IMAGE SEQUENCE MEASURES FOR AUTOMATIC TARGET TRACKING

W.-H. Diao[1, 2], X. Mao[1, *], H.-C. Zheng[1], Y.-L. Xue[1], and V. Gui[3]

[1]School of Electronic and Information Engineering, Beihang University, Beijing 100191, China

[2]Institute of Manned Spacecraft System Engineering, China Academy of Space Technology, Beijing 100094, China

[3]Faculty of Electronics and Telecommunications, Politehnica Timisoara University, Timisoara 300223, Romania

Abstract—In the field of automatic target recognition and tracking, traditional image metrics focus on single images, ignoring the sequence information of multiple images. We show that measures extracted from image sequences are highly relevant concerning the performances of automatic target tracking algorithms. To compensate the current lack of image sequence characterization systems from the perspective of the target tracking difficulties, this paper proposes three new metrics for measuring image sequences: inter-frame change degree of texture, inter-frame change degree of target size and inter-frame change degree of target location. All are based on the fact that inter-frame change is the main cause interfering with target tracking in an image sequence. As image sequences are an important type of data in the field of automatic target recognition and tracking, it can be concluded that the work in this paper is a necessary supplement for the existing image measurement systems. Experimental results reported show that the proposed metrics are valid and useful.

1. INTRODUCTION

An automatic target recognition (ATR) system needs to detect, track and recognize many different military objects. Numerous algorithms [1–4] have been proposed to achieve application-specific

performance requirements. It is well known that performance evaluation is an indispensable part for an effective algorithm. While target detection [5, 6], recognition [7, 8] and tracking algorithms [9, 10] are used under a variety of military tasking scenarios such as air-to-air, air-to-ground, ground-to-ground, etc., extensive testing is particularly important. However, in most cases, algorithm testing is usually performed over a limited set of scenarios and the images database generally only includes a small set of collected images or artificially generated images [11–15]. The limited variability of experimental images leads to the question of how robust the algorithm performance is for all scenario conditions of practical relevance. One way to solve the problem is to correlate the image characteristics with the algorithm performance and measure the images using quantitative metrics.

The image measure methods in the field of automatic target recognition and tracking have different characteristics, in comparison with other image measure algorithms: they are conceived to describe the information interfering with the target in the images. Such characteristics are: target to background contrast (TBC), signal-to-noise ratio (SNR), signal-to-clutter ratio (SCR) and probability of edge (POE). These metrics are based on various theories. However, nearly all of them apply to single images. Since the characteristics of a single image and the characteristics of an image sequence are different, it is not surprising that the metrics for single images are not best suited for measuring image sequences. The existing image sequence measure methods, such as peak signal-to-noise ratio (PSNR) [16, 17], mean square errors (MSE) [18–20], mean structural similarity (MSSM) [21], have been proposed in the fields of image coding and image communication. These metrics, different from that in the field of automatic target tracking, focus on measuring information losses in video compression, therefore they have not been applied to evaluate the performance of automatic target tracking algorithms in image sequences.

Among mathematicians, physicists, and computer scientists, there is a general consensus that the complexity of an object or a system is a measure of the inherent difficulty of performing the tasks associated with it [22]. Thus, in the context of ATR performance, the image measure has to be a description of the inherent difficulty of fulfilling an ATR task for a given image sequence. Based on the above ideas, this paper proposes and evaluates several measures of image sequence characteristics, in view of describing the degree of difficulty of tracking targets in image sequences.

The structure of this paper is as follows. Traditional image metrics are listed and discussed firstly. Then the characteristics of image

sequences and the deficiency of existing metrics are analyzed. Based on the results of this analysis, new metrics suitable for measuring image sequences are proposed. To our best knowledge, it is the first time that inter-frame change of texture, of target size and of target location information are used to characterize image sequences. The proposed metrics are validated by experiments in the next section. Finally, the conclusions are drawn.

2. IMAGE MEASURES

In order to design the performance evaluation system, we need to know how to characterize the image inputted into the ATR system, so that we can relate these characteristics to the performance of the target recognition or the tracking algorithm. So far, many image metrics have been proposed. This paper classifies the current image metrics into two categories in terms of their definition principles: one is based on statistical information, while the other is based on "clutter" theory. In the next two sections, we review some commonly used image metrics. Statistical metrics are reviewed in Section 2.1, while "clutter" metrics are reviewed in Section 2.2.

2.1. Image Measures Based on Statistical Information

Statistical metrics are functions of gray level information in a monochromatic image. Among the ones we have found, some depend on all pixel values in the image, some on local pixel values and others only on pixel values in the target region. Table 1 lists these statistical metrics.

2.1.1. Global Metrics

Some researchers assume that a target recognition algorithm performs best when presented with a highly contrasting target against a uniform, untextured, background. Many of the global gray-level metrics in the literature seem to be designed with that in mind. The standard deviation of an image and the entropy of its histogram are two of the simplest measures of an image [17–19].

Beard et al. [23] defined TBC to be an image metric:

$$TBC = |\mu_T - \mu_B| \tag{1}$$

where μ_T is the mean of the target and μ_B the mean of the background. The author supposes that the bigger the contrast between the target and the background, the easier is to find the true target in the image.

Table 1. Statistical metrics.

Global metrics	
	Image gray-level standard deviation
	Image gray-level entropy
$TBC = |\mu_T - \mu_B|$	Target to background contrast
$SNR = (I_t - I_B)/C_B$	Signal-to-noise ratio
DTC	Degree of target being confused
DTS	Degree of target being shielded
$TIR = (\mu_T - \mu_B)/(\sigma_T + \sigma_B)^{1/2}$	Target interference ratio
$TBIR = (\mu_T - \mu_B)/(\sigma_T\sigma_B)^{1/2}$	Target background interference ratio
$TIR^2 = (\mu_T - \mu_B)^2/(\sigma_T + \sigma_B)$	TIR squared
$TBIR^2 = (\mu_T - \mu_B)^2/\sigma_T\sigma_B$	TBIR squared
$C = \int \ldots \int \min\{f_T(x_1, \ldots, x_n),$ $f_B(x_1, \ldots, x_n)\}dx_1, \ldots, dx_n$	
$ETB = |E(T) - E(B)|$	Target versus background entropy
KSZ	Kolmogorov-Smirnov statistic
Local metrics	
$LTBC$	Local target to background contrast
$LTIR$	Local target interference ratio
$LSNR$	Local signal-to-noise ratio
$LETB$	Local target versus background entropy
$LKSZ$	Local Kolmogorov-Smirnov statistic
Target specific metrics	
TSD	Target standard deviation
ENT	Target entropy
POT	Pixels on target
$P2/A$	Perimeter squared over area
$ATES$	Average target edge strength
ESD	Target edge standard deviation

A classical method to estimate the quality of infrared small target images [24, 25] is to calculate the signal-to-noise ratio (SNR):

$$SNR = \frac{I_T - I_B}{C_B} \qquad (2)$$

where I_T is the mean gray value of the target region, I_B the mean gray value of background, and C_B the background standard deviation. Small target images are defined as images in which the targets' sizes are less than 0.15% of the whole image [26]. According to Equation (2), SNR mainly reflects the information of background standard deviation

and the contrast between target and background. In addition, for this type of images, Mao and Diao proposed two image metrics in view of quantifying the difficulty degree of infrared small target detection [27]. One is the degree of target being confused (DTC), which quantifies the odds of infrared small target images to provide fake targets. The other one is the degree of target being shielded (DTS), which reflects the contribution of the image to shield the target. It has been proved that DTC and DTS are valid and robust for measuring small target images.

Let μ_T and σ_T be the mean and standard deviation of the gray-levels inside the minimum covering rectangle of the target and let μ_B and σ_B be the mean and standard deviation of the gray-level in the background region. Based on these parameters, M. Lahart et al. [28] and F. Sadjadi [29] developed an image metric called target interference ratio (TIR),

$$TIR = (\mu_T - \mu_B)/(\sigma_T + \sigma_B)^{1/2} \tag{3}$$

and a metric called target-background interference ratio ($TBIR$),

$$TBIR = (\mu_T - \mu_B)/(\sigma_T \sigma_B)^{1/2} \tag{4}$$

which favors uniform targets against uniform backgrounds and varies inversely in the standard deviation of both target and background. Based on TIR and $TBIR$, two new measures: TIR squared (TIR^2) and $TBIR$ squared ($TBIR^2$) are developed respectively in [22].

Garlson et al. [30] observed that there must be measureable differences between the feature distributions of the target areas and of the background areas when a feature-based target recognizer works well. The author posited that the extent to which the distributions are not different, determines the complexity of the task. In other words, the area of overlap in the target and background feature distributions is a probabilistic measure of the complexity. The proposed measure is

$$C = \int \ldots \int \min\{f_T(x_1, \ldots, x_n), f_B(x_1, \ldots, x_n)\}dx_1, \ldots, dx_n \tag{5}$$

where f_T and f_B are the target and background distributions respectively. In addition, Beard et al. [23] suggested the use of a Kolmogorov-Smirnov statistic (KSZ), since it is a widely used measure of the difference between two distributions. They also suggested the ratio of target entropy to background entropy (ETB), since high entropy targets should be easily distinguished from low entropy background.

2.1.2. Local Metrics

Considering that not all the pixels in the background will interfere with the target, some local metrics are defined based on the theories of global metrics. The valid background region in local metrics is a rectangular annulus whose inner border coincides with the target rectangle and whose outside rectangle area is twice the area of the target rectangle. Some local metrics are also used commonly, such as *LTBC*, *LTIR*, *LSNR*, *LETB*, and *LKSZ*. The definitions of these metrics are similar with the corresponding global metrics, but with different background region sizes.

2.1.3. Target Specific Metrics

The third category of statistical image metrics are target specific metrics which focus on the target region only. Target standard deviation (*TSD*) is used in [31, 32]. This measure simply computes the standard deviation of the pixel intensity values in the target region and can be denoted by σ_T. Being similar with *TSD*, target entropy (*ENT*) [23] is the entropy of the pixel intensity values in the target region.

Some of the target specific metrics depend on the shape or size of targets in the image. The simplest size measure is pixels on target (*POT*) [32], which could be obtained by counting the number of pixels in a target region. Perimeter squared over area (P^2/A) [32, 33] is the squared value of the target perimeter, divided by the total number of *POT*. The perimeter is computed as the sum of the arc lengths in terms of number of equivalent unit pixels in length.

In addition, some metrics are based on the edges of targets in the image, such as average target edge strength (*ATES*) [34] and target edge standard deviation (*ESD*) [32]. *ATES* characterizes the average magnitudes of target edge gradients. To calculate this feature, the Sobel operator is used to generate the edge map and the average of all edge magnitudes above a minimum threshold is used to compute the average edge magnitude of the target. *ESD* is simply the standard deviation of the edge magnitudes in the target region.

2.2. Image Measures Based on "Clutter" Theory

Researchers made much effort on enhancing target or reducing clutter [35], which is an important preprocessing procedure for automatic target tracking. In fact, tracking a specific target in an image sequence can be viewed as searching a target in clutter. When

Table 2. "Clutter" metrics.

Target independent metrics	
$C = \left[\sum\limits_{i=1}^{N} \sigma_i^2 / N \right]^{1/2}$	Statistical variance
$POE = \frac{1}{N} \sum\limits_{i=1}^{N} POE_{i,T}^2$	Probability of edge
$WPOE = \left(\frac{1}{N} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{l} WPOE_{i,T,j}^2 \right)^{1/2}$	Multiresolution POE
Target dependent metrics	
COM	Co-occurrence matrix
$ICOM$	Improved co-occurrence matrix
TIC	Texture-based Image Clutter
$CRLB$	Cramer-Rao Lower Bound
$DOYLE = [(\mu_T - \mu_B)^2 + k(\sigma_T - \sigma_B)^2]^{\frac{1}{2}}$	Metric proposed by L. Doyle
$TSSIM$	Target structure similarity
SCR	Signal-to-clutter ratio

dealing with the proposal of parameters to characterize automatic target tracking in image sequence, it is important to define "clutter".

In recent years, some researchers suggested that only the information interfering with the target is needed to be measured in images and the mentioned information is usually called clutter information. By removing the target and the system noise, the remaining part of a given image can be broadly called the clutter. A relatively widely accepted definition is that "clutter" is the scenario of image content similar to the target, yet not in the target region [36]. Researchers have considered several approaches to modeling the clutter in an image. We divide them into two categories: target independent metrics and target dependent metrics. Table 2 lists these "clutter" metrics.

2.2.1. Target Independent Metrics

Schmieder and Weathersby considered a metric C for clutter [37]. They proposed to divide the scene into N blocks, where the length of the block in each dimension is twice the size of the target. Then the variance within the block is measured and we can easily get the square root of the average of all the block variances in the image. Hence the

clutter measure is

$$C = \left[\sum_{i=1}^{N} \sigma_i^2 / N\right]^{1/2} \tag{6}$$

where σ_i^2 is the variance within the i-th block and N the number of the blocks.

The probability of edge metric (POE) [38–40] is designed to imitate the human visual system, which is basically a band-pass filter and highly sensitive to image edges. The technical details of this metric are as follows: the given image is divided into blocks twice the size of the target in each dimension firstly; then a DOOG (Difference of Offset Gaussian) [41] operator is applied to each block to simulate one of the channels in preattentive human vision; after that, the net effect is used to enhance the edges; then, the resulting histogram of the processed scene is normalized to values between 0 and 255 for each scene, while the threshold is chosen empirically; finally, the fraction of points which pass the threshold T in the i-th block is computed as $POE_{i,T}$. The probability of the edge metric, POE, is then calculated as

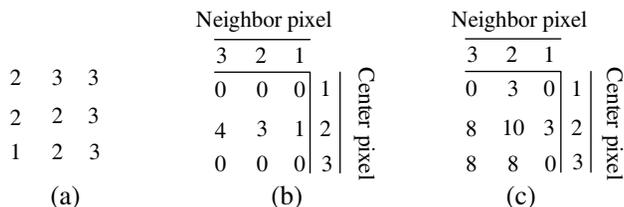$$POE = \left(\frac{1}{N}\sum_{i=1}^{N} POE_{i,T}^2\right)^{\frac{1}{2}} \tag{7}$$

In an application of wavelet technology in signal processing, Meitzler et al. proposed a wavelet-based multiresolution POE, namely $WPOE$ [42]. In order to evaluate the visual system more completely, this metric includes multiresolution levels of the edge point pictures computed from wavelets. The wavelet-based POE is defined as:

$$WPOE = \left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{l} WPOE_{i,T,j}^2\right)^{1/2} = \left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{l}\frac{ep_{i,T,j}^2}{cellpix_{i,j}}\right)^{1/2} \tag{8}$$

where $ep_{i,T,j}$ is the number of edge points above a certain threshold at a certain level and $cellpix$ the number of pixels in a cell at a certain level.

2.2.2. Target Dependent Metrics

Some target dependent metrics mainly consider the texture of the target. One of the best way to measure the texture of a defined image area is by calculating its Markov co-occurrence matrix (COM) [43, 44], which contains information about the intensity-value distribution and the possible transitions among neighbor pixels in the examined image

Neighbor pixel

	3	2	1	
	0	0	0	1
	4	3	1	2
	0	0	0	3

Center pixel

Neighbor pixel

	3	2	1	
	0	3	0	1
	8	10	3	2
	8	8	0	3

Center pixel

2	3	3
2	2	3
1	2	3

(a)   (b)   (c)

Figure 1. Calculation instance of gray-level co-occurrence matrix.

area. In order to illustrate how the COM is calculated, let us look at an example as shown in Fig. 1.

The COM is a square $(N \times N)$ matrix, where N is the number of possible pixel intensity values that occur in the image. In this example, $N = 3$ and the intensities of the pixels are shown in image (a). Each image pixel contributes to the COM according to its neighbor pixels' intensity distribution. For example, considering the center pixel, it has one 1, three 2's, and four 3's as neighbors. Its contribution to the COM can be described as image (b). Following the above rules, the overall COM of this image area (image (a)) is a 3×3 matrix as shown in image (c).

Based on the theory of COM, Aviram and Rotman developed a new texture metric which is called improved co-occurrence matrix ($ICOM$) [45]. This metric incorporates the attributes of global texture matching and of local texture distinctness, by analyzing the texture differences between the target and suspected target areas and also between suspected target areas and their local background. The technical details of the $ICOM$ are somewhat complicated and can be found in reference [45].

An image measure that considers both the clutter and the size of target is derived in [46]. It is named texture-based image clutter (TIC), and defined as

$$TIC = \frac{I(\Delta)}{\Delta} \qquad (9)$$

where Δ is the length of the object side and I the inertia measure, defined as

$$I(\Delta) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-j)^2 P_\Delta(i,j) \qquad (10)$$

In Equation (10), i, j represent the image gray level, $P_\Delta(i,j)$ is the (i, j)th element of P_Δ which is the location operator [43, 47].

Li and Zhang proposed a new TIC measure in [48]. This measure

can be calculated as

$$TIC = \frac{I(\Delta)}{\sum P_T(i, j) \times \Delta} \tag{11}$$

where P_T is the COM of the target region.

He et al. developed a background clutter measure for automatic target recognition, based on the theory of Cramer-Rao lower bound ($CRLB$) [49]. This metric defines the potential of false alarms by determining the correspondence level between a target and background through the application of the $CRLB$.

In addition, some local contrast metrics are usually designed to measure the target's distinctness from its background, assuming that high-contrast values relate to a more detectable target. One typical metric is the $DOYLE$ metric (named after L. Doyle) [45], which combines target and background means and variances of pixel values. The basic form of the $DOYLE$ metric is

$$DOYLE = \left[(\mu_T - \mu_B)^2 + k(\sigma_T - \sigma_B)^2 \right]^{\frac{1}{2}} \tag{12}$$

where μ_T and μ_B are target and background means; σ_T and σ_B are target and background standard deviations; k is a weighting coefficient. The other derivatives of the basic metric are the log $DOYLE$ metric,

$$DOYLE_{\log} = \left[(\log \mu_T - \log \mu_B)^2 + k(\log \sigma_T - \log \sigma_B)^2 \right]^{\frac{1}{2}} \tag{13}$$

and the hybrid $DOYLE$ metric,

$$DOYLE_{\mathrm{hyb}} = \left[(\log \mu_T - \log \mu_B)^2 + k(\sigma_T - \sigma_B)^2 \right]^{\frac{1}{2}} \tag{14}$$

Chang and Zhang [50, 51] developed a metric which is suitable for electro-optical images and named it target structure similarity ($TSSIM$). This measure estimates the degree of clutter in an image by a simple comparison of luminance, contrast and structure between the target and the background areas. Technical details of the $TSSIM$ are rather complicated and can be seen in Chang and Zhang's published papers.

Signal-to-clutter ratio (SCR), which describes the effects of clutter on target acquisition, is a very important concept for image measurement. This metric has been defined in several ways. For example, Wu et al. [52] defined the SCR as

$$SCR = 10 \log \frac{a^2}{\sigma^2} \tag{15}$$

where a^2 is the intensity of target and σ^2 the variance of local background. Meanwhile, Aviram and Rotam [45, 53] defined the SCR as

$$SCR = \frac{DOYLE}{POE} \tag{16}$$

2.3. Defects of Traditional Metrics for Image Sequences

After investigating the existing image metrics, we find that all these metrics focus on single image and do not take account of the characteristics of image sequences. This paper summarizes the defects of traditional metrics for measuring image sequences as follows.

(1) Traditional metrics mainly describe the information of target, background, noise, or the relationship among them in a single image. However, they do not consider the sequence information between different image frames and this will inevitably result in loss of information.

(2) The aim of traditional metrics is to measure the difficulty degree of target detection or recognition in images, while image sequences are generally used for target tracking. Therefore, measuring image sequences using the traditional intra-frame metrics inevitably leads to inaccuracy.

Since the existing metrics are not designed to measure image sequences which are currently used in ATR systems, it can be concluded that the current image measurement systems have an obvious drawback. Therefore, to establish a complete image measurement system, it is necessary to propose new metrics, suitable for image sequence characterization. Generally speaking, the purpose of processing image sequences is to track the target. Hence, quality measurement results of image sequences should reflect the difficulty degree of target tracking. This paper firstly studies the factors interfering with target tracking and then develops three new image sequence metrics.

3. IMAGE SEQUENCE MEASURES

In the process of studying target tracking algorithms, we found that the difficulty of target tracking is mainly caused by the inter-frame image changes, while the complexity within a single image disturbs target tracking only a little. See Fig. 2 for an illustration of the situation. Image (a) shows a real image sequence, while image (b) shows an artificial image sequence. In the first sequence, although the target and background are both very complicated, there is no significant change between different image frames. Experiments have proved that it is very easy to track the target in this image sequence. From the second sequence, we can see that each single image in this image sequence is simple, but inter-frame changes are acute. The textures of background and target, the target size and the target location all vary greatly in different image frames. We utilize several tracking algorithms to track the target in this image sequence, including the Kalman filter, the
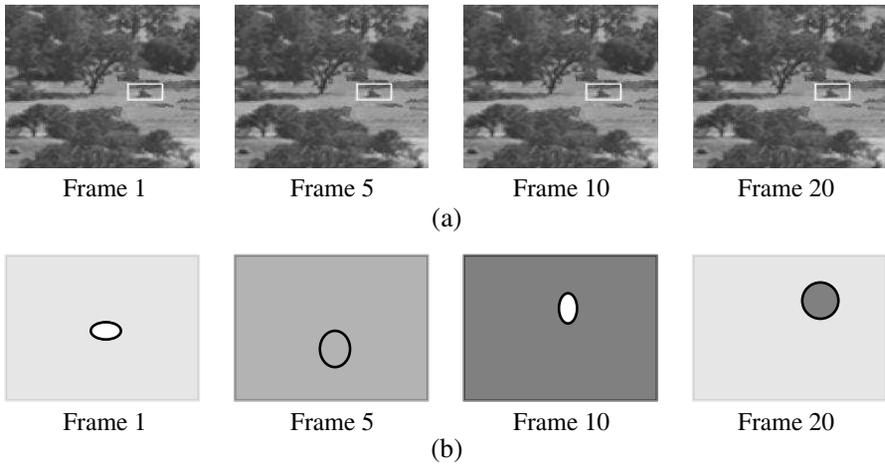
Figure 2. Two groups of image sequences.

Camshift tracker and the Particle filter. The tracking results of the three algorithms are all target loss. This experiment shows that target tracking in this sequence is very difficult.

Based on the above analysis, this paper concludes that inter-frame change is the main factor interfering with target tracking in image sequences. Therefore, it is necessary to exploit the *inter-frame change* information to measure image sequences. Further investigation finds that the regions far from the target can be neglected when measuring the image sequence, as these regions have little interference on target tracking. For example, the woods in the upper left corner of Fig. 2(a) are rather complicated, but they cannot interfere with the tracking task for most tracking algorithms. This article summarizes the factors that affect target tracking into 3 categories, namely:

(1) Inter-frame change of texture (the goal region includes the target and local background only).
(2) Inter-frame change of target size.
(3) Inter-frame change of target location.

The image sequences metrics proposed in this paper are calculated by quantitatively describing the above three factors, and the technical details of them are introduced in following parts.
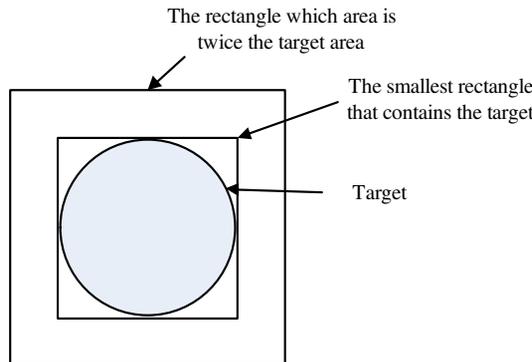
3.1. Inter-frame Change Degree of Texture ($IFCDT$)

Image texture is assumed to be one of the main perceptual cues that dominate visual performance and the term "texture" generally refers

to a repetition of basic texture elements. It is widely accepted that a texture element contains several pixels, whose placement can be periodic, quasiperiodic, or random. Texture characterization methods can be broadly classified into two main categories: statistical and structural. Textures which are random in nature are well suited for statistical characterization, while periodic or deterministic textures are better described using placement rules, which define the adjacency and periodicity of the pixels in the texture elements. As the targets are random and the scenarios are generally natural in the images used for ATR, a natural texture will be more appropriately described using the statistical approach.

As mentioned in Section 2.2.2, the co-occurrence matrix (COM) is a good way to measure the statistical properties of a texture in a defined image area. We utilize this method to describe the texture information of a single image in image sequence. The valid region for calculating inter-frame change of texture includes target region and local background surrounding the target. Specifically, the region is defined as the rectangle which area is twice the target area according to [45, 53] and [54], as shown in Fig. 3.

Further, this paper utilizes the ratio $\frac{\sum |CM_i - CM_j|}{\sum |CM_i + CM_j|}$ to describe the change degree of texture between two different image frames in the same sequence, where CM_i is the COM of the i-th image frame and CM_j is the COM of the j-th image frame. We can see that the above ratio equals to 0 when the two images are exactly the same, while it can equal 1, when the COMs of the two images are totally different.

Figure 3. Valid region for calculating inter-frame change degree of texture.

Then we define the *IFCDT* measure of an image sequence as follows:

$$IFCDT = \frac{1}{N-1} \sum_{i=2}^{N} \left(\frac{\sum |CM_i - CM_{i-1}|}{\sum |CM_i + CM_{i-1}|} \right) \qquad (17)$$

where N is the number of images in given image sequence and $\sum (Matrix)$ the sum of all the elements in the *Matrix*.

Considering that the normalized (to 1) value is usually easier to be accepted and understood than non-normalized values and since the possibility of *IFCDT* to be greater than 1 is very small, this new metric is redefined as:

$$IFCDT = \min \left(\frac{1}{N-1} \sum_{i=2}^{N} \left(\frac{\sum |CM_i - CM_{i-1}|}{\sum |CM_i + CM_{i-1}|} \right), \ 1 \right) \qquad (18)$$

3.2. Inter-frame Change Degree of Target Size (*IFCDTS*)

In general, the shape of the target is irregular. Therefore, this article describes the target size using the length and width of the smallest rectangle which contains the target, as shown in Fig. 3. An image sequence metric focuses on measuring the change degree of target size is developed in this section, and it is calculated as follows:

$$IFCDTS = \frac{1}{N-1} \sum_{i=2}^{N} \left(\frac{|l_i - l_{i-1}|}{l_{i-1}} + \frac{|w_i - w_{i-1}|}{w_{i-1}} \right) \qquad (19)$$

where N is the number of images in the given image sequence, and l_i and w_i are the length and width of the smallest rectangle containing a target in the i-th image frame, respectively. This definition makes sure that the *IFCDTS* equals 0 when the target size between adjacent frames is constant, and equals 1 as the target size change range between adjacent frames is equal to the target size itself. Being similar with *IFCDT*, we force *IFCDTS* values between 0 and 1 and modify Equation (19) to

$$IFCDTS = \min \left(\frac{1}{N-1} \sum_{i=2}^{N} \left(\frac{|l_i - l_{i-1}|}{l_{i-1}} + \frac{|w_i - w_{i-1}|}{w_{i-1}} \right), \ 1 \right) \quad (20)$$

3.3. Inter-frame Change Degree of Target Location (*IFCDTL*)

In this section, an image sequence metric for describing the change degree of target location is proposed. In our experiments, we have found that it is not true that target tracking will always be disturbed
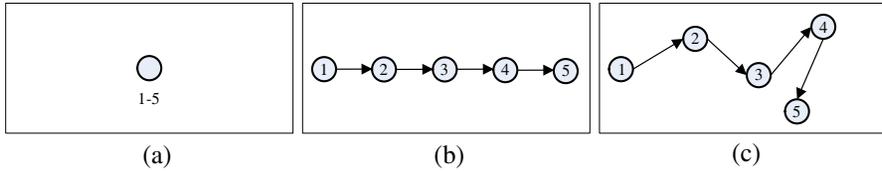
Figure 4. Three cases of target location changes in image sequences.
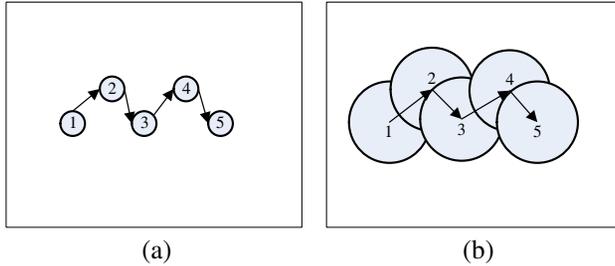
Figure 5. Two groups of image sequences with the same target displacement but different target sizes.

when target location changes in the sequence. See Fig. 4 for an illustration. Image (a) illustrates a scenario where the target location has no change in the whole image sequence; in image (b), the inter-frame change of the target location is regular, namely, uniform linear motion; in image (c), the target location in different image frames does not change regularly, showing irregular trajectories. We have found that, for most tracking algorithms, target location changes as shown in Figs. 4(a) and (b) have little influence on target tracking and tracking accuracy is high. However, when the image sequences meet the condition in Fig. 4(c), target tracking is more difficult, and tracking accuracy is low. Therefore, this paper concludes that the changes of target location do not interfere with target tracking if the target locations in different image frames is static or has uniform, linear motion trajectory. In this case, this new metric should be set to be 0.

In addition, for different image sequences with the same target location change conditions, the target size also has influence on the target tracking result. For example, as shown in Fig. 5, the target displacements in the two groups of image sequences are the same, while the target sizes are different. It is evident that tracking target in sequence Fig. 5(b) is easier than in sequence Fig. 5(a). Hence, this paper adds the target size as a factor for calculating *IFCDTL*.

Let N denote the number of images in a given image sequence, let

vector d_i be the target displacement from frame $i - 1$ to frame i, and let Δ_i denote the size of the target in the i-th image frame. The target size represents the mean area of the smallest rectangle which contains the target. The *IFCDTL* is defined as

$$IFCDTL = \min \left(\frac{1}{N-2} \sum_{i=3}^{N} \frac{|d_i - d_{i-1}|}{\Delta_{i-2}}, \ 1 \right) \qquad (21)$$

From Equation (21), we see that this metric equals to 0 not only when the target location in a given sequence is unchanged, but also when the target in a given sequence shows uniform linear motion, since in this case inter-frame target displacement d_i never changes. We can also see that this metric is set to be 1 in the special case when the target displacement is equal to or greater than the target size. These properties of the proposed metric are in accord with the results of our experiments.

3.4. Calculation Methods of the Three New Metrics When the Target Disappears

Because of being sheltered by background or other reasons, the target may disappear from the field of view. In this particular case, the calculation methods of *IFCDT*, *IFCDTS* and *IFCDTL* defined in Equations (18), (20) and (21) are not valid, because there is no information about target texture, target location and target size, which are essential factors in proposed metrics. We know that target tracking in the images without target information is very difficult. Therefore, these particular images should be assigned to high degrees of difficulty when measuring the image sequence using the new metrics. Taking *IFCDT* as an example, the contribution of the i-th frame image to *IFCDT* is $\frac{\sum |CM_i - CM_{i-1}|}{\sum |CM_i + CM_{i-1}|}$ under normal circumstances. However, for the image in which the target is totally shielded, we set its contribution to *IFCDT* to 1. Similarly, the contributions of the images without target to *IFCDTS* and *IFCDTL* are set to 1 as well.

In the following section, we will show the experiments for validating our image sequence metrics.

4. EXPERIMENTAL RESULTS

Generally speaking, image quality evaluation results should be associated with the ATR algorithm performance and there should be a monotone relationship between a good image measure and the ATR algorithm performance [46, 48]. Therefore, to validate the image

sequence metrics proposed in this paper, real image sequences are used as samples for analyzing the relationship between our metrics and actual performance of tracking algorithms. Specific experimental conditions are as follows.

4.1. Tracking Algorithm Selection

It is widely accepted that automatic target tracking algorithms can be divided into two categories: 'filtering and data association' tracking framework and 'target representation and location' tracking framework [55]. In this experiment, in order to ensure the universality of the selected algorithm, the tracking algorithm used here combines the advantages of the two mentioned tracking frameworks: the Kalman filter is used for location prediction, while the Mean-shift algorithm is used for accurate matching. The procedure of the tracking algorithm is broadly demonstrated as follows.

Step 1: Preprocess the initial frame image and each of the subsequent images, including image denoising and image enhancing and differential operator (Sobel operator) filtering;

Step 2: Extract the edge information of target to generate Eigen template in the preprocessed initial frame image;

Step 3: After tracking in the frame image $i - 1$, predict the start searching location in the frame image i by the Kalman filter;

Step 4: Search the optimal location of the target in frame image i near the start searching location predicted in Step 3 using the Mean-Shift algorithm;

Step 5: Update the Eigen template using the method mentioned in Step 2 every constant period;

Step 6: Advance to the next frame and go to Step 3, until the last frame of the image sequence is reached.

4.2. Experiment Samples

The samples used in the experiment are all real target image sequences, including 21 groups of visual image sequences and 75 groups of infrared image sequences, while each image sequence includes 200 image frames. Figs. 6 and 7 show some samples of them.

4.3. Image Sequence Metrics Selection

We have found that the texture, target size and target location in actual image sequences may be all variable in different image frames,

frame 1   frame 20   frame 40   frame 60
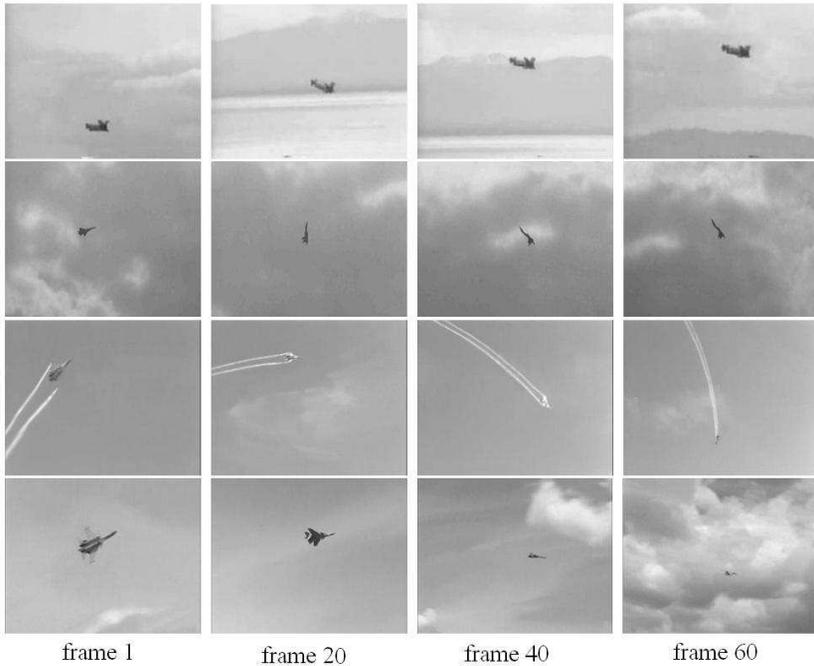
Figure 6. Four groups of visual image sequences in our experiment.

so it is meaningless to describe the relationship between a single metric (one of *IFCDT*, *IFCDTS* and *IFCDTL*) and the tracking result, as the other two metrics also reveal in part the tracking difficulty and destroy the one-to-one relationship between the selected metric and tracking result. In order to solve this problem, an integrated indicator which combines the three measures is proposed. We call it inter-frame change degree (*IFCD*), defined as:

$$IFCD = \sqrt{IFCDT^2 + IFCDTS^2 + IFCDTL^2} \qquad (22)$$

4.4. Tracking Performance Representation

In this experiment, the tracking precision is used to quantify the tracking results in given image sequences, defined as follows:

$$S = \frac{1}{N} \sum_{i=1}^{N} \sqrt{(|S_{oi} - S_{ri}|)^2} \qquad (23)$$

where N is the number of images in tested image sequence, i is current frame number, S_{oi} and S_{ri} are actual target location and target location

| frame 1 | frame 50 | frame 100 | frame 150 |
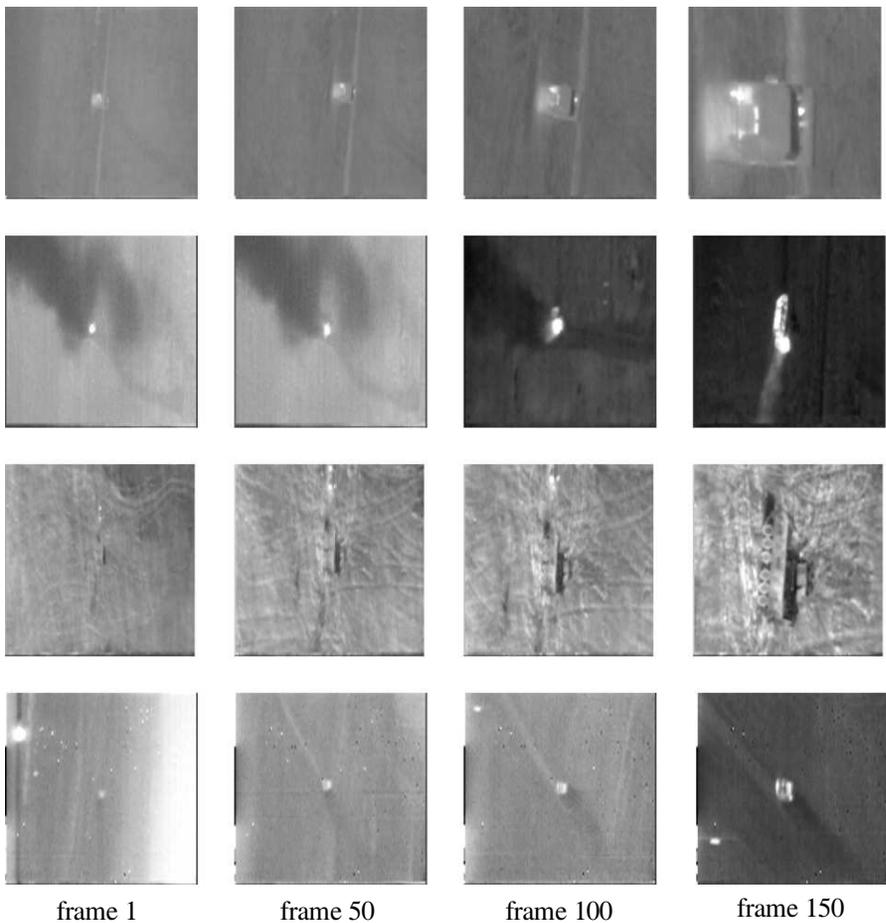
Figure 7. Four groups of infrared image sequences in our experiment.

obtained by tracking algorithm respectively.

Under the above experimental conditions, we get the experimental results of visual image sequences samples. For the tested 21 image sequences, their *IFCD* values are not exceeding 0.8049 and the relationship between *IFCD* values and tracking precisions is shown in Fig. 8. From this curve, we can see that there is a good, monotone, almost linear relationship between the image sequence metric and the actual tracking performance. Meanwhile, this paper calculates the values of two classical traditional image metrics (SNR and TBC) of the tested image sequences, which are used most extensively in the field of automatic target recognition and tracking. The relationship
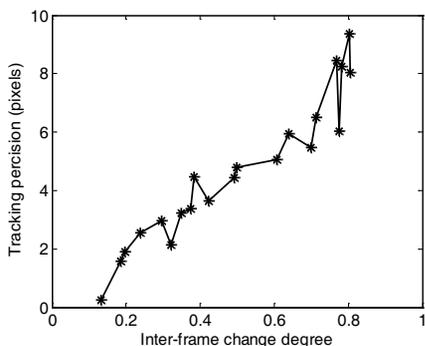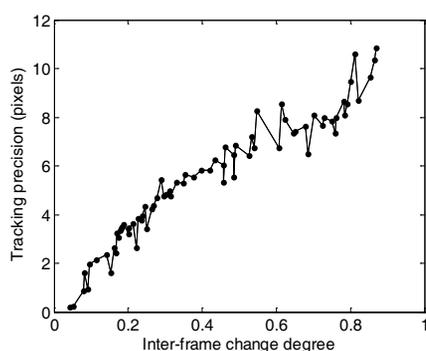
Figure 8. Relationship between tracking precision and inter-frame change degree for visual image sequences.
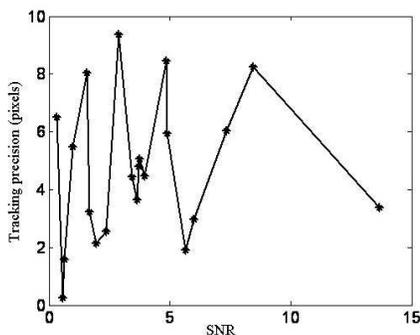
Figure 9. Relationship between tracking precision and SNR for visual image sequences.
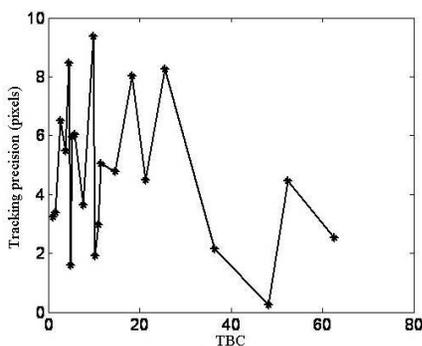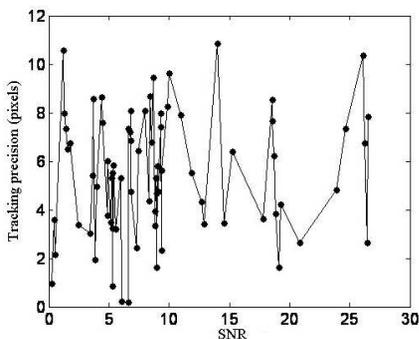

Figure 10. Relationship between tracking precision and TBC for visual image sequences.
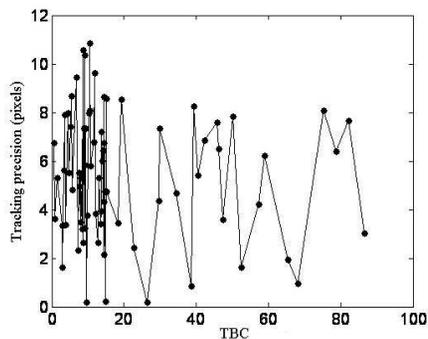

Figure 11. Relationship between tracking precision and inter-frame change degree for infrared image sequences.

between SNR values and tracking precisions is shown in Fig. 9, and the relationship between TBC values and tracking precisions is shown in Fig. 10. It is obvious that the correlations between traditional metrics and actual tracking performance are very poor.

In the second group of experiments, we get 75 sets of results of infrared image sequences samples and the relationship between *IFCD* values and tracking precisions is shown in Fig. 11. We can see that there is also a strong correlation between actual performance and our metric for infrared image sequences. In this experiment, we also

Figure 12. Relationship between tracking precision and SNR for infrared image sequences.

Figure 13. Relationship between tracking precision and TBC for infrared image sequences.

give the relationships of SNR values and TBC values with tracking precisions, as shown in Fig. 12 and Fig. 13. The curves in these two figures illustrate that the traditional metrics cannot describe the difficulty degree of the target tracking in infrared image sequences.

The experimental results show that the greater the *IFCD* value is, the greater the tracking error is and the tracking algorithm may miss the target when the *IFCD* value is high enough. Since the proposed image sequence metrics meet the criterion that there has to be a monotone relationship between a good image measure and ATR algorithm performance, this paper concludes that the new metrics are much more valid to measure the tracking difficulties in image sequences than the old ones.

5. DISCUSSION AND CONCLUSION

This paper studies and discusses thirty-four commonly used image measures and classifies them into two categories in terms of their definition principles: statistical metrics and "clutter" metrics. These metrics have been summarized in Table 1 and Table 2. On this basis, the paper finds that the existing image metrics have obvious deficiencies and are not appropriate to measure image sequences.

To compensate the lack of current image sequence measure systems in the field of automatic target recognition and tracking, this paper analyzes the characteristics of image sequences and concludes that the inter-frame change is the main factor interfering with target tracking. Further, this paper proposes three new measures for image sequences, including the inter-frame change degree of texture

($IFCDT$), the inter-frame change degree of target size ($IFCDTS$), and the inter-frame change degree of target location ($IFCDTL$). We prove by experiments that the new metrics are valid to measure image sequences with the purpose of assessing the difficulties associated with the task of target tracking.

In the future, we will continue to study valid image measure methods according to corresponding image processing purposes (target detection, recognition, and tracking). Further, we will establish an image database for performance evaluation of automatic target recognition algorithms, based on the theory of image measurement.

ACKNOWLEDGMENT

REFERENCES

1. Vicen-Bueno, R., M. Rosa-Zurera, M. P. Jarabo-Amores, et al., "Automatic target detection in simulated ground clutter (Weibull distributed) by multilayer perceptrons in a low-resolution coherent radar," *IET Radar, Sonar and Navigation*, Vol. 4, No. 2, 315–328, 2010.
2. Mušicki, D., "Doppler-aided target tracking in heavy clutter," *Proceedings of the International Conference on Information Fusion*, 1–7, 2010.
3. Dudgeon, D. E. and R. T. Lacoss, "An overview of automatic target recognition," *Lincoln Laboratory Journal*, Vol. 6, No. 1, 3–9, 1993.
4. Fiala, P., T. Jirků, R. Kubásek, P. Drexler, and P. Koňas, "A passive optical location with limited range," *PIERS Online*, Vol. 2, No. 6, 685–688, 2006.
5. Tian, B., D.-Y. Zhu, and Z.-D. Zhu, "A novel moving target detection approach for dual-channel SAR system," *Progress In Electromagnetics Research*, Vol. 115, 191–206, 2011.
6. AlShehri, S. A., S. Khatun, A. B. Jantan, R. S. A. Raja Abdullah, R. Mahmood, and Z. Awang, "Experimental breast tumor detection using NN-based UWB imaging," *Progress In Electromagnetics Research*, Vol. 111, 447–465, 2011.

7. Huang, C.-W. and K.-C. Lee, "Application of ICA technique to PCA based radar target recognition," *Progress In Electromagnetics Research*, Vol. 105, 157–170, 2010.

8. Chang, Y.-L., C.-Y. Chiang, and K.-S. Chen, "SAR image simulation with application to target recognition," *Progress In Electromagnetics Research*, Vol. 119, 35–57, 2011.

9. Wang, X., J.-F. Chen, Z.-G. Shi, and K. S. Chen, "Fuzzy-control-based particle filter for maneuvering target tracking," *Progress In Electromagnetics Research*, Vol. 118, 1–15, 2011.

10. Wang, Q. C., J. Li, M. Zhang, and C. Yang, "H-infinity filter based particle filter for maneuvering target tracking," *Progress In Electromagnetics Research B*, Vol. 30, 103–116, 2011.

11. Chen, Y., G. Chen, R. S. Blum, et al., "Image quality measures for predicting automatic target recognition performance," *Proceedings of the IEEE Aerospace Conference*, 1–8, 2008.

12. Clark, L. G. and V. J. Velten, "Image characterization for automatic target recognition algorithm evaluations," *Optical Engineering*, Vol. 30, No. 2, 147–153, 1991.

13. Liu, R., E. Liu, J. Yang, et al., "Point target detection of infrared images with eigentargets," *Optical Engineering*, Vol. 46, No. 11, 501–503, 2007.

14. Ma, Y. and B. Kong, "A study of object detection based on fuzzy support vector machine and template matching," *Proceedings of the World Congress on Intelligent Control and Automation*, 4137–4140, 2004.

15. Yang, L. and J. Yang, "Detection of small targets with adaptive binarization threshold in infrared video sequences," *Chinese Optics Letters*, Vol. 4, No. 3, 152–154, 2006.

16. Abousleman, G. P., M. W. Marcellin, and B. R. Hunt, "Hyperspectral image compression using entropy-constrained predictive trellis coded quantization," *IEEE Transactions on Image Processing,* Vol. 6, No. 4, 566–573, 1997.

17. Dachasilaruk, S., "Wavelet shrinkage and compression for SAR images," *Proceedings of the International Multi-conference on Systems, Signals and Devices*, 1–6, 2008.

18. Prasantha, H. S., H. L. Shashidhara, and K. N. Balasubramanya Murthy, "Image compression using SVD," *Proceedings of the International Conference on Computational Intelligence and Multimedia Applications*, 143–145, 2007.

19. Terki, N., N. Doghmane, Z. Baarir, et al., "Study of filter effects in wavelet image compression," *Proceedings of the International*

Conference on Information and Communication Technologies: From Theory to Applications, 369–370, 2004.

20. Trieu-Kien, T., J. Jyh-Horng, I. S. Reed, et al., "A fast encoding algorithm for fractal image compression using the DCT inner product," *IEEE Transactions on Image Processing*, Vol. 9, No. 4, 529–535, 2000.

21. Wang, Z., L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Image Communication*, Vol. 19, No. 2, 121–132, 2004.

22. Richard, A. P. I. and N. S. Robin, "Image complexity metrics for automatic target recognizers," *Proceedings of the Automatic Target Recognizer System and Technology Conference*, 1–17, 1990.

23. Beard, J., L. Clark, and V. Velten, "Characterization of ATR performance in relation to image measurements," *ATRWG Report*, AFWAL/AARF, Wright Patterson AFB, OH, 1985.

24. Gao, S. and P.-L. Shui, "Method for moving point target detection in image sequences based on directional cumulation," *Proceedings of the SPIE 6795*, 67952I-1–67952I-6, 2007.

25. Nevis, A., "Image characterization and target recognition the surf zone environment," *Proceedings of the SPIE*, Vol. 2765, 46–58, 1996.

26. Yonoviz, D., "Tunable wavelet target extraction preprocessor," *Proceedings of the SPIE*, Vol. 6569, 1–12, 2007.

27. Mao, X. and W.-H. Diao, "Criterion to evaluate the quality of infrared small target images," *Journal of Infrared, Millimeter and Terahertz Waves*, Vol. 30, No. 1, 56–64, 2009.

28. Lahart, M., T. Jones, and F. Shields, "Trends and capabilities of ATR algorithm technology for ground vehicles," *Proceedings of the IRIS Conference on Targets, Backgrounds, and Discrimination*, Monterey, CA, 1988.

29. Sadjadi, F., "Measures of effectiveness and their use in comparative image fusion analysis," *Proceedings of the IEEE Geoscience and Remote Sensing Symposium*, 3659–3661, 2003.

30. Garlson, J. J., J. B. Jordan, and G. M. Flachs, "Task specific complexity metrics for electronic vision," *Proceedings of the International Conference on Image Processing, Analysis, Measurement, and Quality*, 35–40, 1988.

31. Sadjadi, F. A. and M. E. Bazakos, "Perspective on automatic target recognition evaluation technology," *Optical Engineering*, Vol. 30, No. 2, 183–188, 1991.

32. Loyd, G. C. and J. V. Vincent, "Image characterization for

automatic target recognition algorithm evaluations," *Optical Engineering*, Vol. 30, No. 2, 147–153, 1990.

33. Todt, E. and C. Torras, "Detection of natural landmarks through multiscale opponent features," *Proceedings of the International Conference on Pattern Recognition*, 976–979, 2000.

34. Zhou, C., G. Zhang, and J. Pen, "A general evaluation method for segmentation algorithm based on experimental design methodology," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 258–262, 1995.

35. Vicen-Bueno, R., R. Carrasco-Alvarez, M. Rosa-Zurera, et al., "Sea clutter reduction and target enhancement by neural networks in a marine radar system," *Sensors*, Vol. 9, 1913–1936, 2009.

36. Bhanu, B., "Automatic target recognition: State of the art survey," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 22, No. 4, 364–379, 1986.

37. Schmieder, D. E. and M. R. Weathersby, "Detection performance in clutter with variable resolution," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 19, No. 4, 622–630, 1983.

38. Rotman, S. and M. L. Kowalczyk, "Clutter analysis for modeling and improving human and automatic target acquisition," *Proceedings of the SPIE 2020*, 131–142, 1993.

39. Tidhar, G., G. Reiter, Z. Avital, et al., "Modeling human search and target acquisition performance: IV. Detection probability in the cluttered environment," *Optical Engineering*, Vol. 33, No. 3, 801–808, 1994.

40. Ralph, S. K., J. M. Irvine, M. Snorrason, et al., "An image metric-based ATR performance prediction testbed," *Proceedings of the International Conference on Artificial Intelligence and Pattern Recognition*, 192–197, 2005.

41. Young, R. A., "Simulation of human retinal function with the Gaussian derivative model," *Proceedings of the IEEE Proceedings on Computer Vision and Pattern Recognition*, 564–569, 1986.

42. Meitzler, T. J., R. E. Karlsen, G. R. Gerhart, et al., "Wavelet transforms of cluttered images and their application to computing the probability of detection," *Optical Engineering*, Vol. 35, No. 10, 3019–3025, 1996.

43. Haralick, R. M., K. Shanmugan, and I. Dinstein, "Texture features for image classification," *IEEE Transactions on System, Man and Cybernetics*, 610–621, 1973.

44. Waldman, G., J. Wootton, G. Hobson, et al., "A normalized clutter measure for image," *Computer Vision, Graphics and Image*

Processing, Vol. 42, No. 3, 137–156, 1988.

45. Aviram, G. and S. R. Rotman, "Evaluation of human detection performance of targets and false alarm, using a statistical texture image metrics," *Optical Engineering*, Vol. 39, No. 8, 2285–2295, 2000.

46. Trievdi, M. M. and M. V. Schirvaikar, "Quantitative characterization of image clutter: Problems, progress, and promises," *Proceedings of the International Conference on Characterization, Propagation, and Simulation of Sources and Backgrounds*, 288–299, 1993.

47. Conners, R. and C. Harlow, "A theoretical comparison of texture algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 2, No. 2, 204–222, 1980.

48. Li, M. and G. Zhang, "Image measures for segmentation algorithm evaluation of automatic target recognition system," *Proceedings of the International Symposium on Systems and Control in Aerospace and Astronautics*, 673–679, 2006.

49. He, G., J. Zhang, and H. Chang, "Clutter metric based on the Cramer-Rao lower bound on automatic target recognition," *Applied Optics*, Vol. 47, No. 29, 5534–5540, 2008.

50. Chang, H.-H. and J.-Q. Zhang, "Evaluation of human detection performance using target structure similarity clutter metrics," *Optical Engineering*, Vol. 45, No. 9, 41–47, 2006.

51. Chang, H. and J. Zhang, "New metrics for clutter affecting human target acquisition," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 42, No. 1, 361–368, 2006.

52. Wu, B., H.-B. Ji, and P. Li, "New method for moving dim target detection based on third-order cumulant in infrared image," *Journal of Infrared and Millimeter Waves,* Vol. 25, No. 5, 364–367, 2006.

53. Aviram, G. and S. R. Rotam, "Analyzing the effect of imagery wavelength on the agreement between various image metrics and human detection performance of targets embedded in natural images," *Optical Engineering*, Vol. 40, No. 9, 1877–1884, 2001.

54. Rotman, S. R., D. Hsu, A. Cohen, et al., "Textural metrics for clutter affecting human target acquisition," *Infrared Physics & Technology*, Vol. 37, No. 6, 667–674, 1996.

55. Yang, L., "Study on infrared small target detection and tracking algorithm under complex backgrounds," Ph.D. Thesis, Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, 2006.