



# Infrared and radar fusion detection method based on heterogeneous data preprocessing

Yuan Wei<sup>1,2</sup> · Zhengdong Cheng<sup>1,2</sup> · Bin Zhu<sup>1,2</sup> · Xiang Zhai<sup>1,2</sup> · Hongwei Zhang<sup>1,2</sup>

Received: 15 August 2019 / Accepted: 25 September 2019 / Published online: 8 October 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

The difficulty in infrared and radar fusion target detection lies in the pre-processing steps such as sample reduction and feature reduction of heterogeneous sensors. For the problem of data volume redundancy, HVS sample reduction is proposed to reduce the amount of data in the infrared data set. Infrared and radar joint feature vectors are constructed through nearest neighbor data association. For the problem of feature redundancy, a weighted mutual information feature selection method based on prior information is proposed according to the difference of feature sources. The fusion data set is then used for classification. The experiments show that the HVS-based sample reduction and the prior weighted feature selection achieved a higher detection probability and a shorter detection time, and improved the robustness in the case of large differences between the training set and the test set.

**Keywords** Target detection · Data fusion · Sample reduction · Feature selection

## 1 Introduction

The effectiveness of infrared and radar fusion detection has been proved by massive applications in early warning and guidance. The difficulty of fusion detection lies in the fact that infrared detector and radar are heterogeneous sensors. It is necessary to construct joint feature vectors through preprocessing steps. Additionally, the infrared image provides image information on different spectral ranges (Wen et al. 2018a, b; Jia et al. 2017; Junwei et al. 2013; Wang et al. 2016), and various features of the target can be extracted from the radar echo (Eryildirim and Onaran 2011; Zhai and Jiang 2014). In air defense early warning systems, the preprocessing problems such as data volume redundancy and feature redundancy of infrared and radar fusion target detection are still unsolved hot issues.

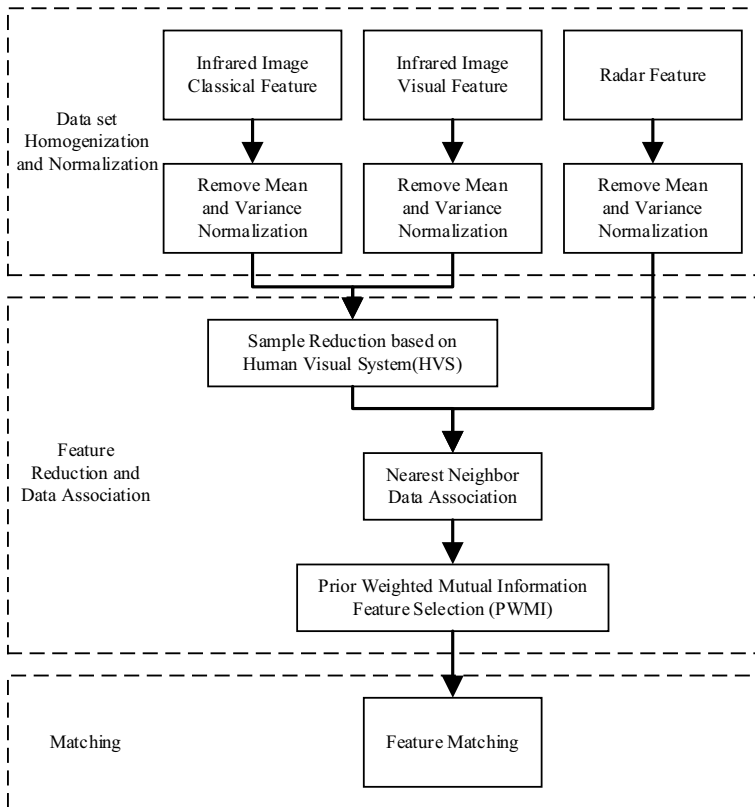
Figure 1 is a flow chart of the infrared and radar fusion detection scheme, which considers the target detection problem as a two-class problem of “target” and “non-target”. The steps of feature fusion include: data set homogenization and normalization, feature

---

✉ Yuan Wei  
weiyuan2550@163.com

<sup>1</sup> State Key Laboratory of Pulsed Power Laser Technology, Hefei 230000, China

<sup>2</sup> National University of Defense Technology, Hefei 230000, China



**Fig. 1** Infrared and radar fusion detection

reduction and data association, and feature matching (Luo et al. 2011). The first two steps are pre-processing, the purpose is to build and optimize the data set for the classifier to perform feature matching. Preprocessing is a bridge connecting classifiers and specific problems.

Farcmeni and Greco (2016) systematically discussed the sample reduction. Zhang et al. (2016) proposed a sample reduction method similar to the filtering method. The experimental results on the face data set show that the sample reduction improves the detection effect. Tang et al. (2017) mentioned the Laplacian embedded sample reduction method to solve the sample problem in cluster analysis. These studies primarily screen unsupervised data sets through statistical theory. To construct a feature dataset of infrared images, the dataset needs to contain interference, clutter, background, and abstract the target into a data vector. Relevant research has not been found.

In terms of feature dimension reduction, Ghojogh et al. (2019) and Li et al. (2018) have made detailed summaries of the related research on feature dimension reduction. Similarly, infrared and millimeter-wave fusion data sets are highly interpretable compared to dictionary-like data sets, which is not considered in previous studies.

In infrared and radar fusion detection, data volume redundancy and feature redundancy are two problems encountered in practical applications, and need to be overcome in the preprocessing stage. For the problem of excessive data volume in the data set from the infrared image,

we propose a sample reduction method based on Human Visual System (HVS) to reduce the volume of data set. By eliminating the background and clutter data that are not helpful in classification, the data set is optimized. A nearest neighbor data association is performed to construct a joint feature vector. For the problem of feature redundancy, a Prior Weighted Mutual Information Feature Selection (PWMI) is performed using the physical meaning of the feature. The experiments show that the joint data set after sample reduction and feature selection can obtain higher detection probability than the unfiltered data set. The application scenarios of infrared and radar fusion target detection are varied. The average performance of PWMI is comparable to that of classical algorithms. And it is superior to other feature selection methods in the case of significant difference between training set and detection set. The fusion detection scheme proposed can effectively solve the problems in the fusion detection and improve its robustness.

## 2 Sample reduction of infrared data sets based on hvs

A single infrared image contains approximately  $480 \times 640$  pixels. If we convert all the pixels into a data set  $\Omega$ , an excessive amount of data will consume a large amount of computing resources, and the redundant information will also affect the detection result. To build a data set based on more images, we must filter  $\Omega$  to obtain a reduced subset. In infrared small target detection, the non-target part contains low-grey pixels, which composed of clear sky and low-brightness clouds, and includes high-grey pixels, which are composed of highlight clouds, buildings, trees, and fake targets. The target itself is a high-grey point, and the false alarm is mainly derived from the background noise and the target of the high grey scale. The background noise of the low grey scale has little influence on the detection of the small target than the former. Low grey noise exacerbates the imbalance between the target and background data, but does not contribute much to the classification. By excluding these low-grey noises to construct a reduced data set  $\Omega^R$ , the efficiency of the algorithm can be improved without reducing the classification accuracy.

The human visual system (HVS) is an efficient screening method. The infrared image is transformed into the LOG scale space, and the greyscale difference mechanism of the HVS is used to suppress the peripheral region of the pixel to be inspected. At the same time, the transformation on multiple scales also utilizes the multiscale mechanism of the HVS to extract the scale of the points of interest.

As Fig. 2 shows, infrared image  $I$  is subjected to LOG scale spatial transformation with scales of  $\sigma_1, \dots, \sigma_N$ , and a plurality of scale spatial transformation maps  $I_1, \dots, I_N$  are obtained. The transformation maps of the respective scales are normalized. For point  $(x, y)$ , the scale  $\sigma_m(x, y)$  of the normalized scale space map sequence that takes the transform value  $I_n(x, y)$  to the maximum value  $I_m(x, y)$  is the scale of the data point. The maximum value of the scale space images  $I_1, \dots, I_N$  is obtained pixel by pixel to obtain a multiscale fused image  $I_m$ .

$$I_m(x, y) = \max_{n=1 \dots N} [I_n(x, y)], \quad n = 1, 2 \dots N \quad (1)$$

$$\sigma_m(x, y) = \arg \max_{n=1 \dots N} [I_n(x, y)], \quad n = 1, 2 \dots N \quad (2)$$

Fast non-maximum suppression is conducted through image meshing.  $I_m$  is divided into image blocks of  $G_{1,1}, G_{1,2} \dots G_{1,g_n}, G_{2,g_n} \dots$  and  $G_{g_m,g_n}$  of  $g_m$  rows and  $g_n$  columns by grids

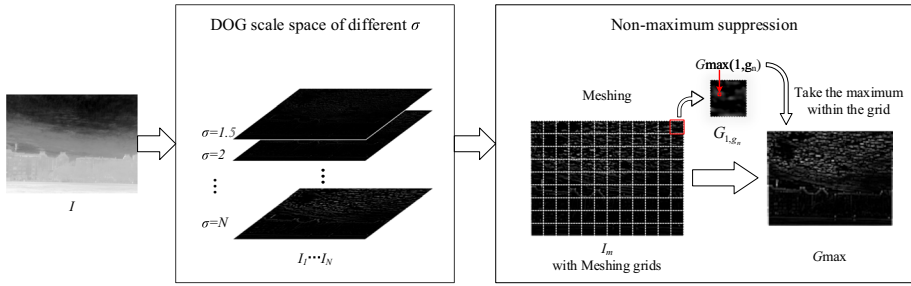


Fig. 2 HVS-based infrared image sample reduction

of a side length  $a$ . For each image block  $G_{i,j}$ , its maximum value  $G_{max}(i,j)$  and coordinates  $(x_{i,j}, y_{i,j})$  are obtained.  $G_{max}$  is the maximum combination of size  $g_m \times g_n$ . For all pixels on  $G_{max}$ , if the pixel is not larger than each value on all of 8 neighbourhoods, the pixel is suppressed to 0. Traverse every grid  $G_{i,j}$  for  $G_{max}(i,j)$ , Combine them to obtain  $G_{max}$ , the maximum map after suppression. As the Fig. 2 shows, the size of  $G_{max}$  is  $g_m \times g_n$ , which contains all the non-zero pixels in  $I_m$ . The non-zero pixels in  $G_{max}$  is the reduced data.

For the pixel  $G_{max}(i,j)$ , the corresponding infrared data point coordinates are  $(x_{i,j}, y_{i,j})$ , the grey value is  $f_{Grey} = I(x_{i,j}, y_{i,j})$ , the scale is  $f_{Scale} = \sigma_m(x_{i,j}, y_{i,j})$ , and the gradient is  $f_{Grad} = func_g(I)(x_{i,j}, y_{i,j})$ , in which  $func_g(\cdot)$  is the gradient function. The infrared classical feature vector is  $(f_{Grey}, f_{Scale}, f_{Grad})$ .

It can be seen from the process of non-maximum suppression that the data point obtained by the suppression is the maximum value among the 9 adjacent grids. For any point, even if the point is located at the edge of the grid, it will be the largest in the vicinity of the range of  $2a \times 2a$ . Figure 3 shows a point located at the edge of the grid of the second row, second column. Even at the edge of the grid, the  $2a \times 2a$  area around the point is still in the suppression range (the adjacent 9 grids).

$2a \times 2a$  is a range that can be represented by a single maximum point. As the Society of Photo-Optical Instrumentation Engineers defined, a small target has a total spatial extent of less than 80 pixels, occupying no larger than  $9 \times 9$  (Chen et al. 2013). To ensure that the suppression result of a target contains only one output, the range of  $2a \times 2a$  should be no less than  $9 \times 9$ . Take  $a = 5$ , thus the target will be extracted into a single point after suppression, at the same time multiple targets will not be confused as a single point.

Algorithm 1 shows the pseudocode for the sample reduction of an infrared image data set.

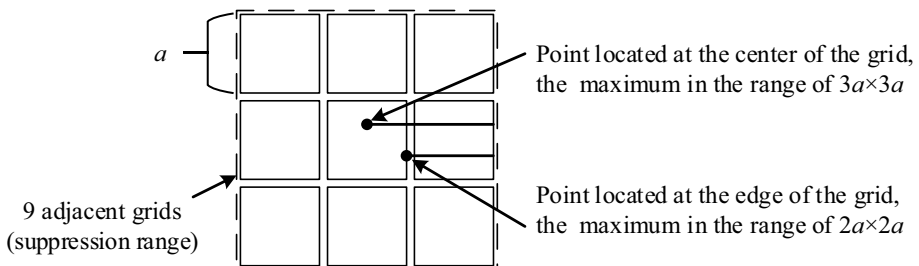


Fig. 3 HVS-based infrared image sample reduction

```

Algorithm 1 Sample reduction of data sets based on HVS
Step 1 Multiscale LOG filtering to obtain multiscale fused images  $I_m$  and scale  $\sigma_m$ .
for  $n = 1 : N$ 
     $I_{LOG} = LOG_{\sigma_n} * I$ 
End
Step 2 Non-maximum suppression
for  $i = 1 : g_m$ 
    for
        if  $G_{max}(i, j) > 0$  and  $\forall G_{max}(i, j) > G_{max}(i \pm 1, j \pm 1)$ 
             $G_{sup}(i, j) = G_{max}(i, j)$ 
        end
    end
end
end
    
```

### 3 Prior weighted mutual information feature selection

Feature selection is used to exclude features that do not contribute to classification from feature vectors and to optimize the feature vectors that contain redundant and repeated information. After data association, the feature vectors are high-dimensional infrared features and radar features. Unlike the dictionary-like features, these features have physical meanings, strong interpretability, and different physical characteristics due to their physical sources and extraction principles. Additionally, there is strong independence between different features. Based on these characteristics, this manuscript proposes a Prior Weighted Mutual Information Feature Selection based on a feature’s prior information.

First, a brief introduction to information and mutual information theory is given. Shannon Information Theory gives the concept of information entropy. The information entropy of random variables  $Y$  is:

$$H(Y) = - \sum_{y=1}^{N_y} P(y) \log P(y) \tag{3}$$

Entropy measures the degree of uncertainty of a single variable  $Y$ , and its value is determined by the distribution of variable  $Y$ . To measure the correlation between two variables,

the concept of mutual information is proposed. The expression of the mutual information of variables  $X$  and  $Y$  is

$$I(X;Y) = - \sum_{x,y} P(x,y) \log \left( \frac{P(x,y)}{P(x)P(y)} \right) \tag{4}$$

in which  $P(x,y)$  is the joint probability density function of variables  $X$  and  $Y$ ,  $P(x)$  and  $P(y)$  are the marginal probability density functions, and the mutual information is the relative entropy between the joint distribution  $P(x,y)$  and the product distribution  $P(x)P(y)$ .

Feature selection based on mutual information is a filtering method. It can simultaneously consider the extent to which the feature itself contributes to the classification and the actual degree of correlation between the features. The mutual information value characterizes the dependency between a single feature and a data class, and is thus the dependency between two features. Based on the combination and iteration of mutual information values, each feature is scored to achieve sorting and optimization of the features. Classical methods include mutual information maximization (MIM) (Lewis 1992), mutual information preference (MIFS), minimum redundancy maximum correlation (mRMR) (Peng et al. 2005), conditional mutual information feature (CIFE) (Lin and Tang 2006), joint mutual information (JMI) (Meyer et al. 2008), etc. Among them, mRMR has been widely praised in the industry and is the most representative classic mutual information feature selection method.

$$mRMR(X_k) = MI(X_k, Y) - \frac{1}{|S|} \sum_{X_j \in S} [MI(X_k, X_j)] \tag{5}$$

mRMR adopts the method of forward search. It gradually adds the feature of the highest score to an empty set and continuously updates the score of the remaining features. Equation (6) is the score of the  $k$ th feature,  $|S|$  is the number of features in the selected feature set  $S$ , and  $\frac{1}{|S|}$  is the factor of correlation between two selected features. By setting a threshold for the feature's score, select the feature that meets the requirements is selected.

Mutual information statistically recognizes the contribution of features to classification and the degree of redundancy relative to other features. Different from the statistical significance emphasized by these methods, infrared and radar features have clear physical meaning and are highly interpretable, and their prior information can be used to assist in feature selection. The statistical knowledge acquired through the data set is subject to the limitations of the data set itself, and is thus easy to overfit. Radar features and infrared features are derived from heterogeneous sensors, and the redundancy of information between them is minimal. The redundancy of information between infrared features based on different visual theories is higher, and the redundancy of information between features with the same principles is the highest.

Therefore, this paper proposes the Prior Weighted Mutual Information Feature Selection method:

$$PWMI(X_k) = MI(X_k, Y) - \frac{\beta(k,j)}{|S|} \sum_{X_j \in S} [MI(X_k, X_j)] \tag{6}$$

where  $\beta(k,j)$  is the weight coefficient of the mutual information value of feature  $X_k$  and feature  $X_j$ .

To elaborate on the application of PWMI, Fig. 4 shows the weight distribution of a feature vector in feature selection. The characteristics adopted in the experiment are divided into three categories: infrared classical features, infrared visual features and radar features. There are 3 infrared classical features: greyscale, scale, and Sobel gradient. According to their sources, they can be divided into three groups: scale, greyscale and gradient. 9 visual features are listed: LoG (Shao et al. 2012), ILOGb (Han et al. 2016), LoPSF (Moradi et al. 2018), EbLDM (Deng et al. 2017), LCM (Han et al. 2019), LGSM (Wei et al. 2016), MLHM (Nie et al. 2018), MPCM (Wei et al. 2016), NWIE (Deng et al. 2016). According to their sources, they can be divided into 4 groups: attention mechanism characteristics, contrast mechanism characteristics (difference), contrast mechanism characteristics (quotient), and edge mechanism. 4 radar characteristics are listed: acceleration, velocity, distance, and RCS. According to their sources, they are divided into three groups: speed, distance, and scattering area. The weight coefficient  $\beta_{k',j'}$  is determined by the categories and group of the two features.

In this example, the weights of the features based on the same principle are  $\beta_{inner}$ . The weights of the features belonging to a large class of features and different groups are  $\beta_{nor}$  (both infrared classical features),  $\beta_{vis}$  (both infrared visual features), and  $\beta_{radar}$  (both radar features). The weights of the two features belonging to the radar feature and the infrared feature are  $\beta_{n_r}$  and  $\beta_{v_r}$ , respectively. The weights of features belonging to the infrared classical features and the visual features are  $\beta_{n_v}$ . The distribution of  $\beta_{k',j'}$  is shown in Fig. 4.

In Eq. (6),  $MI(X_k, X_j)$  is the mutual information value of the features  $X_k$  and  $X_j$ . A high mutual information value indicates a strong correlation in the statistical sense. Additionally,

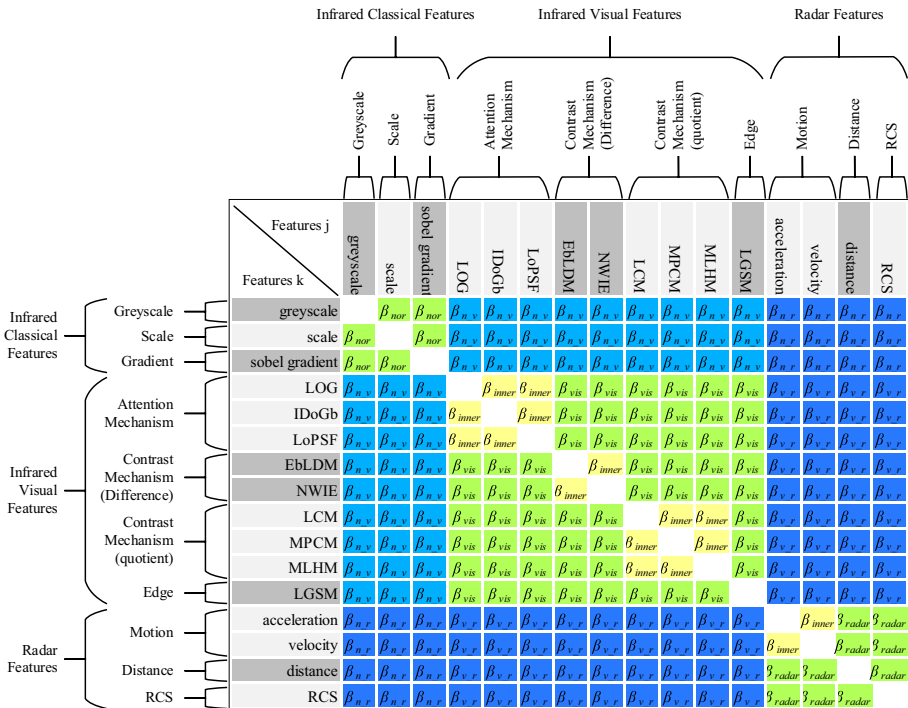


Fig. 4 A prior weight distribution of two feature mutual information

features of the same group characterize the same aspect of the HVS. For example, the infrared LoG feature and the LoPSF feature both reflect the same saliency in terms of the HVS. The information extracted by the two features has a certain overlap, thus the correlation between them is strong.

To minimize redundant information, it is given a higher weight coefficient, making it inferior to other categories. For features belonging to the same category and different groups, the weight coefficient is smaller than the former coefficient. For features belonging to different classes, which are extracted from different sources of information, the degree of redundancy between the two features is minimal, thus giving them the smallest weighting factor.

The characteristic weight  $\beta(k, j)$  of PWMI and the relationships of each value are as follows:

$$\beta(k, j) = \begin{cases} \beta_{n\_m} & X_k \text{ and } X_j \text{ belong to infrared regular feature and radar feature} \\ \beta_{v\_m} & X_k \text{ and } X_j \text{ belong to infrared visual feature and radar feature} \\ \beta_{n\_y} & X_k \text{ and } X_j \text{ belong to infrared regular feature and infrared visual feature} \\ \beta_{nor} & X_k \text{ and } X_j \text{ belong to infrared regular feature but different groups} \\ \beta_{vis} & X_k \text{ and } X_j \text{ belong to infrared visual feature but different groups} \\ \beta_{mmw} & X_k \text{ and } X_j \text{ belong to radar feature but different groups} \\ \beta_{inner} & X_k \text{ and } X_j \text{ belong to the same feature and the same group} \end{cases} \tag{7}$$

$$1 \leq \beta_{n\_y} = \beta_{v\_m} < \beta_{n\_m} < \beta_{nor} = \beta_{vis} = \beta_{mmw} < \beta_{inner} \tag{8}$$

The value of the weight coefficient A in the paper is determined experimentally as:

$$\begin{aligned} \beta_{n\_m} &= \beta_{v\_m} = 1 \\ \beta_{n\_y} &= 1.1 \\ \beta_{nor} &= \beta_{vis} = \beta_{mmw} = 1.3 \\ \beta_{inner} &= 1.5 \end{aligned} \tag{9}$$

## 4 Experiment and analysis

To verify the validity and feasibility of the proposed fusion detection method, the training set and test set are respectively established using the measured data and simulation data under different times and for different scenes, and the RBF-SVM classifier is used for classification. Three experiments of fusion detection and single sensor detection, HVS sample reduction, and PWMI were conducted and the results were analysed.

### 4.1 Data set and experiment settings

The data set used in the experiment is an infrared and radar feature fusion data set. The infrared features are collected by 12 image sequences for different scenes. The collection scene and image characteristics of each sequence are shown in Table 1:

The radar features in the data set are based on the simulated data at different settings. The settings for the simulated data are shown in Table 2:

The data set required for the experiment is constructed by infrared image and simulated radar signals of different scenes. A total of 4 infrared features are included:



**Table 1** Statistics of small target detection

Num	Scene	Background	Small target	False target
1	1	Building	Moving fast, low greyscale	None
2	1	Building	Moving, low greyscale	None
3	2	Trees	Moving, low greyscale	None
4	3	Building, trees, a small amount of clouds	Moving, low greyscale	None
5	3	Very few clouds	Moving fast, high greyscale	Bird
6	3	High brightness clouds	Moving fast	Bird
7	4	Several high brightness clouds	Moving, high greyscale	Plane
8	4	Several high brightness clouds	Almost stationary, low greyscale	None
9	4	High brightness clouds	Moving fast	Plane
10	5	Building, several high brightness clouds, pedestrian	Moving	Plane
11	5	Building, moving pedestrian	Almost stationary, high greyscale	None
12	5	Building, moving vehicle, Moving pedestrian	Moving, high greyscale	None

**Table 2** Radar feature simulation settings

Type	Maximum acceleration (m/s <sup>2</sup> )	Velocity (m/s)	distance (km)	RCS(dBsm) (min/mean/max)
Tactical missile (target)	4–6G	≤ 391.7	5–9	– 30/– 20/10
Medium passenger aircraft	0.39G	51–125	7–20	16.6/20/43.9
Combat aircraft	5G	300–420	5–12	3/10/33
Helicopter	5	0–102	4–10	– 10/10/65
Bird	0.03	0–21	0.01–3	– 33/– 30/– 20

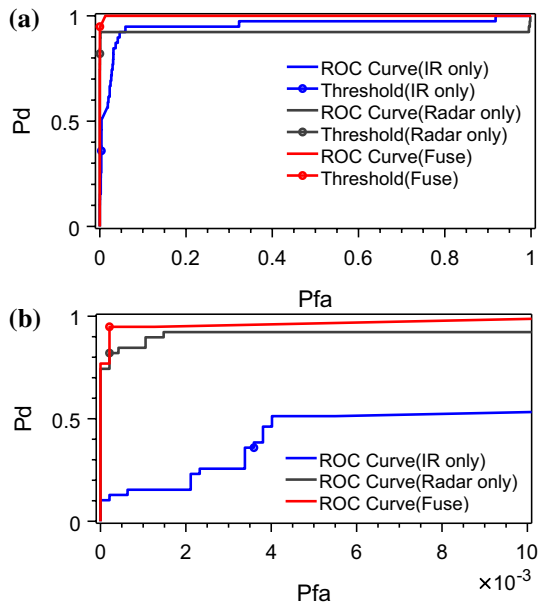
grey scale, scale, and sobel gradient. The 9 visual features are: LoG, IDoGb, LoPSF, EbLDM, LCM, LGSM, MLHM, MPCM, and NWIE, and the 4 radar features are: acceleration, speed, distance, and RCS.

To obtain more reliable classification performance, the experiment uses the cross-validation method to perform tenfold cross-validation on the data set. Each experiment takes turns to use different parts of the data as a training set and test set in a 4:1 ratio. Each experiment will determine the detection probability and false alarm probability of the experiment.

## 4.2 Results and analysis of detection based on infrared features, radar features and all features

To verify the superiority of fusion detection, target detection is conducted based on infrared features, radar features and all features, respectively. The detection results are shown as follows:

**Fig. 5** ROC curves of single sensor detection and fused detection



**Table 3** Single sensor detection and fusion detection results

	Detection based on infrared features	Detection based on radar features	Detection based on all features
$P_d$	0.359	0.8205	0.949
$P_{fa}$	0.00359	0.000211	0.000211
AUC	0.954	0.923	0.999

Figure 5a shows the ROC curve after parameter tuning, where the circular mark is the threshold taken by the SVM classification, and the Fig. 5b is a partial enlargement of the low false alarm probability.

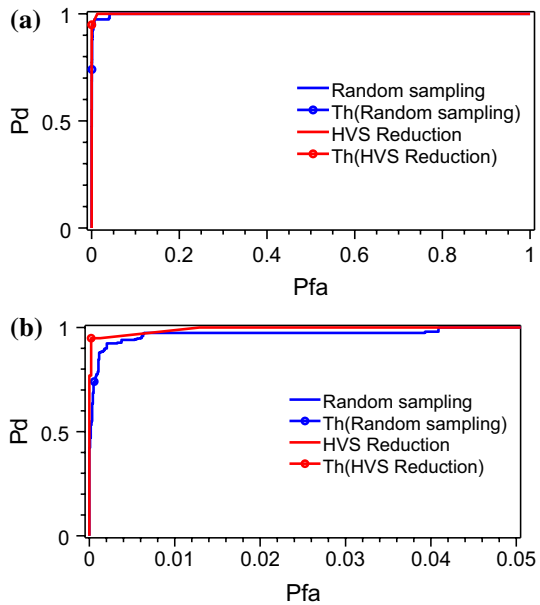
We chose detection probability (Pd) and false alarm probability (Pfa), and Area Under Curve(AUC) as the indicators. Pd and Pfa are classic indicators of target detection. AUC represents the area under the ROC curve. And its value is independent of the threshold, thus it can measure the performance of the classifier more comprehensively. The larger the AUC is, the better the classifier detection effect. The experimental results are shown in Table 3:

When only infrared features or radar features are utilized, the ROC curves are located below the ROC curve of the fusion detection. The circular mark on the ROC curve is the classification threshold, and the SVM’s classification results for the fusion data set are also superior to the single sensor. The AUC value of the fusion detection is higher than the single sensor detection, which also verifies the previous results.

**Table 4** Data set construction and detection results

		Randomly sampled	HVS sample reduction
Data set size	Total number of pixels (width × height × number of images)	640 × 480 × 120	640 × 480 × 120
	Data set size (data volume × number of features)	36864 × 14	25234 × 14
Detection results	Detection probability	0.7403	0.9487
	False alarm probability	0.0005724	0.0002115
	AUC	0.9983	0.9996

**Fig. 6** Detection results with sample reduction and without



**4.3 Results and analysis of fusion detection based on HVS infrared data set sample reduction**

The infrared data set was constructed in a manner of no screening and HVS sample reduction, respectively. In the control group, the image pixels were randomly selected as the data set without screening. In the experimental group, the data set was selected based on the HVS-based data set sample reduction. The parameters of the data set and the corresponding detection results based on it are shown in Table 4:

The amount of data in the randomly sampled data set is adjusted so that the amount of data is roughly equal. The detection results show that the reduced data set has a higher probability of detection. The ROC chart of the detection results of the two data sets is shown below:

Figure 6 shows the ROC plot of the results, where (b) is a partial enlargement of the low false alarm probability portion of (a). The ROC curve based on the HVS data set is higher

than the randomly sampled data set, and has a higher AUC value. Thus, its detection effect is better than the randomly sampled data set.

#### 4.4 Results and analysis of fusion detection based on PWMI feature selection

To verify the selection effect of the PWMI feature selection method, 5 classical feature selection methods [Fisher Score (Duda et al. 2012), MI (Lewis 1992), mRMR (Peng et al. 2005), CIFE (Lin and Tang 2006), and JMI (Meyer et al. 2008) and ILFS (Roffo et al. 2017)] were used as the control group. The results of each classification and the average results of the multiple trials were analysed to verify the effectiveness of the proposed method. 5 partitioning schemes were selected to divide the data set into training set and test set. The detection probabilities and number of selected features of each partitioning scheme and the mean value of each scheme are shown in Table 5:

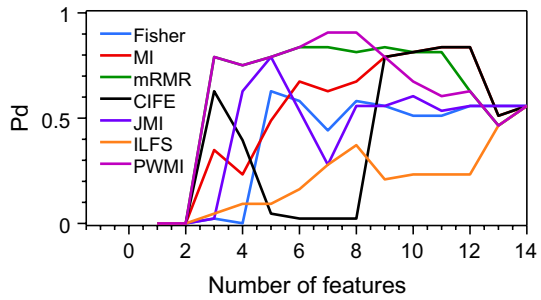
As shown above, the feature selection algorithm with the highest detection probability is marked in bold, and “Ave” represents the average value of the 5 schemes. In the first, third, and fifth experiments, the detection results based on the feature selection data set are superior to the complete data set in the detection probability, and in the second and fourth experiments, the detection probability of the feature selection data set is equivalent to the complete data set. Since some features are lowly correlated with the target, filtering out these features helps improve the detection probability by the classifier. From the average of the detection probability, the feature selection method with the highest detection probability is mRMR, with a detection probability of 0.7574, which is generally better than other feature selection algorithms. The average detection probability of PWMI is 0.7396, which is second only to mRMR.

However, the case of Experiment 1 is quite special. The training set consists of image sequences 1–9, and the test set consists of image sequences 10–12. It can be seen from Table 4 that the training set and the detected concentrated infrared features are obtained for completely different scenes. Using all the features for classification, the detection probability is lower than most of the selected detection probabilities. mRMR is the best-performing feature selection method in many experiments, but in experiment 1 its detection probability is lower than PWMI. Under the condition that the difference between the training set and the test set is large, only the statistical information of the training set is used to select the optimal feature, and the most suitable feature cannot be selected accurately. It is thus necessary to judge the priority of the selection considering the

**Table 5** Detection probability and number of selected features when applying different feature selection methods

Scheme	No feature selection	Fisher score	MI	mRMR	CIFE	JMI	ILFS	PWMI
1	0.5581	0.6279(5)	0.8372(11)	0.8372(5)	0.8372(11)	0.7907(5)	0.3721(8)	<b>0.9302(6)</b>
2	<b>0.6923</b>	0.6667(11)	0.641(8)	<b>0.6923(7)</b>	0.5128(11)	0.5128(6)	0.3333(8)	<b>0.6923(7)</b>
3	0.7907	0.5814(8)	0.8372(11)	<b>0.8837(7)</b>	0.7907(9)	0.5581(8)	0.4651(11)	0.7907(5)
4	<b>0.6667</b>	0.6154(7)	0.5597(11)	<b>0.6667(5)</b>	0.5385(11)	<b>0.6667(6)</b>	0.2564(9)	0.644(6)
5	0.573	<b>0.7813(5)</b>	0.6408(5)	0.7073(5)	0.4418(11)	0.6917(3)	0.2051(8)	0.6408(7)
Ave	0.6561	0.6545	0.7032	<b>0.7574</b>	0.6242	0.644	0.3264	0.7396

**Fig. 7** The influence of the number of selected features on the detection probability



physical meaning of each feature. Figure 7 shows the variation of the detection probability in Experiment 1 with the number of selected features, and Table 6 shows the scores obtained by all feature selection algorithms.

It can be seen that the PWMI and mRMR results are the same in the selection of the first 6 features, and the PWMI is better than the mRMR when the seventh feature is selected. As seen from Table 6, mRMR selects NWIE, and PWMI selects MLHM. The seven choices of MRMR and PWMI are marked by bolding, while their seventh choices are marked by underlining. MLHM belongs to the infrared visual feature of the contrast mechanism characteristic (quotient) group. There are no similar type features for the first 6 features selected. NWIE and EbLDM belong to the contrast mechanism characteristic (difference). The PWMI gives EbLDM a higher weight as punishment, so that the MLHM based on different physical principles is selected as the candidate target. Thus, a better detection probability is achieved under the condition that the training set and the test set have large differences.

**Table 6** Scores of various features by the feature selection algorithms

Feature	Fisher	MI	mRMR	CIFE	JMI	ILFS	PWMI
Greyscale	0.0008(15)	0.0034(14)	-0.3561(16)	0.0088(5)	0.0068(7)	88.4931(10)	-0.3561(16)
Scale	0.1195(12)	0.0003(15)	<b>-0.0039(3)</b>	0.0044(7)	0.0008(15)	63.1153(12)	<b>-0.0043(3)</b>
Sobel gradient	0.0921(13)	0.0095(3)	<b>-0.0138(6)</b>	-0.001(14)	0.0012(14)	66.4545(11)	<b>-0.0319(6)</b>
LOG	3.1365(5)	0.0076(7)	-0.0904(10)	0.0115(4)	0.0069(5)	104.2971(6)	-0.1001(9)
IDoGb	0.4463(10)	0.0058(10)	<b>0.0025(2)</b>	0.0032(8)	0.0032(12)	104.2971(5)	<b>0.0021(2)</b>
LoPSF	0.303(11)	0.0046(12)	<b>-0.0133(5)</b>	0.0012(12)	0.0028(13)	104.296(9)	<b>-0.0197(5)</b>
EbLDM	0.0244(14)	0(16)	<b>-0.0056(4)</b>	0.0031(9)	0.0004(16)	3.3019(14)	<b>-0.0064(4)</b>
NWIE	0.8628(9)	0.0051(11)	<b>-0.0194(7)</b>	0.0029(10)	0.0046(11)	104.2971(4)	-0.1411(10)
LCM	0(16)	0.0092(4)	-0.1529(13)	-0.0008(13)	0.0076(3)	0(16)	-0.1543(11)
MPCM	1.1292(8)	0.0065(8)	-0.3109(15)	0.0138(3)	0.0068(6)	104.2971(3)	-0.3109(15)
MLHM	1.1788(7)	0.0059(9)	-0.0432(8)	0.0047(6)	0.0053(10)	104.2971(2)	<b>-0.0444(7)</b>
LGSM	81.6172(1)	0.004(13)	-0.1877(14)	0.0199(1)	0.0074(4)	9.7636(13)	-0.1957(13)
Acceleration	1.6195(6)	0.0102(2)	-0.0714(9)	-0.0168(16)	0.0077(2)	104.2971(1)	-0.0785(8)
Velocity	4.3806(3)	0.0081(6)	-0.1375(12)	-0.0063(15)	0.0057(9)	104.296(8)	-0.2224(14)
Distance	3.9707(4)	0.0084(5)	-0.1129(11)	0.0013(11)	0.0065(8)	104.2964(7)	-0.1558(12)
RCS	38.5315(2)	0.018(1)	<b>0.018(1)</b>	0.018(2)	0.018(1)	0.3819(15)	<b>0.018(1)</b>

## 5 Conclusion

The performance of infrared and radar fusion target detection is better than that of single sensor, and data preprocessing is the key to data fusion. By reducing the amount of data in the infrared data set, the detection probability is improved. The Prior Weighted Mutual Information Feature Selection method utilizes the physical meaning of the feature itself, and can achieve higher detection probability in the case of large difference of scenes, and the average detection probability is equivalent to the classical algorithm. It effectively improves the performance and robustness of the infrared and radar fusion detection.

**Acknowledgements** This research was funded by the National Natural Science Foundation of China grant numbers 61271376; Natural Science Foundation of Anhui Province grant numbers 1208085MF114.

## Compliance with ethical standards

**Conflicts of interest** The authors declare no conflict of interest.

## References

- Chen, C.L.P., Li, H., Wei, Y., et al.: A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **52**(1), 574–581 (2013)
- Deng, H., Sun, X., Liu, M., et al.: Infrared small-target detection using multiscale grey difference weighted image entropy. *IEEE Trans. Aerosp. Electron. Syst.* **52**(1), 60–72 (2016)
- Deng, H., Sun, X., Liu, M., et al.: Entropy-based window selection for detecting dim and small infrared targets. *Pattern Recogn.* **61**, 66–77 (2017)
- Duda, R.O., Hart, P.E., Stork, D.G., et al.: *Pattern classification*. Wiley, Hoboken (2012)
- Eryildirim, A., Onaran, I., et al.: Pulse Doppler radar target recognition using a two-stage SVM procedure. *IEEE Trans. Aerosp. Electron. Syst.* **47**(2), 1450–1457 (2011)
- Farcomeni, A., Greco, L., et al.: *Robust methods for data reduction*. Chapman and Hall/CRC, London (2016)
- Ghojogh, B., Samad, M.N., Mashhadi, S.A., et al.: Feature selection and feature extraction in pattern analysis: a literature review (2019). arXiv preprint [arXiv:1905.02845](https://arxiv.org/abs/1905.02845)
- Han, J., Liu, S., Qin, G., et al.: A local contrast method combined with adaptive background estimation for infrared small target detection. *IEEE Geosci. Rem. Sens. Lett.* **16**(9), 1442–1446 (2019)
- Han, J., Ma, Y., Huang, J., et al.: An infrared small target detecting algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* **13**(3), 452–456 (2016)
- Jia, J., Wang, Y., Zhuang, X., et al.: High spatial resolution shortwave infrared imaging technology based on time delay and digital accumulation method. *Infrared Phys. Technol.* **81**, 305–312 (2017)
- Junwei, L., Yueming, W., Xizhong, X., et al.: Study on shortwave infrared long-distance imaging performance based on multiband imaging experiments. *Opt. Eng.* **52**(4), 1–9 (2013)
- Lewis, D.D.: Feature selection and feature extraction for text categorization. In: *Proceedings of the Workshop on Speech and Natural Language*. Association for Computational Linguistics, pp. 212–217 (1992)
- Li, J., Cheng, K., Wang, S., et al.: Feature selection: a data perspective. *ACM Comput. Surv. (CSUR)* **50**(6), 94 (2018)
- Lin, D., Tang, X., et al.: Conditional infomax learning: an integrated framework for feature extraction and fusion. In: *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, pp. 68–82 (2006)
- Luo, R.C., Chang, C.C., Lai, C.C., et al.: Multisensor fusion and integration: theories, applications, and its perspectives. *IEEE Sens. J.* **11**(12), 3122–3138 (2011)
- Meyer, P.E., Schretter, C., Bontempi, G., et al.: Information-theoretic feature selection in microarray data using variable complementarity. *IEEE J. Sel. Top. Signal Process.* **2**(3), 261–274 (2008)
- Moradi, S., Moallem, P., Sabahi, M.F., et al.: A false-alarm aware methodology to develop robust and efficient multi-scale infrared small target detection algorithm. *Infrared Phys. Technol.* **89**, 387–397 (2018)
- Nie, J., Qu, S., Wei, Y., et al.: An infrared small target detection method based on multiscale local homogeneity measure. *Infrared Phys. Technol.* **90**, 186–194 (2018)

- Peng, H., Long, F., Ding, C., et al.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 1226–1238 (2005)
- Roffo, G., Melzi, S., Castellani, U., et al.: Infinite latent feature selection: a probabilistic latent graph-based ranking approach. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1398–1406 (2017)
- Shao, X., Fan, H., Lu, G., et al.: An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system. *Infrared Phys. Technol.* **55**(5), 403–408 (2012)
- Tang, M., Nie, F., Jain, R., et al.: A graph regularized dimension reduction method for out-of-sample data. *Neurocomputing* **225**, 58–63 (2017)
- Wang, Y., Xie, F., Wang, J., et al.: Short-wave infrared signature and detection of aircraft in flight based on space-borne hyperspectral imagery. *Chin. Opt. Lett.* **14**(12), 122801–122804 (2016)
- Wei, Y., You, X., Li, H., et al.: Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recogn.* **58**, 216–226 (2016)
- Wen, M., Wei, L., Zhuang, X., et al.: High-sensitivity short-wave infrared technology for thermal imaging. *Infrared Phys. Technol.* **95**, 93–99 (2018a)
- Wen, M., Wang, Y., Yao, Y., et al.: Design and performance of curved prism-based mid-wave infrared hyperspectral imager. *Infrared Phys. Technol.* **95**, 5–11 (2018b)
- Zhai, S., Jiang, T., et al.: Target detection and classification by measuring and processing bistatic UWB radar signal. *Measurement* **47**, 547–557 (2014)
- Zhang, M., He, R., Cao, D., et al.: Simultaneous feature and sample reduction for image-set classification. In: *Thirtieth Aaai Conference on Artificial Intelligence* (2016)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.