doi:10.3788/gzxb20154403.0311001

基于多属性决策的红外序列图像复杂度分析

乔立永[1],徐立新[1],高敏[2]

(1 北京理工大学 机电学院,北京 100081)

(2 军械工程学院 导弹工程系,石家庄 050003)

摘　要:针对红外序列复杂度对目标跟踪性能的影响问题,提出基于多属性决策评估红外序列复杂度.采用修正逼近理想解多属性决策和熵权方法,综合 7 种图像度量尺度,评估红外序列各帧图像复杂度;基于加权和多属性决策及熵权方法,综合 3 种度量尺度,评估红外序列整体复杂度.采用归一化相关模板匹配算法、基本均值偏移算法和方差比算法进行跟踪实验.采用复杂度不同的红外序列,验证提出的红外序列复杂度评估方案的有效性.结果表明:提出的红外序列复杂度评估方案能够真实显示各种红外序列目标跟踪任务困难度的差异,与跟踪性能指标之间的相关性强,并能准确反映目标跟踪任务的主要影响因素.

关键词:红外;序列;复杂度;分析;多属性决策

中图分类号:TP391.4　　　**文献标识码**:A　　　**文章编号**:1004-421(2015)03-0311001-12

Infrared Image Sequence Complexity Analysis Based on Multi-attribute Decision Making

QIAO Li-yong[1], XU Li-xin[1], GAO Min[2]

(1 *School of Mechatronical Engineering, Beijing Institute of Technology, Beijing* 100081, *China*)

(2 *Ordnance Engineering College, Shijiazhuang* 050003, *China*)

Abstract: To analyse the influences of infrared sequence complexity on the target tracking performance, the infrared sequence complexity evaluation had been modeled as a multi-attribute decision making problem. The each frame complexity of the infrared sequence had been evaluated with seven image metrics based on the modified technique for order preference by similarity to ideal solution method and entropy weights. The whole infrared image sequence complexity had been evaluated with three metrics based on weighted summation method and entropy weights. The normalized correlation template matching algorithm, basic mean shift algorithm, and the variance ratio algorithm had been used to implement tracking experiments. Infrared sequences with different complexity had beed used to validate the effectiveness of the presented infrared sequence evaluation method. The experiments showed that: the proposed infrared sequence complexity evaluation solution could truly indicate the differences of the tracking task difficulties for diverse infrared sequences, there was strong correlation with the tracking performance, and could accurately reflect the major influencing factors for target tracking task.

Key words: Infrared; Sequence; Complexity; Analysis; Multi-attribute decision making

OCIS Codes: 110.3080; 110.3925; 100.2960; 100.4995; 100.4999

Foundation item: Ordnance Engineering College's Research Project

First author: QIAO Li-yong (1982－), male, Ph. D candidate, mainly focuses on pattern recognition, automatic target recognition and tracking technology of infrared imaging guidance. Email:lile-116@163.com

Supervisor: XU Li-xin (1969－), male, professor, Ph. D. degree, mainly focuses on pattern recognition, artificial intelligence, complex system theory. Email:lxxu@bit.edu.cn

Received: Sep. 12, 2014; **Accepted**: Dec. 26, 2014

0 Introduction

Complexity is a concept almost applied to all discipline, from sociology, physics, chaos theory to computational complexity. Because there are great diversities in their research subjects and methods, there is not a unified definition for complexity. The concept of image complexity was first proposed by Peters[1], who considered that an image complexity metric should provide a priori estimate of the difficulty of locating a true target in an image. The infrared image sequence complexity metric not only can be used for target tracking algorithms development and performance evaluation for infrared seekers and imaging fire control applications[2], but also can be used for synthetic infrared scene[3], assessing camouflage[4], et al. In this paper, the infrared image sequence complexity is addressed from the perception of target tracking task difficulty. According to Yilmaz[5], the target tracking performance depends on the context/environment in which the tracking is performed and the end use for which the tracking information is being sought. Different tracking algorithms often use different information and paradigms to process incoming images, tasks that are difficult for one algorithm may be easy for another and vice versa. Different tracking algorithms can be compared based on the measures of image complexity, which can quantify the relative ease or difficulty of tracking the target in an infrared image sequence. To facilitate analysis, we mainly consider the image complexity measure for single target tracking task.

At present, some image metrics have been presented for quantifying the scene complexity. The US Army Aviation and Missile Command (AMCOM) had developed a Gray-Level Co-occurrence Matrix based Trackability Metric (GLCM-TM) to predict the performance of auto-trackers for imaging infrared missile seekers[2]. Anderson developed a combined metric to predict auto-detection and tracker performance by a classical neural network approach[6]. Diao proposed Inter-Frame Change Degree of Target (IFCDT) metric, and experiment results showed that there was a good monotone relationship between this image sequence metric and target recognition performance[7]. Scott defined a trackability measure in an information theoretic framework[8]. Qiao established functional models by partial least squares method between three infrared image complexity metrics and two target detection performance metrics simultaneously[9].

As aforementioned, the difficulty of an ATR task depends not only on its input, but also on the type of information it extracts and its method of extracting information[1]. A particular image metric derived from one image feature for an infrared sequence may be not indicative of the target tracking task difficulty for a particular tracking algorithm. Presently, a comprehensive complexity metric constructed with many image metrics via objective fraction weight for characterizing an infrared image sequence does not exist. The Multi-Attribute Decision Making (MADM) methods are frequently used to solve real world problems with multiple, conflicting, and incommensurate attributes[10, 11]. Therefore, the MADM technique provides an effective framework for the comprehensive evaluation of the infrared image sequence complexity. This paper models the infrared image sequence complexity evaluation as a MADM problem, and presents a simple and effective approach to solve this problem.

The infrared image sequence complexity is divided into two aspects in this paper.

First, the single frame integrated complexity of the infrared image sequence has been established with seven image metrics via entropy weights, and the modified technique for order preference by similarity to ideal solution (M-TOPSIS) has been used to rank order frames by their respective integrated complexity.

Second, the whole infrared image sequence complexity has been established on the results of the single frame complexity of the infrared image sequence. Three metrics have been proposed, and the weighted summation method with entropy weights has been used to rank order various infrared image sequences by their respective integrated complexity.

1 Multi-attribute decision making

Multi-attribute decision making refers to making decisions in the presence of multiple, usually conflicting, criteria, and is widely used in ranking or selecting one or more alternatives from a set of available alternatives[10-11]. A MADM problem with a total of m alternatives characterized by n attributes is generally described by an $m\times n$ decision matrix $\boldsymbol{D}$.

$$\boldsymbol{D}=\begin{bmatrix} & M_1 & M_2 & \cdots & M_n \\ S_1 & d_{11} & d_{12} & \cdots & d_{1n} \\ S_2 & d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_m & d_{m1} & d_{m2} & \cdots & d_{mn} \end{bmatrix} \tag{1}$$

where $\boldsymbol{S}=(S_1,S_2,\cdots,S_m)$ are the considered alternatives, $\boldsymbol{M}=(M_1,M_2,\cdots,M_m)$ are the corresponding attributes vector for each alternative, and d_{ij} denotes the attribute value for the ith alternative with respect to the jth attribute.

In addition, the attributes values of the decision matrix need to be normalized to dimensionless values. The most commonly adopted normalization methods adjust attributes scores based on their distance from a maximum and / or minimum value. Other techniques use an ideal point instead of the minimum or maximum value. The ideal point is a value that represents the best possible or most desired outcome for a given attribute. Sometimes the ideal point is only a theoretical concept, never actually attainable in practice.

If the expectancy is larger-the-better, then it can be expressed by

$$v_{ij}=\frac{d_{ij}-\min d_j}{\max d_j-\min d_j} \tag{2}$$

If the expectancy is smaller-the-better, then it can be expressed by

$$v_{ij}=\frac{\max d_j-d_{ij}}{\max d_j-\min d_j} \tag{3}$$

If the expectancy is nominal-the-best (i. e. closer to the desired value), then it can be expressed by

$$v_{ij}=1-\frac{|d_{ij}-d|}{\max|d_{ij}-d|} \tag{4}$$

where d_{ij} is the original attribute value, v_{ij} is the normalized attribute value, $\max(d_j)$ and $\min(d_j)$ are the maximum and minimum of the jth attribute, d is the ideal value of the jth attribute. The value range of the original attribute value d_{ij} is transformed to $[0,1]$ after normalization.

In this paper, the M-TOPSIS method[10] has been employed to evaluate the single frame complexity of the infrared image sequence; the weighted summation method[11] has been employed to evaluate the whole infrared image sequence complexity; the entropy method[10] has been employed to determine assessing weights for these two kinds of complexity evaluation.

The essential ideal of the entropy weight method is that the relative importance of an attribute is directly related to the information transmitted by the attribute relative to the set of attributes under consideration. The greater the dispersion in the evaluations of the objects for a given metric, the more important the metric.

Due to space limitations, the M-TOPSIS method, the weighted summation method, and the entropy weight method will not be described in detail here, and these papers[10-11] can be referenced.

2 The single frame integrated complexity of the infrared image sequence based on MADM

2.1 Establishment of the single frame integrated complexity of the infrared image sequence based on MADM

As aforementioned, the infrared image sequence complexity is addressed from the perception of target tracking task difficulty in this paper. To facilitate analysis and calculation, seven metrics have been proposed to measure image complexity: 1) Intra-Frame Target Texture Distinctness (IFTTD); 2) Intra-Frame Clutter To Signal (Target) Ratio (IFCSR); 3) Intra-Frame Target Occultation Ratio (IFTOR); 4) Inter-Frame Target Texture Variation (IFTTV); 5) Inter-Frame Target Size Variation (IFTSV); 6) Inter-Frame Target Orientation Variation (IFTOV); 7) Inter-Frame Target Location Variation (IFTLV). The former three metrics can be categorized as intra-frame target identifiability, and the other four metrics can be categorized as inter-frame target variation. Certainly, these measurable image complexity metrics are not independent, but interdependent.

Each of the seven metrics only addresses a subset of the single frame infrared image complexity, and they should be combined into a holistic metric, which can provide a comprehensive evaluation of the overall single frame infrared image complexity, and different frames can be ranked in terms of their respective integrated image complexity. The single frame integrated complexity is the combination of intra-frame target identifiability and inter-frame target variation. The logical relation between the single frame integrated complexity and seven complexity metrics is shown in Fig. 1.

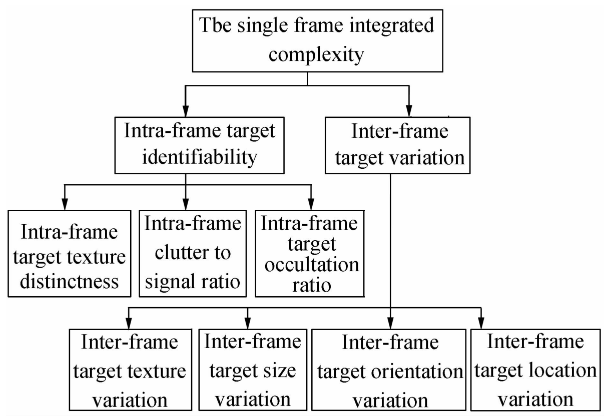


Fig. 1 The logical relation between single frame integrated complexity of infrared sequence and seven complexity metrics

The establishment of the single frame integrated complexity for an infrared image sequence based on the M-TOPSIS method is performed by following steps. 1) Construct the decision matrix with the seven complexity metrics values for every frame of an infrared image sequence; 2) Normalize the image complexity decision matrix; 3) Determine the assessing weights by entropy method; 4) Carry out the residual steps of the M-TOPSIS method; 5) Rank order frames by their respective integrated image complexity.

Because we do not only compare the complexity of each frame in an infrared image sequence, the complexities of all the frames in an infrared image sequence are used to provide a integrated evaluation of the whole infrared image sequence complexity. It must be pointed out that the minimums and maximums for the seven complexity metrics respectively are different among various infrared image sequences. The integrated complexities of different infrared image sequences should be compared on a common basis. Therefore, the seven complexity metrics values in the decision matrix are normalized by Eq. (4), which can use the ideal maximum and minimum for the seven complexity metrics respectively to normalize, as opposed to the maximum and minimum of each complexity metric for an infrared image sequence. The seven complexity metrics and their corresponding ideal maximum and minimum are described in detail as follows.

2.2 Intra-frame target identifiability

Intra-frame target identifiability is referred to the easiness of discriminating the target from background. In this section, three metrics have been introduced to measure the intra-frame target identifiability, including IFTTD metric, IFCSR metric, and IFTOR metric. First, the target and local background window as the reference Region Of Interests (ROIs) are defined for evaluation purpose. To be consistent with the tracking algorithms employed in the experiments in Section 5, the target window is represented by a rectangle bounding box. As is shown in Fig. 2, the target and local background window are at a randomly possible orientation in the digital image coordinate system. The target region is defined as the minimum bounding rectangle which encompasses all pixels classified as target, and the outer rectangle ring around the target gate is defined as the local background region, which is found by scaling the target window by a factor two in each dimension.

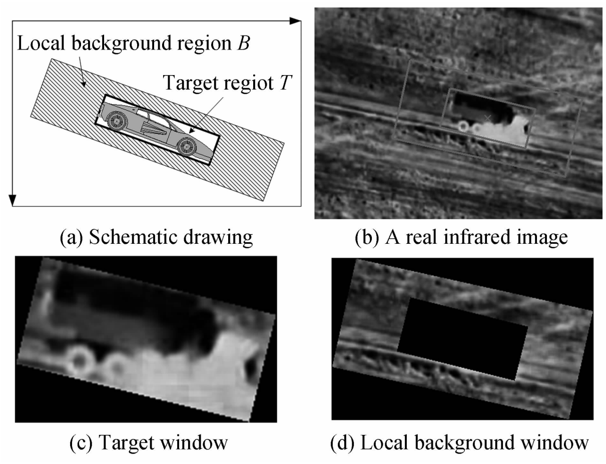


Fig. 2 Target and local background window at a randomly possible orientation

2.2.1 *Intra-frame target texture distinctness*

The Gray Level Co-Occurrence Matrix (GLCM) differences between the target and its local background have been used to measure the intra-frame target texture distinctness. This is an advantage over traditional methods in that metrics based on GLCM not only evaluate the value of pixels in an image but also the relative displacement between pixels of different values.

The IFTTD metric is defined by the Average Co-occurrence Error (ACE) as follows[12].

$$\mathrm{IFTTD}=\frac{1}{N_{\mathrm{GLCM}}}\sum_{d\in D}\sum_{\theta\in\Theta}\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}|P_G^T(i,j|d,\theta)-P_G^B(i,j|d,\theta)| \tag{5}$$

where N_{GLCM} is the total number of the GLCM, G is the number of possible gray levels, $P_G^T(i,j|d,\theta)$ is the joint probability of a pixel of gray level i and a pixel of gray level j given the distance d in direction θ for the target pattern, and $P_G^B(i,j|d,\theta)$ is the corresponding joint probability for the local background pattern. When founding the local background GLCM, all target pixels are excluded from the evaluation and vice-versa.

The dimension of a GLCM is determined by the maximum gray value of the pixel. To implement the ACE, the gray level number of infrared images is set to 32, which decrease the co-occurrence matrix sizes as 32×32; A quantized orientation interval of 45^0 has been employed, corresponding to 0°, 45°, 90° and 135°, each a member of the set Θ, the use of different distance ranges is to obtain both local and global textural information; Last, a range of distance values $d\in D=\{1,3,5,7\}$ has been employed. Thus, the total number of GLCM is $N_{\mathrm{GLCM}}=|D||\Theta|=16$, $|\cdot|$ is the cardinality or the number of elements in the set. Although averaging the total co-occurrence matrix differences inevitably smoothing some spatial features, it does incorporate both local and global textural information constructively.

The target and local background GLCMs for different distance and orientation for the target and local background window in Fig. 2(c)～(d) is shown in Fig. 3. The target GLCM and local background GLCM shown in Fig. 3 suggest that the GLCM formalism can be used to quantify how closely the target matches the local background not only in contrast but also in structure by measuring the degree the similarity between the matrices.

Obviously, the sum of any normalized GLCM equals to 1. According to triangle inequality, for a particular d and θ, the range of $|P_G^T(i,j|d,\theta)-P_G^B(i,j|d,\theta)|$ is $[0,2]$. The minimum 0 and maximum 2 are

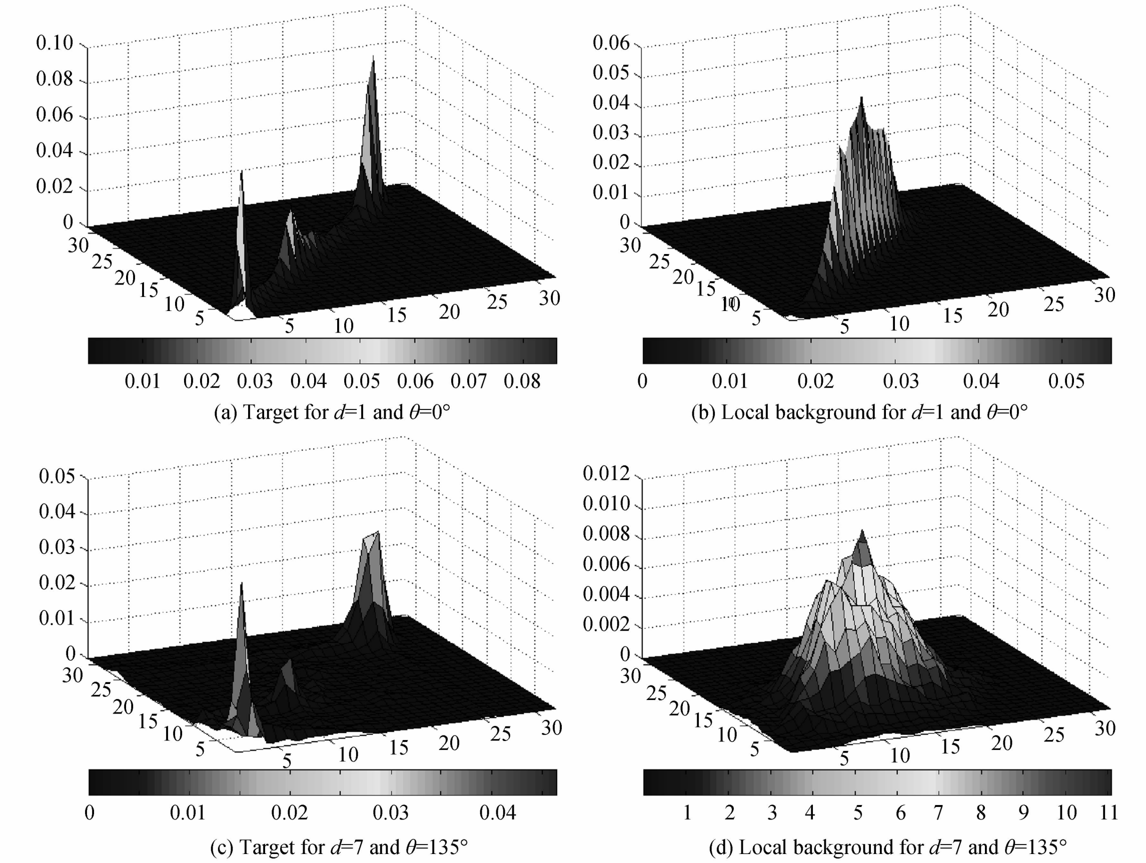


Fig. 3 Target and local background GLCMs for the target and local background window in Fig. 2(c)-(d)

theoretically limits, and will never be actually attainable in practice, because the target GLCM and the local background GLCM could not be entirely same or different in a real infrared scene.

To facilitate the normalization of IFTTD metric values for different frames in an infrared image sequence, the minimum 0 and maximum 2 are used as the common minimum and maximum for the normalization of IFTTD metric values by Eq. for different frames in an infrared image sequence.

2.2.2 *Intra-frame clutter to signal (target) ratio*

The Signal to Clutter Ratio (SCR) metric, which incorporates the difference between the mean target and background intensity, the effects of target structure and the local background structure, is defined as follows[6].

$$\mathrm{SCR}=\frac{\sqrt{(\mu_T-\mu_B)^2+\sigma_T^2}}{\sigma_B} \tag{6}$$

where μ_T is the average value of target intensity, μ_B is the average value of local background intensity, σ_T is the standard deviation of the target intensity, and σ_B is the standard deviation of the local background intensity.

The ideal value of SCR metric is infinite. To facilitate the normalization of SCR metric values for different frames in an infrared image sequence, the inverse of SCR metric has been employed in this paper, which can be called Intra-Frame Clutter to Signal (target) Ratio (IFCSR).

$$\mathrm{IFCSR}=\frac{\sigma_B}{\sqrt{(\mu_T-\mu_B)^2+\sigma_T^2}} \tag{7}$$

The ideal value of IFCSR metric is 0, and the maximum is set to 1, which is a larger value for IFCSR metric. These two values are used as the common minimum and maximum for the normalization of IFCSR metric values by Eq. for different frames in an infrared image sequence.

2.2.3 *Intra-frame target occultation ratio*

A critical problem frequently causing target tracking failure is occlusion, whether it is partial or complete. The more part of the target is occulted, the more difficult is to track the target, and the higher complex is the current frame infrared image. The ratio of the number of the pixels occluded divided by the total number of the pixels on the target[3].

$$\mathrm{IFTOR}=\frac{S_{\mathrm{OTA}}}{S_{\mathrm{FTA}}} \tag{8}$$

where S_{OTA} is the occulted target area, and S_{FTA} is the full target area. In practice, the target area in the previous frame before the target disappears is taken as the full target area.

Obviously, the ideal value of IFTOR metric is 0 when there is no occultation occurs; the nadir value of IFTOR metric is 1 when the target is occulted completely. These two values are used as the common minimum and maximum for the normalization of IFTOR metric values by Eq. for different frames in an infrared image sequence.

2.3 Inter-frame target variation

Inter-frame target variation has important effects on target tracking result[7]. The larger is inter-frame target variation, the more difficult is to track the target, and the higher complex is the current frame infrared image. Four metrics have been employed to measure inter-frame target variation: inter-frame target texture variation, inter-frame target size variation, inter-frame target orientation variation, and inter-frame target location variation.

2.3.1 *Inter-frame target texture variation*

As similar with the IFTTD metric, the Inter-Frame Target Texture Variation (IFTTV) is also measured by the ACE between current frame and previous frame as follows.

$$\mathrm{IFTTV}=\frac{1}{N_{\mathrm{GLCM}}}\sum_{d\in D}\sum_{\theta\in\Theta}\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left|P_G^C(i,j\,|\,d,\theta)-P_G^P(i,j\,|\,d,\theta)\right| \tag{9}$$

where N_{GLCM} is the total number of the GLCM, G is the number of possible gray levels, $P_G^C(i,j\,|\,d,\theta)$ is the joint probability of a pixel of gray level i and a pixel of gray level j given the distance d in direction $\theta$for the target pattern in current frame, and $P_G^P(i,j\,|\,d,\theta)$ is the corresponding joint probability for the target pattern in previous frame.

The configurations of the gray level number of infrared images, orientation interval, and range of distance values for the IFTTV metric are the same with the IFTTD metric. In addition, the common minimum and maximum for the normalization of IFTTV metric values by Eq. for different frames in an infrared image sequence are also the same with the IFTTD metric.

2.3.2 *Inter-frame target size variation*

The Inter-frame Target Size Variation (IFTSV) metric is defined as follows.

$$\mathrm{IFTSV}=\sqrt{\left(\frac{|L_i-L_{i-1}|}{L_{i-1}}\right)^2+\left(\frac{|W_i-W_{i-1}|}{W_{i-1}}\right)^2} \tag{10}$$

where L_i and W_i are the length and width of the target respectively in current frame, and L_{i-1} and W_{i-1} are the length and width of the target respectively in previous frame.

Obviously, the ideal value of IFTSV metric is 0 when the target size remains changeless between consecutive frames; the nadir value of IFTSV metric is $\sqrt{2}$ when the target length and width changes equal to the length and width of the target respectively in previous frame. These two values are used as the common minimum and maximum for the normalization of IFTSV metric values by Eq. for different frames in an infrared image sequence.

2.3.3 *Inter-frame target orientation variation*

The Inter-Frame Target Orientation Variation (IFTOV) metric is defined as follows.

$$\mathrm{IFTOV}=|\phi_i-\phi_{i-1}| \tag{11}$$

where ϕ_i and ϕ_{i-1} are the target orientation angles in current frame and previous frame respectively.

Obviously, the ideal value of IFTOV metric is 0 when the target orientation remains changeless between consecutive frames; the maximum of IFTOV metric does not exist, because the target orientation variation differs greatly from various real infrared image sequences, such as the flying fighter-plane infrared sequence and stationary tank infrared sequence. Therefore, the maximum of IFTOV metric is set as 20°. These two values are used as the common minimum and maximum for the normalization of IFTOV metric values by Eq. for different frames in an infrared image sequence. When the IFTOV metric is larger than 20° for some frames in a infrared image sequence, the normalization result of IFTOV metric value is set as the same with 20°.

2.3.4 *Inter-frame target location variation*

When the target is far from the camera its projection on the image plane is relatively smaller than when it is closer[13]. Therefore, the inter-frame target location variation should be normalized by the target size. Considering the target orientation, the inter-frame target location variation should be calculated in the rotated coordinate system. Thus, the Inter-Frame Target Location Variation (IFTLV) metric is defined as follows.

$$\mathrm{IFTLV}=\sqrt{V_L^2+W_L^2} \tag{12}$$

where

$$V_L=\frac{\cos\phi_{i-1}(x_i-x_{i-1})-\sin\phi_{i-1}(y_i-y_{i-1})}{L_{i-1}/2} \tag{13}$$

$$W_L=\frac{\sin\phi_{i-1}(x_i-x_{i-1})+\cos\phi_{i-1}(y_i-y_{i-1})}{W_{i-1}/2} \tag{14}$$

Obviously, the ideal value of IFTLV metric is 0 when the target remains stationary between consecutive frames; the nadir value of IFTLV metric is $\sqrt{2}$ when the distances that the target moves along the length and along the width equal to the semi-length and semi-width of the target respectively in previous frame. These two values are used as the common minimum and maximum for the normalization of IFTLV metric values by Eq. for different frames in an infrared image sequence.

2.4 Some issues of the seven complexity metrics when the target disappears

When the target is completely occulted or the target leaves out of the image plane, the target disappears from the field of view, and there are some issues to be considered. The target attributes such as image features, size, orientation, and location, et al., can not be obtained, and they only can be estimated when the target disappears. Therefore, when the target disappears from the field of view for some time, the IFTTD metric, IFCSR metric, and IFTOR metric are set to the highest complexity values; the inter-frame target variation metrics are set to the corresponding complexity metrics values of the previous frame before the target disappears completely.

3 The whole infrared image sequence complexity based on MADM

The whole infrared image sequence complexity is established on the weighted summation method. Three metrics have been employed to measure the complexity of the whole infrared image sequence: the Average Complexity of Each Frame (ACEF), the Proportion of High Complexity Frames number to the total frames number (PHCF), the Average Complexity of High Complexity Frames (ACHCF). The average complexity of each frame is defined as the threshold, and the frames whose complexities are above the threshold are defined as high complexity frames.

In addition, the three complexity metrics values in the decision matrix for different infrared image sequences are normalized by Eq., the maximum and minimum of each complexity metric are obtained by comparison of different infrared image sequences.

The entropy method is again used for assessing corresponding weights of the three complexity metrics.

The establishment of the whole infrared image sequence complexity based on the weighted summation method is performed by following steps. 1) Construct the decision matrix with the three complexity metrics values for each infrared sequence; 2) Normalize the image sequence complexity decision matrix; 3) Determine the assessing weights by entropy method; 4) Carry out the residual steps of the weighted summation method; 5) Rank order infrared sequences by their respective integrated complexity.

4 Experiments

4.1 Establishment of experiments

4.1.1 *Infrared image sequences to be tested on*

To illustrate the effectiveness the proposed infrared sequence complexity evaluation solutions, five real infrared image sequences with different complexities have been tested on, including a stationary tank infrared sequence collected from a captive flight test platform provided by AMCOM, three infrared sequences from the VIVID dataset, and a flying fighter plane infrared sequence. The AMCOM closure sequence exemplifies the challenges of target tracking such as poor target visibility, strong ego-motion, small targets, size variations, significant clutter and background noise. The infrared sequences in the VIVID dataset present frequent occlusion by foliage, target exiting and re-appearing, which can be used to examine the performance of track loss of the tracking algorithms. In the flying fighter plane infrared sequence, the Su-27 was flying at the speed of 800～900 km/h. The flying fighter plane infrared sequence exemplifies the challenges of target tracking such as high speed motion, deformation, background clutter, and motion blur.

To compute the complexity metrics of the five infrared image sequences, the ground-truth information of each frame has be obtained manually by stepping frame-by-frame through the recorded sequence.

4.1.2 *Target tracking algorithms and performance evaluation measures*

Because most infrared imaging auto-trackers implemented in systems are either company proprietary or classified, no specific algorithms will be described particularly[6]. These target tracking algorithms in open literature with available source codes are employed and evaluated in this paper, and similar algorithms can be expected in real infrared imaging auto-trackers. The employed target algorithms codes were provide by Robert[14], including Normalized Correlation Template Matching (NCTM), Basic Mean Shift (BMS), and Variance Ratio (VR).

The overlap error metric has been employed to measure the tracking errors[15], which is based on the amount of overlap between the ground truth and tracker boxes. Fig. 4 shows a tracker box overlaid on a ground truth box in an infrared image with each region identified. The area (in pixels) for the truth region, tracker region, and overlap region are defined as NTRU, NTRK, and NOL, respectively. A ratio of NOL to NTRU gives an indication of the overlap compared to the ground truth box.

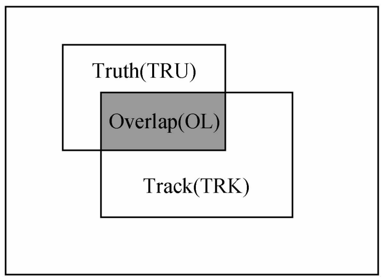


Fig. 4 Track box overlaid on ground truth box

The two ratios are defined as follows.

$$R_{TRU}=N_{OL}/N_{TRU} \tag{15}$$

$$R_{TRK}=N_{OL}/N_{TRK} \tag{16}$$

The overlap error metric E_{OEM} is defined as follow.

$$E_{\mathrm{OEM}}=1-\frac{R_{\mathrm{TRU}}+R_{\mathrm{TRK}}}{2} \tag{17}$$

when the tracker box is exactly overlaid on and is the same size as the truth box, the overlap error metric equals to the minimum 0; when there is no overlap between the tracker box and the truth box, the overlap error metric equals to the maximum 1.

When the overlap error metric E_{OEM} is greater than a threshold T_{OEM}, the tracker is considered to have lost the target. In this paper, the threshold T_{OEM} is set 0.9.

The lost track ratio η is the ratio between the number of frames where the tracker is not successful N_L and the total number of frames in an infrared image sequence.

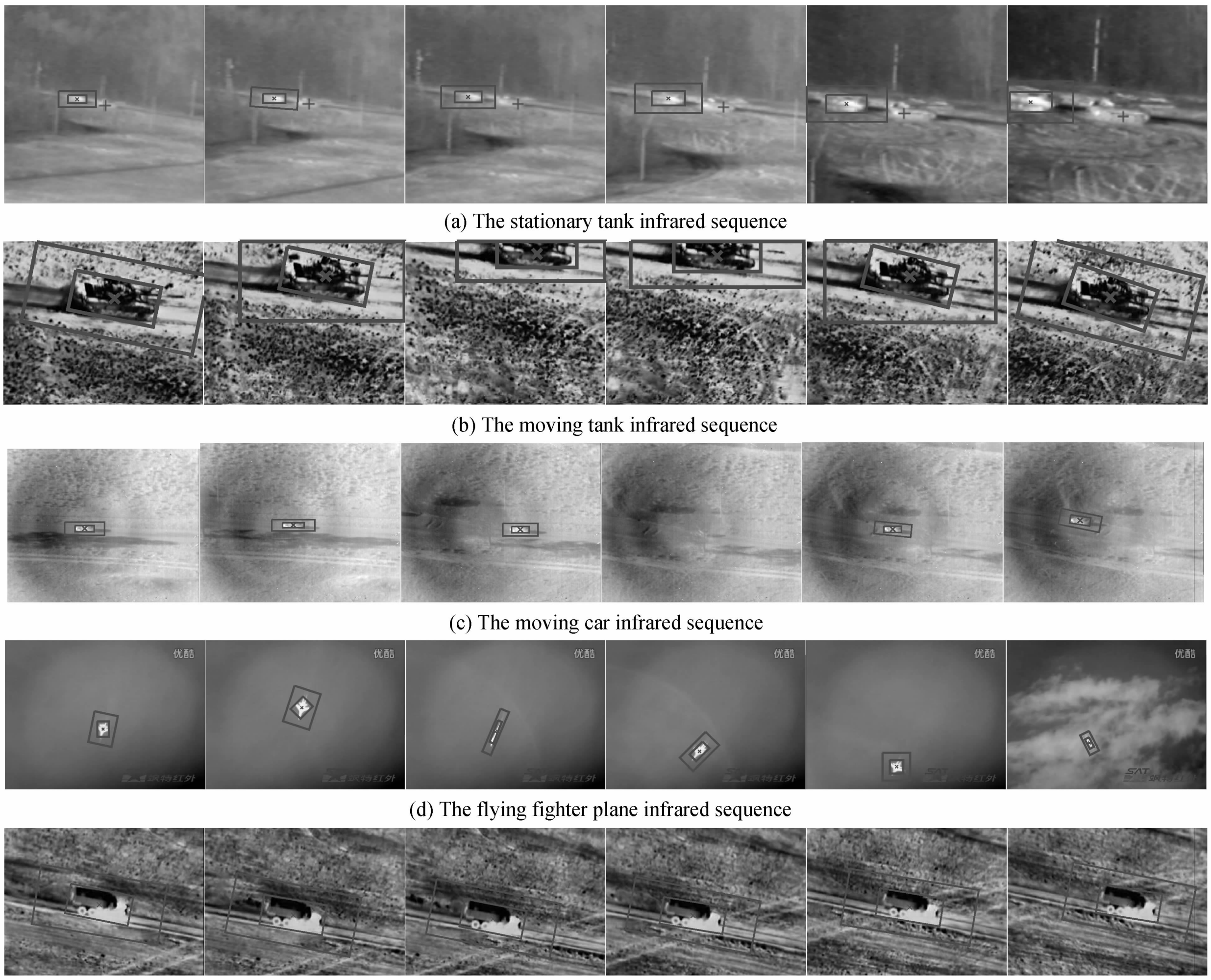
$$\eta=N_{\mathrm{L}}/N_{\mathrm{T}} \tag{18}$$

If the lost track ratio exceeds 10%, then the target in the infrared image sequence is scored as "unsuccessfully tracked" by the tracking algorithm; otherwise, the target in the infrared image sequence is scored as "successfully tracked". $T_{\mathrm{S}}=1$ indicates being successfully tracked ; $T_{\mathrm{S}}=0$ indicates being unsuccessfully tracked.

The average overlap error metric $\overline{E}_{\mathrm{OEM}}$ is calculated for the frames where the tracker is successful. And the performance vector $(T_{\mathrm{S}},\eta,\overline{E}_{\mathrm{OEM}})$ is used to evaluate the tracking algorithm comprehensively. These three performance metrics should be considered sequentially. Only when the values of the infrared image sequence "successfully tracked" indicator T_{S} for two tracking algorithms are the same, the other two performance metrics need to be compared. Likewise, only if the values of the lost track ratio η for two tracking algorithms are comparable, the average overlap error metric $\overline{E}_{\mathrm{OEM}}$ can be compared.

4.2 The single frame integrated complexity of five infrared sequences

The ground truth results for the five infrared image sequences are shown in Fig. 5. For space limitations, the visual tracking results by the three tracking algorithms for the five infrared image sequences have not been listed. The results of the single frame integrated complexity for five infrared sequences are calculated by the solution established in



(a) The stationary tank infrared sequence

(b) The moving tank infrared sequence

(c) The moving car infrared sequence

(d) The flying fighter plane infrared sequence

(e) The moving truck infrared sequence

Fig. 5 The ground truth results for the five infrared image sequences

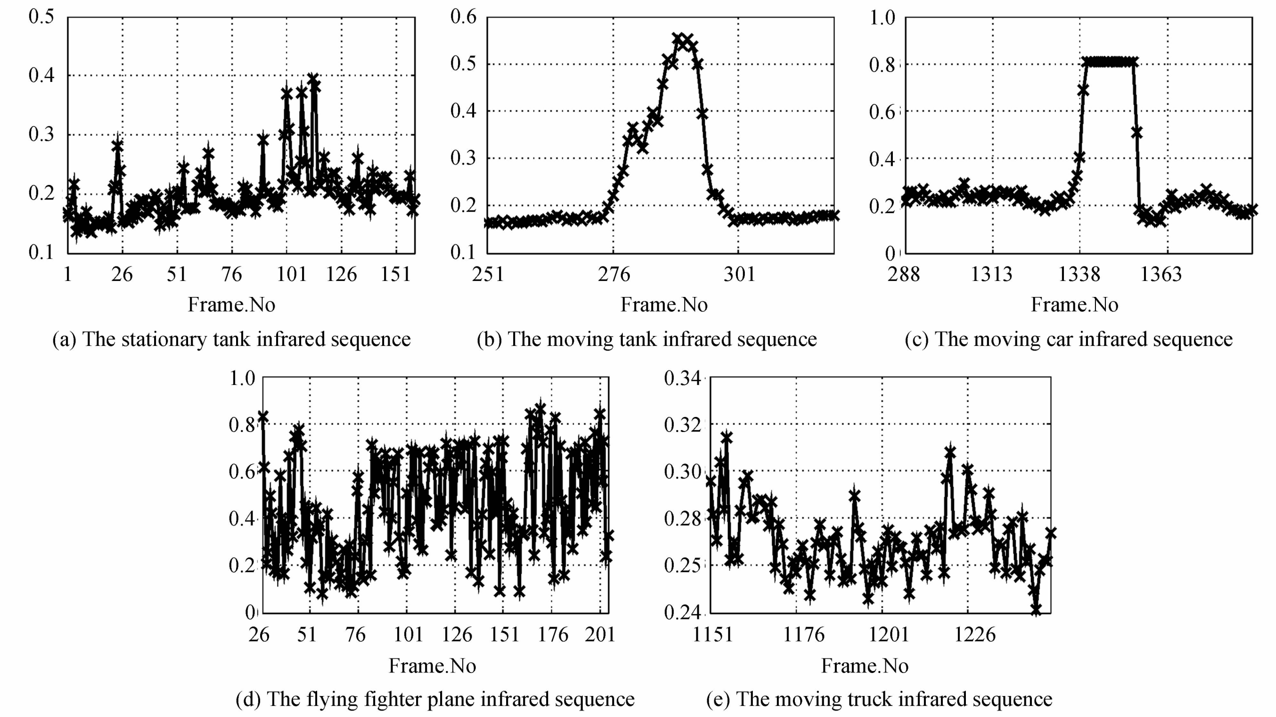


Fig. 6 The single frame integrated complexity of the five infrared sequences

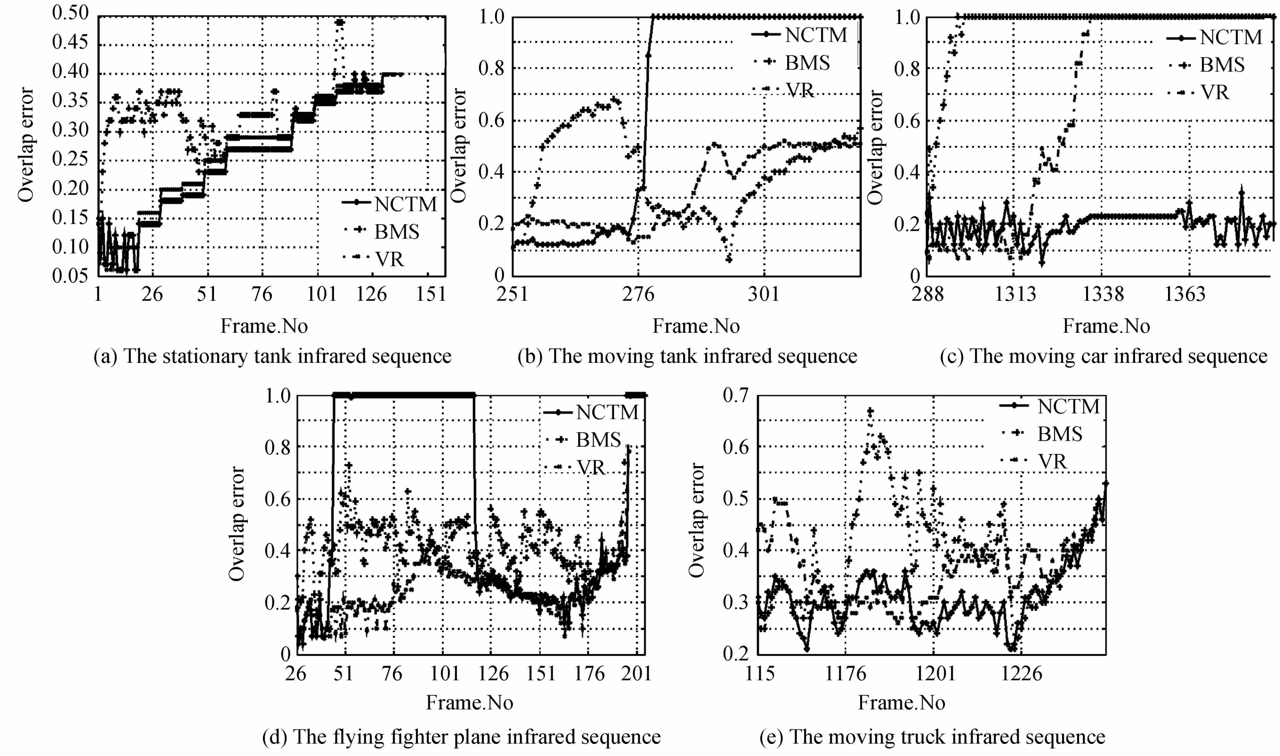


Fig. 7 Tracking errors of the three algorithms for the five infrared sequences

Section 3, and are shown in Fig. 6. The entropy weights of the seven complexity metrics for the single frame integrated complexity of five infrared sequences are listed in Table 1. The overlap errors of the three tracking algorithms for the five infrared image sequences are shown in Fig. 7. According to Table 1, the objective weights of the seven single frame complexity metrics derived by the entropy weight method are significantly different to each other for every infrared image sequence, which reflects the capability of the entropy weight method in reflecting the average intrinsic information generated by the complexity metrics values. This would help identifying the most important complexity metric on which the frames of each infrared image sequence have the most divergent variation.

For the stationary tank infrared sequence, the tank size is becoming larger during the infrared camera closing in on the stationary tank. Due to the motion of camera, there are some variations of the tank location. Because the three tracking algorithms can not adjust the tracking window adaptively, the overlap errors become larger as the camera is closing in on the tank. In addition, the tracking window of the BMS algorithm is frequently deviates from the target window due the similar gray intensity feature in the local background region with the target, which results in larger overlap error across the tracking process. According to Table 1, the IFTLV metric has the highest degree of importance in evaluating the single frame integrated complexity of the stationary tank infrared sequence, followed by the IFTTD metric. The weight for the IFTOR metric is zero, because there is no occultation through out the stationary tank infrared sequence. As is shown in Fig. 6(a), the integrated complexities of the most frames in the stationary tank infrared sequence are below 0.2, and the integrated complexity will be above 0.2 when the target size or location varies.

For the moving tank infrared sequence, the major cause of tracking failures or errors is that parts of the target leave out of the image plane from frame00275 to frame00304, which is considered as occultation. The maximum of occultation ratio occurs at frame00291. As is shown in Fig. 6(b), the integrated complexity of the target becomes larger when parts of the target begins to leave out of the image plane, and vice versa. The integrated complexities of the most frames in the moving tank infrared sequence are below 0.2, and the integrated complexity will be above 0.2 when parts of the target leave out of the image plane. The maximum integrated complexity of the infrared sequence occurs at frame00289, and the integrated complexity at frame00291 is the second largest. According to Table. 1, the IFTOR metric has the highest degree of importance, followed by the IFTCSR metric, which is consistent with our subjective judgment. The NCTM algorithm loses the target at frame00279, and does not reacquire the target until the end frame. It is preposterous for the BMS algorithm that the overlap error decreases when parts of the target leave out of the image plane and the overlap error increases when the target moves into the image plane. The cause for this preposterous situation is that the tracking window of the BMS algorithm is frequently deviates from the target window due to the similar gray intensity feature in the local background with the target before parts of the target leave out the image plane, which results in larger overlap errors; the tracking window adjusts adaptively to be smaller when parts of the target leave out the image plane, which results in smaller overlap errors; the tracking window can not adjusts adaptively to be larger when the target moves into the image plane, which results in larger overlap errors again. The VR algorithm has higher resistance to clutter interference than the NCTM algorithm, and the overlap error is smaller than that of the BMS algorithm before parts of the target leave out of the image plane. As similar as BMS algorithm, the tracking window of the VR algorithm adjusts adaptively to be smaller when parts of the target leave out the image plane, which results in smaller overlap errors; the tracking window of the VR algorithm can not adjusts adaptively to be larger when the target moves into the image plane, which results in larger overlap errors.

For the moving car infrared sequence, the major causes of tracking failures or errors are the clutter in the local background and the complete occultation. The target is occulted partially or or completely by a tree from frame01331 to frame01359. The BMS algorithm loses the target at frame01295 due to the impact of clutter. The VR algorithm loses the target at frame01333 due to the impact of occultation. Because the time that the target is completely occulted is very short, the tracking performance of the NCTM has not been affected. Because when the target disappears from the field of view, the IFTTD metric, IFCSR metric, and the IFTOR metric are set to the highest complexity values, and the inter-frame target variation metrics are set to the corresponding complexity metrics values of the previous frame before the target disappears completely, the integrated complexity maintains the maximum during complete occultation, as is shown in Fig. 6(c). According to Table. 1, the IFTOR metric has the highest degree of importance, followed by the IFTCSR metric, which is also consistent with our subjective judgment.

For the flying fighter-plane infrared sequence, the major causes of tracking failures or errors are high speed motion, deformation, and background clutter. As is shown in Fig. 6(d), the integrated complexities of the most frames in the flying fighter-plane infrared sequence are between 0.2 and 0.8, which are evidently higher than those of the other four infrared sequences. According to Table 1, the IFTLV metric has the highest degree of importance, followed by the IFTOV metric, which is also consistent with our subjective judgment. The NCTM algorithm loses the target at frame00045 due to the high speed motion of the target, reacquires the target at frame00118, and loses the target at frame00198 due the clutter caused by the clouds. The tracking performance of the VR algorithm is better than the BMS algorithm before losing the target. The VR algorithm also loses the target at frame00198, and the BMS algorithm loses the target at frame00196, which are all due to the clutter caused by the clouds.

Table 1 Entropy weights of the seven single frame complexity metrics for the five infrared sequences

Infrared sequences \ Complexity metrics	IFTTD	IFTCSR	IFTOR	IFTTV	IFTSV	IFTOV	IFTLV
Stationary tank	0.228	0.133	0.000	0.116	0.048	0.053	0.422
Moving tank	0.048	0.106	0.746	0.016	0.027	0.022	0.035
Moving car	0.117	0.397	0.414	0.014	0.013	0.009	0.036
Flying fighter plane	0.004	0.017	0.000	0.033	0.008	0.168	0.770
Moving truck	0.100	0.252	0.000	0.065	0.050	0.221	0.312

For the moving truck infrared sequence, there are lots of similar gray intensity pixels in the local background region with the target, which causes the tracking windows of the three algorithms deviate from the target window. As is shown in Fig. 6(e), the integrated complexities of the most frames in the moving truck infrared sequence are between 0.24 and 0.3. According to Table 1, the IFTLV metric has the highest degree of importance, followed by the IFTCSR metric.

Based on the above experiment results and analysis, the weight coefficients produced by the entropy method are divergent for the single frame complexity metrics, which reflects that it can better resolve the inherent conflict between the complexity metrics embedded in this MADM problem, and it can reflect the major influencing factor for target tracking task. The M-TOPSIS method can identify the relevance of the complexity metrics to the tracking task difficulty, which suggests that the M-TOPSIS method can reflect the decision information emitted by the seven complexity metrics effectively. Through the analysis of the relation between the image complexity metrics and tracking performance results by the three algorithms, we can understand their advantages and disadvantages. The NCTM algorithm is more susceptible to high speed motion; the BMS algorithm is more susceptible to distraction by clutter in the local background region; the VR algorithm combined with an online feature selection mechanism, which is based on applying the two class variation ratio to log likelihood distributions computed for a given feature from samples of object and background pixels, has improved the tracking performance of the BMS algorithm. The major factors that influence the tracking performance results of the three algorithms can be summarized as follows: occultation, clutter, size variation, and high speed motion.

4.3 The whole infrared image sequence complexity of five infrared sequences

The comprehensive performance evaluation results of the three tracking algorithms for the five infrared sequences are listed in Table 2, the total complexity metrics values and the total complexity of each infrared sequence are listed in Table 3, and the entropy weights of the total complexity metrics for the five infrared sequences are listed in Table 4.

Table 2 The comprehensive performance evaluation results of the three tracking algorithms for the five infrared sequences

Infrared sequences \ Algorithms	NCTM			BMS			VR		
	T_S	η	$\overline{E}_{OEM}$	T_S	η	$\overline{E}_{OEM}$	T_S	η	$\overline{E}_{OEM}$
Stationary tank	1	0.00	0.28	1	0.00	0.35	1	0.00	0.29
Moving tank	0	0.59	0.18	1	0.00	0.40	1	0.00	0.33
Moving car	1	0.00	0.19	0	0.91	0.56	0	0.54	0.26
Flying fighter plane	0	0.45	0.24	1	0.06	0.42	1	0.04	0.24
Moving truck	1	0.00	0.31	1	0.00	0.40	1	0.00	0.36

Table 3 Comparison of the total complexity of the five infrared sequences

Infrared sequences \ Total complexity metrics	ACEF	PHCF	ACHCF	Total complexity	Rank
Stationary tank	0.200	0.044	0.348	0.033	5
Moving tank	0.240	0.229	0.440	0.258	3
Moving car	0.313	0.190	0.713	0.512	2
Flying fighter plane	0.450	0.733	0.542	0.878	1
Moving truck	0.271	0.080	0.301	0.103	4

Table 4 Entropy weights of the total complexity metrics for the five infrared sequences

ACEF	PHCF	ACHCF
0.287	0.420	0.293

According to the comparison method by the comprehensive tracking performance vector addressed in Section 4.1.2, the tracking performance results of the five infrared sequences by the three algorithms can be ranked. The tracking performance results of the stationary tank infrared sequence by the three tracking algorithms are the best among all the five infrared sequences, followed by the moving truck infrared sequence, the moving tank infrared sequence, the moving car infrared sequence, and the flying fighter plane infrared sequence.

According to the total complexity of the five infrared sequences listed in Table 3, the total complexity of the stationary tank infrared sequence is the least, followed by the moving truck infrared sequence, the moving tank infrared sequence, the moving car infrared sequence, and the flying fighter plane infrared sequence.

According to Table 4, the percentage of high complexity frames metric has the highest degree of importance in evaluating the whole infrared sequence complexity, followed by the average of high complexity frames metric. This is in line with the real situation and the general perception of influencing factors for target tracking task. Usually, tracking failures or errors usually happen at high complexity frames.

Based on the above analysis, the whole infrared sequence complexity evaluation method proposed in this paper could indicate the difference of target tracking task difficulty for various infrared sequences, and the tight correlations between the total complexity of the infrared sequence and the comprehensive tracking performance results have been validated.

5 Conclusions

The infrared image sequence complexity evaluation solutions have been established with MADM technique, which is addressed from the perception of target tracking task difficulty. The M-TOPSIS method combined with entropy weights has been used to establish the single frame integrated complexity of the infrared image sequence with seven image complexity metrics. The weighted summation method combined with entropy weights has been used to establish the whole infrared image sequence complexity with three metrics. The well-known challenging AMCOM closure sequences and VIVID dataset have been used to validate the effectiveness of the proposed solutions. To analyze the relationship between the complexity metrics and the tracker performance, the NCTM algorithm, BMS algorithm, and VR algorithm have been used to process the chosen infrared sequences.

The experiments for the single frame integrated complexity show that the weight coefficients produced by the entropy method are divergent, which ensures that the evaluation result is not affected by the interdependence of complexity metrics. The M-TOPSIS method can identify the relevance of the complexity metrics to the tracking task difficulty. The advantages and disadvantages of the three tracking algorithms have been demonstrated through the relation between the single frame complexity metrics and tracking performance results. The major factors that influence the tracking performance results of the three algorithms can be summarized as follow: occultation, clutter, size variation, and high speed motion.

The experiments for the whole infrared image sequence complexity show that the proposed total complexity solution for the whole infrared image sequence could indicate the difference of target tracking task difficulty for various infrared image sequences, and the tight correlation between the total complexity of the infrared sequence and the comprehensive tracking performance results have been validated.

References

[1] PETERS II R A, STRICKLAND R N. Image complexity metrics for automatic target recognizers[C]. Naval Surface Warfare Center, Silver Spring, MD: Automatic Target Recognizer System and Technology Conference, 1990, 1-17.

[2] BRACKNEY B A, HELTON M K, HAMMON R K. Development of the gray-level co-occurrence matrix target trackability metric for imaging infrared missile systems [C]. Orlando, Florida: SPIE Conference on Infrared imaging Systems: Design, Analysis, Modeling, and Testing IX, 1998, 240-254.

[3] LANDEAU S, DAGOBERT T. Image database generation using image metric constraints: An application within the CALADIOM project[C]. Orlando (Kissimmee), FL: SPIE 6234, Automatic Target Recognition XVI, 2006, 623410.

[4] NYBERG S, BOHMAN L. Assessing camouflage methods using textural features[J]. *Optical Engineering*, 2001, **40** (9): 1869-1876.

[5] YILMAZ A, JAVED O, SHAH M. Object tracking: A survey[J]. *ACM Computing Surveys*, 2006, **38**(4): 1-46.

[6] ANDERSON D R, JIM MOORE J, MONTGOMERY, *et al*. Infrared seeker performance metrics[R], 2003.

[7] DIAO W H, MAO X, ZHENG H C, *et al*. Image sequence measures for automatic target tracking [J]. *Progress In Electromagnetics Research*, 2012, **130**: 447-472.

[8] ACTON S T, AKSEL A. An information theoretic trackability measure[C]. Burlingame, California, USA: SPIE 8296, Computational Imaging X, 82960W, 2012, 1-12.

[9] QIAO L, XU L, GAO M. Influences of infrared image complexity on the target detection performance[J]. *Infrared and Laser Engineering*, 2013, **42**(S1): 253-261.

[10] DENG H, YEH C, WILLIS R J. Inter-company comparison using modified TOPSIS with objective weights[J]. *Computer & Operations Research*, 2000, **27**: 963-973.

[11] CHURCHMAN C W, ACKOFF R L. An approximate measure of value[J]. *Journal of the Operations Research Society of America*, 1954, **2**(2): 172-187.

[12] COPELAND A C, TRIVEDI M M. Signature strength metrics for camouflaged targets corresponding to human perceptual cues [J]. *Optical Engineering*, 1998, **37**(2): 582-591.

[13] MAGGIO E, CAVALLARO A. Accurate appearance-based Bayesian tracking for maneuvering targets [J]. *Journal Computer Vision and Image Understanding*, 2009, **113**(4): 544-555.

[14] COLLINS R, ZHOU X, TEH S K. An open source tracking testbed and evaluation web site [C]. IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, 2005.

[15] LAVELLEE P D, LIGHTFOOT J A, CHAMBLISS M A, *et al*. Methodology for evaluating autonomous IR tracker performance[R]. Houghton, 2002.