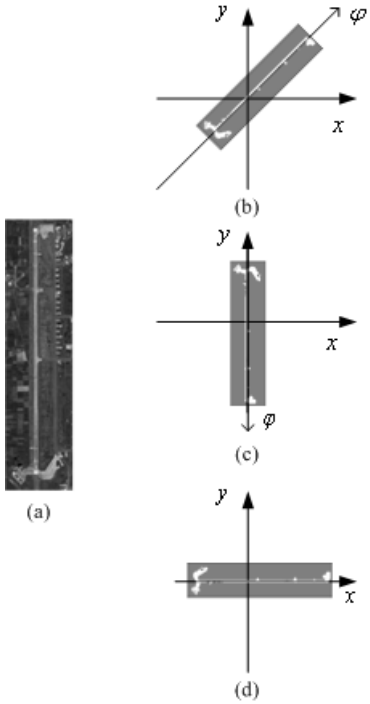


Fig.2. Running example of shape rotation

C. Feature Extraction

Both the local and global factors should be taken into account in feature extraction. We choose features described as follows:

1) Moment: The moment technique is widely used due to its appealing mathematical simplicity and versatility. Hu. M. K testified that seven moments had favorable stability to describe target and invariant property of translation, rotation and scale, which was called Hu's moment invariants [14]. Their two-dimensional $(p+q)$ order moment and central moment are:

$$m_{pq} = \sum_{x=0}^{N}\sum_{y=0}^{M} x^p y^q f(x,y) \tag{12}$$

$$\mu_{pq} = \sum_{x=0}^{N}\sum_{y=0}^{M} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \tag{13}$$

where

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \tag{14}$$

The formulas to compute Hu's moment invariants are constructed by Eq. (13) and can be found in the references. In addition Suk derived formulas of ten affine invariant moments through geometry method [15]. Invariant moments are also good shape descriptors. Hu moments are defined as $\Phi_i (i = 0,1,\cdots,6)$, and affine invariant moments are $\Psi_i (i = 0,1,\cdots,9)$.

2) Block moment of inertia: According to coordinate origin shown in Fig.2 (d) and the largest external rectangular region, the target is decomposed 4 one-level sub-targets, then these one-level targets are decomposed 16 two-level sub-targets. There are 21 target blocks as shown in Fig.3.

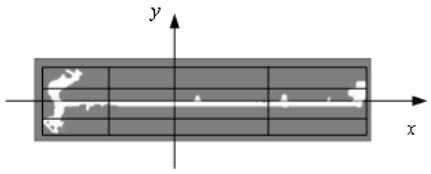


Fig.3. Target decomposition

Normalized moment of inertia (NMI) has good invariance [16], which is defined as:

$$NMI[f(x,y)] = \frac{\sqrt{\sum_x \sum_y [(x-\bar{x})^2 + (y-\bar{y})^2] \times f(x,y)}}{M[f(x,y)]} \tag{15}$$

where

$$M[f(x,y)] = \sum_x \sum_y f(x,y) \tag{16}$$

We define block moment of inertia is $NMI_i \ (i = 0,1,\cdots,20)$.

3) Aspect Ratio: According to the target shape principal orientation shown in Fig.2 (d), the projection of target in the x-axis is x', and in the y-axis is y'. Aspect Ratio is defined as $H = x'/y'$.

4) Compact Degrees: Compact Degrees is the ratio of perimeter and area. Perimeter is the number of edge points of target region, and area is the number of pixels in target region. We define that Compact Degrees is S .

As calculating features mentioned above, there is one eigenvector $T = [\Phi(0), \Phi(1), \cdots, \Phi(6), \Psi(0), \Psi(1), \cdots, \Psi(9), NMI(0), NMI(1), \cdots, NMI(20), H, S]$. Each element is random variable obeyed Guassian distribution. Their mean and standard deviation can describe features quantitatively. Experiments denote that the description performance of low-order moments is better than high-order moments, the Compact Degrees is better than Aspect Ratio. NMI can describe target by multi-level, but excessive levels may increase the amount of information sharply, and lead to general loss of description.

D. Similarity Measure

For two targets M and N, their feature eigenvectors are T_M and T_N , which are normalized as Normal Distribution with $N(0,1)$. The Euclidean distance between two feature eigenvectors of targets is

$$\begin{aligned} Dis(M,N) = & \\ & \omega_\Phi \times \sqrt{\sum_{i=0}^{6} (\Phi_M(i) - \Phi_N(i))^2} \\ & + \omega_\Psi \times \sqrt{\sum_{j=0}^{9} (\Psi_M(j) - \Psi_N(j))^2} \\ & + \omega_{NMI} \times \sqrt{\sum_{k=0}^{20} (NMI_M(k) - NMI_N(k))^2} \\ & + \omega_H \times \sqrt{(H_M - H_N)^2} \\ & + \omega_S \times \sqrt{(S_M - S_N)^2} \end{aligned} \tag{17}$$

ω expresses the weight of features, it is chosen through experiment. The Euclidean distance between the two feature eigenvectors shows the degree of similarity: the smaller the distance, the more similar the two targets.

III. Experiment

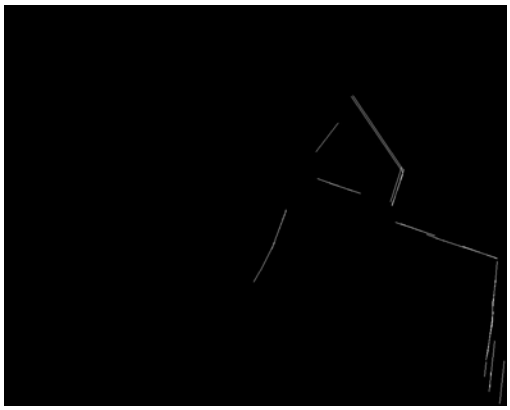
A. ROI Detection

In order to prove the proposed method, the technique was implemented using C++ and tested over a set of remote sensing images of various resolutions, some of which are shown in Fig.1 and Fig.4. Fig.4 shows how to detect harbor region in a panchromatic remote sensing image (1000×800 pixels, 1-meter resolution). Line segments are extracted as shown in Fig.4 (b), where $T_1 = 40$, $T_2 = 20$ and $T_3 = 40$, which are related with resolution and target size. The ROI with harbor can be determined by the parallel line segments formed by breakwater edges and their location, which are semi-enclosed and both sides are all waters.

Table 1 summarizes the ROI detection results obtained on eight groups of images. For each image, true targets were verified by artificial interpretation, such as one airfield in

Fig.1 (a) and one harbor in Fig.4 (a). If the region detected by the proposed method contains true target, the ROI detection is correct detection. False alarm is that the detected region does not contain expected target. Missed alarm is that the region containing target is not detected. As can be seen from Table 1, False alarm rate and missed alarm rate are both below 20%. There are two false alarms: one is a section of highway and the other is a section of river. Strictly speaking, the highway is also an artificial target. Two missed alarms are all harbors. Because harbor model is more difficult to establish, the performance of airfield ROI detection is better than that of harbor. A limitation of the technique is the lack of detection of some harbors, where surrounding buildings impact the line segments extraction. So the next works is to develop better target models and improve the harbor detection accuracy.

(a) Original remote sensing image



(b) line extraction

Fig.4. Running example of harbor detection

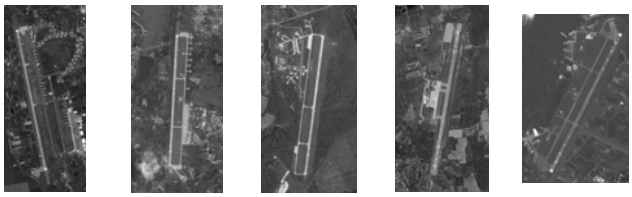
TABLE 1
The Statistics of ROI Detection Results

Image	True target number	Correct detection number	False alarm	Missed alarm
1	2	2	0	0
2	1	1	0	0
3	1	0	1	1
4	1	1	1	0
5	2	2	0	0
6	1	1	0	0
7	2	1	0	1
8	1	1	0	0

B. Similarity measure

Fig.5 (a) and Fig.6 (a) show the detected regions from Fig.1 (a) and Fig.4 (a). After ROI detection, the target is extracted to retrieve in template image library. We establish a template image library with 50 airfields and 50 harbors. Template images are all 1-meter resolution. Due to the impact of various factors like pixel resolution, sensor angle and range, targets have various transformations on template images, such as translation, rotation and scale. Similarity is calculated in accordance with Eq. (17). We choose the weights of features according to experiment and experience: $\omega_{\Phi} = 0.3$, $\omega_{\Psi} = 0.2$, $\omega_{NMI} = 0.1$, $\omega_{H} = 0.1$, $\omega_{S} = 0.3$. Retrieval results are shown in Fig.5 (b) and Fig.6 (b), where there are 5 most similar templates for each target. As can be seen from figures, the results are consistent with human visual experience. Other experiments also prove that the method has good generality and versatility.

(a) The detected airfield



(b) The most similar airfields

Fig.5. Airfield retrieval result

(a) The detected harbor

(b) The most similar harbors

Fig.6. Harbor retrieval result

IV. CONCLUSIONS

The basic goal of this course is to present a procedure to recognize large artificial targets from a remote sensing image by matching linear feature based ROI candidates to template images. We aim to develop a voting procedure for extraction of line segments in images using Hough transform. We believe that the region with linear features is probably the region with target like airfield and harbor. So the region including regular-array straight line features in image is detected and considered as ROI. Target type is identified by

model test. In particular, ROI detection improves accuracy rate and makes retrieval more purposive. Also, we try to show some challenges in the field of multi-levels representation of shape by features, which have favorable stability to describe target and invariant property of translation, rotation and scale. Feature eigenvectors are used to measure similarity. We have implemented and integrated them in a running system which can easily cooperate with each other. The experimental results encompass examples in images with airfield and harbor. As a result, our approach is robust to detection of regular lines, and the power assist of target recognition becomes effective. Future research will be focused on how to improve detection of harbor region and extract better shape descriptors. In addition, we keep exploring ways to reduce computational complexity.

REFERENCES

[1] M. Bober, "MPEG-7 Visual Shape Descriptors", *IEEE Trans. Circuits and Systems for video Technology*, vol. 11(6), pp.716-719, 2001.

[2] T. Sikora, M. Senior. "The MPEG-7 Visual Standard for Content Description-An Overview", *IEEE Trans. Circuits and Systems for video Technology*, vol, 11(6), pp, 696-702, 2001.

[3] J. W. Han, J. G. Han, L. Guo. "A novel image retrieval technique using salient edges", *in Proc. Storage and Retrieval for Media Databases 2002*, San Jose, California, 2002, pp. 69-78.

[4] J. W. Han, L. Guo, "Novel texture-based image retrieval approach using variogram function," in Proc. Internet Multimedia Management Systems III Conference, Boston, 2002, pp. 209-220.

[5] B. M. Mehtre, M. S. Kankankanhalli, W. F. Lee, "Shape measures for content-based image retrieval: a comparison", *Information Processing & Management*, vol. 33(3), pp. 319-337, 1997.

[6] G. Aggarwal, T. V. Ashwin, S. Ghosal, "An image retrieval system with automatic Query modification", *IEEE Trans. Multimedia*, vol. 4(2), pp. 201-214, 2002.

[7] C. Carson, S. Belongie, H. Greenspan, et al, "Blockworld: image segmentation using expectation-maximization and its application to image querying",. *IEEE Trans. on PAMI*, vol. 24(8), pp. 1026-1038, 2002.

[8] F Jing, M. Li, H. J. and Zhang, B. Zhang, "An efficient region-based image Retrieval framework," *in Proc. ACM Multimedia*, Berlin, 2002, pp. 456-465.

[9] M. Araiza-esquivel, S. Guel-sandovli, "Two-channel computer-generated holograms: a simplified method," *Optics and Laser in Engineering*, vol. 39(5/6), pp. 629-634, 2003.

[10] W. Niblack, D. Petkovic, "On Improving the Accuracy of the Hough Transform", *Machine Vision and Applications*, vol. 3(2), pp. 87-106, 1990.

[11] A. F. Leandro, M. O. Manuel, "Real-time line detection through an improved Hough transform voting scheme", *Pattern Recognition*, vol. 41(1), pp. 299-314, 2008.

[12] J Cha, R. H. Cofer, S. P. Kozaitis, "Extended Hough transform for linear feature detection". *Pattern Recognition*, vol. 39(6), pp. 1034-1043, 2006.

[13] P. Common, "Independent component analysis, a new concept?", *Signal Processing, Special Issue on Higher Order Statistics (S0165-1684)*, vol. 36(3), pp. 287-314, 1994.

[14] S. K. Liao, M. Pawlak, "On image analysis by moments", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18(3), pp. 254-26, 1996.

[15] T. Suk, J. Flusser, "Graph Method for Generating Affine Moment Invariants," *in Proc. the 17th International Conference on Pattern Recognition*, Cambridge, 2004, pp.192-195.

[16] X. Yang, G. Fu, D. Miao , et al, "A New Approach to Target Recognition Based on Image NMI Feature ", *Computer Engineering*, vil. 28(6), pp. 149-150, 2002.