A shadow detection of remote sensing images based on statistical texture features

XIA Huaiying, GUO Ping

Image Processing & Pattern Recognition Laboratory, *Beijing Normal University*, *Beijing* 100875, *China*

Abstract: Shadow detection for high spatial resolution remote sensing images is very critical for image segmentation, feature extraction, image matching, automatic target detection and target location. In order to improve the accuracy of shadow detection, we propose a new shadow detection method based on a statistical mixture model, which combines several radial basis function neural networks. Four statistical features, including energy, entropy, contrast and inverse difference moment, extracted from grey level concurrence matrix are used as the model input features. EM-like algorithm is adopted to estimate the model parameters through optimizing the system cost function. Comparative experiments are performed between the Gaussian background model and the histogram threshold method. Experimental results show that higher detection accuracy of the proposed approach is obtained. The proposed method can solve the problem such as high reflective regions and false alarms in the presence of water, as well as the repeated threshold calculation.

Key words: Shadow detection, radial basis function neural network, mixture model, statistical texture feature.

CLC number: TP 751 **Document code:** A

Citation format: Xia H Y and Guo P. 2011. A shadow detection of remote sensing images based on statistical texture features. *Journal of Remote Sensing*, **15**(4): 778–791

1 INTRODUCTION

The shadow on the remote sensing images is a dark area with little information because the solar rays are blocked by tall buildings, trees, mountains and other tall objects. On one hand, the existence of shadow has a negative influence on image processing operations, such as feature extraction, image registration, image retrieval (Bao & Guo, 2006), and image segmentation. Specifically, the existence of shadow has a deep impact on the accuracy of image classification, target detection, target location, which lead these operations become impossible. On the other hand, the shadow region in remote sensing images may provide more about three-dimensional information of objects than the surrounding background, such as buildings. Therefore, it becomes necessary for preprocessing shadow areas. While shadow detection is a key step in the shadow processing, it plays an important role in the follow-up processing operations for remote sensing images.

Existing shadow detection methods for remote sensing images can be divided into two main categories: one is based on model (Finlayson, *et al.*, 2002) which employs *a priori* knowledge of the illumination, DSM data and the 3D geometry of the scene to calculate positions of shadows (Yao & Zhang, 2006; Gils, 2001). However, this method has rarely been used because the prior knowledge is not always available. The other one is based on property which is generally used by analyzing the difference of lightness, geometry structure and color character between shadow regions and non-shadow regions (Gwinner & Schaale, 1997; Wang & Wang 2002; Tsai, 2006). Shadow properties include spectroscopic properties, texture properties and marginal. For example, Jiang and Ward (1992) propose a detection method, mainly based on lower intensity of shadow region than non-shadow region while geometric properties are used in Funka-Lea G and Bajcsy (1995). Other properties, such as histogram threshold, homomorphic filtering and color invariance in shadow region were also reported (Highnam & Brady, 1997; Elemadmia, *et al.,* 2003; Salvador, *et al.,* 2004). With the theoretical foundation of those methods, researchers in China have proposed a number of different detection methods, such as shadow detection of color image based on RGB color space (Wang, *et al.*, 2002), total variation model that based on K-L transform (Wang, *et al.,* 2004). Xu and Xu(2003) proposed automated extraction of shadows using very-high reso-

Received: 2010-05-07; **Accepted:** 2010-08-19

Foundation: National Nature Science Foundation China (No. 90820010, No. 60911130513)

First author biography: XIA Huaiying (1984—), female, master student, she graduated from Beijing Normal University. Her main research field is remote sensing imagery. 2 papers published and indexed by EI. E-mail: xia_huaiying@163.com

Corresponding author: GUO Ping (1957—), male, professor, IEEE senior member. His research interests are computational intelligence, pattern recognition, image processing, software reliability engineering, and spectra analysis. He has published more than 200 academic papers. E-mail: pguo@ieee.org

lution spatial satellite images based on normalized processing and texture analysis. Guo, *et al*. (2006) proposed multispectral detection method; Xia, *et al*. (2009) proposed a detection method based on Affinity Propagation algorithm. Ji and Yuan (2007) proposed hange detection of man-made objects; Bao, *et al*. (2010) proposed a detection method based on green component in RGB color space and intensity component in HSI color space.

The above analyzed methods play an important role in certain fields of applications. However, there is no common detection method for all remote sensing images yet, because of the complexity of the formation mechanism of the shadows. For examples, the detection method based on the property of color invariance has more restrictive conditions on the scene, which supposes that shadow region is flat surface texture with single light. Detection method based on threshold segmentation has the problem of high reflective regions and false alarms in the presence of water, as well as repeated threshold calculation. Furthermore, it is only for suitable for grayscale images. Detection methods based on homomorphic filtering may mistakenly detect useful information as shadow regions, for most of the information concentrated in the low frequency, as well as the parameters of low-pass filter need be decided by experience. Besides, high complexity of the algorithms of conversion between spatial and frequency may be resource intensive. Many other detection methods also have the problem of subjectivity of threshold selection and repeated experiment verifications.

In perspective of recognition research, shadow detection can be treated as a problem of image classification and target recognition, which is to partition an image into several regions. The clustering algorithm is most widely used algorithm in pattern recognition, and it can be applied to shadow detection. The goal of clustering is to group image pixels together that exhibits some type of similarities such as color, texture, or brightness to form shadowed regions and no-shadowed regions. The clustering algorithm is based on pixel similarity, and the locations of boundaries between regions come naturally to the human observer. In order to improve the accuracy of shadow detection, we propose a new shadow detection method on considering a statistical mixture model which combines several radial basis function neural network (RBFNN). Experimental results show that higher detection accuracy of the proposed method is obtained. And the proposed method can solve the problem such as high reflective regions and false alarms in the presence of water, as well as the repeated threshold calculation.

2 BACKGROUND KNOWLEDGE

2.1 Gaussian RBF neural network

RBF neural network is an important artificial neural network. Due to its nonlinear processing capabilities, simple structure and fast training speed, it has become a widely used network model. The architecture of RBFNN is shown in Fig. 1, and the mathematical description is as follows.

$$g(\boldsymbol{X}, \boldsymbol{W}) = \sum_{i=1}^{K_j} w_i \phi(\|\boldsymbol{X} - \mu_i\|) + w_0 \tag{1}$$

where $\boldsymbol{X}$ is the input of the network, g($\boldsymbol{X}$, $\boldsymbol{W}$) is the output of the network and $\boldsymbol{W}$=(w_i|i=1, 2, ..., K_j) is the weight, μ_i is the centre value of the radial basis function, w_0 is a bias constant, $\phi(\|x - \mu_i\|)$ is the radial basis function. When radial basis function adopts Gaussian function, RBF neural network can be expressed as following:

$$g(\boldsymbol{X}, \boldsymbol{W}) = \sum_{i=1}^{K_j} w_i \exp\left[-\frac{\|\boldsymbol{X} - \mu_i\|^2}{2\gamma_i^2}\right] + w_0 \tag{2}$$

The training of the RBF neural network includes identification of the center μ of the Gaussian kernel function, the variance γ and the weights $\boldsymbol{W}$ of the output.

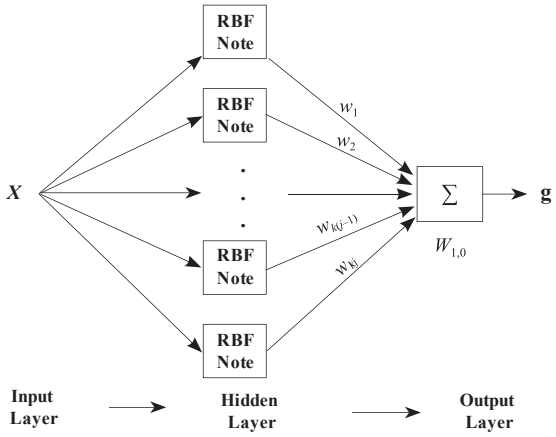


Fig. 1 Architecture of the RBFNN

2.2 Mixture model of RBF neural network

Given the data sets $D = \{\boldsymbol{x}_i\}_{i=1}^N$ to be classified are assumed to be samples from a mixture of H Gaussian densities with joint probability density of which the mathematical expressions are shown as

$$p(\boldsymbol{Z} \mid \boldsymbol{X}, \Theta) = \sum_{j=1}^{H} P(j) p(\boldsymbol{Z} \mid \boldsymbol{X}, \Theta_j)$$
$$\text{with } p(j) \geqslant 0 \text{ and } \sum_{j=1}^{H} p(j) = 1 \tag{3}$$

where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the input vector and output vector in feature space, and the dimension are d_x and d_z respectively. $\boldsymbol{\Theta}$ represents the parameter group. H is the number of component in the mixed model. $P(j)$ is the parameter of mixture weight, and each component is composed of K_j radial basis functions. Fig. 2 shows the architecture of the mixture model.

Assume each component is in the form of Gaussian function,

$$p(Z \mid X, \Theta_j) = G(z, g_j(X, W), \sigma_j^2)$$
$$= \frac{1}{(2\pi\sigma_j^2)^{d_z/2}} \times \exp\{-\frac{1}{2\sigma_j^2}\|z - g_j(X, W)\|^2\} \tag{4}$$

where $gj(\boldsymbol{X},\boldsymbol{W})$ is considered as radial basis function (RBF) neural network.

The cost function of the system, which also called the error function, can be written as

$$E = -l(\boldsymbol{\Theta}) = -\sum_{i=1}^{N} \ln p(z_i \mid x_i, \boldsymbol{\Theta})$$
$$= -\sum_{i=1}^{N} \ln \sum_{j=1}^{H} p(j) p(z_i \mid x_i, \boldsymbol{\Theta}_j) \tag{5}$$

where $l(\boldsymbol{\Theta})$ is the log-likelihood function of the system, $\boldsymbol{\Theta}$={σ, $\boldsymbol{W}$, γ, μ}.

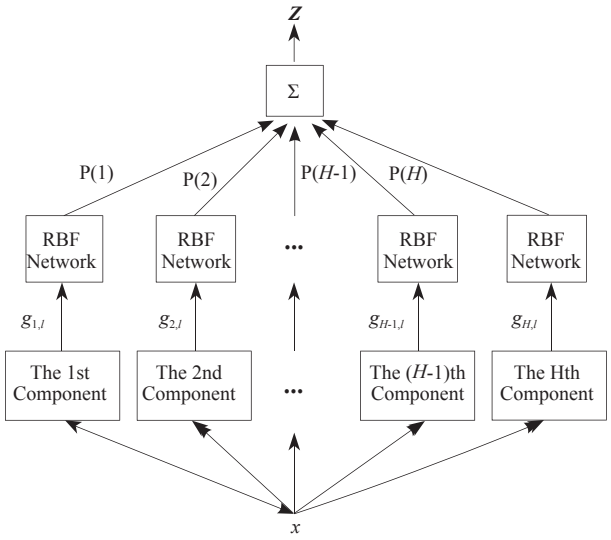


Fig. 2 Architecture of Mixture RBFNN

2.3 EM-LIKE learning algorithm

Given the learning samples, the parameters in mixed RBFNN classifier can be estimated by maximizing the log-likelihood function, which can be described as the following steps (Guo, *et al.*, 2003):

E-Step: Fix up the parameter $P(j)^{\text{old}}$ and Θ^{old}, and compute the variable $h(j, n)$ according to the following equation:

$$h(j,n)=\frac{P(j)^{\text{old}}G(z_n,g_j(x_n,W^{\text{old}}),(\sigma_j^2)^{\text{old}})}{\sum_{i=1}^{H}P(i)^{\text{old}}G(z_n,g_i(x_n,W^{\text{old}}),(\sigma_i^2)^{\text{old}}))} \tag{6}$$

M-Step: Compute the new estimates of the parameters $P(j)^{\text{new}}$ and Θ^{new} according to the following formulae:

$$\begin{aligned} P(j)^{\text{new}} &= \frac{1}{N}\sum_{n=1}^{N}h(j,n) \\ \Theta_j^{\text{new}} &= \Theta_j^{\text{old}}+\eta\nabla l(\Theta)\big|_{\Theta=\Theta^{\text{old}}} \end{aligned} \tag{7}$$

Apply the EM-like algorithm above to estimate parameters, and the expected output of the mixture model can be summarized as Eq. (8):

$$\begin{aligned} \langle \boldsymbol{Z}\rangle &= \int \boldsymbol{Z}p(Z\mid \boldsymbol{X},\boldsymbol{\Theta})\mathrm{d}\boldsymbol{Z} \\ &= \int \boldsymbol{Z}\sum_{j=1}^{H}P(j)G(\boldsymbol{Z},g_j(\boldsymbol{X},W_j),\sigma_j^2)\mathrm{d}\boldsymbol{Z} \\ &= \sum_{j=1}^{H}P(j)g_j(\boldsymbol{X}_j,\boldsymbol{W}_j) \end{aligned} \tag{8}$$

2.4 Texture feature

Texture feature of remote sensing image represents the shape, homogeneity, orientation and intensity between different categories of landmarks and spatial relations. Gray Level Co-occurrence Matrix (GLCM), is one of the most known texture analysis methods to estimate image properties related to second-order statistics. There are four statistical features commonly used, including Energy, Entropy, Contrast and Inverse Difference Moment (Baraldi & Parmiggiani, 1995).

Energy is also called Angular Second Moment and uniformity is a measure of textural uniformity of an image. Energy reaches its highest value when gray level distribution has either a constant or a periodic form.

$$\text{Energy}=\sum_{i=1}^{N}\sum_{j=1}^{N}P_d^2(i,j) \tag{9}$$

Entropy measures the disorder of an image and it achieves its largest value when all elements in GLCM are equal. When the image is not texturally uniform, many GLCM elements have a very small value which implies that entropy is very large. Therefore, entropy is inversely proportional to GLCM energy.

$$\text{Entropy}=-\sum_{i=1}^{N}\sum_{j=1}^{N}P_d(i,j)\log P_d(i,j) \tag{10}$$

Contrast reflects the clarity and the depth extent of grooves. The deeper of texture grooves, the greater of the contrast, and the clearer of the visual effects. On the contrary, the smaller of the contrast, the shallower of the grooves, and the more fuzzy of the effect.

$$\text{Contrast}=\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2P_d(i,j) \tag{11}$$

Inverse difference moment (IDM) measures image homogeneity. This parameter achieves its largest value when most of the occurrences in GLCM are concentrated near the main diagonal. IDM is inversely proportional to GLCM.

$$\text{IDM}=\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{P_d(i,j)}{|i-j|^2} \tag{12}$$

3 SHADOW DETECTION

3.1 Analysis of principles

3.1.1 Analysis of the texture feature

Through extensive experiment studies, it can be concluded that entropy and contrast statistical features of shadow region decrease gradually while region of non shadow has a fluctuated response. Intensity of shadow region is very close and has little changes, thus the energy of the shadow regions is higher than that of non-shadow regions. Fig. 3 shows one of the remote sensing images which have many shadows. Fig.4 and Fig.5 show the comparative analysis of four statistical characteristics between shadow regions and non-shadow regions of Fig. 3.

At present, many shadow detection methods are based on the spectral characteristics of the shadow, setting one or more thresholds using the lower intensity of shadow regions. However, it is difficult to distinguish similar targets such as water bodies, green grasses, and the blue roof as well as the process of repeated threshold calculation. Through the analysis between shadow and non-shaded areas on the texture analysis, we conclude that it will be easier to separate by using the feature of the entropy, energy, contrast and inverse difference moment.

Fig. 3 Orignal remote sensing image

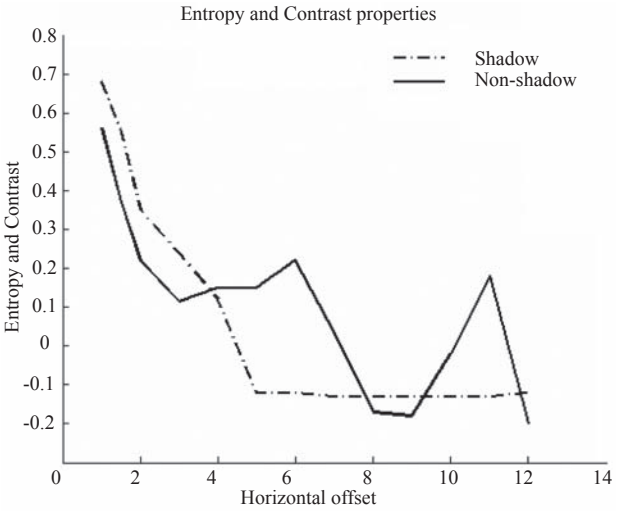


Fig. 4 Analysis of entropy and contrast between shadow and non-shadow region

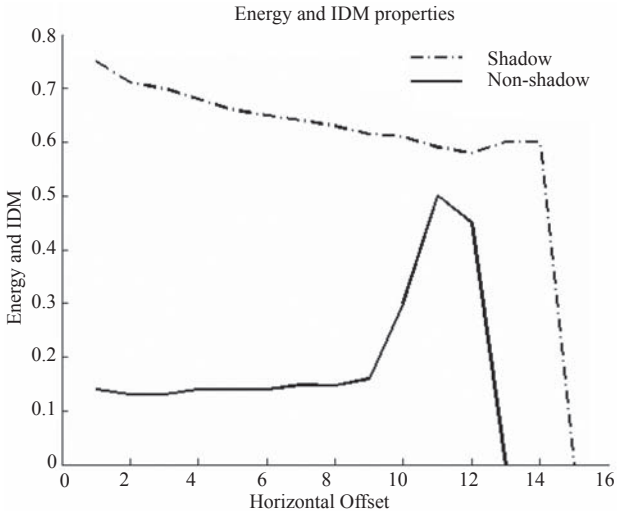


Fig. 5 Analysis of energy and IDM between shadow and non-shadow region

3.1.2 The Statistical mixture model

The basic idea of Gaussian background model is to regard the natural landscape and artificial surface features as background (Ji, *et al.*, 2007) and detect shadow region by considering them to be consistent with Gaussian distribution, which is not the property of the shaded area. However, a simple Gaussian model cannot express the natural background and artificial surface features appropriately. Furthermore, high-resolution remote sensing image is different from the nature image which contains a lot of surface features, and the traditional Gaussian model cannot describe high resolution remote sensing images well. In theory, multi-component compositions of the mixture model can describe the complicated distribution as well as high scalability. By adjusting the parameters of the mixture model, the mixture model can meet the requirements to describe the actual high-resolution remote sensing images. As the RBF network has the strong nonlinear and localized capabilities, it further enhances the remote sensing image classification and shadow detection accuracy.

3.2 Construction of mixture model

The structure of the mixture RBFNN we designed is shown in Fig. 2, in which RBF network is adopted as the component of the mixture model.

(1) Input of the mixture model: Four statistical features of an image were extracted as input vector: $\boldsymbol{X}=\{\boldsymbol{x}_{\text{Energy}}, \boldsymbol{x}_{\text{Entropy}}, \boldsymbol{x}_{\text{Contrast}}, \boldsymbol{x}_{\text{IDL}}\}$. When calculating GLCM and the texture feature, we decomposed the image into a set of small pixel region. The size of the set is very important to the accuracy of the classification. We tried to set the sizes of window to 8×8, 6×6, 4×4 and 2×2 respectively, and compared the accuracy of classification of test samples. As the process based on 4×4 window has a higher computing precision and a faster computing speed, the size of 4×4 window finally was adopted to extract features. When use different directions, it has little effect. In our experiment, 45 degree direction was used.

(2) Expected output: The expected output of the hybrid model includes shadow regions, border regions and non-shadow regions. We use one-of-*K* encoding method: $\boldsymbol{Z}_1=(1,0,0)^{\text{T}}$ represents for class one that belongs to shadow region,; $\boldsymbol{Z}_2=(0,1,0)^{\text{T}}$ represents for class two that belongs to border region; and $\boldsymbol{Z}_3=(0,0,1)^{\text{T}}$ represents for class three that belongs to non-shadow region.

(3) Actual output: Actual output is a probability vector $\boldsymbol{P}=\{p_1, p_2, p_3\}$, p_i denotes the probability that $\boldsymbol{X}$ belongs to class i (i=1, 2, 3). If p_j is the largest in $\boldsymbol{P}$, $\boldsymbol{X}$ will be divided to class j.

3.3 Experimental procedure

3.3.1 Training Sample Selection

The training data set used in our experiments was downloaded from Google Earth without particular selection. We selected three training classes as following, C1: shadow, C2: The junctione of shadow and non shadow, C3: non-shadow, and there are fifteen selected traning images for each class.

3.3.2 Parameter Initialization

As discussed above, eight parameters are included in the experiment: $\{H, K_j, P(j), \sigma, \boldsymbol{W}, \gamma, \mu, \eta\}$. The initial value of these parameters can have great effect on the accuracy of classification.

(1) The number of component H in the mixed model is set to 5, that is H=5;

(2) The node number K_j of the middle layer for RBF neural network are 11, 12, 13, 14, 15, respectively, that is K_1=11, K_2=12, K_3=13, K_4=14, K_5=15;

(3) The initial probabilities $P(j)$ of the five components are set to the average value 0.2, $P(j)$=1/H (j=1, 2, …, 5);

(4) The variance σ_j^2 is set to 9 as the initialization value which is an experience value;

(5) The initial RBF weight vector is set as expression (13):

$$w(j,i,l)=0.25+0.75\times Z(j,l) \tag{13}$$

j=1,…, H, i=1,…, K_j, l=1,…, d_z, $\boldsymbol{Z}(j,l)$ is a column vector;

(6) The mean μ of each RBF network is calculated by K-means clustering algorithm, and the specific process can be described as follows:

step1 The mean of j^{th} RBF network $\mu_{j,i}$ is selected randomly;

step2 By the minimum Euclidean distance principle, x_n will be assigned to the center $\mu_{j,i}$;

step3 Calculate the mean of class i and update the cluster center $\mu_{j,i}$;

step4 Loop iteration, until all cluster centers no longer change, and then acquire all of the initial mean $\mu_{j,i}$, (j=1,⋯, H, i=1,⋯, K_j);

(7) The variance of j^{th} RBF network is calculated as:

$$(\gamma_j^2)_i = \left| x_{j,i} - \mu_{j,i} \right|^2 \tag{14}$$

(8) The learning factor η is an import parameter in the model. If being poorly set, the EM algorithm may not get a convergence result, or the result may locate in a local minimum or maximum point. With experiments, we find that for the multi-source data, the learning factor should be a relatively small value. Otherwise, the algorithm may not converge. In our experiments, we have tried several values and find that the best classification result can be obtained when η =0.5.

3.3.3 Description of Training process

The EM-like learning algorithm is applied to train the RBFNN model. The detailed steps are shown in subsection 2.3. In our experiments the training procedure terminates and reaches the extremes just after several loops and this takes less than a few seconds, which can be described in Fig. 6. Then we can obtain the parameters Θ={ $\boldsymbol{W}$, σ^2, γ, μ}of mixture model.

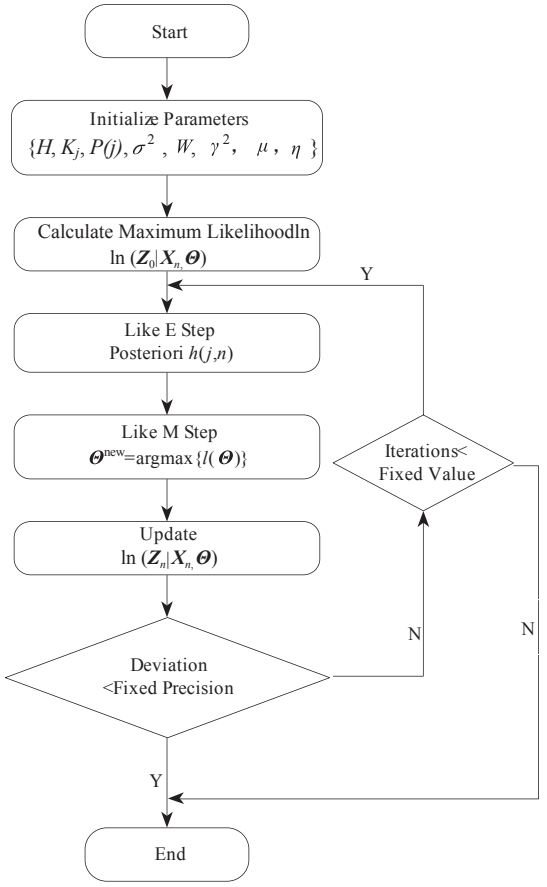


Fig. 6 Trainning flowchart

4 EXPERIMENTAL ANALYSIS

We run a series of experiments of shadow detection based on the proposed method. The computer configuration is listed as following: Memory 2 G, CPU Core II 1.8 G. The remote sensing images with heavy shadows for experiments are downloaded from Google Earth, and the sizes of those images are not changed. In this study, only detection results of three images are listed. Fig. 7(a), Fig. 8(a) and Fig. 9(a) are examples of three images with heavy shadows, respectively. We denote them as Img1 (200×200), Img2 (230×200) and Img3 (230×150). It can be seen that the original remote sensing image has heavy shadowed buildings, high reflective regions, and dark green lawns. In order to observe the effect of detection and compensate the information of shadow regions later, the pixel with grey value of 255 will represent shadow regions, and the pixel value which is original represents that the pixel belongs to no-shadowed region. We also compare them with the Gaussian background and the histogram threshold method. Subfigure (b) in Fig. 7, Fig. 8 and Fig. 9 are the results of applying SMM-RBFNN method respectively. Subfigure (c) shows the results using the Gaussian background method and subfigure (d) is the detection result using the histogram threshold method.

For a more objective analysis of test results, suggested from the analytical methods used in the literature (Martel-Brinson N, *et al.*, 2005), we mainly analyze three types of data TN, FN, FS, where TN expresses the number of correct recognition of the shadow region, FN expresses the number that mistakenly identify the non-shadow area as shadow area, and FS expresses the number of missed the shadow region. Two indicators for evaluating the detection accuracy are calculated, one is the correct detection rate DR, the other is the missed rate FR. DR and FR are defined as:

$$\begin{aligned} \mathrm{DR} &= \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}} \times 100\% \\ \mathrm{FR} &= \frac{\mathrm{FS}}{\mathrm{TN}+\mathrm{FN}+\mathrm{FS}} \times 100\% \end{aligned} \tag{15}$$

A good shadow detection method should satisfy the DR as large as possible, while keep the FR as small as possible.

We calculate the two indices for the SMM-RBFNN method, the Gaussian background model method and the histogram threshold method according to Eq. (15), which are illustrated in Table 1.

We can draw the following conclusions according to results shown in Table 1:

(1) Histogram threshold has the problem of false dismissals, and it is easy to detect dark roads, blue water which also exhibits higher hue and lower luminance as the same as shadow regions as shadow region;

(2) The accuracy of the Gaussian background method is better than that of the histogram method. However, it has the problem of dismissals that to detect similar areas such as dark roads, blue water and green vegetation as shadow areas;

(3) The average DR of the proposed method is 92% and FR is about 7%. The comparison results show that the proposed method can distinguish dark objects from shadows, and the shape of the segmented shadows is preserved well. It has a higher detection precision than that of Gaussian background method and histogram threshold method.

Threshold method is mainly suitable for gray images. Meanwhile, there is no universal method to determine the threshold for all images. Besides, repeated experiments show that the proposed method has better stability than the Gaussian background method.

In contrast to the Gaussian background model, which has the problem of dismissals, we observe that the proposed method can solve the problem of false dismissals, and it improves the accuracy of shadow detection. Thus, SMM-RBFNN test method shows better detection accuracy than the Gaussian background method and the histogram threshold. As long as to set a good initial parameter

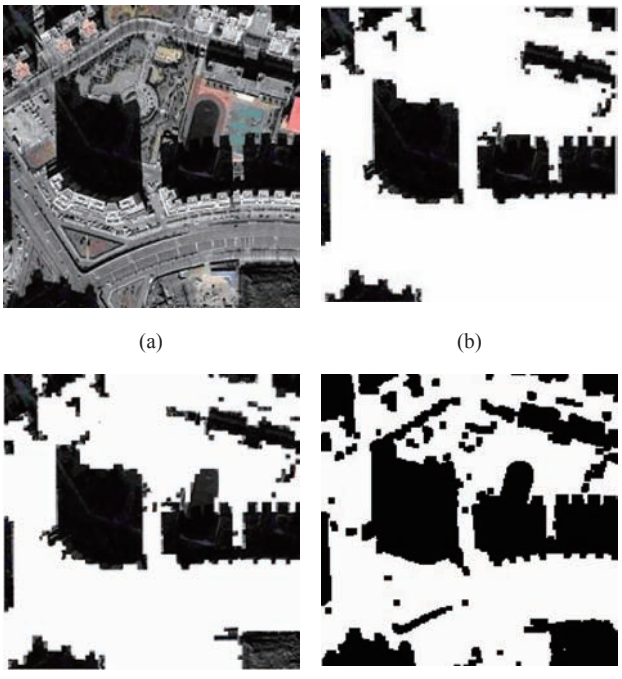
values, the detection process can be run automatically to achieve a high standard of automated inspection.

Meanwhile, there is no universal method to determine threshold for all images when use the histogram threshold. Furthermore, histogram threshold needs to remove the isolated points by using mathematical morphology.

The reason that mixture model is able to obtain high detection accuracy is incorporation of the mixture model to describe high resolution remote sensing image. Therefore, it is more suitable to keep in accordance with actual distribution, as well as the good scalability of the RBF network.

Table 1 Analysis of two indices

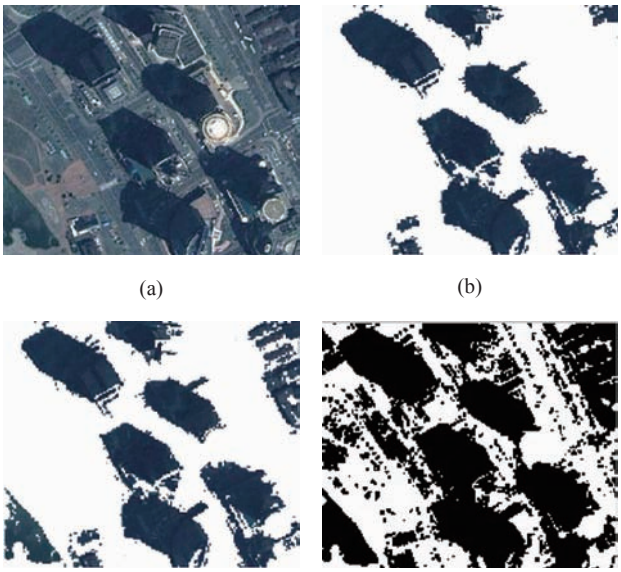
Method	Image	DR	FR	Test Result
SMM-RBFNN	*Img*1	92.86%	6.67%	Very Good
	*Img*2	91.67%	7.69%	Very Good
	*Img*3	92.62%	6.78%	Very Good
Gaussian Background	*Img*1	81.25%	11.11%	General
	*Img*2	83.33%	14.28%	General
	*Img*3	88.27%	10.74%	General
Histogram Threshold	*Img*1	68.42%	9.52%	Bad
	*Img*2	75.00%	14.28%	Bad
	*Img*3	74.17%	13.27%	Bad



(a) (b)

(c) (d)

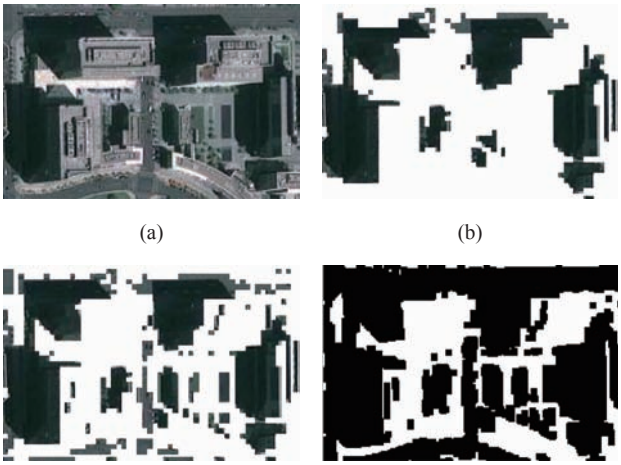
Fig. 7 Img1 detection results
(a) Orignal image, (b) Result of SMM-RBFNN,
(c) Result of Gaussian background, (d) Histogram threshold



(a) (b)

(c) (d)

Fig. 8 Img2 detection results
(a) Orignal image, (b)Result of SMM-RBFNN,
(c) Result of Gaussian background, (d) Histogram threshold



(a) (b)

(c) (d)

Fig. 9 Img3 detection results.
(a) Orignal image, (b)Result of SMM-RBFNN,
(c)Result of Gaussian background, (d) Histogram threshold

In principle, by adjusting the mixture model parameters (number of mixture components H, the number of radial basis function K_j in RBF network, K_j generally depends on the number of training samples N, usually take $K_j=\{2/3\times N+j\}(j=1,\cdots, H)$ that can satisfy the requirements of specific problems, which improves the output accuracy of combination of the network model.

5 CONCLUSION

In this paper, a statistical mixed model based on radial basis function neural network is applied in detection of shadows of remote sensing images. By analyzing the properties of shadow and non-shadow areas, four statistical features, including energy, entropy, contrast and inverse difference moment, extracted from grey level concurrence matrix are used as the model input features. EM-

like algorithm is adopted to estimate the model parameters through optimizing the system cost function. The comparison results show that the proposed method can distinguish dark objects from shadows, and the shape of the segmented shadows is preserved well. It has a higher detection precision than that of the Gaussian background and the histogram threshold method. Mixture model is more suitable to express the actual distribution of the complex than the simple Gaussian model. Compared with the Gaussian background and the traditional histogram threshold method, mixture model shows higher shadow detection accuracy. At the same time, this method does not require assumptions, it leads much broader applications.

In this paper, we use four texture features as input vector and it works well for shadow detection. The combination of texture and spectral features is our next research goal. Besides, the number of RBF network K_j, and the learning factor of EM-Like algorithm both have an important impact on detection accuracy. Thus, in the future work, we will focus on how to find a more efficient initialization method instead of selecting initial values by experience.

Acknowledgements: The authors would like to express their thanks for Mr. Hu Rukun's help in most of experiment work. Prof. Guo is the author to whom all correspondence should be addressed.

REFERENCES

Bao H Y, Li Y and Yin Y Y. 2010. The study on shadow detection and shadow elimination in the urban aerial image. *Remote Sensing Information*, (1): 44–47

Bao Q and Guo P. 2006. Comparative studies on similarity measures for remote sensing image retrieval based on histogram. *Journal of Remote Sensing*, **10**(6): 893–900

Baraldi A and Parmiggiani F. 1995. An investigation of the textural characteristics associated with gray Level concurrence matrix statistical parameters. *IEEE Transaction on Geo-science and Remote Sensing*, **33**(2): 293–304

Etemadnia H and Alsharif M R. 2003. Automatic image shadow identification using LPF in homomorphic processing system. Proceedings of VII Digital Image Computing: Techniques and Applications, Sydney: Committee of DICTA: 429–438

Finlayson G D, Hordley S D and Drew M S. 2002. Removing shadows from images. *Lecture Notes in Computer Science*, **2535**: 129–132

Funka-Lea G and Bajcsy R. 1995. Combining color and geometry for the active, visual recognition of shadows. Proceedings of fifth International Donference on Computer Vision: 203–209

Gils P T. 2001. Remote sensing and cast shadows in mountainous terrain. *Photogrammetric Engineering and Remote Sensing*, **67**(7): 833–839

Guo J H, Tian Q J and Wu J Z. 2006. Study on multispectral detection shadow areas and a theoretical model of removing shadows from remote sensing images. *Journal of Remote Sensing*, 10(2): 151–159

Guo P, Qin D M, Hu Z Y and Du W M. 2003. A novel method for spectral signal pattern recognition. *Spectroscopy and Spectral Analysis*, **23**(4): 811–815

Gwinner K and Schaale M. 1997. A case study on the influence of shadows and shading on multispectral airborne imaging data. Presented at the third International Airbome Remote Sensing Conference and Exhibiton. Copenhagen, Demark

Highnam R and Brady M. 1997. Model-based image enhancement of far infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(4): 410-415

Jiang C and Ward M O. 1992. Shadow identification. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition: 606–612

Ji S P and Yuan X X. 2007. A method for shadow detection and change detection of man-made objects. *Journal of Remote Sensing*, **11**(3): 323–329

Martel-Brinson N and Saccharine A. 2005. Moving cast shadow detection from a Gaussian mixture shadow model. IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 643–648

Musavi M T, Ahmed W, Chan K H, Faris K B and Hummels D M. 1992. On the training of radial basis function classifiers. *Neural Networks*, **5**(4): 595–603

Salvador E, Cavallaro A and Ebrahimi T. 2004. Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding*, **95**(2): 238–259.

Tsai V J D. 2006. A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Transactions on Geosciences and Remote Sensing*, **44**(6): 1661–1671

Wang J L and Wang S G. 2002. A shadow detection method of color image based on RGB color space. *Information Technology*, **26**(12): 7-8

Wang S G, Wang J L and Guo L Y. 2004. Shadow Detection of color aerial images based on K-L transformation. *Journal of geomatics*, **29**(2):21–23

Xia H Y, Chen X Y and Guo P. 2009. A shadow detection method for remote sensing images using affinity propagation algorithm. Proceedings 2009 IEEE International Conference on Systems, Man and Cybernetics: 3116–3121

Xu M Z and Yu Z H. 2003. Automated extraction of shadows in very-high resolution spatial satellite images. *Wtusm Bulletin of Science and Technology*, **28**(1): 202–222

Yao J and Zhang Z F. 2006. Hierarchical shadow detection for color aerial images. *Computer Vision and Image Understanding*, **102**(1): 602–609

基于统计混合模型的遥感影像阴影检测

夏怀英，郭平

北京师范大学 图形图像与模式识别实验室，北京 100875

摘　要：为提高阴影检测精度，提出一种新的遥感影像阴影检测方法—将径向基函数神经网络构建的混合模型(称作SMM-RBFNN)应用于遥感影像阴影检测。灰度共生矩阵中的能量、熵、对比度和逆差矩4种统计特征量作为混合模型的输入特征矢量，采用类“期望–最大化”算法(类EM)进行参数估计，训练检测器实现阴影检测。对多幅带有浓厚阴影的遥感影像进行实验，结果表明所提出的方法明显优于传统的高斯背景法和直方图阈值法，能够较好地解决强反射性地物漏检和水体错检问题，能够克服基于阈值思想的检测法需要反复实验选取阈值的缺点。

关键词：阴影检测，径向基函数神经网络，混合模型，纹理特征

中图分类号：TP751　　**文献标志码**：A

引用格式：夏怀英，郭平. 2011. 基于统计混合模型的遥感影像阴影检测. 遥感学报，**15**(4): 778–791

Xia H Y and Guo P. 2011. A shadow detection of remote sensing images based on statistical texture features. *Journal of Remote Sensing*, **15**(4): 778–791

1 引　言

阴影主要是高大建筑物、树木和山脉等对太阳光线的遮挡，在相应位置的影像内产生较暗的区域。阴影的存在一方面对很多图像处理操作造成很大影响，如图像配准、图像分割、特征提取、目标变化检测和定位以及图像检索(包倩和郭平，2006)等。另一方面可以利用阴影提取建筑物的三维信息，因此非常有必要对遥感影像中的阴影进行预处理。阴影检测是阴影处理的重要步骤之一，提高阴影检测精度对后续的遥感图像处理具有重要作用。

现有的遥感影像阴影检测方法大致可分为两大类：一类是基于模型(Finlayson 等，2002)的方法。该类方法是根据相关场景、遮挡物体的几何形状、DSM数据以及太阳照射方位、传感器等相关参数来计算阴影区域(Yao和Zhang，2006；Gils，2001)。但这些信息在实际应用中并不容易得到，有很大的局限性；另一类方法是基于阴影的性质，主要是用阈值分割的思想进行检测，根据阴影区域的共性与非阴影区域的差别提取阴影(Gwinner 和Schaale， 1997; 王军利和王树根， 2002；Tsai，2006)。阴影性质主要包括光谱性质、纹理性质和边缘性质。Jiang和Ward(1992)提出了一种检测方法，主要根据阴影区域的亮度要比非阴影区域低的性质；Funka-Lea G和Bajcsy(1995)提出基于阴影几何性质的检测方法；Highnam 和Brady(1997)提出基于直方图阈值的阴影检测方法；Elemadmia和Alsharif (2003)提出了基于同态滤波的检测方法；Salvador等人(2004)提出了一种基于阴影区域的色彩不变性的检测方法。国内许多学者在这些方法的基础上，提出了一些不同的检测方法，如王树根等提出基于彩色空间的色彩转换方法(王军利和王树根，2002)，整体变分模型法，基于K-L变换的检测方法(王树根 等，2004)；许妙忠和余志惠(2003)提出归一化处理、纹理分析的方法；虢建宏等(2006)提出多波段检测方法；Xia等人(2009)提出基于Affinity Propagation聚类算法的阴影检测方法；季顺平和袁修孝(2007)提出改进的高斯背景模型检测方法；鲍海英等(2010)提出结合GRB空间G分量和HSI空间亮度分

收稿日期：2010-05-07；**修订日期**：2010-08-19

基金项目：国家自然科学基金 (编号：90820010；编号：60911130513)

第一作者简介：夏怀英(1984—)，女，北京师范大学信息科学与技术学院硕士研究生，研究方向为模式识别和遥感图像处理，发表EI检索论文2篇。E-mail: xia_huaiying@163.com。

通信作者：郭平，E-mail: pguo@ieee.org。

量的阈值检测法。

这些方法在一些特定领域得到了很好地应用，但是由于阴影形成机理的复杂性，至今还没有一种通用的检测方法适用所有的遥感影像。如基于色彩不变的检测方法对场景限制条件较多，假设阴影区域是非纹理的平坦表面，必须是单一光照；基于直方图的检测方法存在强反射性地形、地物漏检和水体错检的缺陷，而且只适合灰度图像；基于同态滤波的方法有可能将有用信息检测为阴影，因为一幅图像上的大部分信息集中在低频部分，还有低通滤波器的参数要人为设置，空域和频域转换导致算法的时间复杂度高，资源占有率高。其他很多方法也存在阈值选择的主观性和反复实验的缺点。

阴影检测在某种意义上可以看作是分类和目标识别的过程，可以根据遥感影像的光谱、纹理等特征将阴影作为单独一类识别出来。为了提高阴影检测精度，本文提出基于径向基函数神经网络(Radial basis function Neural Network，RBFNN)组成的混合模型作为分类器应用于遥感影像阴影检测的方法。实验证实，所提出的方法能够较好的解决强反射性地物漏检和水体错检的问题，提高了阴影检测的精度，克服了基于阈值分割检测方法中阈值需要反复实验才能确定的缺点。

2 背景知识

2.1 高斯核RBF神经网络

RBF神经网络(RBFNN)是一种重要的人工神经网络，由于具有非线性处理能力，以及结构简单、训练速度快而逐渐成为一种用途广泛的网络模型。RBFNN结构如图1所示(Musavi 等，1992)，其数学描述为：

$$g(\boldsymbol{X},\boldsymbol{W})=\sum_{i=1}^{K_j} w_i\phi(\|\boldsymbol{X}-\mu_i\|)+w_0 \tag{1}$$

式中，$\boldsymbol{X}$，$g(\boldsymbol{X},\boldsymbol{W})$，$\boldsymbol{W}=(w_i|i=1,2,\ldots,K_j)$分别为网络的输入、输出与权重，$\mu_i$是径向基函数的中心，$w_0$是偏置神经元，$\phi(\|x-\mu_i\|)$是一个径向基函数。取径向基函数为高斯型时，RBF网络的数学模型可以表示为：

$$g(\boldsymbol{X},\boldsymbol{W})=\sum_{i=1}^{K_j} w_i\exp\left[-\frac{\|\boldsymbol{X}-\mu_i\|^2}{2\gamma_i^2}\right]+w_0 \tag{2}$$

对RBF神经网络的训练包括确定高斯核函数的中心μ、宽度即方差γ和输出单元的权值w。

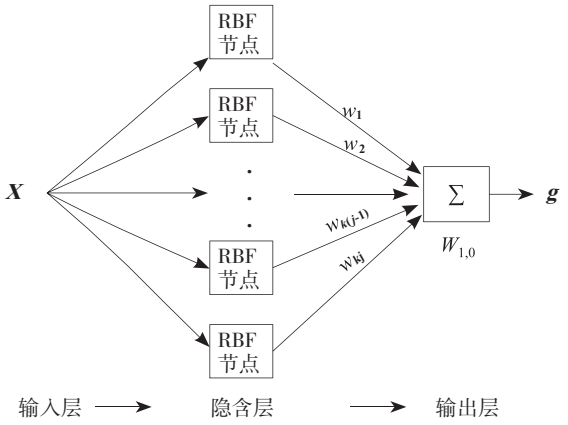


图1 RBF神经网络结构图

2.2 RBF神经网络混合模型

混合模型假设有H个单模型混合组成，各个模型具有不同的权值，即分类器是由H个RBFNN构成的混合模型。其结构可用图2表示。高斯混合模型的联合概率密度可以表示为(郭平 等，2003)：

$$p(\boldsymbol{Z}\mid\boldsymbol{X},\boldsymbol{\Theta})=\sum_{j=1}^{H}P(j)p(\boldsymbol{Z}\mid\boldsymbol{X},\boldsymbol{\Theta}_j) \tag{3}$$

式中，$p(j)\geqslant 0$，$\sum_{j=1}^{H}p(j)=1$，$\boldsymbol{X}$，$\boldsymbol{Z}$分别为输入矢量和输出矢量，其维数分别为$\boldsymbol{d}_x$和$\boldsymbol{d}_z$，Θ代表参数组。H为混合模型中的分量个数，$\boldsymbol{P(j)}$是混合权重参数。混合模型中每个分量都服从高斯分布：

$$\begin{aligned}p(Z\mid X,\Theta_j)&=G(z,g_j(X,W),\sigma_j^2)\\&=\frac{1}{(2\pi\sigma_j^2)^{d_z/2}}\times\exp\{-\frac{1}{2\sigma_j^2}\|z-g_j(X,W)\|^2\}\end{aligned} \tag{4}$$

映射函数$g_j(\boldsymbol{X},\boldsymbol{W})$为高斯核RBF神经网络。系统的代价函数或者称作误差函数有如下形式：

$$\begin{aligned}E=-l(\boldsymbol{\Theta})&=-\sum_{i=1}^{N}\ln p(z_i\mid x_i,\boldsymbol{\Theta})\\&=-\sum_{i=1}^{N}\ln\sum_{j=1}^{H}p(j)p(z_i\mid x_i,\boldsymbol{\Theta}_j)\end{aligned} \tag{5}$$

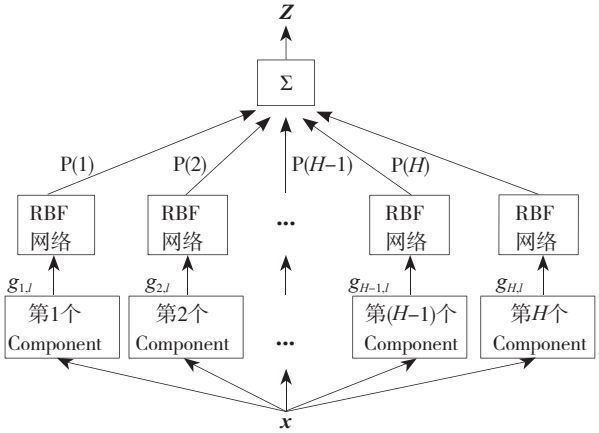


图2 统计混合模型结构图

2.3 类EM参数估计

对参数$\boldsymbol{\Theta}$进行最大似然估计，对系统代价函数(5)采用类EM算法估计(Guo 等， 2003)：

E-Step：固定$P(j)$old和$\boldsymbol{\Theta}$old，按照式(6)计算后验概率：

$$h(j,n)=\frac{P(j)^{\text{old}}G(z_n,g_j(x_n,W^{\text{old}}),(\sigma_j^2)^{\text{old}})}{\sum_{i=1}^{H}P(i)^{\text{old}}G(z_n,g_i(x_n,W^{\text{old}}),(\sigma_i^2)^{\text{old}}))} \tag{6}$$

M-Step：按照式(7)计算参数$P(j)_{\text{new}}$和$\boldsymbol{\Theta}_{\text{new}}$，在用最大似然估计时，沿着梯度方向移动。

$$\begin{aligned} P(j)^{\text{new}} &= \frac{1}{N}\sum_{n=1}^{N}h(j,n) \\ \Theta_j^{\text{new}} &= \Theta_j^{\text{old}}+\eta\nabla l(\Theta)\big|_{\Theta=\Theta^{\text{old}}} \end{aligned} \tag{7}$$

采用上面的类EM算法估计参数后，混合模型的期望输出为：

$$\begin{aligned} \langle \boldsymbol{Z}\rangle &= \int \boldsymbol{Z}p(Z\mid \boldsymbol{X},\boldsymbol{\Theta})\mathrm{d}\boldsymbol{Z} \\ &= \int \boldsymbol{Z}\sum_{j=1}^{H}P(j)G(\boldsymbol{Z},g_j(\boldsymbol{X},W_j),\sigma_j^2)\mathrm{d}\boldsymbol{Z} \\ &= \sum_{j=1}^{H}P(j)g_j(\boldsymbol{X}_j,\boldsymbol{W}_j) \end{aligned} \tag{8}$$

2.4 纹理特征

遥感图像的纹理特征表现地物的形状、均质程度、方位以及不同类地物间的亮度关系和空间关系。灰度共生矩阵Gray Level Co-occurrence Matrix，GLCM)是最常用的纹理分析方法之一。常用的纹理指数包括能量指数(Energy)、熵指数(Entropy)、对比度(Contrast)和逆差矩(Inverse Difference Moment)(Baraldi 和Parmiggiani， 1995)：

能量(Energy)：又称角二阶矩，反映了图像灰度分布均匀程度和纹理粗细度，当一个区域中的灰度值非常接近或者变化不大时，能量值达到一个最大值。

$$\text{Energy}=\sum_{i=1}^{N}\sum_{j=1}^{N}P_d^2(i,j) \tag{9}$$

熵(Entropy)：反应图像中纹理的非均匀程度和复杂程度，当GLCM中所有值都相等时，熵的值达到最大。

$$\text{Entropy}=-\sum_{i=1}^{N}\sum_{j=1}^{N}P_d(i,j)\log P_d(i,j) \tag{10}$$

对比度(Contrast)：反映了图像的清晰度和纹理沟纹深浅的程度。纹理沟纹越深，其对比度越大，视觉效果越清晰；反之，对比度小，则沟纹浅，效果模糊。

$$\text{Contrast}=\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2P_d(i,j) \tag{11}$$

逆差距(IDM)：度量图像纹理局部变化的多少，反映图像纹理的同质性，如果图像不同区域之间缺少变化，即局部均匀，则逆差矩的值就比较大。

$$\text{IDM}=\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{P_d(i,j)}{|i-j|^2} \tag{12}$$

3 基于统计混合模型的阴影检测

3.1 原理分析

3.1.1 纹理特征分析

通过对不同图像的阴影和非阴影区域多次实验发现，阴影区域的熵和对比度呈一个递减的趋势，而非阴影区域的熵和对比度具有周期性波动的特性，阴影区域灰度值非常接近而且变化不大，所以阴影区域的能量具有比非阴影偏高的特性。图4和图5是对图3中遥感图像的阴影和非阴影区域的4个统计特性的比较分析。

图3 原始遥感影像

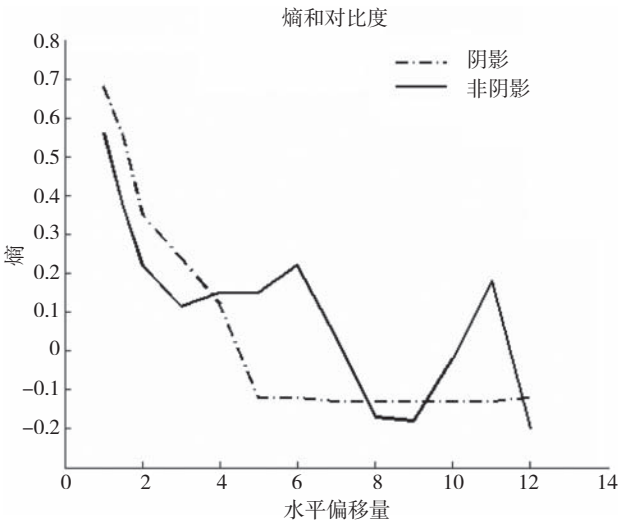


图4 阴影和非阴影区域的熵和对比度分析

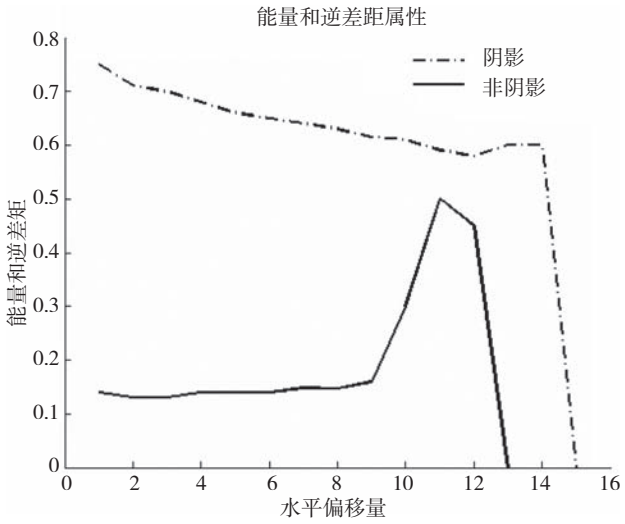


图5 阴影和非阴影区域的能量和逆差矩分析

目前很多阴影检测方法都是基于阴影的光谱特性的，利用阴影区域具有较低的灰度值设置一个或多个阈值，但是影像上和阴影相近的目标如水体、绿色草地和蓝色屋顶等很难区分，在阈值选取时通过反复实验获得。通过上述对阴影和非阴影区域纹理特征分析可知，熵、能量、对比度和逆差矩作为阴影检测的特征更加具有可分性。

3.1.2 统计混合模型

基于高斯背景模型的检测方法(季顺平和袁修孝，2007)的基本思想是把自然景物和人工地物都看作背景，认为其符合高斯分布，而阴影区域则不符合这种分布，从而将阴影检测出来。但事实上，单纯高斯模型并不能很好的表达自然背景和人工地物，高分辨率遥感图像不同于自然图像，通常包含很多地物，每一种地物类型可能服从不同的高斯分布，所以传统的高斯模型并不能很好的表示高分辨率遥感影像。从理论上，多分量组成的混合模型可以描述任何复杂的分布，并且具有较强的推广能力，通过调节混合模型的参数就可以满足描述实际高分辨率遥感图像。由于RBF网络具有很强的非线性与局域化的能力，从而进一步提高遥感影像分类及阴影检测的精确度。

3.2 混合模型构建

本文采用RBF网络作为构建混合模型的分量。

(1)混合模型的输入：四维特征向量即$\boldsymbol{X}=\{\boldsymbol{x}_{\text{Energy}}, \boldsymbol{x}_{\text{Entropy}}, \boldsymbol{x}_{\text{Contrast}}, \boldsymbol{x}_{\text{IDL}}\}$。在计算灰度共生矩阵及其纹理指数的时候，本文采用分块的方式，对每一小块分别提取4种纹理指数，作为一个样本点。在对图像分块时，窗口的大小是一个非常重要的因素，实验中分别选取了8×8，6×6，4×4和2×2的窗口大小。通过分析比较，发现窗口大小为4×4时可以获得比较好的检测效果，而采用不同的方向影响不大，本文实验中采用的方向是45°。

(2)期望输出：混合模型的期望输出为阴影、阴影与非阴影交界和非阴影三大类。期望输出$\boldsymbol{Z}$采用**K**之一的编码方式，属于阴影区域，$\boldsymbol{Z}_1=(1, 0, 0)^{\mathrm{T}}$；属于边界区域，$\boldsymbol{Z}_2=(0, 1, 0)^{\mathrm{T}}$；属于非阴影区域，$\boldsymbol{Z}_3=(0, 0, 1)^{\mathrm{T}}$。

(3)实际输出：实际输出是一个概率向量$\boldsymbol{P}=\{p_1, p_2, p_3\}$，$p_i$表示$\boldsymbol{X}$属于第$i$类的概率($i$=1，2，3)，如果$\boldsymbol{P}$中$p_j$分量相对最大，则将对应的输入样本$\boldsymbol{X}$划分到第$j$类。

3.3 实验流程

3.3.1 训练样本选取

实验中采用的训练样本影像是从Google Earth上截取的。选择了3类训练样本，第1类为阴影，第2类为阴影与非阴影交界区，第3类是非阴影。对于每一类，在实验中选取了15幅样本。

3.3.2 参数初始化

模型中共有8个参数需要初始化，即{H，K_j，$P(j)$，σ^2，$\boldsymbol{W}$，γ^2，μ，η}：

(1)实验中采用5个RBF网络组成的混合模型，即H=5；

(2)RBF节点数分别为11、12、13、14和15，即K_1=11，K_2=12，K_3=13，K_4=14，K_5=15；

(3)混合模型中各个RBF网络权重的初始值设为$P(j)=1/H(j=1, \cdots, 5)$，即$P(j)$=0.2；

(4)各个RBF网络的方差σ^2设为9，这是一个根据多次实验获得的经验参数；

(5)RBF网络的权值初始化为：

$$w(j,i,l)=0.25+0.75\times Z(j,l) \tag{13}$$

j=1，…，H，i=1，…，K_j，1=1，…，d_z，$\boldsymbol{Z}(j, l)$是一个列向量；

(6)每个RBF网络中的各个RBF节点的均值μ通过K均值聚类算法进行初始化，具体过程可以描述如下：

步骤1 第j个RBF网络的RBF节点的均值$\mu_{j,i}$，随机选取K_j个数据作为聚类中心；

步骤2 按欧式距离最小原则，将待分数据点x_n分配给中心为$\mu_{j,i}$的第i类；

步骤3 计算第i类的均值并更新聚类中心$\mu_{j,i}$；

步骤4 循环迭代，直到各类聚类中心不再变化，得到各类的初始均值$\mu_{j,i}$(j=1，…，H，i=1，…，K_j)；

(7)第j个RBF网络第i个节点的方差：

$$(\gamma_j^2)_i = \left| x_{j,i} - \mu_{j,i} \right|^2 \tag{14}$$

(8) 学习因子η是一个非常重要的参数，如果选取不当，类EM算法就得不到最好的收敛结果，或者收敛于局部最小值或者收敛于局部最大值，通过多次实验分析，发现η应该取一个很小的数，在实验中根据经验选择η= – 0.5。

3.3.3 训练流程描述

参数初始化好后，按照2.3节中的类EM算法进行参数估计，训练流程如图6所示。经过不断迭代学习获得混合模型的参数$\boldsymbol{\Theta}$={$\boldsymbol{W}$，σ^2，γ，μ}。

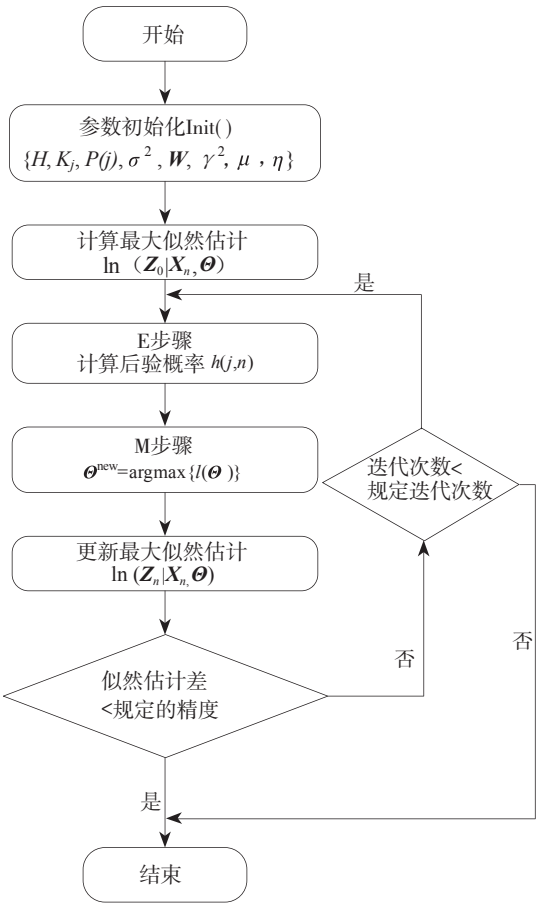


图6 混合模型训练流程图

4 实验结果与分析

利用本文提出的方法，进行了许多实际阴影检测实验。实验采用的机器内存为2 G，CPU双核1.8 GHz，采用从Google Earth上随机下载的含有浓厚阴影的遥感图像，实验过程中没有对图像的尺寸进行裁剪操作。由于篇幅限制，下面列出了其中三幅图像的检测结果。图7(a)、图8(a)和图9(a)是3幅带有阴影的图像，分别称作Img1(200 × 200像元)、Img2(230 × 200像元)和Img3(230 × 150像元)。可以看出，原图中包含大量阴影区域以及与阴影色彩相近的绿色草地、水体等。为了便于观察检测效果，我们将检测到的非阴影区域用白色表示，检测到的阴影区域保持原来的颜色。并和高斯背景法和传统的直方图阈值法进行了对比分析。图7(b)、图8(b)和图9(b)是SMM-RBFNN法检测结果，图7(c)、图8(c)和图9(c)是高斯背景模型法检测结果，图7(d)、图8(d)和图9(d)是传统直方图阈值法的检测结果。

为了更加客观的分析检测结果，借鉴了文献(Martel-Brinson等，2005)所采用的分析方法，主要分析TN、FN、FS 3类数据。其中TN表示正确识别的阴影区域数目，FN表示将非阴影区域误识别为阴影区域数目，FS表示漏检的阴影区域数目。通过两个指标来评价检测精度，一个是阴影区域的正确检测率DR，一个是阴影区域的漏检率FR。DR和FR定义为：

$$\begin{aligned} \mathrm{DR} &= \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}} \times 100\% \\ \mathrm{FR} &= \frac{\mathrm{FS}}{\mathrm{TN}+\mathrm{FN}+\mathrm{FS}} \times 100\% \end{aligned} \tag{15}$$

一个好的阴影检测方法应该满足DR尽量大，同时FR尽量小。

根据式(15)计算了SMM-RBFNN法、高斯背景模型法和直方图阈值法的DR和FR两种评价指标(表1)。

表1 DR和FR两种评价指标分析

方 法	影 像	DR指标	FR指标	检测效果
SMM-RBFNN	Img1	92.86%	6.67%	很好
	Img2	91.67%	7.69%	很好
	Img3	92.62%	6.78%	很好
高斯背景法	Img1	81.25%	11.11%	一般
	Img2	83.33%	14.28%	一般
	Img3	88.27%	10.74%	一般
直方图阈值法	Img1	68.42%	9.52%	不好
	Img2	75.00%	14.28%	不好
	Img3	74.17%	13.27%	不好

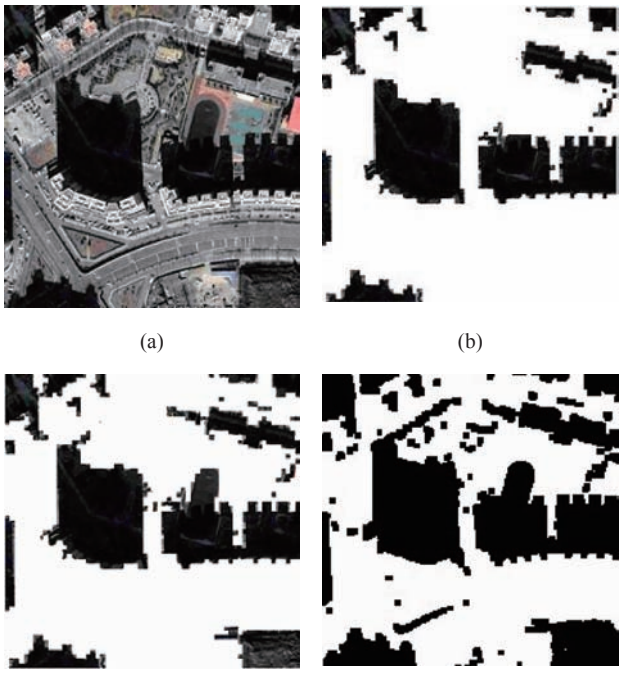
根据表1所示的3种检测方法的DR和FR指标以及人眼目视判断可以得出如下结论：

(1)直方图阈值法易将影像中与阴影区域色彩相近的绿色草地、灰度值较低的水体、路面等错误识别为阴影；

(2)高斯背景法检测精度较直方图阈值法要好，但是也会将与阴影灰度值相近的区域如绿色植被错误

识别为阴影；

(3)SMM-RBFNN方法的DR值平均在92%左右，FR值平均为7%左右，说明SMM-RBFNN检测法具有较高的正确检测率，较低的误检率和漏检率，检测效果很好。

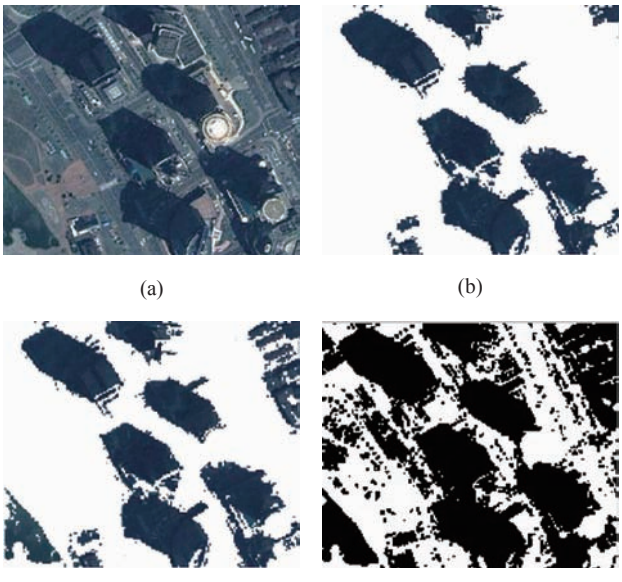


(a) (b)

(c) (d)

图7 Img1检测结果

(a)原图；(b)SMM-RBFNN法检测结果；
(c)高斯背景模型检测结果；(d)直方图阈值检测结果

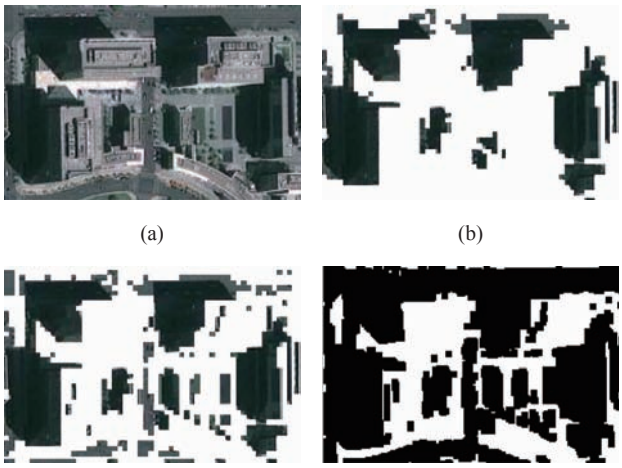


(a) (b)

(c) (d)

图8 Img2检测结果

(a)原图；(b)SMM-RBFNN法检测结果；
(c)高斯背景模型检测结果；(d)直方图阈值法检测结果



(a) (b)

(c) (d)

图9 Img3检测结果

(a)原图；(b)SMM-RBFNN法检测结果；
(c)高斯背景模型检测结果；(d)直方图阈值检测结果

由此可见，SMM-RBFNN检测法不但是有效的，而且在检测精度上明显优于高斯背景法和直方图阈值法。只要确定好初始参数值，整个检测流程就不再需要人工干预，达到了较高的自动化检测标准，而高斯背景模型则易将绿色草地、强反射性地物和水体等错误识别为阴影区域；传统的阈值法需要设置多个阈值剔除与阴影区域相近的物体，同时还需要数学形态学处理去除孤立点。混合模型之所以能够获得较高的检测精度，主要是用混合模型表示高分辨率遥感图像更加符合实际情况，所采用的RBF网络分量具有很好的扩展能力。原则上通过调节混合模型的参数(混合分量的个数H、RBF中径向基函数的个数K_j，K_j一般取决于训练样本的数目N，通常情况下取$K_j=\{2/3\times N+j\}(j=1,\cdots,H)$)可以满足具体问题要求，从而提高了组合网络模型的输出精度。

5 结 论

提出了将径向基函数神经网络构建的统计混合模型应用于遥感影像阴影检测，通过分析阴影区域和非阴影区域纹理统计特性，综合运用了基于灰度共生矩阵的4种统计特性即能量指数、熵指数、对比度和逆差矩作为相似性度量指标。对混合模型进行参数估计时，采用了类期望-最大化算法，通过大量实验证明，该方法可以非常准确的检测出阴影，能够克服基于阈值思想的检测方法中阈值选择的主观性和反复实验的缺点，很好的解决强反射性地形、地物漏检和水体错检的问题。混合

模型能够表示实际中的复杂分布情况，比单纯高斯模型更加适合描述高分辨率遥感图像阴影检测问题。与传统的高斯背景法和直方图阈值法进行对比实验发现，混合模型具有更高的阴影检测精度。同时，本文方法不需要假设条件，具有更加广泛的应用范围。

本文采用纹理特征作为输入特征矢量，而纹理和光谱特征的组合是我们下一步的研究目标。另外，混合模型的参数初始值是经过多次实验获得的经验值，还没有一个统一的方法初始化参数。RBF网络的节点数K_j、类EM算法的学习因子η等参数对检测精度影响很大，因此，如何选取更加合理的参数有待进一步深入研究。

志　谢　在此衷心的感谢胡汝坤同学在实验上提供的帮助。

REFERENCES

Bao H Y, Li Y and Yin Y Y. 2010. The study on shadow detection and shadow elimination in the urban aerial image. *Remote Sensing Information*, (1): 44–47

Bao Q and Guo P. 2006. Comparative studies on similarity measures for remote sensing image retrieval based on histogram. *Journal of Remote Sensing*, **10**(6): 893–900

Baraldi A and Parmiggiani F. 1995. An investigation of the textural characteristics associated with gray Level concurrence matrix statistical parameters. *IEEE Transaction on Geo-science and Remote Sensing*, **33**(2): 293–304

Etemadnia H and Alsharif M R. 2003. Automatic image shadow identification using LPF in homomorphic processing system. Proceedings of VII Digital Image Computing: Techniques and Applications, Sydney: Committee of DICTA: 429–438

Finlayson G D, Hordley S D and Drew M S. 2002. Removing shadows from images. *Lecture Notes in Computer Science*, **2535**: 129–132

Funka-Lea G and Bajcsy R. 1995. Combining color and geometry for the active, visual recognition of shadows. Proceedings of fifth International Donference on Computer Vision: 203–209

Gils P T. 2001. Remote sensing and cast shadows in mountainous terrain. *Photogrammetric Engineering and Remote Sensing*, **67**(7): 833–839

Guo J H, Tian Q J and Wu J Z. 2006. Study on multispectral detection shadow areas and a theoretical model of removing shadows from remote sensing images. *Journal of Remote Sensing*, 10(2): 151–159

Guo P, Qin D M, Hu Z Y and Du W M. 2003. A novel method for spectral signal pattern recognition. *Spectroscopy and Spectral Analysis*, **23**(4): 811–815

Gwinner K and Schaale M. 1997. A case study on the influence of shadows and shading on multispectral airborne imaging data. Presented at the third International Airbome Remote Sensing Conference and Exhibiton. Copenhagen, Demark

Highnam R and Brady M. 1997. Model-based image enhancement of far infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(4): 410-415

Jiang C and Ward M O. 1992. Shadow identification. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition: 606–612

Ji S P and Yuan X X. 2007. A method for shadow detection and change detection of man-made objects. *Journal of Remote Sensing*, **11**(3): 323–329

Martel-Brinson N and Saccharine A. 2005. Moving cast shadow detection from a Gaussian mixture shadow model. IEEE Computer Society Conference on Computer Vision and Pattern Recognition: 643–648

Musavi M T, Ahmed W, Chan K H, Faris K B and Hummels D M. 1992. On the training of radial basis function classifiers. *Neural Networks*, **5**(4): 595–603

Salvador E, Cavallaro A and Ebrahimi T. 2004. Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding*, **95**(2): 238–259.

Tsai V J D. 2006. A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Transactions on Geosciences and Remote Sensing*, **44**(6): 1661–1671

Wang J L and Wang S G. 2002. A shadow detection method of color image based on RGB color space. *Information Technology*, **26**(12): 7-8

Wang S G, Wang J L and Guo L Y. 2004. Shadow Detection of color aerial images based on K-L transformation. *Journal of geomatics*, **29**(2):21–23

Xia H Y, Chen X Y and Guo P. 2009. A shadow detection method for remote sensing images using affinity propagation algorithm. Proceedings 2009 IEEE International Conference on Systems, Man and Cybernetics: 3116–3121

Xu M Z and Yu Z H. 2003. Automated extraction of shadows in very-high resolution spatial satellite images. *Wtusm Bulletin of Science and Technology*, **28**(1): 202–222

Yao J and Zhang Z F. 2006. Hierarchical shadow detection for color aerial images. *Computer Vision and Image Understanding*, **102**(1): 602–609

附中文参考文献

鲍海英，李艳，尹永宜．2010．城市航空影像的阴影检测和阴影消除方法研究．遥感应用，(1)：44–47

包倩，郭平．2006. 基于直方图的遥感图像相似性检索方法比较. 遥感学报，**10**(6)：893–900

虢建宏，田庆久，吴昀昭．2006．遥感影像多波段检测与去除理论模型研究．遥感学报，**10**(2)：151–159

郭平，覃冬梅，胡占义，杜为民．2003．一种光谱识别的新方法．光谱学与光谱分析，**23**(4)：811–815

季顺平，袁修孝．2007．一种基于阴影检测的建筑物变化检测方法．遥感学报，**11**(3)：323–329

王军利．王树根．2002．一种基于RGB彩色空间的影像阴影检测方法．信息技术，**26**(12)：7–8

王树根，王军利，郭丽艳．2004．基于K-L变换的航空影像阴影检测．测绘信息与工程，**29**(2)：21–23

许妙忠，余志惠．2003．高分辨率卫星影像中阴影的自动提取与处理．测绘信息与工程，**28**(1)：202–222