Exo-atmospheric target discrimination using probabilistic neural network

Jianlai Wang (王暕来)[*] and Chunling Yang (杨春玲)

School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China

[*]Corresponding author: wjl820518@sina.com

Received November 26, 2010; accepted February 24, 2011; posted online May 26, 2011

Exo-atmospheric targets are especially difficult to distinguish using currently available techniques, because all target parts follow the same spatial trajectory. The feasibility of distinguishing multiple type components of exo-atmospheric targets is demonstrated by applying the probabilistic neural network. Differences in thermal behavior and time-varying signals of space-objects are analyzed during the selection of features used as inputs of the neural network. A novel multi-colorimetric technology is introduced to measure precisely the temporal evolutional characteristics of temperature and emissivity-area products. To test the effectiveness of the recognition algorithm, the results obtained from a set of synthetic multispectral data set are presented and discussed. These results indicate that the discrimination algorithm can obtain a remarkable success rate.

OCIS codes: 010.0280, 100.4996, 300.6340.

doi: 10.3788/COL201109.070101.

In response to a threat missile, an interceptor missile with a kinetic warhead has been launched with the intention of intercepting and killing the lethal reentry vehicle (RV) in the exo-atmosphere before it reaches its target. Thus, in theater ballistic missile defense (TBMD), it is critical for the interceptor to discriminate the RV from the booster parts and other decoys at distances of hundreds of kilometers. Exo-atmospheric targets are especially difficult to distinguish using currently available techniques because all target parts follow the same trajectory during the exo-atmospheric portion of the flight[1,2]. The infrared (IR) sensor views the threat for about 10 s and starts at about 30 s before intercept. There is a need to use data from the first 5 s of the 10-s window to discriminate the RV. The threat complex is about 100 km away from the IR sensor 30 s before intercept. Within this range, even relatively large pieces are represented as dots on the IR image. In addition, discrimination techniques that use the difference in size as a feature do not work either.

Artificial neural networks (ANNs) currently comprise a hot research area in many real world applications, including image processing[3,4], classification[5], prediction[6], and so on. In recent years, methods using time-delay neural network (TDNN)[7] and its variants, such as adaptive time-delay neural network (ATNN)[8], finite-time impulse response (FIR) TDNN[9], and time-delay recurrent neural network (TDRNN)[10], among others, have also achieved ideal results in exo-atmospheric target discrimination.

Multispectral sensor data show great potential for use in automatic target recognition (ATR) because they provide both temporal and spectral features about the targets; thus, they can be of tremendous aid in discrimination[11,12]. In this letter, we propose an algorithm for target discrimination based on multispectral sensor data, using not only the temporal features but also the spectral characteristics. We use a neural network, which is a variant of the probabilistic neural network (PNN) which is particularly appropriate for pattern classification[13].

We first analyze the thermal behavior of objects in space and their equilibrium temperatures. The equilibrium temperature can be obtained by solving the heat transfer equation, which is given by

$$mc\frac{\mathrm{d}T}{\mathrm{d}t} + Q_\mathrm{E} = Q_\mathrm{A} + Q_\mathrm{I}, \qquad (1)$$

where the mass and heat capacity of the object are denoted by m and c, respectively; Q_E and Q_A are the power emitted and absorbed by the object, respectively.

For an object illuminated by the sun, it is well known that

$$Q_\mathrm{E} = A_\mathrm{S}\varepsilon(\lambda, T)\sigma T^4 \qquad (2)$$

and

$$Q_\mathrm{A} = \alpha(S + S_\mathrm{R})A_\mathrm{C} + \varepsilon(\lambda, T)EA_\mathrm{C}, \qquad (3)$$

where A_S and A_C are the surface and average cross-sectional areas of the object, respectively; $\varepsilon(\lambda, T)$ is the emissivity averaged over the infrared band; σ is the Stefan-Boltzmann constant (5.67×10^{-8} W/(m$^2 \cdot$ K^4)); S is the solar flux (1360 W/m^2); S_R is the solar flux reflected from the Earth (known as the albedo flux and typically taken to be equal to $0.3S$); E is the Earth's IR flux (about 240 W/m^2); α is the absorptivity averaged over the visible and near infrared (NIR) band. By combining Eqs. (1)−(3), we can finally obtain an equation for the equilibrium temperature, T_eq, as

$$T_\mathrm{eq} = \left\{ \frac{A_\mathrm{C}}{A_\mathrm{S}} \left[\frac{\alpha}{\varepsilon(\lambda, T)} \frac{(S + S_\mathrm{R})}{\sigma} + \frac{E}{\sigma} \right] \right\}^{1/4}. \qquad (4)$$

Thus, depending on the surface coating used, objects in space during the daytime can have widely varying equilibrium temperatures. In fact, one can obtain any equilibrium temperature for a sphere ranging from 227 to 540 K using more than one surface coating.

For an object in space in the shadow of the Earth, the

situation is considerably different. Since both the solar flux, S, and the albedo, S_R, of the Earth are zero, in this case the equilibrium temperature can be reduced to

$$T_\text{eq} = \left(\frac{A_\text{C}}{A_\text{S}} \frac{E}{\sigma} \right)^{1/4}. \qquad (5)$$

In practice, an object on a ballistic trajectory in space may or may not reach its equilibrium temperature. How quickly it does so depends on the difference between its initial temperature at release and its equilibrium temperature. It also depends on its emissivity and thermal mass. For a heavy object, such as a warhead, the time to reach equilibrium can be many hours, much longer than the actual time it spends in space. On the other hand, for very light objects, such as balloon decoys, equilibrium could be reached in a matter of minutes.

Furthermore, for an object at an absolute temperature T, the total power $P_\text{S}(\lambda, T)$ received by the sensor in a small bandwidth $\Delta\lambda$ centered at wavelength λ is approximately given by

$$P_\text{S}(\lambda, T) = \frac{\tau(\lambda) A_\text{O} A_\text{C}}{\pi R^2} \varepsilon(\lambda, T) M(\lambda, T) \Delta\lambda, \qquad (6)$$

where $\tau(\lambda)$ is the spectral atmospheric transmissivity, A_O is the receiving area of the sensor optics, R is the range of object to the aperture of a system, and $\varepsilon(\lambda, T)$ and $M(\lambda, T)$ represent the emissivity and spectral radiant exitance for the wavelength λ at temperature T, respectively. According to Planck's formula, $M(\lambda, T)$ can be described with a relation stated as

$$M(\lambda, T) = \frac{c_1}{\lambda^5} (\mathrm{e}^{c_2/\lambda T} - 1)^{-1}, \qquad (7)$$

where c_1 and c_2 are radiation constants given by 3.7418×10^8 W·m^{-2}·μm^4 and 1.4388×10^4 μm·K, respectively.

This is achieved based on the fact that the IR signal from a space-object varies as a function of the projected area observed by the sensor. As each space-object rotates, precesses, and/or tumbles, the projected area seen by the sensor changes, resulting in a signal. This signal is a point of light with a time-changing intensity for each object. Although missile fragment and warhead possess similar temperatures, the variation range of emissivity-area product, $\varepsilon(\lambda, T) A_\text{C}$, of the former can be larger than that of the latter due to its tumble or precession because it is not spin-stabilized.

As analyzed above, the feature selected for discrimination is concluded in Table 1. We use a PNN to identify the targets. The general PNN is a direct neural-network implementation of the Parzen non-parametric probability density function (PDF) estimation and the Bayesian classification rule.

Considering the general case being evaluated in this work, one is faced with the problem of classifying a given d-dimensional input vector (the feature vector) $\mathbf{x}$ in one of K classes $c_1, c_2, \cdots, c_K$. For the "0−1" cost function, the Bayesian classifier leads to the maximum a posteriori (MAP) classifier, i.e.,

$$C(\mathbf{x}) = \arg\ \max p(\mathbf{x}|c_i) P(c_i), \quad i = 1, 2, \cdots, K, \qquad (8)$$

where $C(\mathbf{x})$ is the class of input $\mathbf{x}$ belonging to $\{c_i, i = 1, 2, \cdots, K\}$, and $P(c_i)$ and $p(\mathbf{x}|c_i)$ are the *a priori* probability and a conditional PDF for class c_i, respectively. The key issue for the implementation of this Bayesian classifier is to extract the conditional distributions from the training data set. When a Gaussian kernel is adopted, the Parzen PDF estimator can be represented by

$$p(\mathbf{x}|c_i) = \frac{1}{N_i (2\pi)^{d/2} \sigma^d} \sum_{j=1}^{N_i} \exp\left[-\frac{(\mathbf{x} - \mathbf{x}_i^{(j)})^\text{T} (\mathbf{x} - \mathbf{x}_i^{(j)})}{2\sigma^2} \right], \qquad (9)$$

where N_i is the number of samples in the training set belonging to class c_i, $\mathbf{x}_i^{(j)}$ represents the jth sample belonging to class c_i, and σ is called the smoothing parameter.

PNN is a direct implementation of the above estimator. It consists of three feedforward layers: input layer, pattern layer, and summation layer (Fig. 1).

The input layer works as a distribution mechanism and receives input components from the data set. Therefore, the number of nodes in this layer is equal to the dimension of the input vector. All of these nodes are fully connected with the nodes in the pattern layer, representing all instances of the dataset. The pattern layer consists of K pools of the pattern neurons. In each pool, there are N_i number of pattern neurons, $i = 1, 2, \cdots, K$. Since PNN is applicable to general classification problems, assuming that the feature vector to be classified must belong to one of these known classifications, the absolute probabilistic value of each classification is not important. In addition, only the relative value needs to be considered. Hence, for the input feature vector $\mathbf{x}$, the output of each pattern neuron is given by

$$f(\mathbf{x}; \sigma) = \frac{1}{N_i} \exp\left[-\frac{(\mathbf{x} - \mathbf{x}_i^{(j)})^\text{T} (\mathbf{x} - \mathbf{x}_i^{(j)})}{2\sigma^2} \right], \qquad (10)$$

where $\mathbf{x}_i^{(j)}$ is the weight vector of the jth neuron in the ith pool, and the nonlinear function, $f(\cdot)$, represents the activation function of the neurons. In total, there are K neurons in the summation layer where the ith neuron, $i = 1, 2, \cdots, K$, forms the weighed sum of all the outputs from the ith pool in the pattern layer. The weights of the summation layer are determined by the decision cost

Table 1. Variation Range of the Characteristic Parameters of Objects in Space

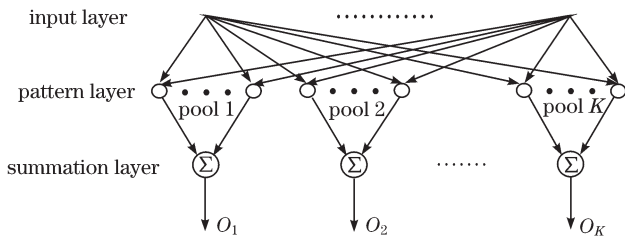| Object Type	Release Temperature	Temperature Variation Range	Emissivity-Area Product Variation Range
Warhead	About 700 K	Slow	Small
Missile Fragment	About 600 K	Fast	Large
Balloon Decoy	About 300 K	Fast	Small

Fig. 1. Structure of the PNN.

function and the *a priori* class distribution. For the "0-1" cost function and the uniform *a priori* class distribution, the weights will be one for all the neurons in the summation layer. For the input pattern **x** of the unknown class, the final decision can be attained through a simple comparison of all the outputs given by

$$\mathbf{x} \in c_k, \quad \text{if } O_k > O_i, \quad i \neq k, \quad i,k \in \{1,2,\cdots,K\}. \tag{11}$$

The emissivities of real objects are functions of wavelength, temperature, and surface condition. Therefore, any method for the radiometric determination of temperature that does not account for emissivity can produce significant errors.

Ratio or two-color pyrometers can circumvent the emissivity measurement issue in certain specific cases. The two-color method uses an approximation of the Planck relation called the Wien radiation relation given as

$$M(\lambda,T) \approx \frac{c_1}{\lambda^5} e^{-c_2/\lambda T}. \tag{12}$$

Assuming $\tau(\lambda_i) = \tau(\lambda_j)$ and $\Delta\lambda_i = \Delta\lambda_j$, the object temperature T can be solved using the radiation ratio of two different wavelengths as follows:

$$\frac{1}{T} = \frac{\ln\frac{P_S(\lambda_i,T)}{P_S(\lambda_j,T)} - 5\ln\frac{\lambda_j}{\lambda_i} - \ln\frac{\varepsilon(\lambda_i,T)}{\varepsilon(\lambda_j,T)}}{c_2\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_i}\right)}. \tag{13}$$

If, in addition, we suppose $\varepsilon(\lambda_i,T) = \varepsilon(\lambda_j,T)$ (i.e., for a gray body), the object temperature T can be approximately estimated by the color temperature $T_C(\lambda_i,\lambda_j)$, which is given by

$$\frac{1}{T_C(\lambda_i,\lambda_j)} = \frac{\ln\frac{P_S(\lambda_i,T)}{P_S(\lambda_j,T)} - 5\ln\frac{\lambda_j}{\lambda_i}}{c_2\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_i}\right)}. \tag{14}$$

The accuracy of this measurement depends not only on the signal-to-noise ratio (SNR), but also on how much variation exists in the emissivity of the target over the spectral ranges being used and the amount of temperature variation over the target.

To further increase the accuracy of the measurement, the emissivity must be modeled better. This can be achieved by measuring the spectral radiance at a larger (> 2) number of wavelengths. By combining Eqs. (13) and (14), we can have

$$\ln\frac{\varepsilon(\lambda_i,T)}{\varepsilon(\lambda_j,T)} - c_2\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_j}\right)\frac{1}{T} = -c_2\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_j}\right)\frac{1}{T_C(\lambda_i,\lambda_j)}. \tag{15}$$

In most cases, emissivity can be adequately represented by a smooth function of wavelength having a number of undetermined parameters. Thus, we can model the natural logarithm transformation of the ratio between $\varepsilon(\lambda_i,T)$ and $\varepsilon(\lambda_j,T)$ as a polynomial function:

$$\ln\frac{\varepsilon(\lambda_i,T)}{\varepsilon(\lambda_j,T)} = \sum_{l=0}^{L-2} a_l(\lambda_i - \lambda_j)^l. \tag{16}$$

It can be easily proven that $a_0 = 0$, so that

$$a_1(\lambda_i - \lambda_j) + \cdots a_{L-2}(\lambda_i - \lambda_j)^{L-2} - c_2\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_j}\right)\frac{1}{T}$$

$$= -c_2\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_j}\right)\frac{1}{T_C(\lambda_i,\lambda_j)}. \tag{17}$$

Here, the unknown parameters are the a_l ($1 \leqslant l \leqslant L-2$) coefficients and the absolute temperature T. For a number of L detection bands, we can have $L(L-1)/2$ radiation ratios. The overdetermined system of equations is inconsistent and generally has no solution. More precisely, we can minimize the square root of the sum of the squares of the error as

$$f_E = \| \mathbf{Ax} - \mathbf{b} \|_2, \tag{18}$$

where

$$\mathbf{A} = \begin{bmatrix} (\lambda_1 - \lambda_2) & \cdots & (\lambda_1 - \lambda_2)^{L-2} & -c_2\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \\ \vdots & \vdots & \vdots & \vdots \\ (\lambda_{L-1} - \lambda_L) & \cdots & (\lambda_{L-1} - \lambda_L)^{L-2} & -c_2\left(\frac{1}{\lambda_{L-1}} - \frac{1}{\lambda_L}\right) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a_1 & \cdots & a_{L-2} & 1/T \end{bmatrix}^T, \tag{19}$$

$$\mathbf{b} = -c_2 \begin{bmatrix} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)\frac{1}{T_C(\lambda_1,\lambda_2)} & \cdots & \left(\frac{1}{\lambda_{L-1}} - \frac{1}{\lambda_L}\right)\frac{1}{T_C(\lambda_{L-1},\lambda_L)} \end{bmatrix}^T.$$

Using least squares estimation approach, the following equation is obtained:

$$\mathbf{x}_{LS} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{20}$$

In this manner, more precise determination of tempera-

ture can be made by reducing the uncertainties due to the unknown emissivity of the object.

Once we estimate the temperature of an object, it is also possible to estimate $\varepsilon(\lambda_i,T)A_C$. Using Eq. (6), we obtain an equation for the emissivity-area product of the

object as a function of measured or estimated values. This is given by

$$\varepsilon(\lambda_i, T)A_{\mathrm{C}} = \frac{\pi R^2 P_{\mathrm{S}}(\lambda_i, T)}{\tau(\lambda_i)A_{\mathrm{O}}M(\lambda_i, T)\Delta\lambda_i}. \quad (21)$$

Properly evaluating the target discrimination algorithm requires multispectral IR data based on target features, such as dynamics, surface emissivities, and temperatures. Unfortunately, these data are not available. Thus, the creation of simulated data based upon certain standard models for IR signal intensities in the bands of interest is necessary. One standard model, referred to as the Bhattacharyya optical decoy evaluation (BODE) model, produces the time series of the IR measurements and includes long-term time trends.

According to a slightly simplified second-order form of the model, the relative measured IR intensity, $J_i(t)$, for the ith waveband is given by

$$J_i(t) = A_i + B_i t + C_i t^2 + S_i \sin(\omega t + \phi) + n_i(t), \quad (22)$$

where A_i indicates an average intensity, B_i is a linear intensity change with time, and C_i is a quadratic intensity change with time; S_i represents the sinusoidal amplitudes; the frequency ω is assumed to be the same for all bands (for a given object) and indicates the coning or precession rate of the object; the phase ϕ is the same for all bands and is not considered physically significant for discrimination; $n_i(t)$ is considered to be zero mean, additive noise. SNR for the ith waveband is defined as the ratio of the square of the modulating sinusoid amplitude to the variance of the additive Gaussian noise. This is given by

$$\mathrm{SNR}_i = \frac{S_i^2}{E[n_i^2(t)]}. \quad (23)$$

The number of spectral bands adopted during the whole test was 12, which was evenly distributed in the range of $4-15$ μm.

According to the BODE model and the parameters in Table 2, the intensity history of three types of space-objects for the ith waveband with $\mathrm{SNR}_i = 1$ was obtained, as shown in Fig. 2.

Thus, in order to reduce the dimensionality of the input vectors, we decided to adopt the temperature and emissivity-area products of the time-varying object as characteristics.

The dynamic temperature of the space-object can be easily extracted from the simulated data using the method proposed above. Figure 3 shows the comparison
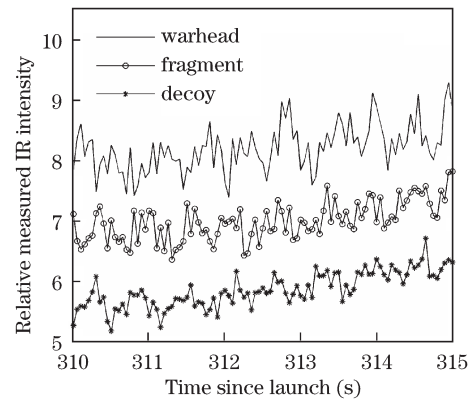

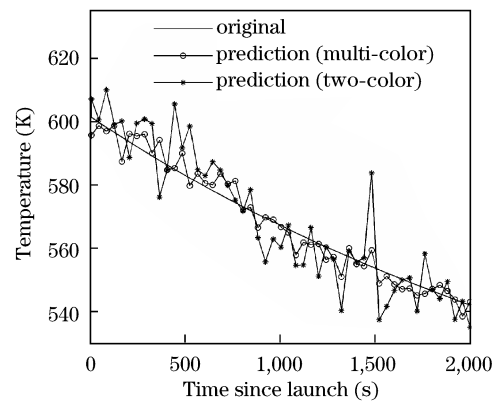Fig. 2. Intensity history for three types of space-objects.

Fig. 3. Comparison between the original and predicted temperatures.
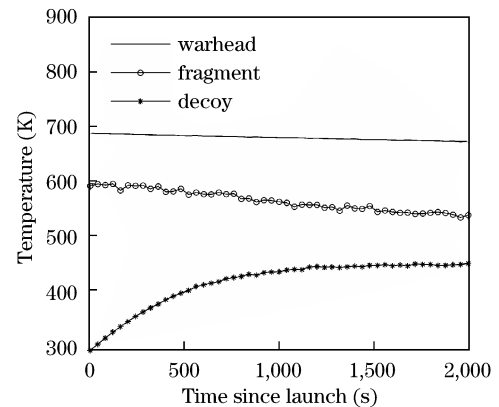

Fig. 4. PDFs corresponding to the temperature for the three types of space-objects.

Table 2. Basic Parameters of Space-Objects

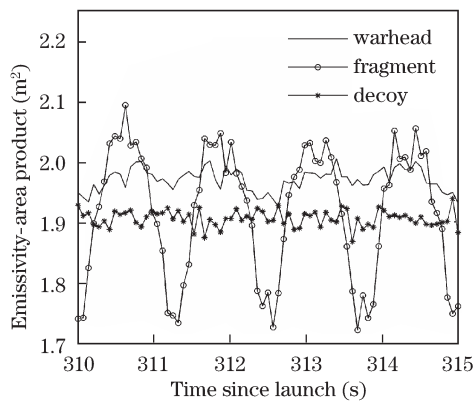| Object Type	Warhead	Missile Fragment	Balloon Decoy I	Balloon Decoy II
Release Temperature (K)	700	600	300	300
Shape	Cone	Cylinder	Sphere	Sphere
Surface Coating	Aluminum	Aluminum	Aluminum	Mylar
Density (g/m^3)	$2.7{\times}10^6$	$2.7{\times}10^6$	$2.7{\times}10^6$	$1.39{\times}10^6$
Specific Heat (J/(g·K))	0.904	0.904	0.904	1.15
Emissivity	0.036	0.036	0.036	0.5
Absorptivity	0.192	0.192	0.192	0.17

Fig. 5. PDFs corresponding to the emissivity-area products of the three types of space-objects.

Table 3. Comparison of Identification Results Between Two Methods

SNR	1			2			3		
Object	W	F	D	W	F	D	W	F	D
$R_{\mathrm{PNN}}(\%)$	73.8	70.1	69.4	87.6	84.1	84.7	99.1	98.4	98.2
$R_{\mathrm{BP}}(\%)$	64.2	61.3	63.2	72.7	71.2	73.4	89.5	86.3	88.4

* R_{PNN} and R_{BP} are average success rates using PNN and BP, respectively; W, F, and D stand for warhead, missile fragment, and balloon decoy, respectively.

between the original and predicted temperatures determined by two different means; the former was obtained by the multi-color method proposed above, and the latter by the conventional two-color pyrometers. It can easily be observed that the prediction error of the former method is smaller. Thus, in this manner, more precise determination of temperature can be made.

In order to appreciate the discriminating power of the available variables, we constructed the PDFs for each of the two dynamic features conditioned on different object types $p(\mathbf{x}_j|c_i)$, where $j = 1-2$ and $i = 1-3$, as shown in Figs. 4 and 5, respectively. As these graphs show, both of these two dynamic features constitute very good discriminants.

The whole data set we used for the present study consisted of 100 signatures for each type of objects. Data were gathered for 5 s in 0.5-s increments (11 time steps) starting, as noted earlier, 30 s before intercept. These were then encoded onto the neural network for each object in the scene. All of these were used in the construction of the PDFs for each of the dynamic features. The comparison of identification results between PNN and back-propagation (BP) under different single-band SNRs are shown in Table 3.

The results confirm that the algorithm proposed in this letter can obtain better performance on target discrimination with low SNR than the conventional BP neural network.

In conclusion, an algorithm for exo-atmospheric target discrimination from a sequence of multispectral IR data has been proposed. The proposed method can be applied in TBMD applications. We use a PNN that, when provided with dynamic characteristics extracted from a multispectral IR data, outputs the probability of an object belonging to any one of a number of possible classes. The PNN is superior to others that simply provide a yes or no decision or even some "confidence index", although its relation to the actual probability is unknown. Indeed, the information supplied by the PNN is based on Bayes theory, which provides an optimal and sound mathematical basis for decision-making. Another major advantage of PNNs is their short training time. They are ready to perform their classification task after being shown, only once, a set of exemplars of objects of the various classes to be learned. Results demonstrate that the trained neural networks are able to successfully identify warheads from other missile parts and decoys.

References

1. T. A. Postol, "Explanation of why the sensor in the exoatmospheric kill vehicle (EKV) cannot reliably discriminate decoys from warheads", http://fas.org/spp/starwars/program/news00/postol_atta.pdf (May 11, 2000).
2. T. A. Postol, "Technical discussion of the misinterpreted results of the IFA-1A experiment due to tampering with the data and analysis and errors in the interpretation of the data", http://fas.org/spp/starwars/program/news00/postol_attb.pdf (May 11, 2000).
3. H. Zhang, Kybernetes **39**, 1366 (2010).
4. X. Huang, R. Zhang, and J. Pang, Chin. Opt. Lett. **7**, 1004 (2009).
5. J. Wang, G. Chen, T. Zhu, S. Gao, B. Wei, and L. Bi, Chin. Opt. Lett. **7**, 1058 (2009).
6. R. M. Balabin and E. I. Lomakina, J. Chem. Phys. **131**, 074104 (2009).
7. C. L. Resch, Johns Hopkins APL Technical Digest **19**, 315 (1998).
8. D.-T. Lin, J. Inform. Sci. Eng. **20**, 959 (2004).
9. B. Zhang, "Point target detection and recognition algorithms in optical image terminal homing system" (in Chinese) PhD. Thesis (National University of Defense Technology, Changsha, 2005).
10. T. Liu, J. Lv, S. Wang, and H. Lu, J. Electron. Inform. Technol. (in Chinese) **32**, 80 (2010).
11. L. Cao, C. Wu, Q. Hou, and W. Zhang, Opt. Tech. (in Chinese) **36**, 145 (2010).
12. S. Li, B. Zhang, L. Gao, and M. Peng, Acta Opt. Sin. (in Chinese) **30**, 2116 (2010).