A Method for Extracting and Selecting Features of Thermal Infrared Remote Sensing Images of Harbor Targets

Lan MA1,*, Ting JIANG1 and Xiao-yong CHEN2,a

1Remote Sensing Engineering, Institute of Surveying and Mapping, Zhengzhou, 450000, China
2Remote Sensing Engineering, Nanjing Normal University, Nanjing, 210097, China
*yjszx_ma@163.com
ajszx_chen@163.com
*Corresponding author

Keywords: Feature extraction, Feature selection, TIR images, Harbor targets, Classifier, TIR remote sensing.

Abstract. To improve the inadequate effectiveness of thermal infrared (TIR) image recognition due to low contrast, striping noise and low spatial resolution, this paper conducted research with respect to harbor targets. It adopted sample features like texture, geometry, shape and statistics, selected the best combination of image features with an attribute selection evaluator, and chose the best classifier and the best classifier parameters according to the accuracy of recognition. This process can generate the best features for the classification of TIR images with a certain degree of robustness. The results of the experiment indicate that 1) the libSVM (a support vector machine), after parameter optimization, is the most accurate classifier and the best result can reach 82.4137%; 2) the one-dimensional statistical features of pixel value, mean value, variance, mid-value, and gray level histogram are more meaningful for the recognition; and 3) daytime images can be more accurately classified than nighttime ones.

Introduction

Thermal infrared (TIR) remote sensing is widely used for military purposes thanks to its capability of obtaining such information as the temperatures and states of surface features at a great distance and around the clock. The extraction and selection of the features of infrared remote sensing images and the classification of such images are key steps for the recognition of TIR targets; both the classification and recognition algorithm and the selection of features, through their interaction, can influence the accuracy of the classification and recognition of targets. As one of the much-debated challenges, this issue has received a good deal of attention and research.

Target features refer to the measurable function values of targets. Commonly used image features include color, edge, texture, and statistical features[1]. As a global feature, color is insensitive to variations in the direction and size of images or image regions and, consequently, unfit for an effective capture of local features of the targets in an image. For TIR images, the gray level of image pixels, which is a function of temperature, is characterized by low signal-to-noise ratio and blurred edges; this makes it considerably more difficult to extract this feature from an image for distinction between the targets and their backgrounds, and requires other types of features for a further description of those targets. Edge features, or a set of pixels showing abrupt changes in gray level, can be extracted with the gradient method or in other ways. As a global feature, texture reflects the pattern of gray level variation in the sub-regions of an image, i.e. the distribution of gray levels or directional information shown on the surface of the targets; textural structures typically vary from one image region to another[2]. Statistical features mainly refer to the mean value, variance, contrast and distribution of the gray levels of targets. There is a great variety of commonly used image target features, which can be divided into two broad categories 1) global features and 2) local features according to the scope of the region they describe[3]. The applicability of each method also differs accordingly.
Target expression involves the creation of a referential model that represents the target in the image region, or the selection of features. J. S. Desilva has extracted seven Hu invariant moments from infrared images of ship targets and achieved the automatic recognition of ship targets with the nearest neighbor classifier[4]. Paul Withagen has extracted from forward-looking infrared images 32 features of ship targets, including radiation, location, shape, and invariant moments. Then he analyzed and extracted features in terms of Mahalanobis distance and nearest neighbor measurement. Finally, he achieved the recognition of six categories of ships through a combination of a linear classifier, a nearest neighbor classifier, and a quadratic nonlinear classifier[5]. Jorge Alves has extracted the histograms of the edges of infrared ship targets, the features of which are combined with moment features for the recognition of ship targets through a neural network[6]. These feature selection methods, which are mostly based on particular conditions, such as infrared image sequences and close-distance infrared images, are not highly applicable to infrared remote sensing images with low spatial resolution.

Due to the combined effect of sensors, surroundings and other factors, TIR remote sensing images are generally characterized by low signal-to-noise ratio, low contrast between target and background, and low spatial resolution. The extraction and selection methods listed above are not quite applicable to harbor images, which feature complicated backgrounds, a multitude of targets, and indistinct textural information. In this paper, therefore, we have selected sample features like texture, geometry, shape and statistics according to the characteristics of TIR remote sensing images, chosen the optimum group of image features with the attribute selection evaluator, and opted for the best classifier and classifier parameters according to the accuracy of recognition, in order to classify TIR images more effectively.

The Extraction and Selection of the Features of Infrared Remote Sensing Images

Feature Extraction Method

Hessian Eigenvalues

The Hessian matrix is a square matrix of second-order partial derivatives that describes a local curvature function. It is widely used in image feature extraction operators. For instance, it is used as the input image in the SURF feature extraction operator. The Hessian matrix is as follows:

\[ H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix} \]  

(1)

Where,

\[ I_{xx} = G_{xx} * \text{Img} \]
\[ I_{yy} = G_{yy} * \text{Img} \]
\[ I_{xy} = G_{xy} * \text{Img} \]  

(2)

Here, \( \text{Img} \) means the original image. Calculate the second-order partial derivatives of \( x \) and \( y \) in the function \( g(x, y) \):

\[ G_{xx} = \frac{\partial^2 g(x, y)}{\partial x^2} \]
\[ G_{yy} = \frac{\partial^2 g(x, y)}{\partial y^2} \]
\[ G_{xy} = \frac{\partial^2 g(x, y)}{\partial x \partial y} \]  

(3)

g(\( x, y \)) is a Gaussian function,
\[ g(x, y) = \frac{1}{2\pi\delta^2} \exp\left(-\frac{x^2 + y^2}{2\delta^2}\right) \] (4)

Calculate the approximate value of the Hessian matrix determinant\[7\], and use it as the Hessian eigenvalue of this experiment:

\[ \det(H_{\text{approx}}) = I_{xx}I_{yy} - (0.9I_{xy})^2 \] (5)

**Local Entropy**

The entropy of an image, which is a measure of the uncertainty of the spatial distribution of its gray levels\[8\], is defined as follows:

\[ H = -\sum_{i=1}^{256} p_i \log p_i \] (6)

\( p_i \) is the probability of the number of pixels at the gray level in question in the total number of pixels. It is called local entropy when what is to be solved is the information entropy of a local image window. The local entropy of an image reflects the statistical features of the spatial distribution of image energy[9]. It has a certain degree of filtration, rotational invariance, and strength of local information variation.

**Fractal Dimension Image**

According to the fractal theory, if we move a sliding window over the original image and calculate the fractal dimension (FD) of the window as the FD of the central pixel in that window, i.e. the FD of each pixel in the image, we can repeat the process and find out the FD image. In an FD image, the FD values of man-made targets are higher than those of the natural background, which can serve to distinguish between man-made and natural targets.

To be more specific, we may define a sliding window with the size \( m \times n \) (which matches the size of the target to be examined), scan an image with that window from left to right and from top to bottom, and use the FD calculated from all the pixels in the window as the FD of the central pixel. Let \((i,j)\) be the central pixel in the window,

\[ F(k) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left| f(i, j) - f(i + k, j) \right|^2 + \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left| f(i, j) - f(i, j + k) \right|^2 \]

\[ = \frac{(m-k)n + (n-k)m}{(m-k)n + (n-k)m} \] (7)

In which \( k \) is the step length and \( C \) is a constant. From the formula above we can obtain \( H_{i,j} \):

\[ \log F_{i,j}(k) = 2H_{i,j} \log k + \log C \] (8)

And eventually the FD of each pixel in the image[10]:

\[ FD_{i,j} = 3 - H_{i,j} \] (9)

In this experiment, we have selected a 5×5 neighborhood, with 2 as the value of \( k \) and 1,000 as the value of \( C \).

**Third-order Cumulant**

High-order statistics is widely used in signal processing. In particular, the third-order cumulant is the principal instrument of analysis for the processing of non-Gaussian signals. The skewness (\( S \)) obtained through the calculation of third-order cumulants, also known as third-order characterization, is often used to measure the degree to which a stochastic process deviates from normal/Gaussian distribution[11]; zero \( S \) value means compliance with Gaussian distribution. This statistic can be used to measure the Gaussian features of a TIR image.
\[ \hat{S} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i - \eta}{\sigma} \right)^3 \]  \hspace{1em} (10)

In this formula, \( \eta \) is the mean value of the image and \( \sigma \) is the standard deviation.

**Other Statistical Feature Values of an Image**

All the first-order statistics here [12] are obtained from the histogram of an image. Let \( r_k \) be the kth gray level of a gray image; \( P(r_k) \) represents the frequency of the appearance of \( r_k \), i.e. the value of the gray histogram.

1. **Mo1**
   The mean value of this first-order statistic is defined as:
   \[ Mo1 = \sum_{k=1}^{256} r_k P(r_k) \]  \hspace{1em} (11)

2. **To1**
   Local features of an image can be expressed as the nth-order moment of the mean value:
   \[ \mu_n = \sum_{k=1}^{256} (r_k - Mo1)^n P(r_k) \]  \hspace{1em} (12)

   When \( n = 3 \),
   \[ To1 = \sum_{k=1}^{256} (r_k - Mo1)^3 P(r_k) \]  \hspace{1em} (13)

3. **Co1**
   From this first-order statistic we can obtain the contrast feature:
   \[ Co1 = \sqrt[3]{\sum_{k=1}^{256} (r_k - Mo1)^2 P(r_k)} \]  \hspace{1em} (14)

4. **So1**
   The smoothness of the variation in the gray scale information of an image can be calculated with the following formula:
   \[ So1 = 1 - \frac{1}{(1 + Co1^2)} \]  \hspace{1em} (15)

5. **Uo1**
   Under the condition of continuous radiation from surface features, the evenness of a TIR image reflects its uniformity:
   \[ Uo1 = \sum_{k=1}^{256} [p(r_k)]^2 \]  \hspace{1em} (16)

To extract local features of an image, we have employed 25 features, including texture (the five features of a gray level co-occurrence matrix[13]), gray level (gray level of the current pixel, mean value of neighborhood pixels, mid-value of pixels, and variance), geometry (Hessian eigenvalues), one-dimensional statistical features (local entropy, FD image, third-order cumulate, Mo1, To1, Co1, So1, and Uo1), and shape (the seven features of invariant moments[14]).

**Feature Selection Method**

Since irrelevant attributes have a negative impact on most machine learning schemes, the algorithm of machine learning requires, first of all, the selection of the attributes most suitable for decision-
making. Automatic selection is quite useful because manual selection[15], though it is the best approach, is tedious and prone to error. There are two ways to select attribute subsets:

1. The filter method. Before the start of learning, filter the attribute set to generate a most promising subset;
2. The wrapper method. The learning method is “wrapped” in the selection process, and the algorithm for the eventual machine learning will be used to evaluate the subset.

Automatic attribute selection requires the establishment of the attribute evaluator and the search method[16]. There are three common attribute evaluators—InfoGainAttributeEval, CfsSubsetEval, and WrapperSubsetEval. The first two are filter methods, and the last one is a wrapping method. InfoGainAttributeEval evaluates attributes by measuring their information gain corresponding to a category. CfsSubsetEval, which evaluates the predictive power of each of the attributes and the redundancy between them, is prone to select attributes with high relevance to the category attribute but low relevance between them. WrapperSubsetEval uses a classifier to evaluate an attribute set and, for each subset, estimates the accuracy of the learning plan through cross-validation; this evaluator involves a relatively great deal of computation. Best-first search, which executes retrospective greedy hill-climbing, can search forward from an empty attribute set, or backward from a universal set, or in both directions from a middle point. Ranker is not an attribute search method, but a scheme for ranking individual attributes.

Since we can see the category label to be used during the selection of attributes when using methods like cross-validation, the use of the testing data would affect the construction of the classification model and lead to a biased estimation of accuracy. To avoid this, WEKA’s meta-learning machine AttributeSelectedClassifier is often used for attribute selection. Apart from selecting the attribute evaluator and the search method, this approach also requires the setting of the base classifier and its parameters.

Selection of the Classifier

Classifiers

The C4.5 algorithm, which was developed by Professor Ross Quinlan as an improvement upon the ID3 algorithm in 1993, is one of the Decision Tree classification algorithms. It can process both continuous and discrete data and, using information gain rate as the attribute-selecting criterion of the Decision Tree, reduce over-fitting by pruning the spanning tree. This Decision Tree classifier, which is an eager learner, learns the training data and obtains the classification model before classifying unknown data. Eager learners are contrasted by lazy learners, the most typical of which are the Nearest Neighbor (NN) approaches, including the KNN algorithm. Bayesian networks, also known as belief networks, are a type of probability network and one of the most effective theoretical models for the expression and deduction of uncertain knowledge. The NaiveBayes algorithm assumes that the attributes are independent to each other. Support Vector Machines (SVM), which can simultaneously minimize empirical error and maximize geometric edges, have a good track record in its application to classification. In particular, the libSVM, an SVM library developed by Prof. Chih-Jen Lin from Taiwan in 2001, is a fast-calculating and open-source software. In the experiment, we have selected four common algorithms as classifiers, namely C4.5, KNN, NaiveBayes and libSVM, for the purpose of comparison between different classification algorithms.

The Selection of Classifier Parameters

Many of the parameters of classifier algorithms can affect the result of classification, but manual adjustment would be tedious and inefficient; the selection of parameters based on test data would cause errors in performance estimation. WEKA’s meta-learning machine offers two principal ways of parameter adjustment, i.e. two classifiers—CVParameterSelection and GridSearch. The former can search for the best parameter setting by optimizing the accuracy of the cross-validation of training data. This approach can optimize a base classifier with any number of parameters.
However, when there is a multitude of parameters, the efficiency of search can be affected by the excessively large number of parameter combinations; besides, it can only optimize the direct parameters of a base classifier, with no effect on embedded parameters. The latter is generally suitable for optimizing two parameters, especially embedded ones. In the libSVM classifier, for the cost parameter and gamma parameter of the RBF kernel function[17], let cstep and gstep be the step length of cost and gamma respectively during parameter optimization, then the cost values would be $2^{\text{costmin}}$, $2^{(\text{costmin}+\text{cstep})}$, \ldots, and $2^{\text{costmax}}$, and the gamma values would be $2^{\text{gammamin}}$, $2^{(\text{gammamin}+\text{gstep})}$, \ldots, and $2^{\text{gammamax}}$. The default values are cstep=1, gstep=1. Therefore, cost and gamma are embedded parameters, and the GridSearch classifier would be the best choice for parameter optimization.

The Experiment and Analysis of the Results

Based on the algorithms described above, we conducted an experiment in Matlab 2012a on the Windows 7 platform. We selected two TIR remote sensing images of target harbors for testing (Im1, obtained during the day and Im2, during the night), with a spatial resolution of 15 m. First of all, we conducted manual classification for the two panchromatic images corresponding to the two selected imaged. The classified index images feature five categories—ships (#1, in red), vegetation (#2, in green), buildings (#3, in blue), water (#4, in yellow), and others (#0, in black). Then, based on the corresponding panchromatic images, we conducted polynomial registration for Im1 and Im2 with ENVI. Finally, we superimposed the AOI (area of interest) of manual classification[18] on the registered TIR images to generate the classified index images. The original TIR images and the classified index images are shown below in Fig. 1.

![Figure 1](image)

Figure 1 The original TIR images and the manually classified index images: (a) is Im1, (b) is its manually classified index version; (c) is Im2, and (d) is its manually classified index version.
Visual observation shows that the manually classified index images are not quite accurate in the separation of targets (e.g. ships). Since the original TIR images have relatively low spatial resolution, which makes it difficult to tell the targets apart visually, we have adopted the simultaneously obtained panchromatic images for detailed classification. In comparison with the panchromatic images, the targets in the TIR images are blurrier and, due to thermal diffusion, are slightly larger than the real contours. As a result, the manually classified index images deviate slightly from visual observation.

In the experiment, with the classified index versions of Im1 and Im2 as the input data, we formed a sub-image in the neighborhood of the current pixel, selected local features in the sub-image. Then, using these features and the classified index image values as the sample data, we conducted the classification experiment with SVM and other machine learning methods[19]. In this experiment, the size of the sub-image was 5×5.

Next, the experiment was conducted in terms of the following:

1. Selecting the best combination of image features based on different features’ influence on the accuracy of recognition;
2. Selecting the best classifier parameters based on different parameters’ influence on the accuracy of recognition;
3. Selecting the best classifier based on different classifier’s influence on the accuracy of recognition;
4. Experimenting with night infrared images to obtain the best combination of image features, classifier parameters and recognition accuracy for such images.

### Selecting the Best Combination of Image Features

The sequence and labels of the feature attributes in the samples are shown in Table 1:

<table>
<thead>
<tr>
<th>Feature label</th>
<th>f1</th>
<th>f2</th>
<th>f3</th>
<th>f4</th>
<th>f5</th>
<th>f6</th>
<th>f7</th>
<th>f8</th>
<th>f9</th>
<th>f10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
<td>Five features of the gray level cooccurrence matrix</td>
<td>Pixel value</td>
<td>Mean value</td>
<td>Mean value</td>
<td>Variance</td>
<td>Local entropy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>----------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature label</td>
<td>f11</td>
<td>f12</td>
<td>f13</td>
<td>f14</td>
<td>f15</td>
<td>f16</td>
<td>f17</td>
<td>f18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Hessian eigenvalues</td>
<td>Mel</td>
<td>Tol</td>
<td>Col</td>
<td>Sol</td>
<td>Uo1</td>
<td>FD image</td>
<td>Third-order cumulant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>-------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature label</td>
<td>f19</td>
<td>f20</td>
<td>f21</td>
<td>f22</td>
<td>f23</td>
<td>f24</td>
<td>f25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Seven features of invariant moments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For Im1 and its corresponding index image (Fig. 1 (a) and Fig. 1 (b)), the sample data were generated with the method described above. Twenty-five features and one type of samples were used, and the number of the samples was 12,288. Three evaluators—InfoGainAttributeEval, CfsSubsetEval, and WrapperSubsetEval were selected with WEKA’s automatic attribute selection evaluator function. In particular, C4.5 was selected as the base classifier (with default values as the parameters). The results thus obtained are shown in Table 2.
Table 2. The accuracy of classification for the sample generated from Im1 in different classifiers.

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute sub-set</th>
<th>Classification algorithm</th>
<th>Evaluation strategy</th>
<th>Accuracy of classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Universal set</td>
<td>C4.5</td>
<td>10-fold</td>
<td>79.8177%</td>
</tr>
<tr>
<td>2</td>
<td>f6, f8, f9, f12(CfsSubsetEval)</td>
<td>C4.5</td>
<td>10-fold</td>
<td>79.5980%</td>
</tr>
<tr>
<td>3</td>
<td>f2, f3, f5, f6, f7, f22(WrapperSubsetEval)</td>
<td>C4.5</td>
<td>10-fold</td>
<td>80.5908%</td>
</tr>
</tbody>
</table>

The attribute selection method shows that, for the classifier C4.5, the evaluator WrapperSubsetEval is the most effective for the classification of TIR remote sensing images.

Then we conducted attribute selection with WEKA’s AttributeSelectedClassifier. We set NaiveBayes as the base classifier and chose InfoGainAttributeEval, CfsSubsetEval and WrapperSubsetEval as the evaluator, using ten-fold cross validation. The results are shown in Table 3.

Table 3. Accuracy of classification achieved with different attribute selection algorithms in the NaiveBayes classifier.

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute subset</th>
<th>Classification algorithm</th>
<th>Evaluation strategy</th>
<th>Accuracy of classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>InfoGainAttributeEval</td>
<td>NaiveBayes</td>
<td>10-fold</td>
<td>78.0762%</td>
</tr>
<tr>
<td>2</td>
<td>f6, f8, f9, f12(CfsSubsetEval)</td>
<td>NaiveBayes</td>
<td>10-fold</td>
<td>77.7425%</td>
</tr>
<tr>
<td>3</td>
<td>f3, f4, f6, f7, f8, f20(WrapperSubsetEval)</td>
<td>NaiveBayes</td>
<td>10-fold</td>
<td>76.8880%</td>
</tr>
</tbody>
</table>

The experiment shows that, when WEKA’s AttributeSelectedClassifier is used for attribute selection, the evaluator InfoGainAttributeEval is the most effective for the classifier NaiveBayes. Both NaiveBayes have selected the features f6, f8, f9 and f12 through the evaluator CfsSubsetEval, and both have selected f3, f6 and f7 through WrapperSubsetEval.

Selecting the Best Parameters

We conducted parameter adjustment with WEKA’s classifier CVParameterSelection, first choosing KNN, and then C4.5, as the base classifier. The results are shown in Table 4.

Table 4. The accuracy of classification with different ways of parameter adjustment.

<table>
<thead>
<tr>
<th></th>
<th>KNN</th>
<th>C4.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum parameter</td>
<td>K=10</td>
<td>M=7, C=0.1</td>
</tr>
<tr>
<td>Accuracy of classification</td>
<td>79.1992%</td>
<td>81.193%</td>
</tr>
</tbody>
</table>

A comparison between Table 4 and Table 2 shows that, after parameter optimization, the classifier C4.5 became much more effective.

The classifier CVParameterSelection is not suitable for optimizing the parameters of libSVM, C(cost) and G(gamma), which are not direct parameters of a base classifier. Therefore, we chose the classifier GridSearch for parameter adjustment (XBase=2, Xmin=-3, XMax=8; YBase=2, YMin=-3, YMax=8).
Comparison shows that, with libSVM as the base classifier, the accuracy of classification was considerably increased after parameter adjustment by GridSearch.

**Selecting the Best Classifier**

The following accuracy of classification was obtained for the sample under the conditions of all attributes, optimum classifier parameters and 10-fold cross validation, by the classifiers C4.5, KNN, NaiveBayes and libSVM in the Decision Tree via WEKA’s data mining platform:

<table>
<thead>
<tr>
<th>Classifier</th>
<th>C4.5</th>
<th>KNN</th>
<th>NaiveBayes</th>
<th>libSVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of classification (%)</td>
<td>81.193</td>
<td>79.1992</td>
<td>78.0762</td>
<td>82.0882</td>
</tr>
</tbody>
</table>

The results of the experiment shows that, under similar conditions, NaiveBayes offers the least accuracy of classification whereas libSVM is capable of the best for TIR remote sensing images. In the subsequent experiments, libSVM would be used as the classifier to test the attribute selection method and the optimum classifier parameters.

Next, we used AttributeSelectedClassifier and chose libSVM as the base classifier (C=256, G=1), adopting InfoGainAttributeEval, CfsSubsetEval and WrapperSubsetEval as the evaluator and 10-fold cross validation. The results obtained are shown in Table 7.

Table 7. The accuracy of classification with libSVM’s different attribute selection methods.

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute subset</th>
<th>Classification algorithm</th>
<th>Assessment strategy</th>
<th>Classification accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>f8, f6, f7, f12, f14, f9, f10, f24, f23, f25, f22, f20, f21, f5, f4, f3, f18, f2</td>
<td>libSVM</td>
<td>10-fold</td>
<td>82.4137%</td>
</tr>
<tr>
<td></td>
<td>(InfoGainAttributeEval)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>f6, f8, f9, f12</td>
<td>libSVM</td>
<td>10-fold</td>
<td>80.0618%</td>
</tr>
<tr>
<td></td>
<td>(CfsSubsetEval)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>f2, f4, f6, f7, f8, f10, f12, f14, f18, f20, f22, f23, f24, f25</td>
<td>libSVM</td>
<td>10-fold</td>
<td>82.3324%</td>
</tr>
<tr>
<td></td>
<td>(WrapperSubsetEval)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This experiment shows that, for libSVM, attribute selection with WEKA’s AttributeSelectedClassifier can result in a higher accuracy of classification than when no attribute selection method is used (the accuracy of classification is 82.0882% with all attributes). In particular, InfoGainAttributeEval has proven to be the most effective. The classifiers C4.5,
NaiveBayes and libSVM have all selected the features f6, f8, f9 and f12 via CfsSubsetEval; the features f2, f4, f6, f7, f8, f20, f22 which were selected by C4.5/NaiveBayes, were also chosen via WrapperSubsetEval.

Finally, the attributes selected via WrapperSubsetEval: f2, f4, f6, f7, f8, f10, f12, f14, f18, f20, f22, f23, f24, f25 were subjected to parameter adjustment by the classifier GridSearch: XBase=2, XMin=-3, XMax=8; YBase=2, YMin=-3, YMax=8. The optimized parameters were C=256, G=1, with the accuracy of classification at 82.3324%, in accord with the results of parameter optimization in 4.2.

**Experiment with Nighttime Infrared Image**

For Im2 and its corresponding index image (Fig. 1 (c) and Fig. 1 (d)), sample data were generated with the method described above. We used 25 features and 1 type of sample, the number of samples being 12,768. We conducted parameter adjustment with the classifier GridSearch, and the optimized parameters are: C=4, G=16. Using these parameters (C=4, G=16), we conducted attribute selection with AttributeSelectedClassifier via the evaluator InfoGainAttributeEval/CfsSubsetEval/WrapperSubsetEval and 10-fold cross validation. The results obtained are shown in Table 8.

<table>
<thead>
<tr>
<th>No.</th>
<th>Attribute subset</th>
<th>Classification algorithm</th>
<th>Evaluation strategy</th>
<th>Accuracy of classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>f7, f14, f12, f8, f6, f23, f25, f24, f22, f21, f20, f10, f9, f19, f5, f2, f3, f4, f18</td>
<td>libSVM</td>
<td>10-fold</td>
<td>80.8036%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>f6, f8, f9, f10, f12, f20, f21, f22, f24</td>
<td>libSVM</td>
<td>10-fold</td>
<td>79.8716%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>f4, f5, f6, f7, f8, f9, f14, f18, f19, f20, f21, f22, f24, f25</td>
<td>libSVM</td>
<td>10-fold</td>
<td>80.4276%</td>
</tr>
</tbody>
</table>

According to Table 7 and Table 8, when libSVM is used as the base classifier, attribute selection via InfoGainAttributeEval shows that, for daytime and nighttime TIR images, the features f7 (mean value), f14 (Co1), f12 (Mol), f8 (mid-value) and f6 (pixel value) contribute more to the accuracy of classification; CfsSubsetEval can select f6 (pixel value), f8 (mid-value), f9 (variance) and f12 (Mol) for both daytime and nighttime TIR images; WrapperSubsetEval can select f4 (the fourth feature in the gray level co-occurrence matrix), f6 (pixel value), f7 (mean value), f8 (mid-value), f14 (Co1), f18 (third-order cumulate), f20 (the second feature of invariant moments), f22 (the fourth feature of invariant moments), f24 (the sixth feature of invariant moments), and f25 (the seventh feature of invariant moments) for both daytime and nighttime TIR images. As in the case of daytime TIR remote sensing images, attribute selection with InfoGainAttributeEval offers the highest accuracy of classification. Under the same conditions, such accuracy is also higher for daytime TIR images than nighttime ones because the former are of a higher quality (e.g. with a higher contrast).

**Conclusion**

To improve the inadequate effectiveness of thermal infrared (TIR) image recognition due to low contrast, striping noise and low spatial resolution, we have conducted research with respect to
harbor targets and developed a method for extracting and selecting features of TIR remote sensing images of such targets. We selected sample features like texture, geometry, shape and statistics according to the characteristics of TIR remote sensing images; we chose the best combination of image features with AttributeSelectedClassifier's evaluators—InfoGainAttributeEval, CfsSubsetEval and WrapperSubsetEval; and we conducted parameter adjustment with the classifiers CVParameterSelection and GridSearch, and selected the optimum classifier parameters. Then, applying the classifiers C4.5, KNN, NaiveBayes and libSVM to classification and recognition, we chose libSVM as the best classifier. Finally, we conducted classification experiment for daytime and nighttime TIR images of harbor targets with the classifier libSVM and its optimum parameter and feature selection strategy.

The method described in this paper can easily extract target features from TIR images and obtain a combination of features that can lead to the highest accuracy of classification; besides, when the image is not of a high quality, this method can offer a selection of features with a relatively great robustness. In this paper, with respect to the key issue of feature selection for the recognition of TIR remote sensing images of harbor targets, we experimented with the proposed method and arrived at the following conclusion: the classifier libSVM offers the highest accuracy of classification for TIR images (in this paper, the experimental maximum can reach 82.4137%); parameter adjustment can considerably increase such accuracy; the one-dimensional statistical features of pixel value, mean value, variance, mid-value, and gray level histogram are more meaningful for the recognition of TIR remote sensing images of harbor targets; attribute selection with InfoGainAttributeEval offers the highest accuracy of classification; daytime images can be more accurately classified than nighttime ones under the same conditions. However, the method presented in this paper also requires further research in terms of the following: features are only selected with respect to the proposed feature extraction method; due to the use of the results of manual classification as the sample priors, the accuracy of the sample would depend on that of manual classification and considerably affect the accuracy of classification.

References


