Pixel-Wise Signal-to-Noise Ratio: A Novel Metric for Quantifying the Detectability of Targets in Infrared Images

Seyit Tunç[1] , Hakkı Alparslan Ilgın[2*]

[1] Roketsan, Ankara 06780, Turkey

[2] Department of Electrical and Electronics Engineering, Ankara University, Ankara 06830, Turkey

Corresponding Author Email: ilgin@eng.ankara.edu.tr

ABSTRACT

In this paper, a new Signal-to-Noise Ratio (SNR) metric is proposed to quantify the detectability of targets in infrared (IR) images. The proposed metric is based on the contrast between the target and the background, which is consistent with human perception in terms of distinguishing the target from the background, rather than the raw intensity values of the target. In the contrast calculation, individual contribution of each pixel value of the target is considered in the proposed metric, whereas the mean or a single representative raw intensity value of the target is taken into account in the existing metrics. As subjective evaluations are the most precise tools for distinguishing the target from the background, SNR metrics used for IR images are expected to be as consistent as possible with the human visual system. That is, due to its high contrast sensitivity, the human visual system responds to stimuli by cognitively distinguishing the target from the background. Therefore, human perception-inspired target distinguishability metrics aim to quantify the target detectability consistent with the human visual system, which is capable of distinguishing very small differences in contrast. Extensive performance evaluation tests on well-known IR image datasets, VIVID, SENSIAC and AMCOM, and synthetic image sets demonstrate that the proposed pixel-wise SNR metric quantifies target distinguishability from the background more consistently with subjective evaluations than other SNR metrics. Furthermore, the proposed metric is always robust even when the other metrics fail to accurately quantify target distinguishability.

1. INTRODUCTION

Signal-to-Noise Ratio (SNR) quantifies signal or desired information with respect to noise or unwanted information. In infrared (IR) images, where the distinguishability of a target from the background is examined, the desired information is the target, and the unwanted information is the background. Since target distinguishability assessments are based on human perception, SNR metrics are expected to quantify IR images as consistently as possible with subjective assessment results.

Quantifying target-background distinguishability in IR images by means of an objective metric is an essential task for evaluating object or target detection algorithms in remote sensing and reconnaissance applications. For this aim, various SNR metrics are used to interpret IR images in terms of target distinguishability from the background [1-3]. SNR metrics should accurately quantify this distinction consistent with the human visual system as subjective tests are taken as the reference gold standard for SNR metrics [4]. As being objective measures, SNR metrics consider two important aspects of an IR image that affect quantification success: (i) target saliency, and (ii) background complexity. Performance of a detection algorithm can be evaluated by comparing the SNR values regarding the target distinguishability of the input and enhanced output images [5]. Accordingly, the image enhancement for target distinguishability can be measured by SNR gain, which is the ratio of the output SNR to the input

SNR [5, 6]. In addition to SNR gain, background suppression factor (BSF) is also used to evaluate the performance of a detection algorithm [7]. However, BSF quantifies the enhancement via background suppression, and does not contain any information about the target. Therefore, SNR gain is more commonly used than BSF to quantify enhancement in IR images. However, the reliability of the SNR gain depends on the accuracy of the SNR metric. If the SNR metric does not accurately quantify the target-background distinguishability, all results based on the SNR metric will be adversely affected. Even though different SNR metrics are used to quantify IR images for the detectability of targets, accuracy of these metrics has not been discussed.

SNR metrics used in the literature to quantify IR images in terms of target detectability can be grouped into three main categories. In the first category, defined as raw SNR, metrics are given as the ratio of the target to the standard deviation of the background [1, 8]. Since these metrics do not take into account the contrast of the target with respect to the background [1, 9], they may not accurately quantify target distinguishability under different brightness conditions. In order to overcome this disadvantage, metrics that consider the contrast of the target against the background are used [2, 10]. These metrics, which are considered in the second category called contrast-to-noise ratio (CNR), can give inconsistent results with the human visual system when the target is not uniform. The third category of the metrics is called local signal-to-background ratio (LSBR), because the contrast

calculation uses a limited region around the target instead of the whole background [3]. LSBR can give different values depending on the selection of the local region, which may result in inconsistency with subjective evaluations. Consequently, in a target and background condition for which one metric is consistent with the human visual system, another metric may yield a less consistent or completely inconsistent result. In this context, the main problem with the SNR metrics that quantify target distinguishability in IR images is that they do not always give results consistent with subjective evaluations due to the specified disadvantages.

In this paper, considering the drawbacks of the existing metrics, a novel SNR metric consistent with the human visual system is proposed to quantify the distinguishability of the targets in IR images. SNR metrics are compared with the mean opinion scores of subjective evaluations to check the consistency of the proposed and other metrics with the human visual system. Well-known IR image datasets and synthetic image sets were used in the evaluations. Test results show that the proposed metric is more consistent with the subjective evaluations compared to other SNR metrics. Also, the proposed metric is always robust in quantifying distinguishability of different types of targets with various background conditions, whereas other metrics fail with certain types of targets and background conditions.

The organization of this article is as follows. Current SNR metrics and the proposed pixel-wise SNR metric are explained in Section 2 and 3, respectively. Experimental results and related discussions are given in Section 4. Conclusion is given in the last section of the paper.

2. EXISTING SNR METRICS

In this section, properties and drawbacks of SNR metrics used to quantify the detectability of targets in IR images are explained. These metrics can be grouped into three main categories as follows:

(i) Raw SNR: Metrics that calculate the ratio of the target signal to the standard deviation of the background are expressed by the following equation:

$$rawSNR = \frac{I_T}{\sigma_B} \tag{1}$$

where, I_T and σ_B represent intensity value of the target, and the standard deviation of the background, respectively [1]. In some studies, background is considered as the entire image if the target is very small [11], whereas it is mostly taken as the area other than the target especially when the target is relatively large [12]. In some studies, raw SNR is calculated as the second power of Eq. (1) [8], which is also given in dB as follows [9]:

$$rawSNR_{dB} = 10log\left(\frac{I_T}{\sigma_B}\right)^2 \tag{2}$$

Different definitions for intensity value of the target I_T in Eq. (1) and (2) have been given in the literature. For one pixel-targets, target's own intensity value is considered as the target signal [11]. For a larger target, maximum [13] or mean [14] intensity value of the target pixels is commonly used to represent the target. Various raw SNR metrics that use the mean or standard deviation of the target are utilized in the

paper of Usamentiaga et al. [15] for defect quantification in IR images. SNR is also referred to as Signal-to-Clutter Ratio (SCR) in some studies, and calculated by various equations [16, 17].

Despite its widespread use, raw SNR has two main disadvantages. The first one is that raw SNR value of an IR image with high-intensity background can be large even if the target is dim or barely distinguishable from the background. As the second disadvantage, brighter images can have larger raw SNR values than darker images even if the target distinguishability and contrast is low.

(ii) Contrast-to-Noise Ratio (CNR): Metrics that consider the contrast of the target against the background are called CNR and can generally be described by either:

$$CNR = \frac{|I_T - \mu_B|}{\sigma_B} \tag{3}$$

or

$$CNR_{dB} = 10log\left(\frac{|I_T - \mu_B|}{\sigma_B}\right)^2 \tag{4}$$

where, μ_B is the mean intensity value of the background. If there is a one-pixel target in the IR image, I_T represents the target [10]. If the target consists of more than one pixel, I_T represents the minimum [18], maximum [19], or mean [20] intensity value of the target. In the literature, the mean value is mostly used because it represents all pixel values of the target [20]. However, in case of bimodal target, i.e., if some parts of the target are brighter and other parts are darker than the background, intensity of the target I_T becomes low because of the mean calculation. Consequently, quantifying the detectability of the target by CNR becomes unreliable because of the contradiction between low CNR value and high distinguishability score obtained from subjective evaluations.

(iii) Local Signal-to-Background Ratio (LSBR): Another commonly used specific SNR metric is LSBR [3], which is given as:

$$LSBR = 10log\left(\frac{\sum_{i=L_x}^{H_x}\sum_{j=L_y}^{H_y}(I(i,j) - \mu_R)^2}{\sigma_R^2}\right) \tag{5}$$

This metric is calculated for a region whose boundaries are given by the coordinates (L_x, L_y) to (H_x, H_y) in image I. In this equation, mean intensity value and standard deviation of the region pixels are μ_R, and σ_R, respectively. There are several studies in the literature that use LSBR metric. For example, it is used in the analysis of IR dim small targets in the study of Bai and Zhou [21] and in the evaluation of small target enhancement obtained by IR sensor in the study of Bai [22]. The result of LSBR is closely related to target size, since the mean is not calculated for the target. Additionally, the clutter of the region around the target, not the whole background, is taken into account. Therefore, LSBR can yield different values based on the region around the target. Since LSBR does not consider the contrast of target pixels against the background, effect of the pixels of the target on the contrast calculation is reduced. Besides, the background complexity is not reflected in LSBR, as only the surrounding area of the target is used in LSBR calculation.

3. PROPOSED SNR METRIC: PIXEL-WISE SIGNAL-TO-NOISE RATIO

Since contrast is a decisive parameter in target detection [23], it is essential that an SNR metric takes contrast into account. This is due to the fact that perception is associated with contrast, which simulates the human visual system [24]. High-contrast regions appear more salient than their surroundings [25], even when the background clutter is high. According to a psychological study [26], if there is more than one stimulus inside a neuron's perception area, the neuron tends to choose stimuli with the highest contrast. For this reason, target detection capability of a human is directly related to target's contrast to the background rather than target's brightness [27]. However, since raw SNR does not consider contrast, it often does not give results compatible with the human visual system. Also, although CNR is based on contrast calculation, there is no contribution to the contrast by the diversity of the target pixels against the background, as contrast calculation in CNR is obtained by subtracting the background average from the target pixels' maximum, minimum, or average value. As a result, CNR may give inconsistent results with the human visual system for nonuniform targets. Also, quantification of LSBR is limited to the area around the target.

Considering the disadvantages of the existing metrics, a new pixel-based SNR (pwSNR) metric has been proposed, the consistency of which with the human visual system is shown by the subjective evaluations. The proposed metric is given as:

$$pwSNR = \frac{E[|T(i,j) - \mu_B|]}{\sigma_B} \quad (6)$$

or

$$pwSNR_{dB} = 10log\left(\frac{E[|T(i,j) - \mu_B|]}{\sigma_B}\right)^2 \quad (7)$$

where, $T(i,j)$ are the target pixel intensities located in rows i, and columns j. $E[.]$ denotes the expected value. Unlike the contrast calculation in CNR, in the proposed metric, contrast between the target and the background is obtained in each target pixel separately, and then the average is calculated. Therefore, the advantage of the proposed metric is that it quantifies the distinguishability of the target against the background with the contribution of the bright and dark parts of the target to the contrast individually. In addition, the results are not dependent on the intensity change of the background or the image. The proposed and existing SNR metrics are briefly presented in Table 1.

As summarized in Table 1, only the intensity of the target is taken into account, along with the standard deviation of the background, to calculate raw SNR. In other words, SNR does not evaluate target pixels against the background and is vulnerable to mean changes. As for CNR, it does not calculate individual contribution of each target pixel to the contrast. Therefore, structure of the target is not taken into account in CNR. For example, when computing mean of a nonuniform target with different temperatures in an IR image, brighter and darker areas cancel each other's contribution to the contrast calculation in CNR, which results in low contrast and subsequently low CNR value even though the human perception of the contrast of the target is high. LSBR is related to the surrounding region around the target, without considering target structure and standard deviation of the background. On the other hand, in pwSNR, since the mean of the background is firstly subtracted from each pixel of the target and then absolute values of the resulting pixels are obtained, brighter and darker parts of the target pixels do not cancel each other's contribution to the contrast of the target against the background.

Consistency of the target distinguishability quantification of the proposed metric with subjective evaluations and disadvantages of aforementioned SNR metrics are demonstrated by the experiments in the next section.

Table 1. Existing and proposed SNR definitions

Definition	Formulation	Explanation
rawSNR	$\dfrac{I_T}{\sigma_B}$	Problem of quantifying target distinguishability from the background accurately, since only intensity of the target is considered.
CNR	$\dfrac{|I_T - \mu_B|}{\sigma_B}$	Quantifies target detectability more reasonable than raw SNR since background mean is also taken into consideration together with the target intensity. However, it is weak for nonuniform targets.
LSBR	$10log\left(\dfrac{\sum_{i=L_x}^{H_x}\sum_{j=L_y}^{H_y}(I(i,j) - \mu_R)^2}{\sigma_R^2}\right)$	Considers only the local clutter, and results are closely related to the local region size.
pwSNR (Proposed)	$\dfrac{E[|T(i,j) - \mu_B|]}{\sigma_B}$	Considers the contrast between the background and each pixel of the target instead of raw intensity of the target or contrast between the mean intensity values of the background and the target. It is always robust and consistent with subjective evaluations.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

The consistency of the proposed and other metrics with the human visual system was compared using real-world IR image datasets and synthetic image sets. For this aim, the mean opinion scores of subjective evaluations for target-background distinguishability were used as reference. A subject group of 20 engineering graduates familiar with IR images participated in the subjective evaluations.

Real-world IR images from three different well-known IR image datasets, VIVID, SENSIAC and AMCOM, were used in the evaluations. Also, 8-bit synthetic images with a resolution of 128×128 pixels were generated for the evaluations. Two types of patterns, uniform and nonuniform, were used as targets in the synthetic images, since both target types are common in IR images. There is a four-bar pattern among the nonuniform targets used in the evaluations, as it has been used in some studies on IR images before [28, 29]. It was defined in the study of Ratches et al. [28] for Minimum Resolvable Temperature Difference (MRTD) studies in IR

images, and a similar pattern was also used in a psychological study for visual saliency [29]. It represents a target with low and high temperature regions. Uniform pattern represents a target whose pixels have the same or close intensity values, i.e., temperatures. An 8-bit grayscale synthetic image can be modeled as:

$$I = B + T \qquad (8)$$

where, I, B, and T are image, background and target, respectively. An image consisting of background and a target is shown in Figure 1. In this study, background is modeled using Gaussian noise in the synthetic images.

In the following subsections, disadvantages of the SNR metrics described in the previous sections and consistency of the proposed metric with the subjective evaluations are demonstrated by experiments.
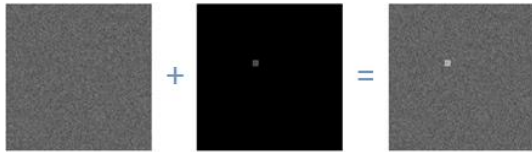
Figure 1. Synthetic image model: Background (left), target (middle) and resulting image (right)

4.1 Examples regarding the inconsistency of the existing metrics with subjective evaluations

The main disadvantage of the existing metrics is that they do not give consistent results with the subjective evaluations in all target and background conditions. As mentioned in Sec. 2, raw SNR cannot accurately quantify target distinguishability in situations where it is difficult to subjectively distinguish the target from the high-intensity background. An example for this case is given by the images shown in Figure 2.
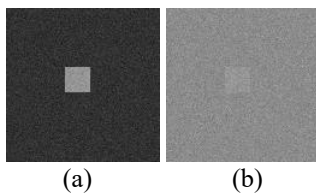

Figure 2. Same target with different backgrounds

In the subjective evaluations of these two images, all subjects agreed that the distinguishability of the target in Figure 2 (a) is higher than the target in Figure 2 (b). The targets in both images are the same with the mean intensity value of 149.9, while the mean intensity value of the background on the left and on the right are 50.1 and 140.0, respectively. Standard deviation of both backgrounds is the same, which is 8.07. By using Eq. (1) with I_T being the mean intensity value of the target, raw SNR of both images is the same with the value of 18.58. Therefore, raw SNR is not consistent with the subjective evaluation results, as it cannot quantify target-background distinguishability correctly. That is, since the target intensity and background variance of both images are the same, raw SNR values calculated for both images are also the same. In Table 2, results of all SNR metrics for the images in Figure 2 (a) and (b) are given. As seen in this table, all

metrics, except raw SNR, are consistent with the subjective evaluation results on the target distinguishability of the images, i.e., values of all metrics except raw SNR for Figure 2 (a) are larger than those for Figure 2 (b).

Table 2. Results of the SNR metrics for the images in Figure 2 (a) and (b)

Image	raw SNR	CNR	LSBR	pwSNR
Figure 2 (a)	18.58	10.28	1021.2	10.28
Figure 2 (b)	18.58	1.18	633.4	1.32

The second example regarding the inconsistency of raw SNR is given in Figure 3. Properties of the images in Figure 3 are given in Table 3, where, μ_T, μ_B, and μ are the mean intensity of the target, background, and whole image, respectively.
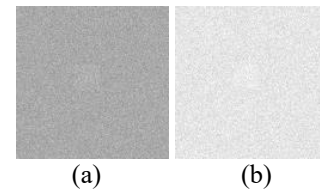
Figure 3. Images with different brightness

In this example, it was shown that brighter images can have larger raw SNR values than darker ones even if target distinguishability and contrast is low. Images in Figure 3 have the same target distinguishability from the background according to the subjective evaluations, and also have the same contrast. However, raw SNR of the left and right images are 24.95 and 31.74, respectively, Therefore, raw SNR results are not consistent with the subjective evaluations.

Table 3. Properties of the images in Figure 3

Image	μ_T	μ_B	μ
Figure 3 (a)	174.7	169.9	170.1
Figure 3 (b)	234.6	229.9	230.0

On the other hand, as shown in Table 4, pwSNR values for both images are nearly equal, which is the most consistent result with the subjective evaluations.

Table 4. Results of the SNR metrics for the images in Figure 3 (a) and (b)

Image	raw SNR	CNR	LSBR	pwSNR
Figure 3 (a)	24.95	0.65	590	0.97
Figure 3 (b)	31.74	0.64	460	0.96

In the next example, the most common disadvantage of CNR is considered. That is, although CNR takes into account the contrast, it may give inconsistent results with subjective evaluations for bimodal targets.
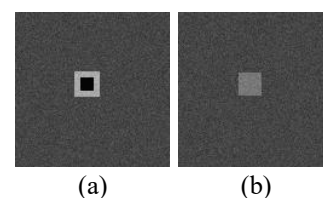
Figure 4. Bimodal and monomodal targets

In Figure 4, both images have the same background with a mean of 70.06. Also, the target size in both images is the same, but the means of the inner and outer parts of the bimodal target in the left image are 5.0, and 169.69, respectively. Target mean of the right image is 119.99. In the subjective evaluations, all subjects agreed that the target of the image in Figure 4 (a) is more distinguishable than the one in Figure 4 (b). However, CNR value of the image with bimodal target is 5.23, whereas CNR value of the right image is 5.97 as shown in Table 5. Therefore, target distinguishability is not accurately quantified by CNR, as it is inconsistent with the subjective evaluation results.

Table 5. Results of the SNR metrics for the images in Figure 4 (a) and (b)

Image	Raw SNR	CNR	LSBR	pwSNR
Figure 4 (a)	20.10	5.23	1297	9.36
Figure 4 (b)	17.00	5.97	777	5.97

## 4.2 Performance evaluations with synthetic and real-world image sets

In this section, performance evaluations are carried out through image sets with five different target distinguishability levels. Initial assessments have three sets of five images each. Experiments were then conducted for three sets of 25 images each. These experiments are described respectively in the next two subsections.

4.2.1 Experiment for three sets of five images each

In this experiment, two synthetic and one real-world image sets are used. These sets, each consisting of five images, are shown in Figure 5. In this figure, each image set is shown in a column from top to bottom with increasing target distinguishability based on subjective evaluations. Image set with uniform target, image set with four-bar target, real-world IR image set, and ground truth set of the real-world IR image set are shown in Figure 5 (a), (b), (c) and (d), respectively. Images with 320×256-pixel resolution from VIVID dataset are seen in the first and the fourth row, while a 640×512-pixel image from SENSIAC dataset is shown in the second row of Figure 5 (c). Images from AMCOM dataset with 128×128-pixel resolution are placed in the third and the last row of Figure 5 (c).

For subjective evaluations, subjects were shown the test images of the uniform, four-bar and real-world image sets in Figure 5 (a), (b) and (c), respectively, and asked to score target distinguishability for each set. Hard-to-detect targets have the lowest score of 1, while others have increasing scores by one according to the distinguishability up to the score of 5. Finally, for each set, the mean of the scores given by the subjects to each image was calculated.

The mean opinion scores of the subjective evaluations for the three image sets in Figure 5 are shown at the left column of Figure 6 (a, b, c). In this figure, the image numbers in the horizontal axis correspond to the images shown in Figure 5 from top to bottom.
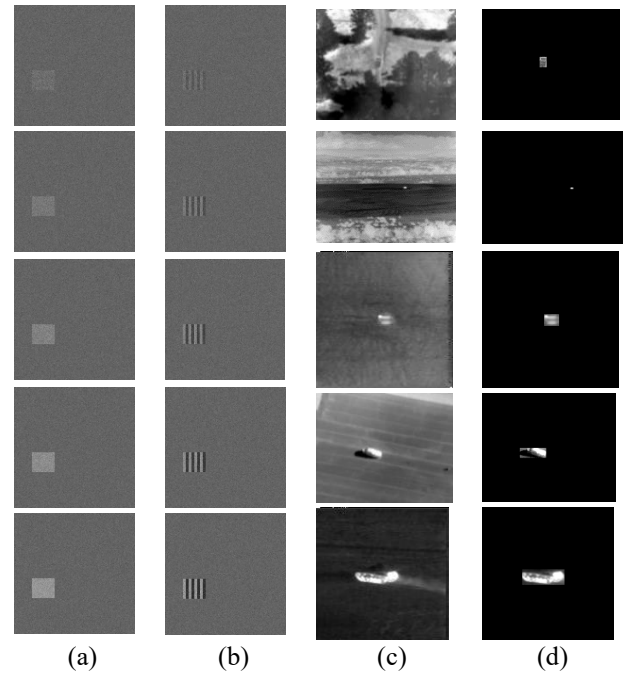
Figure 5. Image sets with five different target-background distinguishability levels used to evaluate SNR metrics. a) Synthetic uniform target image set, b) synthetic four-bar target image set, c) real-world IR image set, d) ground truth of the images in (c)
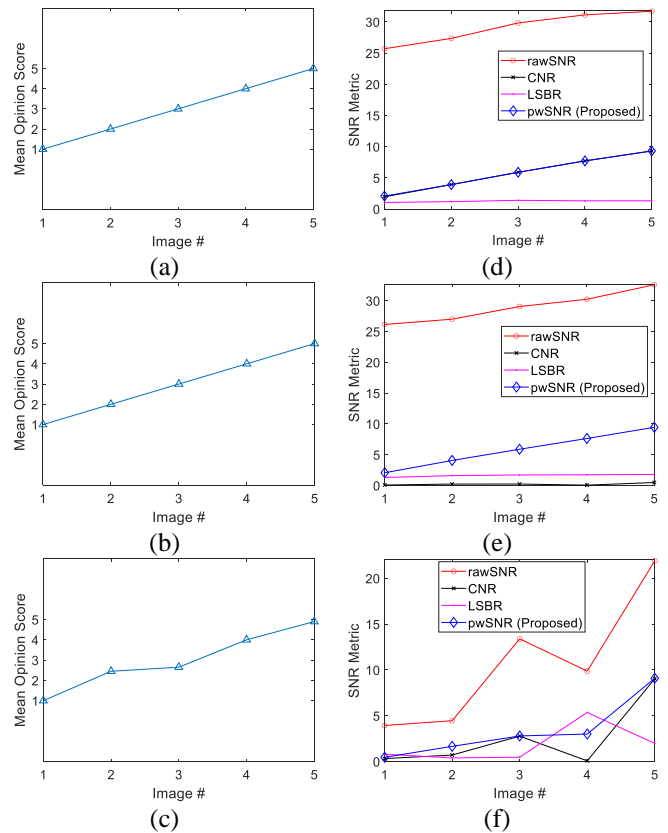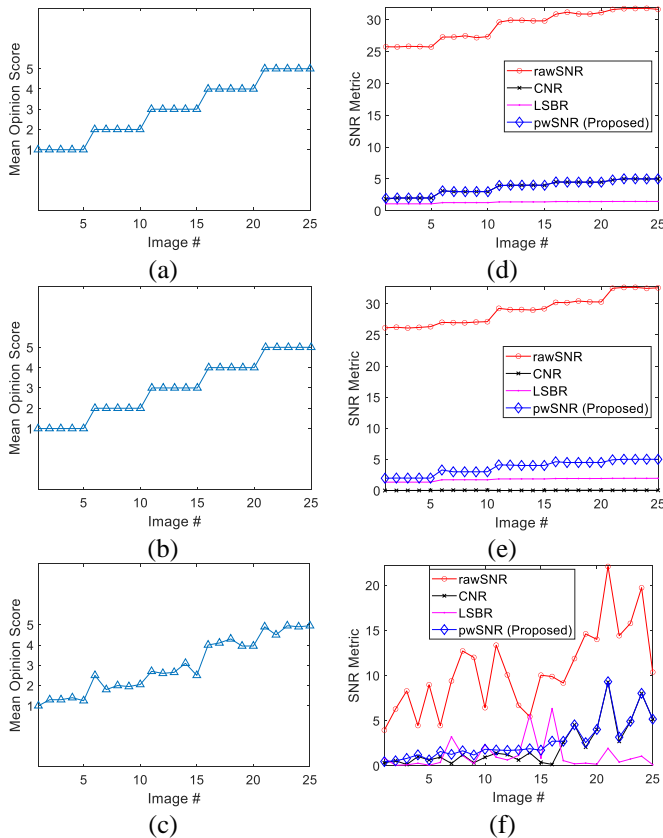

Figure 6. Results of subjective evaluations and SNR metrics for the experiments with uniform, four-bar and real-world IR image sets with 5 images each. Left column from top to bottom: Mean opinion scores of the sets in Figure 5 (a), (b) and (c), respectively. Right column from top to bottom: SNR metric results of the sets in Figure 5 (a), (b) and (c), respectively

All subjects gave the same scores to the same images for the distinguishability of the synthetic uniform targets in Figure 5 (a), and the mean opinion scores for this set is given in Figure 6 (a). Only three of the subjects have different opinions for synthetic four-bar targets in Figure 5 (b), for which mean opinion scores are given in Figure 6 (b). Also, mean opinion scores of the real-world IR images in Figure 5 (c) are shown in Figure 6 (c).

Results of SNR metrics for the image sets in Figure 5 are given in Figure 6 (d), (e), and (f), next to the mean opinion scores. LSBR results are divided by 1000 in order to be shown together with the other metrics. When compared to the mean opinion scores at the left column of Figure 6, it is seen that the most consistent metric is pwSNR for all image sets. On the other hand, other metrics are less consistent with the subjective evaluations, especially for real-world image set. For uniform target, CNR and the proposed metric pwSNR in Figure 6 (d) have close values and are consistent with the mean opinion scores in Figure 6 (a). However, for four-bar image set, CNR in Figure 6 (e) is inconsistent with the mean opinion scores in Figure 6 (b), since the four-bar target is nonuniform. In this case, CNR is close to zero for all five images in Figure 5 (b).

In the next subsection, experimental studies were conducted using a larger data set in which the proposed metric was compared with other metrics.

4.2.2 Experiment for three sets of 25 images each

Figure 7. Results of subjective evaluations and SNR metrics for the experiments with uniform, four-bar and real-world IR image sets with 25 images each. Left column from top to bottom: (a) Mean opinion scores of uniform target image set, (b) Mean opinion scores of four-bar target image set (c) Mean opinion scores of real-world image set. Right column from top to bottom: Corresponding SNR metric results of the left column

A second experiment similar to the first one but this time with 25 images in each of the three sets was carried out. Images in each set in this experiment also have five different target distinguishability levels. Mean opinion score results for three sets of 25 images are given at the left column of Figure 7 for uniform, four bar and real-world image sets in (a), (b) and (c), respectively. Results of SNR metrics are given at the right column of Figure 7, in (d), (e) and (f).

As in the first experiment, the pwSNR results for all three data sets are consistent with the mean opinion scores of subjective evaluations. CNR and the proposed metric give close results if the target is uniform as can be seen in Figure 7 (d), and both are consistent with the scores in Figure 7 (a). On the other hand, CNR in Figure 7 (e) is inconsistent with the mean opinion scores in Figure 7 (b), since four-bar target is bimodal. It is also not successful in quantifying the detectability of nonuniform targets in real-world images. For real world images, raw SNR is also less consistent with the mean opinion scores since it does not consider the contrast. LSBR, on the other hand, is more relevant to the target size rather than the distinguishability. As can be seen from Figure 7 (c) and (f), LSBR is the least consistent metric with the mean opinion scores.

In this section, the consistency of the SNR metrics with the subjective evaluations has been visually evaluated using mean opinion score and SNR graphics. In the next section, this consistency is evaluated quantitatively using correlation coefficients.

4.3 Quantitative analysis of consistency between SNR metrics and subjective evaluations through correlation coefficients

In this section, consistency between SNR metrics and the mean opinion scores is evaluated using Pearson, Kendall and Spearman correlation coefficients [29]. Pearson correlation coefficient is a strength measure of a linear relationship between two variables. It is also called Pearson product-moment correlation coefficient or bivariate correlation. Kendall rank correlation coefficient is a special case of a more general correlation coefficient and commonly referred to as Kendal's tau coefficient. Spearman rank correlation coefficient or Spearman's rho is another measure of rank correlation [30]. The value of all three correlation coefficients ranges between -1 and 1. In other words, the coefficient is 1 if the correlation is the highest, 0 if there is no correlation, and -1 if there is a negative correlation between the variables being compared.

The correlation coefficients between the SNR metrics and the mean opinion scores are given in Tables 6, 7, and 8 for uniform synthetic, four-bar synthetic, and real-world IR image sets in Figure 5, respectively. The correlation coefficients are calculated between the mean opinion scores at the left column and the values of the SNR metrics at the corresponding right column of Figure 6. As can be seen from these tables, the correlation coefficients for pwSNR takes the highest value or is equal to 1 for all image sets. This shows the superiority of the proposed pwSNR over the other metrics for all three sets, and its consistency with the human visual system. On the other hand, other metrics have lower correlations with the mean opinion scores compared to pwSNR, especially for real-world images. Also note that, in Tables 7 and 8, the correlation coefficients for CNR are among the lowest, as CNR is disadvantageous for nonuniform targets.

Table 6. Correlation coefficients between the mean opinion scores in Figure 6 (a) and SNR metrics in Figure 6 (d) for uniform target image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.9784	1	1
CNR	0.9991	1	1
LSBR	0.7847	0.6	0.7
pwSNR	0.9994	1	1

Table 7. Correlation coefficients between the mean opinion scores in Figure 6 (b) and SNR metrics in Figure 6 (e) for four-bar image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.9893	1	1
CNR	0.5910	0.4	0.4
LSBR	0.9091	1	1
pwSNR	0.9997	1	1

Table 8. Correlation coefficients between the mean opinion scores in Figure 6 (c) and SNR metrics in Figure 6 (f) for real-world IR image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.8216	0.8	0.9
CNR	0.6719	0.4	0.4
LSBR	0.5674	0.4	0.6
pwSNR	0.8754	1	1

Table 9. Correlation coefficients between the mean opinion scores in Figure 7 (a) and SNR metrics in Figure 7 (d) of uniform target image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.9794	0.9129	0.9806
CNR	0.9815	0.9129	0.9806
LSBR	0.9339	0.9129	0.9806
pwSNR	0.9823	0.9129	0.9806

Table 10. Correlation coefficients between the mean opinion scores in Figure 7 (b) and SNR metrics in Figure 7 (e) of four-bar image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.9861	0.8165	0.9549
CNR	0.4903	0.2381	0.3409
LSBR	0.8974	0.8573	0.9661
pwSNR	0.9781	0.8505	0.9649

Table 11. Correlation coefficients between the mean opinion scores in Figure 7 (c) and SNR metrics in Figure 7 (f) of real-world IR image set

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.6890	0.4992	0.6723
CNR	0.7927	0.6600	0.7985
LSBR	0.1110	0.1374	0.1878
pwSNR	0.8428	0.8677	0.9694

In Tables 9, 10 and 11, correlation coefficients calculated between the mean opinion scores at the left column and SNR metrics at the corresponding right column of Figure 7 are given. In these tables, correlation coefficients show that pwSNR is the most consistent metric with the subjective evaluations for the three sets of 25 images each as well. The other metrics cannot accurately quantify the target

detectability for all cases, especially for nonuniform targets and real-world images.

In the next subsection, we extend the experiments using 120 real-world IR images to generalize the consistency of the proposed metric with the subjective evaluations.

4.4 Experiments for extended real-world image dataset

We performed an extensive experiment for 120 real-world IR images from AMCOM dataset to examine the reliability and validity of the proposed metric. For the dataset, subjects are asked to grade distinguishability of the targets in IR images within sets consist of five images each. The scoring strategy is the same as in the previous experiments, and is from 1 to 5. After computing the mean opinion scores of the subjective results and the mean of each SNR metric, we calculated correlation coefficients shown in Table 12.

Table 12. Correlation coefficients for 120 real-word IR images from AMCOM dataset

Metric	Pearson's	Kendall's	Spearman's
raw SNR	0.6743	0.5277	0.6577
CNR	0.8437	0.7703	0.8540
LSBR	0.7367	0.6860	0.7766
pwSNR	0.8500	0.7869	0.8665

As seen in Table 12, correlation coefficients between the proposed metric pwSNR and the subjective results are the highest. This shows that, as in the other experiments, the most consistent metric with the human visual system is pwSNR.

As a result, experimental studies have shown that the proposed metric, pwSNR, is superior to other metrics. In summary, in the first part of the experimental studies, it was shown with specific examples under which conditions the existing metrics have disadvantages. In the subsequent experiment, the consistency of the metrics with human perception was tested with three data sets, two synthetic and one real-world, consisting of five images each with five different levels of detectability. In these experiments, the proposed method gave the most consistent results with the mean opinion scores, as shown in the tables and graphs. In the extended experiment, the results of the metrics for three sets of 25 images each with five detectability levels were compared with the mean opinion scores, and it was again graphically and quantitatively shown that the proposed metric gave closest results to human perception. In addition, in the next two experiments, the correlation coefficients between the metrics and the mean opinion scores were calculated, and the best results were obtained for the proposed metric. Finally, an experiment was performed for a large real-world data set and the correlation coefficients between mean opinion scores and SNR metrics was calculated and it was seen that the highest values were obtained for pwSNR.

5. CONCLUSIONS

In this paper, a new Signal-to-Noise Ratio metric, pwSNR, which is consistent with the human visual system, is proposed to quantify target distinguishability from the background in IR images. Various well-known real-world IR image datasets as well as synthetic image sets have been used in the experiments to evaluate the consistency of the SNR metrics with the mean opinion scores of the subjective evaluations by means of

correlation coefficients. Disadvantages of the existing SNR metrics used to quantify target distinguishability in IR images have also been shown. Extensive experiments using different datasets show that the proposed metric is more consistent with the human visual system than the other metrics. Moreover, while the other metrics cannot accurately quantify the target distinguishability for certain types of targets and background conditions, the proposed metric is always robust and give consistent results with subjective evaluations.

REFERENCES

[1] Liu, R., Lu, Y., Gong, C., Liu, Y. (2012). Infrared point target detection with improved template matching. Infrared Physics & Technology, 55(4): 380-387. https://doi.org/10.1016/j.infrared.2012.01.006

[2] Timischl, F. (2015). The contrast-to-noise ratio for image quality evaluation in scanning electron microscopy. Scanning, 37(1): 54-62. https://doi.org/10.1002/sca.21179

[3] Soni, T., Zeidler, J.R., Ku, W.H. (1993). Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data. IEEE Transactions on Image processing, 2(3): 327-340. https://doi.org/10.1109/83.236534

[4] Behzadpour, M., Ghanbari, M. (2023). Improving precision of objective image/video quality meters. Multimedia Tools and Applications, 82(3): 4465-4478. https://doi.org/10.1007/s11042-022-13416-8

[5] Qi, S., Ma, J., Tao, C., Yang, C., Tian, J. (2012). A robust directional saliency-based method for infrared small-target detection under various complex backgrounds. IEEE Geoscience and Remote Sensing Letters, 10(3): 495-499. https://doi.org/10.1109/LGRS.2012.2211094

[6] Zhao, M., Li, L., Li, W., Tao, R., Li, L., Zhang, W. (2020). Infrared small-target detection based on multiple morphological profiles. IEEE Transactions on Geoscience and Remote Sensing, 59(7): 6077-6091. https://doi.org/10.1109/TGRS.2020.3022863

[7] Yang, L., Yang, J., Yang, K. (2004). Adaptive detection for infrared small target under sea-sky complex background. Electronics Letters, 40(17): 1. https://doi.org/10.1049/el:20045204

[8] Zhang, G., Yang, C., Zhang, Y. (2013). Track-before-detect method based on particle filters for target detection. Journal of Computational Information Systems, 9(18): 7281-7289. https://doi.org/10.12733/jcis7587

[9] Lai, J., Ford, J.J., O'Shea, P., Walker, R. (2008). Hidden Markov model filter banks for dim target detection from image sequences. In 2008 Digital Image Computing: Techniques and Applications, pp. 312-319. https://doi.org/10.1109/DICTA.2008.61

[10] Rozovskii, B.L., Petrov, A. (1999). Optimal nonlinear filtering for track-before-detect in IR image sequences. In Signal and Data Processing of Small Targets 1999, 3809: 152-163. https://doi.org/10.1117/12.364016

[11] Gao, C., Zhang, T., Li, Q. (2012). Small infrared target detection using sparse ring representation. IEEE Aerospace and Electronic Systems Magazine, 27(3): 21-30. https://doi.org/10.1109/MAES.2012.6196254

[12] Diao, W.H., Mao, X., Zheng, H.C., Xue, Y.L., Gui, V. (2012). Image sequence measures for automatic target tracking. Progress In Electromagnetics Research, 130:

447-472. https://doi.org/10.2528/pier12050810

[13] Lai, J., Ford, J., O'Shea, P., Walker, R., Bosse, M. (2008). A study of morphological pre-processing approaches for track-before-detect dim target detection. In Proceedings of the 2008 Australasian Conference on Robotics & Automation, pp. 1-10.

[14] Fan, X.Y., Pan, B.C., Liang, J., Huang, Y.H. (2012). A new preprocessing algorithm for detection of a small dim target in an IR image squence. In 2012 International Conference on Wavelet Analysis and Pattern Recognition, pp. 58-61. https://doi.org/10.1109/ICWAPR.2012.6294755

[15] Usamentiaga, R., Ibarra-Castanedo, C., Maldague, X. (2018). More than fifty shades of grey: Quantitative characterization of defects and interpretation using SNR and CNR. Journal of Nondestructive Evaluation, 37: 1-17. https://doi.org/10.1007/s10921-018-0479-z

[16] Wu, B., Ji, H. B., Guo, H. (2007). Moving dim target detection of infrared image based on improved power-law detector. In 2007 International Conference on Wavelet Analysis and Pattern Recognition, pp. 1085-1089. https://doi.org/10.1109/ICWAPR.2007.4421594

[17] Hsieh, T.H., Chou, C.L., Lan, Y.P., Ting, P.H., Lin, C.T. (2021). Fast and robust infrared image small target detection based on the convolution of layered gradient kernel. IEEE Access, 9: 94889-94900. https://doi.org/10.1109/ACCESS.2021.3089376

[18] Zeng, M., Li, J., Peng, Z. (2006). The design of top-hat morphological filter and application to infrared target detection. Infrared Physics & Technology, 48(1): 67-76. https://doi.org/10.1016/j.infrared.2005.04.006

[19] Srivastava, H.B., Kumar, V., Verma, H.K., Sundaram, S.S. (2009). Image pre-processing algorithms for detection of small/point airborne targets. Defence Science Journal, 59(2): 166-174. https://doi.org/10.14429/dsj.59.1505

[20] Diao, W.H., Mao, X., Gui, V. (2011). Metrics for performance evaluation of preprocessing algorithms in infrared small target images. Progress In Electromagnetics Research, 115: 35-53. https://doi.org/10.2528/PIER11012412

[21] Bai, X., Zhou, F. (2010). Analysis of new top-hat transformation and the application for infrared dim small target detection. Pattern Recognition, 43(6): 2145-2156. https://doi.org/10.1016/j.patcog.2009.12.023

[22] Bai, X. (2014). Morphological center operator for enhancing small target obtained by infrared imaging sensor. Optik, 125(14): 3697-3701. https://doi.org/10.1016/j.ijleo.2014.01.130

[23] Wang, D., Lai, R., Guan, J. (2021). Target attention deep neural network for infrared image enhancement. Infrared Physics & Technology, 115: 103690. https://doi.org/10.1016/j.infrared.2021.103690

[24] Rosin, P.L. (2009). A simple method for detecting salient regions. Pattern Recognition, 42(11): 2363-2371. https://doi.org/10.1016/j.patcog.2009.04.021

[25] Reinagel, P., Zador, A.M. (1999). Natural scene statistics at the centre of gaze. Network: Computation in Neural Systems, 10(4): 341. https://doi.org/10.1088/0954-898X_10_4_304.

[26] Reynolds, J.H., Desimone, R. (2003). Interacting roles of attention and visual salience in V4. Neuron, 37(5): 853-863. https://doi.org/10.1016/S0896-6273(03)00097-7

[27] Li, S., Li, C., Yang, X., Zhang, K., Yin, J. (2020).

Infrared dim target detection method inspired by human vision system. Optik, 206: 164167. https://doi.org/10.1016/j.ijleo.2020.164167

[28] Ratches, J.A., Vollmerhausen, R.H., Driggers, R.G. (2001). Target acquisition performance modeling of infrared imaging systems: past, present, and future. IEEE Sensors Journal, 1(1): 31-40.

https://doi.org/10.1109/JSEN.2001.923585

[29] Bilgen, M. (1999). Target detectability in acoustic elastography. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 46(5): 1128-1133. https://doi.org/10.1109/58.796118

[30] Chen, P., Popovich, P. (2002). Correlation: Parametric and Nonparametric Measures. Sage Publications.