An efficient enhanced automatic target recognition approach

Ramy M Bahy

Military Technical Research Center, Cairo, Egypt

Email: ramymbahy@yahoo.com

Abstract. This paper presents a new efficient enhanced automatic target recognition approach based on hybrid feature and intensity based methods. The proposed approach is based on comparison and matching between the acquired target image and the stored database images through edge detection, morphological operation, image transformation, image histogram matching and mutual information for both target and stored database images. Sobel operator is used to detect edges in both acquired target image and the stored database images. Hough transform is used as an image transformation which groups the extracted edge points in edged images to an appropriate set of lines which helps in features extraction and matching processes in both of target and stored database images. This gives an initial indication about realization and recognition between target image and its corresponding database image. Comparisons of image histogram and mutual information for both target and database images are finally used to emphasize the recognition of the target image and its verification with its corresponding database image. The proposed recognition approach passed through four stages which are: edge detection by Sobel edge detector, thinning as a morphological operation, Hough transformation, matching process using median deviation and finally histogram matching and measuring the mutual information between target and the available database images. The experimental results proved that the proposed hybrid target recognition approach gives a more accurate and successful recognition rate than other techniques which depend on using feature based methods or intensity based methods.

1. Introduction

Automatic target recognition (ATR) is a real-time or near-real-time image/signal-understanding system. An ATR is presented with a stream of data. It outputs a list of the targets that it has detected and recognized in the data provided to it. A complete ATR system can also perform other functions such as image stabilization, preprocessing, mosaicking, target tracking, activity recognition, multi-sensor fusion, sensor/platform control, and data packaging for transmission or display.

ATR can be used as a generic term to cover a broad range of military data exploitation technologies and tasks. These include image fusion, target tracking, minefield detection, as well as technologies for specific missions such as persistent surveillance and suppression of enemy air defenses. The term can be broadened to cover homeland security tasks such as border monitoring, building protection, and airport security. It can include environmental efforts such as detection of fires, whales, radioactive material, and gas plumes. Commercial applications similar to the military ATR problem are grouped under the name video analytics. These include parking lot security, speed cameras, and advanced signage. Internet companies are making huge investments in image-based search engines and face recognition. Industrial automation and medical applications of machine vision and pattern recognition use the same basic technology [1].
Target recognition generally consists of three aspects: feature extraction, similarity measurement and best match search. Feature extraction refers to features extracting (e.g., contours, edges and corners) from the original image. Then, a criterion (similarity measurement) is needed to judge how similar a candidate feature in the original image to the predefined target feature. Target detection is an important step before recognition step. In order for detection of targets to be automated, a training database needs to be created. This is usually done using experimental data collected when the target is known, and is then stored for use by the ATR algorithm. [2][3].

Some other applications such as detection and recognition of rows in agricultural crop images acquired by remote sensing from a UAV are introduced in [4]. Detection of rows in crops planted as rows is a fundamental to site specific management of agricultural farms [5]. The main purpose is to detect, track and finally recognize the targets throughout the scene or scenario. In spite of much research on the subject in recent years, the problem remains challenging mainly due to unknown and changing number of targets; noisy; missing observations and interaction of multiple targets [6].

This work is considered as an enhancement of our previously published work [7]. The pre-published ATR approach was mainly using feature based methods for target recognition. This work focuses on ATR using hybrid feature and intensity based method. The main advantage of feature based approaches is the robustness against illumination changes but at the same time these approaches have some drawbacks such as difficulty of features extraction in poor quality images. Moreover, the detected feature sets in the target and database images must have enough common elements, even in situations when the images do not cover exactly the same scene or when there are object occlusions or other unexpected changes. Intensity based methods have great advantages and overcome the drawbacks of feature based methods by comparing and matching the intensity differences of image pairs under a similarity metric such as mutual information or cross correlation regardless of the quality of image pairs. However, the intensity based methods have some disadvantages such as sensitivity to illumination changes and they may fail in target recognition in case of less overlapped area between target and database images [8].

Therefore, in this work an efficient enhanced automatic target recognition (EEATR) approach is proposed which is based on both feature and intensity based methods to overcome the drawbacks of individually feature or intensity based methods.

The proposed EEATR approach is considered a hybrid ATR approach. Whereas, its steps include edge detection using Sobel operator and detection of straight lines using Hough transform (HT) in both of target image and database images. These steps are considered as feature based procedures. Moreover, the proposed EEATR approach includes intensity based steps which are represented in comparison of images histogram and measuring the mutual information between target and database images. The use of HT facilitates the matching process between the target image and its corresponding database image. Only straight lines in both target and database images are compared together. Furthermore, the process of comparison between the histogram of target image and the histogram of stored database images and the process of evaluation and measuring the mutual information (MI) between target and stored database images are used to emphasize the realization and the recognition of target image and its corresponding database image.

This paper is organized as follows: Section 2 illustrates a literature survey on the related work. Section 3 introduces and explains the proposed efficient enhanced automatic target recognition approach (EEATR). Section 4 shows the experimental work and results. Finally, the paper is concluded in section 5.
2. Related work
Many approaches were presented to deal with the problem of ATR; most of these approaches begin with target detection in images. One of the commonly used method, is background subtraction (BS) which is used for moving objects detection when the camera is fixed. In this method, a reference image is taken and chosen as the background and then compared to the next image, thus detecting the changes occurring at the reference background [9] [10] [11]. A generalized gradient model which uses the relationship between a series of two-dimensional images and the speed is known as an optical flow subtraction [12]. Detection of the differences between the sequential video is known as a temporal difference method [13]. Another edge-based segmentation method is considered as an active contour model [14] [15] which is based on the framing of the edges of an object with a closed frame by the energy function within the object area [16]. Some other primary image processing methods used in target recognition such as mean shift algorithm by creating the histogram of a moving object as an appearance feature and a similarity measurement to find the best match for the moving object in the next frames [17]. Potential based methods originated from the edge potential function (EPF) model [18]. EPF based method has been applied to many real world problems, e.g., ground target detection [19] [20], ear detection [21], aperture radar scene matching [22] and video retrieval [23] [24]. Also, atomic potential matching (APM) model overcomes the drawbacks of the conventional EPF model by combining the concept of EPF with a powerful matching tool based on genetic algorithms (GAs). EPF can be easily calculated starting from an edge map and provide a kind of attractive pattern for a matching contour, which is conveniently exploited by GAs [25].

Some other automatic target recognition techniques are used to accurately recognize small and irregularly shaped targets by using a set of two-dimensional (2D) views of the object to get the three-dimensional (3D) modeling of the targets [26]. Other recent techniques for target recognition use neural networks to detect deep targets based on deep convolutional neural network (DCNN) [27].

3. The proposed efficient enhanced automatic target recognition
The proposed recognition technique passed through four stages which are: edge detection by Sobel edge detector, thinning as a morphological operation, HT, matching process using median deviation and finally histogram matching and measuring the MI between target and the available database images. The purpose of using these stages is described in details as follows:

The Sobel operator is selected as an edge detection method because it has the ability to detect only strong edges [28] in the image which is more suitable to our case of study. Whereas, there is no need to detect weak edges as the Canny edge detector operates. The thinning is used as a morphological operation which makes the detected edges in images as sharp and straight as possible which helps the HT in the next stage to work successfully and group the extracted edge points to an appropriate set of lines. HT as mentioned before helps in detection of straight lines in both of target image and database images. This also, helps the matching process in the next stage to work successfully with minimum processing time. Matching process is applied only on straight lines and it is not applied on all detected edges. Histogram comparisons and matching is used to find the best matching of the intensity levels between target and stored database images. Besides, MI is used as a similarity measure between target and database images.

The stages of histogram matching and MI measurement are not relying on the previously mentioned four stages. So, they can be performed in parallel with the other four stages to reduce the required processing time and the computational complexity. MI makes few assumptions about how the intensity values of the images are related. It only relies on statistical dependence. Therefore, the MI can be used as a similarity measurement even if the target image suffers from some artifacts such as noise and blurring effects.

Furthermore, histogram matching and MI measurement are used as a confirmation method to emphasize and finally realize the matching between the target image and its corresponding database image after initial recognition by matching the detected straight lines after applying HT in both target and data base images.
3.1. Implementation of proposed efficient enhanced automatic target recognition approach

The implementation of the proposed EATR approach is explained in detail in the following steps:

1. Apply Sobel edge detector on both target image \((im_T)\) and the database images \((im_{Db})\). Where, \(n\) is the number of the available database images. The output from this step will be an edged target image \((im_{T-edge})\) and edged database images \((im_{Db-edge})\).

2. Apply thinning process on the edged images \((im_{T-edge} \& im_{Db-edge})\). This step is important to make the detected edges in images as sharp and straight as possible. Thinning is a morphological operation that is used to remove selected foreground pixels from binary images, somewhat like erosion or opening. This step it is commonly used to tidy up the output of edge detectors by reducing all lines to single pixel thickness [29].

3. Transform the output images from step 2 into HT space to group the extracted edge points to an appropriate set of lines. HT has the ability to group pixels which belong to the same line by using a two dimensional array, called an accumulator, to detect the existence of a straight line [30]. Therefore, HT is used to detect the straight lines in images.

4. A median deviation value (M) of a 3×3 window size around the edged pixels \((xi)\) is calculated for each edged pixels on the straight lines in both target and database images. Where, \(i\) is the number of pixels which belong to the same straight line. Then, there will be an array \((A1)\) contains the median values of the line pixels in target image. Whereas, \(A1=[M_1,M_2,\ldots,M_n]\), where, \(n\) is the total number of edged pixels on all straight lines in target image. Another cell array \((A2)\) represents the median values of the line pixels in database images.

\[
A2 = [a_1, a_2, a_3,\ldots, a_s]
\]

Where, \(n\) is the number of database images and \(a_s = [M_1,M_2,\ldots,M_s]\), where, \(s\) is the total number of edged pixels on all straight lines for each database image. Each sub-array contains the median values of the line pixels in each of database images. In this step, two or three edge thinned lines in the target image are used for comparing and matching with their corresponding lines in database images in order to reduce the processing time of this step.

5. A normalized cross-correlation (NCC) formula is used to find the best match between median values in the array \((A1)\) which represents the line pixels in target image and the cell array \((A2)\) which represents the median values of the line pixels in database images. NCC is used to find the best match between a small size template and larger array by using a certain selected threshold value. \(A1\) is considered as a template image and the larger array size will be \(A2\). The NCC is calculated according to the following formula [31].

\[
NCC = \frac{\sum_{i,j}[A_1(x,y) - \bar{A}_1][A_2(x-u,y-v) - \bar{A}_2]}{\left(\sum_{i,j}[A_1(x,y) - \bar{A}_1]^2\right)^{0.5} \left(\sum_{i,j}[A_2(x-u,y-v) - \bar{A}_2]^2\right)^{0.5}}
\]  (1)

Whereas, \(\bar{A}_1\) is the mean of the template image and \(\bar{A}_2\) is the mean of \(A2(x,y)\) in the region under the template.

6. The maximum of NCC gives the indication about the best correlation between the values in array \(A1\) and the sub-array \((a_n)\) in \(A2\). Then, \(Max (NCC)\) refers to a certain \(a_n\) and \(n\) refers to the index number of the database image which matched with the target image. So, the target image is initially recognized with its corresponding database image. The output from this step is an initially recognized target image with its corresponding database image \((im_{DbDet})\).

7. An histogram matching is performed between the initially recognized target image from step (6) and its corresponding database image. If the matching percentage between pixel values in both of initially recognized target image and its corresponding database image is greater than 70 % then the target image is almost recognized with its database image.
(8) The mutual information (MI) between both of target and all available database images is computed. Moreover, the MI between the target and database image which has best histogram matching as described in the previous step is also computed. MI similarity measurement is originally used to emphasize and finally realize the recognition between target image and its corresponding database image. MI is calculated according to the following equation [32]:

$$MI(i_m, i_D) = \sum_{i_1=1}^{L} \sum_{i_2=1}^{L} h_{IATR}(i_1, i_2) \log_2 \frac{h_{IATR}(i_1, i_2)}{h_{IATR}(i_1)h_{IATR}(i_2)}$$ (2)

MI is used to evaluate the correlative performances of the target image and the database image. Where, $L$ is the number of histogram bins ($L$=256 in case of gray scale images), $h_{IATR}(i_1, i_2)$ indicates the normalized joint gray level histogram of images $i_m$ and $i_D$. Also, $h_{IATR}(i_1)$, $h_{IATR}(i_2)$ indicate the normalized marginal histograms of the two target image and database images respectively. The MI values between the target image and the database images are recorded.

(8) If the highest MI value is the value of MI between the database image and the target image which is recognized in step (7) then target recognition is realized.

(9) Go to step (5) and change the threshold value if the highest MI value is not the value of MI between the target image and the recognized database image which is determined in step (7).

4. Experimental work and results

A well-known UC Merced Land Use Dataset [33] is selected as database images to prove the effectiveness and accuracy of the proposed EEATR approach. UC Merced Land Use Dataset consists of 21 class land use image dataset (buildings, airplanes, forest,…etc.) meant for research purposes. There are 100 images for each class. Each image measures 256x256 pixels. So, the experimental work was done on about 2100 different images. The images were manually extracted from larger images size from the USGS National Map Urban Area Imagery collection for various urban areas [34]. The pixel resolution of this public domain imagery is 1 foot.

A set of 21 different types of images $I = \{im1, im2, \ldots, im21\}$ with size 256x256 pixels is used to represent the target images. Every image is selected from a set of 100 images which represents different classes of the Merced dataset (database images).

To prove the effectiveness and robustness of our proposed hybrid EEATR approach, the results are compared with our previously published work [7]. The previously published work was based only on features extraction and comparing these features together in both target and database images to find the best match (feature based technique). Moreover, the results are compared with one of the recent techniques which is based on image segmentation and equalizing the image intensities for target and database images (intensity based method) [35].

Table 1 shows that the proposed EEATR approach succeeded to recognize 20 target images, While, feature based ATR approach (FATR) and the intensity based ATR approach (IATR) succeeded to recognize only 15 and 10 images respectively. So, the proposed hybrid approach which is mainly combining the advantage of feature and intensity based approaches gives a more accurate recognition rate than other techniques which are only based on feature or intensity based approaches.

The structure of the proposed EEATR approach is illustrated in figure 1.

| Table 1. Comparison among the proposed EEATR, FATR and IATR approaches. |
|-----------------------------|----------------|----------------|
| Comparison item             | Proposed EEATR approach | FATR approach | IATR approach |
| No. of Succeed recognition  | 20              | 15            | 10           |
| No. of False recognition    | 1               | 6             | 11           |
Figure 2(a), figure 2(b) and figure 2(c) show some of the database images which represent some types of airplanes selected from 100 airplane images as an example selected from one class of Merced dataset images and figure 2 (d) shows one of the target images.

The complete procedures of the proposed target recognition technique are explained in detail in the figures by applying the pre-mentioned steps in the previous section on images shown in figure 2 as an example of our work. Figure 3(a), figure 3(b) and figure 3(c) show the edged database images \( \text{im}_{D1}\)-edge, \( \text{im}_{D2}\)-edge and \( \text{im}_{D3}\)-edge respectively, and figure 3(d) shows the edged target image \( \text{im}_{1}\)-edge after applying the Sobel edge detector.

![Figure 1](image.png)

**Figure 1.** The structure of the proposed ATR approach.
Figure 4(a), figure 4(b) and figure 4(c) show the edged database images $im_{D1-edge}$, $im_{D2-edge}$ and $im_{D3-edge}$ respectively, and figure 4(d) shows the target image ($im_{1-edge}$) after applying the Sobel edge detector + thinning + Hough transform (HT). As shown in figure 4, it is noticed that the straight lines in both the database images and target image are detected due to the HT effect and assigned by colored lines. These straight lines are considered as the extracted features which are used for matching between database and target images. By applying the steps from 4 to 6 on the proposed target recognition approach as in the pre-described section, the target image ($im_{1}$) is initially realized and recognized with its corresponding database image ($im_{D1}$). The histogram matching and MI between the target images ($im_{1}$) and the 100 airplanes database images ($im_{D1}, im_{D2}, \ldots, im_{D100}$) which represent one class from UC Merced Land Use Dataset are measured to verify and emphasize the recognition between the target image ($im_{1}$) and its corresponding database image ($im_{D1}$). As shown in figure 5, the highest value of MI (1.7) is verified between the target image ($im_{1}$) and the database image ($im_{D1}$). Whereas, the other values that represent the MI between $im_{1}$ and the remaining database images are lower than MI ($im_{1}, im_{D1}$). So, MI ($im_{1}, im_{D1}$) > MI ($im_{1}, im_{Dn}$) where n≠1.

---

**Figure 2.** (a), (b) and (c) database images $im_{D1}$, $im_{D2}$ and $im_{D3}$ respectively (d) target image $im_{1}$.
Figure 3. (a), (b) and (c) edged database images $im_{D1\_edge}$, $im_{D2\_edge}$ and $im_{D3\_edge}$ respectively (d) target image $im_{1\_edge}$.

Figure 4. (a), (b) and (c) edged database images $im_{D1\_edge}$, $im_{D2\_edge}$ and $im_{D3\_edge}$ respectively (d) edged target image $im_{1\_edge}$ after applying thinning and HT processes.
Figure 5. MI between target image \((im_T)\) and database images \((im_{Dn})\).

5. Conclusion
The proposed ATR approach is not only depending on feature matching or intensities matching as many other methods but also, makes a verification and emphasizing the matching process by using both of histogram matching and mutual information. Moreover, according to the results shown in the indicated graph in figure 5, it is proved that the target image is completely recognized with its corresponding database image. Because, the highest value of MI is verified between the target image \((im_1)\) and its corresponding database image \((im_{D1})\). Besides, the results in Table1 proved that the proposed enhanced efficient target recognition approach gives better performance and gives more recognition rate than some of other recent techniques which is only based on feature or intensity based approaches.

References
[1] Bruce J S 2018 Automatic Target Recognition (washington: SPIE Press) 3rd Edition TT118 ISBN: 9781510618565
[2] Yebes J J, Bergasa L M and Garcia M 2015 Sensors 4 pp 9228–50
[3] Li B and Yao Y 2014 Engineering Application- Artificial Intelligent 35 pp 14–25
[4] Ramesh K N, Chandrika N, Omkar S N, M B Meenavathi and Rekha V 2016 Int. journal of Image Graphics and Signal Processing 11 pp 25–31
[5] Mishra K and Saroha P 2016 Int. Journal of Image Graphic and Signal Processing 5 pp 76-82
[6] Tunç A, Emre T and Sinan Öz 2015 Proc. Int. Conf. on Computer Vision Imaging and Computer Graphics pp 240–55
[7] Ramy M B 2018 Int. Journal of Image Graphic and Signal Processing 3 pp 18–24
[8] Fitzpatrick J, Hill D L and Maurer C R 2000 Handbook of Medical Imaging Medical Image Processing and Analysis SPIE press 2 pp 447–512
[9] Sang H and Xu C 2013 Moving Object Detection Based on Background Subtraction of Block Updates Proc. Int. Conf. on Intelligent Networks and Intelligent Systems (ICINIS)
[10] Lucia M and Alfredo P 2018 Journal of Imaging 7 pp 1–27
[11] Saravanakumar S, Vadivel A and Ahmed C G 2010 Multiple human object tracking using background subtraction and shadow removal techniques images Int. Conf. on Signal and Image Processing pp 79–84
[12] Attila S and Kristóf K 2018 Shape Recognition Based on Projected Edges and Global Statistical Features Mathematical Problems in Engineering 2018

[13] Chuang C H , Chao Y L and Li Z P 2016 Moving object segmentation and tracking using active contour and color classification models Proc. of IEEE Int. Symp. pp 1–8

[14] Kass M, Witkin A and Terzopoulos D 1988 Int. Journal of Computer Vision 1 pp 321–31

[15] Maziere M, Chassaingl F, Garrido L and Sal embie P 2000 Segmentation and tracking of video objects for a content-based video indexing context Proc. of IEEE Int. Conf. pp 1191–4

[16] Özgen N 2008 Computer based target tracking 2008 M.Sc. Thesis (Gazi University Institute of Science and Technology)

[17] Wang L W and Qin J L 2007 Study on moving object tracking algorithm in video images 8th Int. Conf. on Electronic Measurement and Instruments pp 1–4

[18] Dao M S, Natale F G and Massa A 2003 Edge potential functions and genetic algorithms for shape-based image retrieval Proc. of Int. Conf. on Image Processing pp 1165–73

[19] Li C and Duan H 2014 Aerospace Technology Science and Technology 39 pp 352–60

[20] Xu C and Duan H 2010 Pattern Recognition Letters 13 pp 1759–72

[21] Battisti F, Carli M, De Natale and Neri A 2012 Proc. of SPIE Int. Society for Optical Engineering 18 pp 1243-50

[22] Wang Y and Yin J 2014 Intelligent search optimized edge potential function (EPF) approach to synthetic aperture radar (SAR) scene matching IEEE Congress on Evolutionary Computation (CEC) pp 2124–31

[23] Dao M S, DeNatale F G and Massa A 2005 MPEG-4 video retrieval using video-objects and edge potential functions Springer Advances in Multimedia Information Processing pp 550–7

[24] Dao M S De Natale F G and Massa A 2003 Proc. of Int. Symp. on Intelligent Multimedia Video and Speech Processing ICIP 3 pp 712–29

[25] Li B, Cao H, Hu M and Zhou C 2015 Proc. of 8th Int. Symp. on Computational Intelligence and Design (ISCID ) 1 pp 1–4

[26] Olson C F and Huttonlocher D P 2015 IEEE Transactions on image processing 6 pp103–13

[27] Nasser M 2018 Deeptarget: An automatic target recognition using deep convolutional neural networks IEEE Transactions on Aerospace and Electronic Systems p 99

[28] Weiping Y, Xuezhi W, Moran B, Wheaton A and Cooley N 2012 Elsevier Computers and Electrical Engineering 38 pp 1213–22

[29] Parker J R 2010 Algorithms for image processing and computer vision (2nd edition Wiley computer publishing)

[30] Bahy R M, Salama G I and Mahmoud T A 2012 Registration of multi-focus images using Hough transform Proc. of 29th National Radio Science Conference (NRSC) (Cairo, Egypt) pp 279–84

[31] https://www.mathworks.com/help/images/ref/normxcorr2D 28/12/2019

[32] Shams R and Hartley R 2007 Image registration in Hough space using gradient of images Proc. of IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP) pp 226–32

[33] h ttp://vision.ucmerced.edu/datasets/landuse.html 28/12/2019

[34] https://nationalmap.gov 28/12/2019

[35] Jian T, Xiangtao F, Shenghua W and Shenghua R 2018 Sensor 18 pp 3019–27