Shape feature representation of ground objects from high-resolution remotely sensed imagery base on Fourier Descriptors

WAN Wei[1], FENG Xuezhi[1], XIAO Pengfeng[1], ZHAO Limin[2]

1. *Department of Geographical Information Sciences, Nanjing University, Jiangsu Nanjing* 210093, *China*;
2. *International Institute for Earth System Science, Nanjing University, Jiangsu Nanjing* 210093, *China*

Abstract: The traditional Fourier Descriptors (FDs) are normalized in this paper to make it independent of translation, rotation and scale changes. Four typical objects i.e. building, paddy, road and river are selected and their boundaries are expressed as sequences of complex numbers. FDs are obtained through one-dimensional Fourier transform. The characteristics of the frequency spectrum, contribution rate and the shape reconstruction are analyzed. The results show that the different frequency ranges have different contribution rates; the Direct Component (DC) reaches a proportion of more than 70%; the Low Frequency (LF) and High Frequency (HF) totally reach 7%–24% while the Medium Frequency (MF) merely 2%–4%. The LF components (descriptors 1—5) make a commendable reconstruction of objects' shape and these descriptors are applied to the object-oriented classification. The overall classification accuracy is 98.48% with a Kappa coefficient 0.9714.

Key words: shape feature, Fourier Descriptors, high-resolution, remotely sensed imagery

CLC number: TP751 **Document code:** A

Citation format: Wan W, Feng X Z, Xiao P F and Zhao L M. 2011. Shape feature representation of ground objects from high-resolution remotely sensed imagery base on Fourier Descriptors. *Journal of Remote Sensing*, **15**(1): 73–87

1 INTRODUCTION

High resolution remotelysensed image segmentation techniques automatically classify the neighboring pixels into the meaningful regions based on homogeneity or heterogeneity criteria respectively. These meaningful regions are called image objects. Multi-features such as spectrum, texture, shape, and topology are needed to be extracted and represented so that they can be used for improving the accuracy of the succeeding object-oriented classification. Shape is an important visual feature among the multi-features obtained from the image objects (Pierce, *et al.*, 1994; Luo, *et al.*, 1999; Benz, *et al.*, 2004; Huang, *et al.*, 2007; Zhang, *et al.*, 2008). Methods of shape presentation in current research or softwares are mainly based on some indexes such as area, aspect ratio, compactness ratio (Chen, *et al.*, 2006), roundness and convexity (Werff & Meer, 2008), shape index (Wei, *et al.*, 1999, Chen, *et al.*, 2004).

Fourier Descriptors (FDs), advanced by Cosgriff in 1960, is a method that belongs to contour-based shape techniques. FDs refers to the coefficients of Fourier transform of objects' boundaries (Wang, *et al.*, 2002). The boundary of an object can be counted as a discrete complex function of periodic signal, and one-dimensional Fourier transform can convert this function from spatial domain in to frequency domain. Since it was used for shape description of planar closed curve, lots of applications have been attempted in many different fields. The examples of FDs used in image processing field include hand gesture recognition (Feng, *et al.*, 2003; Liu, *et al.*, 2005; Ge, *et al.*, 2005), gait recognition (Stuart, *et al.*, 2003; Cheng, *et al.*, 2008), image matching (Govindu, *et al.*, 1998; Duan, *et al.*, 2008), image retrieval and classification (Rui, *et al.*, 1998; Wong, *et al.*, 2007; Yadav, *et al.*, 2007). However, the rare research can be seen for applying this method for shape representation of remotely sensed imagery.

The high resolution IKONOS multispectral imagery is used as data source. After normalization, FDs are used to represent the shape feature of the four kinds of objects as follows: building, paddy, road and river. Characteristics of the frequency spectrum, contribution rate and shape reconstruction are analyzed. The shape parameters finally used in the object-oriented classification based on decision tree and high accuracy is obtained.

2 THE PRINCIPLE OF FDs

A digital boundary with K points is shown in Fig.1. Take an arbitrary point as starting point, the coordinates (x_0, y_0), (x_1, y_1), (x_2, y_2), $\cdots$, (x_{k-1}, y_{k-1}) represent a series of boundary points

Received: 2010-01-18; **Accepted**: 2010-04-02

Foundation: The National High Technology Research and Development Program of China (863 Program) (No. 2008AA12Z106); The National Natural Science Foundation of China (No. 40801166); The Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 200802841012).

First author biography: WAN Wei(1986—), female, MDs candidate in Nanjing University, majors in digital image processing on mote sensing. E-mail: wanwei_4824@163.com

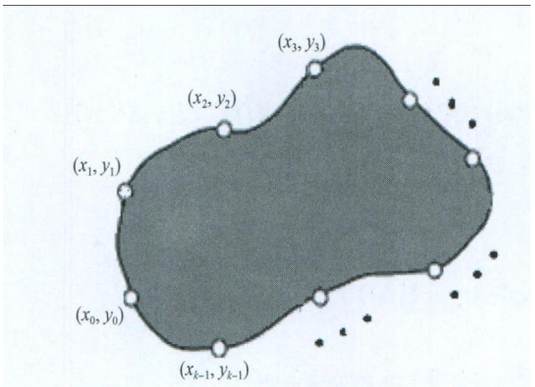

Fig. 1 A digital boundary

and these coordinates can be expressed as $x(k)=x_k$, $y(k)=y_k$. Therefore, the digital boundary itself can be represented as the sequence of coordinate $s(k)=[x(k), y(k)]$, for $k=0,1,2,\cdots,K-1$. Moreover, each coordinate pair could be treated as a complex number so that $s(k)=x(k)+jy(k)$, for $k=0,1,2,\cdots,K-1$. A 2-D problem is reduced to a 1-D one.

According to the principle of Fourier transform for single variable discrete function (Gonzalez and Woods, 2002), the Discrete Fourier Transform (DFT) of $s(k)$ is:

$$a(u)=\frac{1}{K}\sum_{k=0}^{K-1}s(k)e^{-j2\pi uk/K} \tag{1}$$

where $u=0,1,2,\cdots,K-1$. The complex coefficients $a(u)$ are called the FDs of the boundary. Likewise, the inverse Fourier transform of $a(u)$ is:

$$s(k)=\sum_{u=0}^{K-1}a(u)e^{j2\pi uk/K} \tag{2}$$

where $k=0,1,2,\cdots,K-1$. Furthermore, suppose that instead of all the Fourier coefficients, only the first Lcoefficients are used, and the inverse Fourier transform is the following expression:

$$\hat{s}(k)=\sum_{u=0}^{L-1}a(u)e^{j2\pi uk/K} \tag{3}$$

where $k=0,1,2,\cdots,K-1$. Although there are merely L terms to calculate each element of $s(k)$, the range of value of k remains to be 0 to $K-1$, that is to say, the number of points in the approximate boundary is the same as in the original one, but as for the reconstruction of each point, only the first L coefficients are chosen. The more coefficients used, the more approximate for the landscape to the original boundary because lower frequency descriptors represent general information of the shape and the higher frequency the details in Fourier domain.

3 NORMALIZATION OF FDs

A good shape descriptor must be insensitive to translation, rotation and scale changes (Zahn & Roskies, 1972), and this anti-sensitivity is especially important for shape feature representation of ground objects from remotely sensed imagery. Though they are similar in shape, the different objects with the same ground feature may have different scales or directions. Cornproving roads and buildings, roads are similar-appearing but different in directions, buildings are mostly rectangle-like but the size and directions of those rectangles are always different. Therefore, FDs should be normalized in order to make it insensitive to translation, rotation and scale changes. Fig.2 are used to explain the translation, scale changes and rotation of the boundary in Fig.1.

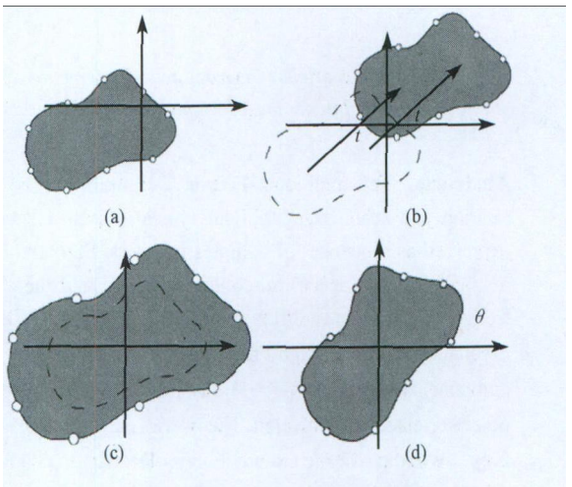

Fig. 2 Translation, scale changes and rotation
(a) Original shape; (b) Translation; (c) Scale changes; (d) Rotation

As shown in Fig.2(b), $s(k)$ changes into $t+s(k)$ after t units ($t=\delta x+j\delta y$) transfering, and the Eq.(1) changes into the following equation:

$$\tilde{a}(u)=\frac{1}{K}\sum_{k=0}^{K-1}[t+s(k)]\times e^{-j2\pi uk/K} \tag{4}$$

Eq.(4) can be reduced to Eq. (5):

$$\tilde{a}(u)=a(u)+\sum_{k=0}^{K-1}t\times e^{-j2\pi uk/K} \tag{5}$$

where $u=0$, $\tilde{a}(0)=a(0)+\sum_{k=0}^{K-1}t\times e\times 0=Kt$, while for $u>0$, $\tilde{a}(u)=a(u)+\sum_{k=0}^{K-1}t\times e^{-j2\pi uk/K}=a(u)+0=a(u)$. It can be seen from this that translation is only related to $u=0$, so the 0th coefficient or rather the Direct component (DC) should be excluded to eliminate the effect of translation.

As shown in Fig.2(c), the original shape is enlarged for m times, or $s(k)$ is multiplied by factor m, then Eq.(1) changes into the following equation:

$$\begin{aligned}\tilde{a}(u)&=\frac{1}{K}\sum_{k=0}^{K-1}[m\times s(k)]\times e^{-j2\pi uk/K}\\&=m\times\frac{1}{K}\sum_{k=0}^{K-1}s(k)\times e^{-j2\pi uk/K}\\&=m\times a(u)\end{aligned} \tag{6}$$

It can be seen from Eq.(6) that $a(u)$ has been enlarged for m times. Here, suppose there is another shape which is multiplied by factor n, then $\tilde{\tilde{a}}(u)=n\times a(u)$, so $\frac{\tilde{a}(u)}{a(u)}=$

$\frac{|\tilde{a}(0)|}{|a(0)|}$ and $\frac{\tilde{\tilde{a}}(u)}{a(u)}=\frac{|\tilde{\tilde{a}}(0)|}{|a(0)|}$, and then $\frac{\tilde{a}(u)}{|\tilde{a}(0)|}=\frac{\tilde{\tilde{a}}(u)}{|\tilde{\tilde{a}}(0)|}$.

Therefore, in order to make it independent of scale changes, FDs should be divided by the absolute value of the 0th coefficient.

As shown in Fig.2d, $s(k)$ changes into $e^{j\theta}s(k)$ after angle θ rotation. According to the property of one-dimensional Fourier transform, only the phase information is influenced by rotation, so after doing modular operation to all the coefficients, the effect of rotation can be eliminated. The following research is based on the normalized FDs that have been dealt with through the above three steps.

4 SHAPE FEATURE ANALYSIS OF GROUND OBJECTS BASED ON FDs

Four objects i.e. building, paddy, road and river are selected as the research objects, the boundary of each object are obtained by image segmentation. Fig.3 shows the boundary contours of the four objects. Shape feature analysis is done through the following three aspects: Characteristics of the frequency spectrum being obtained using FDs, the contribution rate of FDs with different frequency ranges, and shape reconstruction with a certain coefficient.

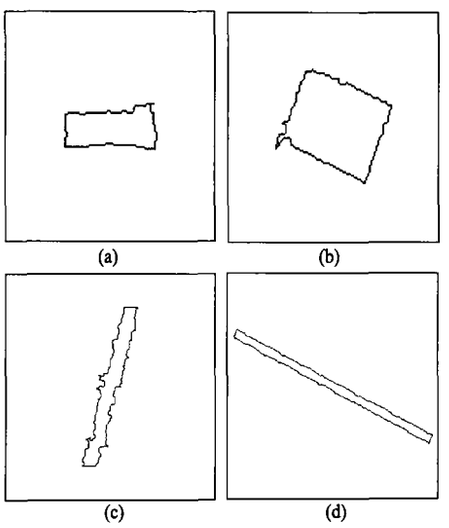

Fig. 3 Boundaries of the four objects
(a) Building; (b) Paddy; (c) Road; (d) River

4.1 Characteristics of the frequency spectrum

According to the principle of FDs and the process of normalization explained above, the boundary coordinates of the four objects are represented as sequences of complex numbers, which are then transformed using DFT. These complex coefficients are all called the original FDs, which are dealt with using the following three steps to make them insensitive to translation, scale changes and rotation as follows.

Step 1 The original complex coefficients are being MOD get their magnitudes so that they are independent of rotation;

Step 2 The magnitudes obtained through step 1 are normalized by dividing the 0th Fourier coefficient;

Step 3 The 0th Fourier coefficient is thrown away and only coefficients from the 1th were chosen. Here, the value of the 0th coefficient equals to 1 because of the normalization in step 2.

According to Eq.(3), suppose L indicates the first L terms of FDs (L begins with 0), take L as horizontal axes and magnitude as vertical axes, the frequency spectrums of the four objects' boundaries are drawn in Fig.4. It should be noted that the 0th coefficients are still in the spectrums, as is for the purpose of making it more clearly for readers to understand. Fig.5 shows the comparison between the four when only the first 1—15 terms are used. Some characteristics can be obtained according to Fig.4 and Fig.5:

(1) The consistency between Low Frequency component (LF), Medium Frequency component (MF) and High Frequency component (HF): It can be seen from Fig.4 that the distribution of LF, MF and HF of the four objects are roughly the same. LF is mainly located in $1\leqslant L\leqslant 5$, which can also be verified from Fig.5; After that, with the increase of L, the magnitudes equal to almost zero until the last several nonzero coefficients of high order, and this Almost Zero part is called MF, whose contribution to shape feature is very few; HF is mainly located in $K-6\leqslant L\leqslant K-2$, that is to say, the last 2–6 terms. In addition, there is a cocked "tail" in the spectrum when $L=K-1$, which has nothing to do with object's shape but owing to the influence of the zig-zag of raster image. So it is just thrown away in the following analysis.

(2) The similarity between spectrums of shape-identical objects: Building and paddy are both rectangle-like, it can be seen from Fig.4(a) and Fig.4(b) that the numbers of these two boundaries are both between 0 and 210, and the maximum magnitudes are both between 0 and 0.05 except the DC components; Road and river are both slender in shape and have greater elongation, the numbers of boundaries are between 0 and 510, and the maximum magnitudes are both between 0.14 and 0.2 except the DC components.

(3) The unlikeness between spectrums of spectrally identical objects: A common challenge in remote sensing is the classification of objects spectrally similar but represents physically different types of ground cover, such as building & road, or lake & river. Take building & road as an example, it can be known from the above analysis that the two have obvious differences in the numbers of the boundaries, or in the magnitudes. Therefore, it is feasible to distinguish the spectrally identical objects using shape-based method.

(4) The reducibleness of the spectrum: it has already been mentioned in Section 2 that only the low-order coefficients are enough to capture the gross shape. Meanwhile, Fig.3 shows that the shape of image objects are often very complex with lots of details. There are hundreds of points to form the boundaries,

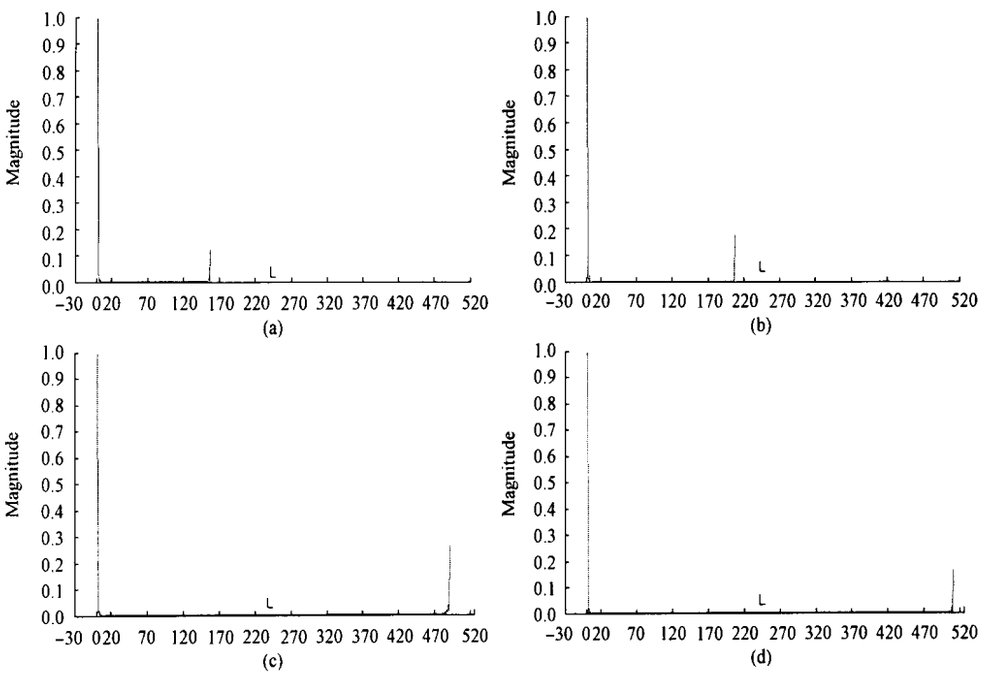

Fig. 4 Fourier Spectrums of the four objects
(a) Building; (b) Paddy; (c) Road; (d) River

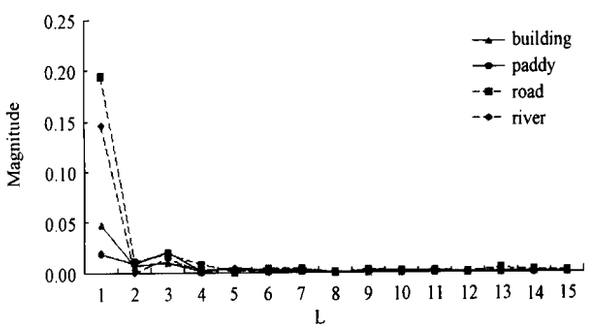

Fig. 5 The first 1—15 FDs of the four objects

however the general shape turns out to be regular. For example, roads are slender while buildings are mostly rectangle-like; rivers are curving and slender while lakes are mainly short and ellipse-like. Paddies are usually regular rectangle-like and have large areas. As a result, the numbers of FDs can be reduced and only the gross information is used to distinguish different kinds of ground truth.

4.2 Contribution rate

The frequency spectrums of the four kinds can be divided into four frequency ranges according to the characteristics: DC (L=0), LF ($1 \leqslant L \leqslant 5$), MF ($6 \leqslant L \leqslant K-7$) and HF ($K-6 \leqslant L \leqslant K-2$). The Contribution rate (Cr) of FDs with the different frequency ranges can be defined as Eq.(7):

$$Cr_L = MagZ_L \Big/ \sum_{L=0}^{K-2} MagZ_L \tag{7}$$

where Cr_L represents the Cr of the L th FD while $MagZ_L$ represents its magnitude, $0 \leqslant L \leqslant K-2$.

Table 1 The contribution rates of different frequency ranges

Frequency \ FDs \ Cr/% \ object		building	paddy	road	river
DC	0	88.73	87.57	71.87	80.71
LF	1	4.23	1.68	13.39	11.82
	2	0.60	0.85	0.76	0.00
	3	0.91	1.82	1.32	1.16
	4	0.14	0.12	0.54	0.03
	5	0.41	0.32	0.02	0.33
	total	6.29	4.79	16.03	13.34
MF	6—(K-7)	2.85	4.97	4.60	3.11
HF	K-6	0.17	0.28	0.47	0.04
	K-5	0.06	0.51	1.39	0.76
	K-4	0.18	0.10	0.82	0.03
	K-3	1.41	0.55	3.64	1.99
	K-2	0.31	1.23	1.18	0.02
	total	2.13	2.67	7.50	2.84

The calculated Cr is shown in Table 1. In the table, The DC component has the highest Cr with more than 70% for each object, and the Cr of building has even reached 88.73%; The LF and HF components have the second highest Cr with 7%—24%, take road as an example, whose Cr of LF is 16.03% and HF 7.5%; On the other hand, The MF component get the lowest Cr with merely 2%—4%, despite of its hundreds of points. In addition, because of its relation to translation of the shape, the DC component is neglected, and the HF component is also neglected due to its only representing the details of the shape. It can be seen from Fig.5 that only the LF component is enough to distinguish those spectrally identical objects, that is to say, only the LF is applied to represent the shape feature of ground object from high-resolution remotely sensed imagery.

4.3 Shape reconstruction

The results of shape reconstruction with different values of L are shown in Fig.6. As for Fig.6(a), 6(b), 6(c) and 6(d), the first figure in the top left corner stands for the original shape, and the succeeding nine figures stand for the results of shape reconstruction when L equals 0, 1, 2, 3, 4, 5, $(K-2)/4$, $(K-2)/2$ and $(K-2)$, respectively. The results show that all of them become perfect circles when $L=0$ because the DC components represent mean values. The four reconstructed shapes become smooth ellipses when $L=1$ with different flattening, building and paddy have a smaller flattening than road or river; With the gradually increasing of L, the reconstructed shapes become closer to real ones, and they have got a perfect alikeness till $L=5$ except for some details. The results of details reconstruction using MF and HF are shown when L equals $(K-2)/4$, $(K-2)/2$ and $(K-2)$, respectively, and the recurrence of shapeoccurs when $L= K-2$.

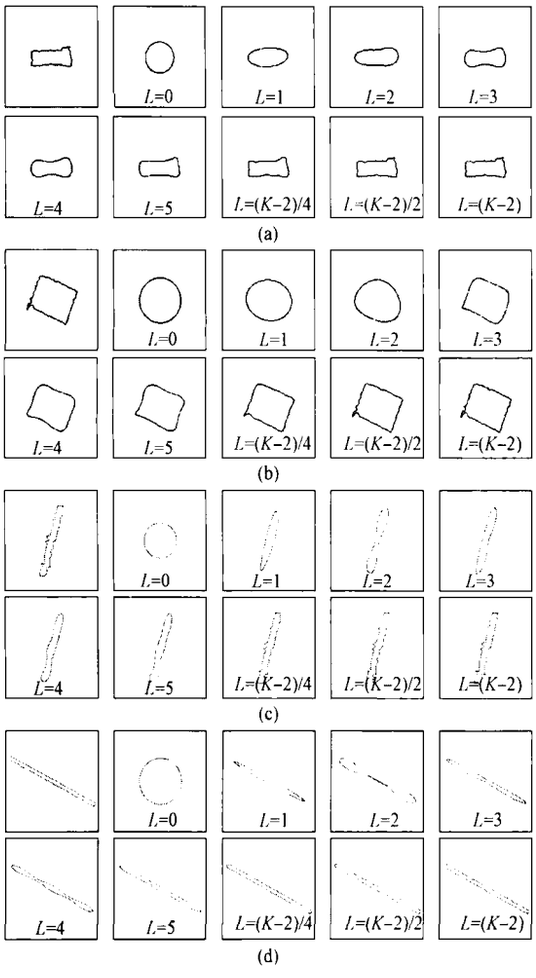

Fig. 6 Shape reconstruction with different values of L
(a) Building; (b) Paddy; (c) Road; (d) River

5 APPLICATION IN OBJECT-ORIENTED CLASSIFICATION

5.1 Classification of IKONOS multispectral image

The very high resolution IKONOS multispectral image with acquisition date of May 12, 2000 is considered, which covers some area of Nanjing (118°46'E, 32°03'N), Jiangsu province, China. The product number is 21249 and the size of the multispectral data is 1751 columns by 1751 rows pixels. The image has a good quality with cloudless condition. A test area of 476×443 pixels is chosen for classification based on decision tree. Fig.7 is pseudo color image and Fig.8 is the segmented image.

Convolve the segmented image with each band of the multispectral image so as to obtain the range of each object in each band, and then the mean gray values of each object in each band is calculated. The results are stretched into 8bit images. The image filled by mean gray values is shown in Fig.10. In the original order, band blue, green, red and NIR are marked as band G1, G2, G3 and G4, respectively; The FDs of each object in Fig.8 are calculated and the 1—5th FDs that stand for LF are

Fig. 7 The original IKONOS image

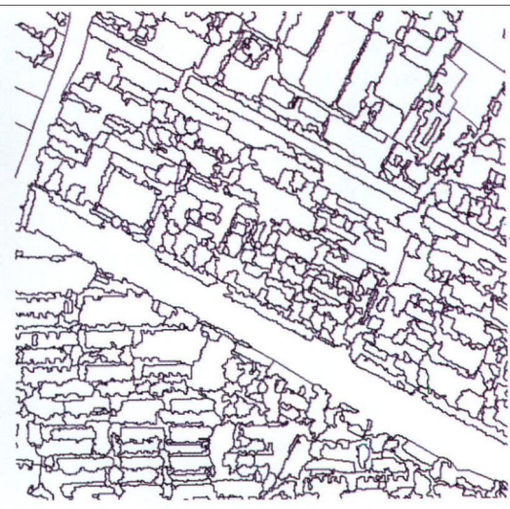

Fig. 8 Results of image segmentation

intercepted. These FDs are then evaluated to each object to form another five 8bit images, which are marked as band S_1, S_2, S_3, S_4 and S_5; Band G1, G2, G3, G4, S_1, S_2, S_3, S_4 and S_5 are stacked to form a new image which is used to build the decision tree (Fig.9). Here, shadows of buildings are divided into an independent class because of their spectrally similarity to water.

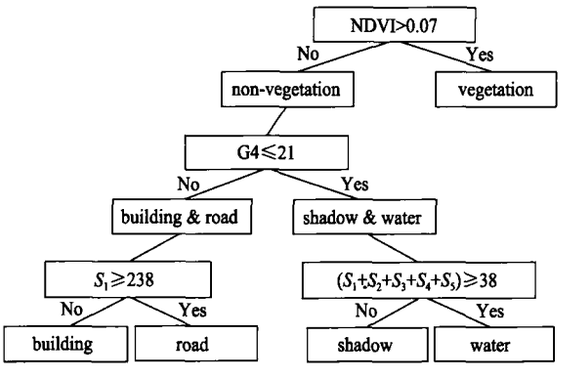

Fig. 9 The decision tree

Firstly, the image is divided into two classes based on NDVI——vegetation and non-vegetation; Then, the non-vegetation areas are divided into building & road, shadow & water based on values of NIR (G4); Finally, the spectrally identical objects i.e. building & road, shadow & water are distinguished by using the five bands S_1, S_2, S_3, S_4 and S_5. It needs to be pointed out that the threshold values in the decision tree are all stretched into 0—255 and cannot be used directly in other images or test areas. The final classification results are shown in Fig.11. Colors representing different classes are explained in Fig.9. For comparison, another classification model that involves only the mean gray value is established as a control group, whose clustering results are shown in Fig.12. It can be seen from Fig.12 that building & road, shadow & water cannot be classified correctly.

5.2 Assessment of accuracy

Qualitatively speaking, the spectrally identical objects are mainly clustered with the help of FDs in the classification. Because of its insensitivity to translation, rotation and scale changes, the normalized FDs can well describe shape feature despite of the objects' difference in location, size or direction. For example, although different buildings have different sizes, geopolitical locations and arrays, they are all parallel to regular rectangles of a certain aspect ratio, and this coherence can be better represented using FDs. Besides, from the reconstruction results of building and road in Fig.6, we can see that both of the reconstructed images are parallel to regular rectangles when L=5, but the both can be clustered respectively according to their different aspect ratios, which can be expressed in band S_1 of the decision tree in Fig.9. In Fig.11, some wrongly classified objects are marked out with the white squares. The far right square shows an object of shadow that is wrongly classified into water because of its shape-alike with river. In the same way, a water object is classified into shadow because it is split when choosing the test area, also a building object in the far left square is wrongly classified into road because of its slender in shape.

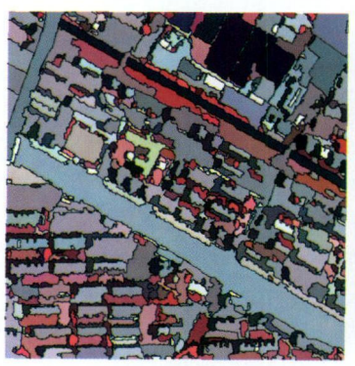

Fig. 10 Image filled by the mean gray values

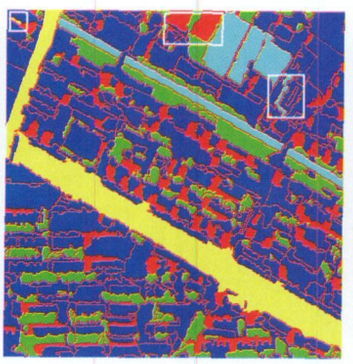

Fig. 11 Classification results

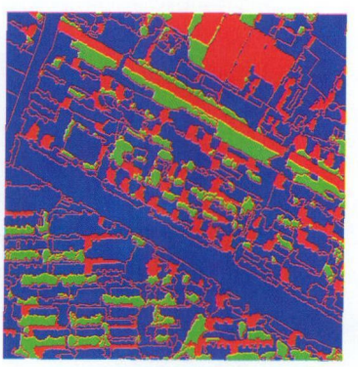

Fig. 12 Classification results of the control group

Quantitatively speaking, the assessment was done using confusion matrix and image object is take as minimum assessment unit. The confusion matrix and the Kappa coefficient are shown in Table 2: the overall classification accuracy is 98.48% and the Kappa coefficient is as high as 0.9714, just as the analysis above, this is an accuracy that models relying only on spectral feature could not achieve. Thus it can be seen that the accuracy of the classification can be better improved when adding shape information into classification model, and the FDs of LF component can well represent the shape feature of ground objects in remotely-sensed imagery.

Table 2 Confusion matrix and Kappa coefficient

the ground truth / classification results	vegetation	building	road	water	shadow	producer accuracy/%
vegetation	184	3	0	0	0	98.40
building	5	542	0	0	0	99.09
road	0	1	2	0	0	66.67
water	0	0	0	6	2	75.00
shadow	0	0	0	2	111	98.23
user accuracy/%	97.35	99.27	100	75.00	98.23	—
overall classification accuracy =845/858 = 98.48%						
Kappa =0.9714						

6 CONCLUSION

In conclusion, FDs is introduced to represent shape feature of ground objects from the high-resolution remotely sensed imagery, after normalization, shape feature representation is discussed in the Fourier frequency domain. Building, paddy, road and river are chosen, and characteristics of the frequency spectrum, contribution rate and shape reconstruction are analyzed. Finally the 1—5th FDs that stand for LF are applied to the object-oriented classification of the test area. The conclusions are as follows: (1) the DC reaches the highest contribution rate to shape with a proportion of more than70%, the LF and HF totally reach 7%—24% while the MF merely 2%—4%; (2) Only the LF components (descriptors 1—5) is enough to make a commendable reconstruction of objects' shape. The DC component is thrown away because of its relation to translation, and the HF and MF components are also not adopted; (3) The 1—5th FDs are applied to the classification model using IKONOS multispectral image of Nanjing, and the overall classification accuracy is 98.48% with a Kappa coefficient 0.9714. Applications of shape feature representation using FDs in object-oriented classification can be further explored.

REFERENCES

Benz U C, Hofmann P, Willhauck G, Lingenfelder I and Heynen M. 2004. Multi-resolution object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, **58**: 239–258

Chen F S, Fu C M and Huang C L. 2003. Hand gesture recognition using a real—time tracking method and hidden markov models. *Image and Vision Computing*, **21**: 745–758

Chen Q X, Luo J C, Zhou C H, Zheng J, Lu X J and Shen Z F. 2004. Classification of Remotely sensed imagery using multi-features based approach. *Journal of Remote Sensing*, **8**(3): 239–245

Chen Y H, Feng T, Shi P J and Wang J F. 2006. Classification of remote sensing image based on object oriented and class rules. *Geomatics and Information Science of Wuhan University*, **31**(4): 316– 320

Cheng Q, Zhuang L J and Fu B. 2008. Gait recognition based on fourier descriptors and artificial neural network. *Journal of wuhan University of Technology*, **30**(1): 126–129

Cosgriff R L. 1960. Identification of shape. *Ohio State Univ. Res. Foundation, Columbus Rep*, 820–11

Duan W, Kuester F, Gaudiot J L and Hammami O. 2008. Automatic object and image alignment using Fourier Descriptors. *Image and Vision Computing*, **26**(9): 1196–1206

Ge Y, Guo X W and Wang L Q. 2005. The application of fourier descriptors to the recognition of alphabet gesture. *Computer Applications and Software*, **22**(6): 12–13

Gonzalez R C and Woods R E. 2002. Digital image processing (Second Edition). Beijing: Publishing House of Electronics Industry: 656–657

Govindu V, Shekhar C and Chellapa R. 1998. Using geometric properties for correspondence-less image alignment. *Proceedings of the International Conference on Pattern Recognition*, **1**: 37–41

Huang X, Zhang L P and Li P X. 2007. Classification of high spatial resolution remotely sensed imagery based on the fusion of spectral and shape features. *Journal of Remote Sensing*, **11**(2): 193– 200

Liu Y, Teng X L and Liu C Q. 2005. Hand gesture recognition based on fourier descriptors with complex backgrounds. *Computer Simulation*, **22**(12): 158–161

Luo J C, Liang Y and Zhou C H. 1999. Scale space based hierarchical clustering method and its application to remotely sensed data classification. *Acta Geodaetica et Cartographica Sinica*, **28**(4): 319–324

Pierce L E, Ulaby F T and Sarabandi K. 1994. Knowledge-based classification of polarimetric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **32** (5): 1081–1086

Rui Y, Huang T S, Ortega M and Mehrotra S. 1998. Relevance Feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Video Technology*

Wang T, Liu W Y, Sun J G and Zhang H J. 2002. Using fourier descriptors to recognize object's shape. *Journal of Computer and Development*, **39**(12): 1714–1719

Wei F M, Li X W, Gu X F, Yu T and Sun Y. 2008. Shape-based classification of "spectral similar" objects in remote sensing image processing. The 14th *National Academic Conference of Image and Graphics*, 466–470

Werff H M A and Meer F D. 2008. Shape-based classification of spectrally identical objects. *ISPRS Journal of Photogrammetry & Remote Sensing*, **63**: 251–258

Wong, W and Shih F Y. 2007. Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps. *Information Sciences*, **177**: 1878–1891

Yadav R B, Nishchal N K, Gupta A K and Rastogi V K. 2007. Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study. *Optics and Lasers in Engineering*, **45**: 695–708

Zahn C T and Roskies R Z. 1972. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, **21**(3): 269–281

Zhang X Y, Yang M H, Liu C J and Wang Z X. 2008. The application of the remote sensing classification method based on object-oriented in the second land use survey. *Remote Sensing Information*, **24**(5): 36–39

基于傅里叶描述子的高分辨率遥感图像地物形状特征表达

万玮[1], 冯学智[1], 肖鹏峰[1], 赵利民[2]

1. 南京大学 地理信息科学系, 江苏 南京 210093;
2. 南京大学 国际地球系统科学研究所, 江苏 南京 210093

摘 要: 本文在对傅里叶描述子进行归一化的基础上, 将该方法引入地物轮廓的形状特征描述中, 针对建筑物、农田、道路和河道4种典型地物, 分别从谱线特征、不同频段描述子对形状特征的贡献率、形状重构三个方面进行分析, 结果表明, 在谱线图中, 直流分量对形状特征的贡献率在70%以上, 低频和高频成分共占7%—24%左右, 中频成分的贡献率只有2%—4%左右, 仅低频成分(第1—5项)便能够很好地进行地物形状重构。最后将第1—5项描述子应用到基于决策树的面向对象分类中, 得出实验区总体分类精度为98.48%, Kappa系数为0.9714。傅里叶描述子的方法能够很好的表达高分辨率遥感图像的地物形状特征。

关键词: 形状特征, 傅里叶描述子, 高分辨率, 遥感图像

中图分类号: TP751 **文献标志码:** A

引用格式: 万玮, 冯学智, 肖鹏峰, 赵利民. 2011. 基于傅里叶描述子的高分辨率遥感图像地物形状特征表达. 遥感学报, **15**(1): 73-87

Wan W, Feng X Z, Xiao P F and Zhao L M. 2011. Shape feature representation of ground objects from high-resolution remotely sensed imagery base on Fourier Descriptors. *Journal of Remote Sensing*, **15**(1): 73-87

1 引 言

高分辨率遥感图像经过分割形成地物对象。在对地物对象特征的描述中, 形状特征与光谱特征、纹理特征一样, 是非常重要的指标。在面向对象分类中, 形状特征主要是与光谱特征, 拓扑特征和纹理特征等相结合以辅助分类, 从而提高分类精度(Pierce 等, 1994; 骆建承等, 1999; Benz 等, 2004; 黄昕等, 2007; 张秀英等, 2008)。现有的研究和面向对象分类软件对形状的表达主要是基于一些形状测度指标, 如面积、长宽比、紧凑度(陈云洁等, 2006)、圆度和凸度(Werff 和 Meer, 2008)、形状指数(魏飞鸣等, 1999, 陈晓秋等, 2004)等。

傅里叶描述子(Fourier Descriptors, FDs)是Cosgriff 在 1960 年提出的一种形状特征描述方法, 它是指物体形状边界曲线的傅里叶变换系数, 是物体边界曲线信号的频域分析结果, 该方法是基于轮廓的形状特征描述方法之一(王涛等, 2002)。区域的轮廓特征可看作是一个离散的复数周期信号函数, 傅里叶变换可以将轮廓函数从空域变换到频域进行处理, 其理论基础是一维傅里叶变换。自 Zahn 和 Roskies(1972)将其应用于平面闭合曲线的形状特征描述以来, 许多专家学者已经在很多领域开展了相关的应用研究。如基于傅里叶描述子进行手势识别(Feng 等, 2003; 刘寅等, 2005; 葛元等, 2005), 步态识别(Stuart 等, 2003; 程琼等, 2008), 图像匹配(Govindu 等, 1998; Duan 等, 2008), 图像检索和分类(Rui 等, 1998; Wong 等, 2007; Yadav 等, 2007)等。但是, 将傅里叶描述子应用到遥感图像形状特征表达的研究较为少见。

本文以 IKONOS 多光谱图像为数据源, 在阐述傅里叶描述子的原理并对其进行归一化的基础上,

收稿日期: 2010-01-18; **修订日期:** 2010-04-02

基金项目: 国家高技术研究发展计划(863 计划)(编号: 2008AA12Z106); 国家自然科学基金项目(编号: 40801166); 高等学校博士学科点专项科研基金新教师课题(编号: 200802841012)。

第一作者简介: 万玮(1986—), 女, 南京大学摄影测量与遥感专业硕士研究生, 主要研究方向为遥感数字图像处理。E-mail: wanwei_4824@163.com。

使用该方法描述建筑物、农田、道路和河道 4 种典型地物轮廓，分别从谱线特征、不同频段描述子对形状特征的贡献率、形状重构 3 个方面进行分析，最后尝试将该方法得到的形状信息应用到基于决策树的面向对象分类中，达到了较高的分类精度。

2 傅里叶描述子的基本原理

图 1 为单个形状的 K 点数字轮廓，以任意点(x_0, y_0)为起点，坐标对(x_0, y_0), (x_1, y_1), (x_2, y_2), …, (x_{k-1}, y_{k-1})表示轮廓的一系列点，这些点的坐标可以表示为: $x(k)=x_k$, $y(k)=y_k$。因此，轮廓可以表示为坐标序列 $s(k)=[x(k), y(k)]$, $k=0,1,2,\cdots, K-1$。在这里，每对坐标都可以看作一个复数，即 $s(k)=x(k)+jy(k)$, $k=0,1, 2,\cdots, K-1$。这样，原来的坐标序列得到了重新的解释，将二维形状转化为一维信号，从而简化了问题。

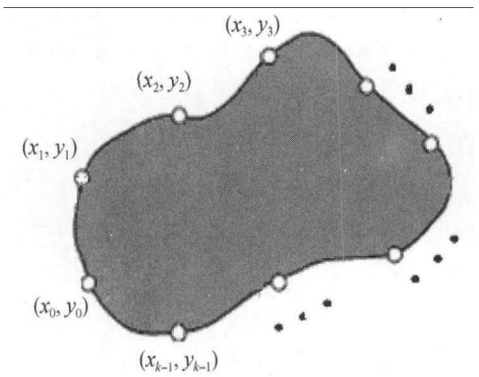

图 1 形状的数字轮廓及其复数序列

根据单变量离散函数的傅里叶变换原理(Gonzalez 和 Woods, 2002)，可以得出 $s(k)$ 的傅里叶变换式:

$$a(u)=\frac{1}{K}\sum_{k=0}^{K-1}s(k)\mathrm{e}^{-j2\pi uk/K} \tag{1}$$

式中，$u=0,1,2,\cdots, K-1$，复系数 $a(u)$即为轮廓的傅里叶描述子。同样，根据一维傅里叶逆变换原理，可以得到 $a(u)$的傅里叶逆变换 $s(k)$如:

$$s(k)=\sum_{u=0}^{K-1}a(u)\mathrm{e}^{j2\pi uk/K} \tag{2}$$

式中，$k=0,1,2,\cdots, K-1$。进一步来说，在傅里叶逆变换时，假设不使用所有的傅里叶系数，而只使用前 L 个系数，即在公式(2)中令 $a(u)=0$, $u>L-1$，结果为 $s(k)$的近似值 $\hat{s}(k)$:

$$\hat{s}(k)=\sum_{u=0}^{L-1}a(u)\mathrm{e}^{j2\pi uk/K} \tag{3}$$

式中，$k=0,1,2,\cdots, K-1$。这里，虽然只有 L 项系数用于计算 $s(k)$的每个元素 $\hat{s}(k)$，但 k 值仍然从 0 取到 $K-1$ 的值，换句话说，在近似轮廓中仍存在与原轮廓同样数目的点，但重构每个点时并不使用同样多的项，只取前面 L 项系数作为形状重构的傅里叶描述子。使用的项数越多，轮廓恢复越近似于原轮廓，这是因为在傅里叶变换中，低频成分描述对象的整体形状，而高频成分则刻画形状的细节部分。

3 傅里叶描述子的归一化

一个好的形状描述子必须对形状的平移、旋转和尺度变化不敏感(Zahn和Roskies, 1972)，这种抗敏感性的要求在遥感图像地物形状特征表达中尤为重要，比如，各种道路虽形状均为细长形，但走向各不相同，而对建筑物来说，虽是近似的规则矩形，但矩形的大小和方向也不尽相同等，所以，在遥感图像中定义的傅里叶描述子必须保证其具有平移、缩放和旋转不变性，因此本文根据其原理对傅里叶描述子做归一化处理。图 2(b), 2(c), 2(d)分别为图 1 中形状的平移、缩放和旋转示意图。

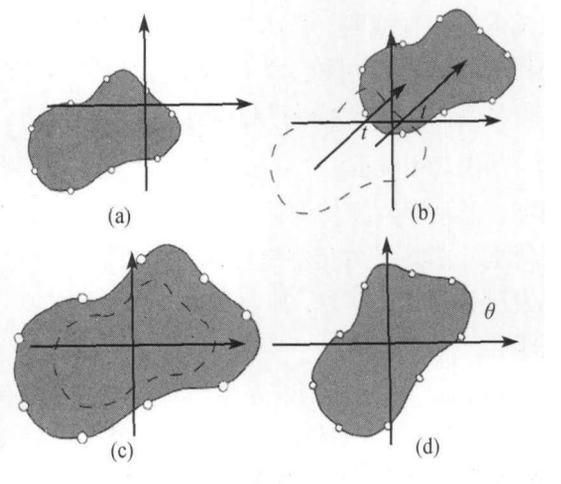

图 2 平移、缩放和旋转示意图
(a) 原形状; (b) 平移; (c) 缩放; (d) 旋转

如图 2(b)所示，将 $s(k)$平移 t 个单位后，即变为 $t+s(k)$，其中 $t=\delta x+j\delta y$。则原傅里叶变换式(1)变为:

$$\tilde{a}(u)=\frac{1}{K}\sum_{k=0}^{K-1}[t+s(k)]\times\mathrm{e}^{-j2\pi uk/K} \tag{4}$$

将式(4)化简，得式(5):

$$\tilde{a}(u)=a(u)+\sum_{k=0}^{K-1}t\times\mathrm{e}^{-j2\pi uk/K} \tag{5}$$

当 $u=0$ 时，$\tilde{a}(0)=a(0)+\sum_{k=0}^{K-1}t\times e\times 0=Kt$ 。

当 $u>0$ 时,

$$\tilde{a}(u)=a(u)+\sum_{k=0}^{K-1}t\times e^{-j2\pi uk/K}=a(u)+0=a(u),$$

可见, 平移的信息仅和 $u=0$ 时相关, 若要消除平移对形状特征表达的影响, 只需舍去傅里叶变换后第 0 项的系数, 即舍去其直流分量(Direct Component, DC)即可。

如图 2(c)所示, 将原形状按照比例放大 m 倍, 即将 $s(k)$乘以一个系数 m, 原傅里叶变换式(1)可按照如下推导过程做相应改变:

$$\begin{aligned}\tilde{a}(u)&=\frac{1}{K}\sum_{k=0}^{K-1}[m\times s(k)]\times e^{-j2\pi uk/K}\\&=m\times\frac{1}{K}\sum_{k=0}^{K-1}s(k)\times e^{-j2\pi uk/K}\\&=m\times a(u)\end{aligned}\tag{6}$$

可见, 改变后的值为原值的 m 倍, 为了方便说明, 这里假设另外有一形状为原形状放大 n 倍的结果, 则 $\tilde{\tilde{a}}(u)=n\times a(u)$, 所以有 $\frac{\tilde{a}(u)}{a(u)}=\frac{|\tilde{a}(0)|}{|a(0)|}$ 和 $\frac{\tilde{\tilde{a}}(u)}{a(u)}=\frac{|\tilde{\tilde{a}}(0)|}{|a(0)|}$, 进而可以得到 $\frac{\tilde{a}(u)}{|\tilde{a}(0)|}=\frac{\tilde{\tilde{a}}(u)}{|\tilde{\tilde{a}}(0)|}$。因此, 为解决缩放不变性问题, 可以将对象傅里叶变换后得到的傅里叶描述子均除以第 0 项的系数的绝对值, 这样便将其归一化成不受形状缩放影响的标准值。

如图 2(d)所示, 将 $s(k)$旋转 θ角度, 得到 $e^{j\theta}s(k)$, 根据一维傅里叶变换的性质可知, 形状的旋转仅影响傅里叶描述子的相位信息, 因此, 在应用中只需舍去所有的相位信息, 即对所有系数求模, 取其幅值即可。

通过以上 3 个步骤的处理, 将傅里叶描述子进行归一化, 使其具有平移、缩放和旋转不变性, 下文的分析和应用均基于归一化后的描述子。

4 基于傅里叶描述子的地物形状特征分析

研究选取建筑物、农田、道路、河道 4 种典型地物作为研究对象, 分别提取其轮廓线, 这里, 研究对象的轮廓线是指图像经过分割得到的地物对象的边界轮廓, 图3 为 4 种典型地物的边界轮廓图。首先分析由傅里叶描述子得出的谱线特征, 然后分析不同频段描述子对形状特征的贡献率, 最后分析特定系数下的地物形状的重构情况。

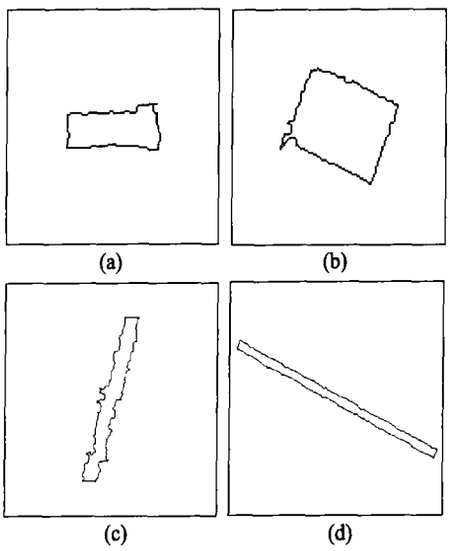

图 3 4 种典型地物的边界轮廓

(a) 建筑物; (b) 农田; (c) 道路; (d) 河道

4.1 谱线特征

根据傅里叶描述子的原理以及归一化过程, 分别将建筑物、农田、道路、河道 4 种典型地物的轮廓线坐标用复数序列表示, 对复数序列做一维傅里叶变换, 得到一系列复系数, 即为初始的傅里叶描述子, 然后对其进行平移、缩放和旋转不变性处理: (1)初始的傅里叶描述子取模, 即求其幅值, 完成旋转不变性处理; (2)对(1)中得到的幅值分别除以第 0 项的值, 即对幅值进行归一化, 完成缩放不变性处理; (3)舍去第 0 项的值即舍去直流分量, 完成平移不变性处理。由于(2)种对其进行了归一化, 因此第 0 项的值均为 1。

将处理后的傅里叶描述子和其对应的幅值作为坐标轴, 绘制 4 种地物轮廓的谱线图(图 4)。为方便下文研究不同傅里叶描述子对形状特征的贡献率, 谱线图中仍然将第 0 项(即直流分量)的值加入其中, 便于观察各系数及其对应幅值的特点。图 5 为取前面 15 项系数时 4 种地物谱线特征的对比图, 根据图 4 和图 5 中谱线的特点, 可以得出以下几个规律:

(1) 低频、中频和高频成分的一致性: 从图 4 可以得到典型地物形状的频谱一致性规律, 即除直流分量以外, 4 种地物的低频、中频和高频成分分布基本相同, 低频成分大致位于 $1\leqslant L\leqslant 5$ 之间, 即第 1–5 项系数之间, 这一点从图5 中也可以得到验证; 往后随着 L 的增加, 其各项幅值几乎均为零, 如图 5 中的第 6—15 项的幅值, 直到最后出现几项幅值不为 0 的高阶系数为止, 这部分即为中频成分, 可见它对地物形状特征的贡献是微乎其微的; 高频成分大致位于 $K-6\leqslant L\leqslant K-2$ 之间, 即倒数第 2—6 项的值。

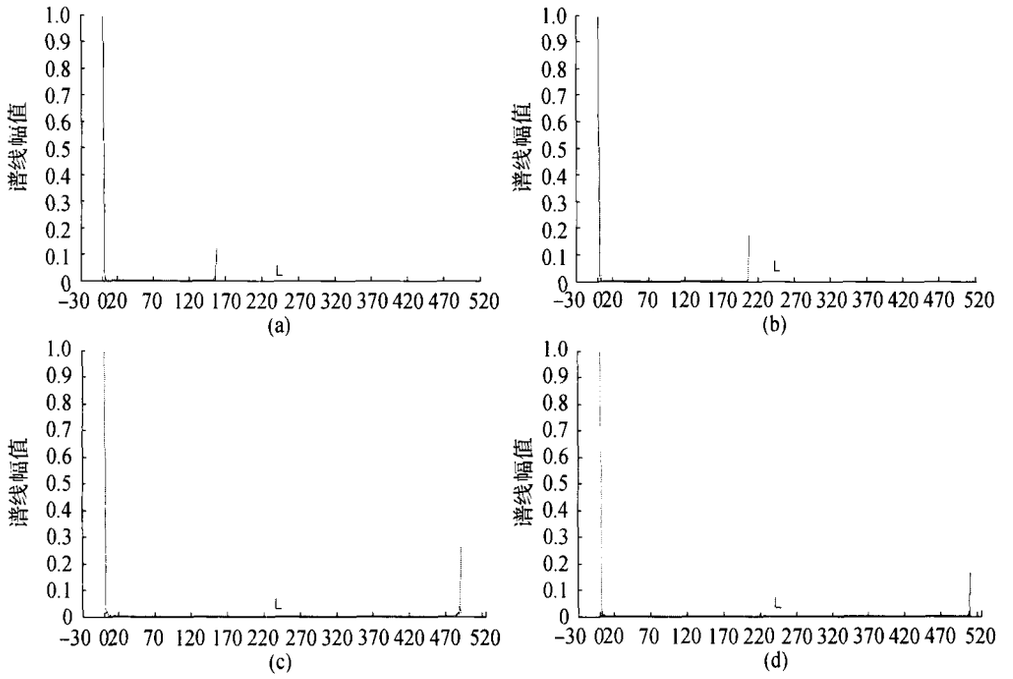

图 4 典型地物及其谱线特征图

(a) 建筑物; (b) 农田; (c) 道路; (d) 河道

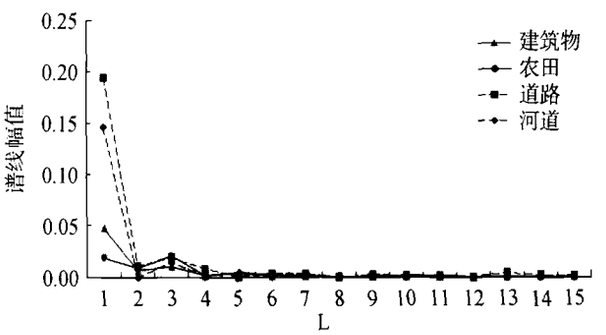

图 5 4 种地物谱线特征比较

需要说明的是, 从图中看出最后一项(即 $L=K-1$)的值很大, 致使谱线图出现一条翘起的"尾巴", 这项应是栅格图形锯齿状结构所形成的高频成分, 与地物形状本身无关, 处理过程中将其舍去。

(2) 形状类似的不同地物谱线特征的相似性: 形状类似的地物其谱线走势基本一致, 如图 5 中的建筑物和农田、道路和河道。建筑物和农田均近似呈规则的矩形, 从图 4(a)和 4(b)中看出二者轮廓线上点个数在 0—210 之间, 除直流分量幅值为 1 外, 谱线最大幅值在 0—0.05之间; 道路和河道对象均呈细长的形状, 其伸长度较大, 二者轮廓线上点个数在 0—510 之间, 除直流分量幅值为 1 外, 谱线最大幅值在 0.14—0.2 之间。

(3) 光谱类似的不同地物谱线特征的相异性: 遥感图像上存在"同谱异物"的现象, 比如建筑物和道路, 池塘和河道等等。从图 4 和图 5 中的建筑和道路谱线走势可以看出, 二者在轮廓线点个数、谱线幅值等方面均有明显的差别。由此可见, 在光谱特征类似的情况下, 借助形状特征来区分地物是可行的。

(4) 谱线的可简化性: 在前面的第 2 节中已经提到, 在傅里叶逆变换时, 可以只取前面几项的系数进行傅里叶逆变换, 便可以得到整体形状而忽略细节部分。从图 3 中可以看出, 遥感图像上地物的形状往往比较复杂, 细节成分偏多, 因而带有很多高频信息, 描述轮廓线的点个数几乎上百个, 但是整体形状却比较规则, 比如道路和河道呈细长形, 而建筑物和农田呈规则的矩形等。因此, 可以考虑减少傅里叶描述子的个数, 简化计算, 只用整体形状来区分不同的地物类型。

4.2 贡献率

对于建筑物、农田、道路和河道 4 种地物类型来说, 根据傅里叶描述子的谱线特征可以将其划分为直流分量(L=0)、低频成分($1\leqslant L\leqslant 5$)、中频成分($6\leqslant L\leqslant K-7$)和高频成分($K-6\leqslant L\leqslant K-2$)4 个频段, 若定义傅里叶描述子对地物形状特征的贡献率(Cr, Contribution rate)如式(7):

$$Cr_L = MagZ_L / \sum_{L=0}^{K-2} MagZ_L \tag{7}$$

式中, Cr_L代表第 L 个描述子的贡献率, $MagZ_L$代表第 L 个描述子对应的幅值, $0\leqslant L\leqslant K-2$。

表 1 给出了根据式(7)计算的各频段傅里叶描述子对地物形状特征的贡献率, 从表中可以看出: 直

流分量的贡献率最大, 4 种地物的均在 70%以上, 其中建筑物的高达 88.73%; 低频和高频两个频段是除直流分量以外贡献率较高的两个部分, 仅 1—5 阶和 K-6—K-2 这 10 项系数就占有了 7%—24%左右的信息量, 其中道路的低频成分贡献率高达 16.03%, 高频成分也占了 7.5%; 相对而言, 中频部分的贡献率则非常低, 中频部分描述子的数目多达一百甚至几百, 而其总贡献率却只有 2%—4%左右, 每个描述子对应的幅值都几乎为 0, 因此在应用中通常将这部分忽略不计。另外, 由于直流分量与形状的平移信息有关, 在实际应用中为确保描述子的平移不变性而将其忽略, 高频成分仅代表形状的细节部分, 从前面图 5 中可以看出仅低频成分便足以区分"同谱异物"现象, 再结合光谱信息便可以很好地对分割后的对象进行分类, 因此高频成分也不予采用。由此可以得出, 基于傅里叶描述子来描述遥感图像中地物轮廓的形状, 真正用到的是其低频成分。

表 1 各频段傅里叶描述子对地物形状特征的贡献率

频段	描述子	建筑物	农田	道路	河道
直流分量	0	88.73	87.57	71.87	80.71
低频	1	4.23	1.68	13.39	11.82
	2	0.60	0.85	0.76	0.00
	3	0.91	1.82	1.32	1.16
	4	0.14	0.12	0.54	0.03
	5	0.41	0.32	0.02	0.33
	总计	6.29	4.79	16.03	13.34
中频	6—(K-7)	2.85	4.97	4.60	3.11
高频	K-6	0.17	0.28	0.47	0.04
	K-5	0.06	0.51	1.39	0.76
	K-4	0.18	0.10	0.82	0.03
	K-3	1.41	0.55	3.64	1.99
	K-2	0.31	1.23	1.18	0.02
	总计	2.13	2.67	7.50	2.84

4.3 形状重构

4.1 和 4.2 两节针对 4 种典型地物, 从谱线特征分析和描述子对形状的贡献率两个方面做了详细阐述, 本节结合遥感图像中地物形状的特点, 对不同数量傅里叶描述子对同一地物形状的重构情况进行分析。4 种地物类型在不同 L 值时的形状重构图如图 6 所示, 对于(a)、(b)、(c)和(d)来说, 左上角第 1 幅图是地物的原始形状, 后面 9 幅图依次是当 L 取 0, 1, 2, 3, 4, 5, (K-2)/4, (K-2)/2 和 K-2 时的形状重构结果。

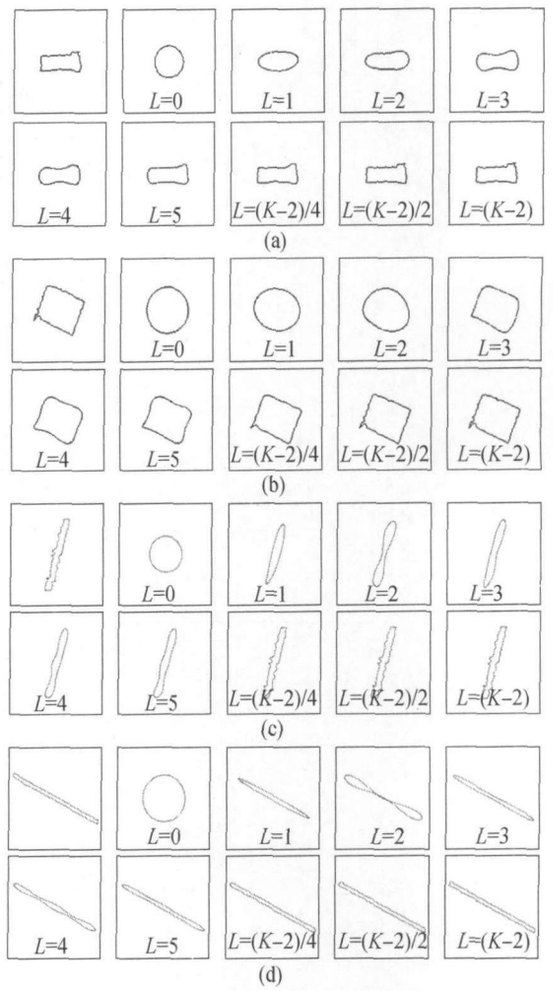

图 6 取不同数量傅里叶描述子时的形状重构
(a) 建筑物; (b) 农田; (c) 道路; (d) 河道

从图中可以看出: 当 L 取 0 时, 4 种地物的形状重构结果均为正圆形, 这是因为其直流分量代表均值信息; 当 L 取 1 时, 4 种地物的形状重构结果均为平滑的椭圆形, 不同的是, 不同地物类型椭圆的扁率有差别, 建筑物和农田的扁率较小, 而道路和河道的则较细长; 随着 L 取值的增加, 重构结果越来越接近地物的真实形状, 可以看出当 L 取 5 时, 除了没有细节信息之外, 重构结果已经接近地物的真实轮廓, 这也是对前面论述的遥感图像形状描述仅采用低频信息的一个验证; 当 L 取(K-2)/4, (K-2)/2 和(K-2)时, 可以体现中、高频部分对形状细节的恢复情况, 当 L 取(K-2)时得到的重构结果即为地物原始形状的再现。

5 在面向对象分类中的应用

5.1 IKONOS 多光谱图像分类试验

采用 2000-05-12 获取的 IKONOS 多光谱图像, 范围覆盖江苏省南京市的部分区域, 产品编号为

21249，获取时间为上午 10 点 30 分左右，波段大小为 1751×1751 像元。图像质量良好，基本无云，选取 476×443 像元的实验区进行基于决策树的分类实验。图 7 为实验区的假彩色合成图像，图 8 为多尺度分割结果，分割效果较理想，地物轮廓线清晰可见。

将分割结果图像分别与多光谱图像的每个波段做卷积，依次获取每个波段每个对象的范围，进而计算各对象各波段的灰度均值，并将其拉伸为 8bit 图像，图 10 为灰度均值填充得到的图像，按照原始图像中蓝、绿、红和近红外的顺序，分别记为波段 G1, G2, G3, G4；计算各对象边界轮廓线的傅里叶描述子，截取其低频成分即第 1—5 项的幅值，分别赋给各对象并将其拉伸为 8bit 图像，构建 5 个新的波段，记为 S_1, S_2, S_3, S_4, S_5；将 G1, G2, G3, G4, S_1, S_2, S_3, S_4, S_5 叠加成一幅新的图像，构建分类决策树，如图 9 所示。由于实验区图像中建筑物受阴影的阴影比较明显，且阴影与水体的光谱特征比较接近，因此分类过程中考虑将阴影对象归为一类。

图 7 IKONOS 原始图像

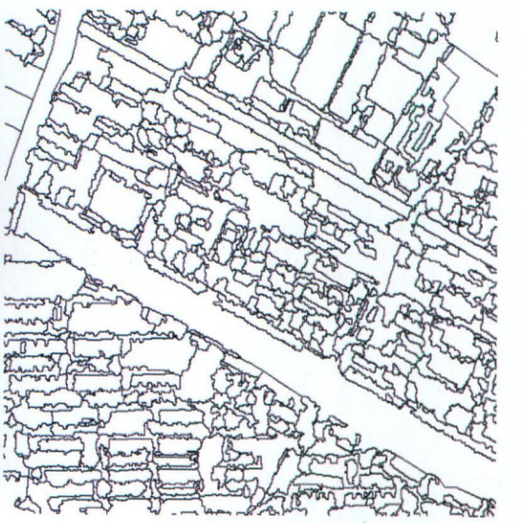

图 8 多尺度分割结果

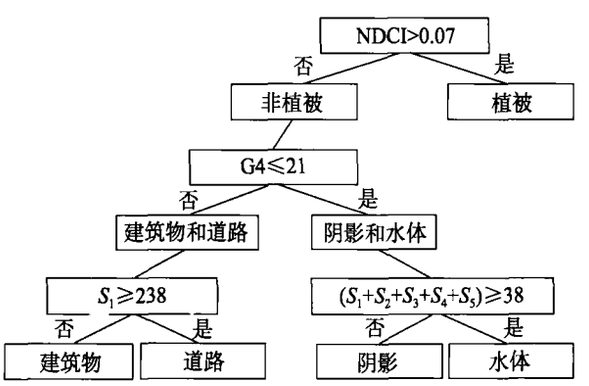

图 9 面向对象分类的决策树

首先根据 NDVI 的值将图像分为植被区和非植被区，然后根据近红外即 G4 波段的值进一步分为建筑物和道路、阴影和水体，最后根据形状特征波段 S_1, S_2, S_3, S_4, S_5 来区分光谱特征相似的建筑物和道路、水体和阴影。这里各阶段所取的阈值均为拉伸到 0—255 范围内的值，对于其他的影像和实验区该值仅供参考。图 11 最终分类结果，各颜色代表的类别如图 9 的决策树所示。为更好地突出形状特征对分类的贡献，研究将仅利用灰度均值进行分类的结果作为对照，图 12 为分类对照图。从图中可以看出，仅采用灰度均值作为分类参数时，道路和建筑物、水体和阴影均不能得到有效地区分。

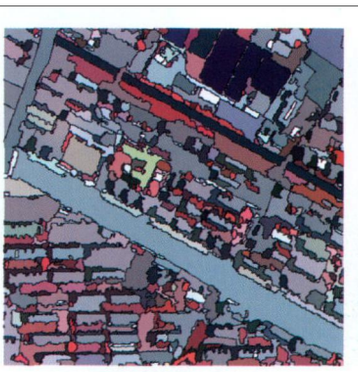

图 10 灰度均值填充结果

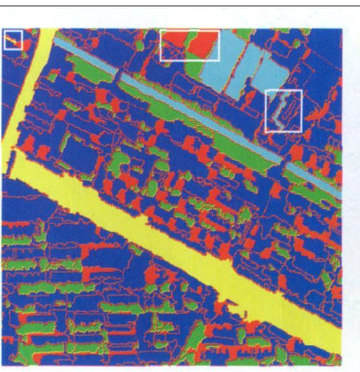

图 11 分类结果

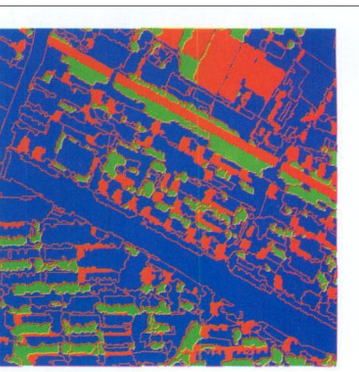

图 12 分类结果对照

5.2 精度评价

定性地来说，根据光谱特征将实验区遥感图像分为植被、建筑物和道路、阴影和水体等3类后，主要是基于构建的形状描述子来进一步区分"同谱异类"的地物类别。由于归一化后的傅里叶描述子具有很好的平移、缩放和旋转不变性，同类地物对象虽在位置、大小和方向等方面有所不同，但并不影响描述子对同类地物形状一致性的表达。比如，虽然各建筑物对象在影像中的位置不同，对象的大小和排列各异，但它们均近似于一定长宽比的规则矩形，这种一致性可以通过描述子很好地表达出来。另外，从图 6 的建筑物和道路的形状重构结果可以看出，当L取 5 时建筑物和道路的重构结果均近似于规则矩形，但根据矩形的长宽比差异便可以有效地区分这两类地物，在图9的决策树中则表现为波段S_1的差异。图 11 中用白色方框标出了个别明显的错分现象，最右边的方框里一个阴影对象被错分为水体，这是由于该对象形状细长，与河道的形状类似而错分；同样，中间方框里两个水体对象因实验区截取时被割裂而错分为阴影，左边方框里一个建筑物对象因形状细长而被错分为道路。

采用混淆矩阵法对分类结果的精度进行评价，以图像对象为最小评价单位，表 2 为分类混淆矩阵及 Kappa 系数。从表中可以看出：实验区总体分类精度为98.48%，Kappa系数为0.9714；由于加入了形状特征，使道路和建筑物、水体和阴影这些光谱特征相似的区域得到了很好的区分。由此可见，在分类模型中加入形状特征可以提高分类精度，而低频成分的傅里叶描述子能够很好地描述遥感图像中地物轮廓的形状。

表 2 分类混淆矩阵及 Kappa 系数

地物类型 / 分类结果	植被	建筑物	道路	水体	阴影	生产者精度/%
植被	184	3	0	0	0	98.40
建筑物	5	542	0	0	0	99.09
道路	0	1	2	0	0	66.67
水体	0	0	0	6	2	75.00
阴影	0	0	0	2	111	98.23
用户精度/%	97.35	99.27	100	75.00	98.23	—
总体分类精度=845/858 = 98.48%						
Kappa系数=0.9714						

6 结 论

本文将傅里叶描述子的思想引入遥感图像的形状特征表达中，在对其进行归一化处理的基础上，从频域的角度探索高分辨率遥感图像的形状特征表达方法，针对建筑物、农田、道路和河道 4 种典型地物，分别从谱线特征、不同频段描述子对形状特征的贡献率、形状重构3个方面进行分析，将第1—5项描述子应用到基于决策树的图像分类中。研究结论如下：(1)在所有的傅里叶描述子中，直流分量对形状特征的贡献率最大，在 70%以上，其次是低频和高频成分，共占 7%—24%左右，中频成分贡献率只有 2%—4%，几乎可以忽略；(2)基于傅里叶描述子来表达遥感图像中地物轮廓的形状，仅低频成分(第 1—5 项)便能够很好地进行地物形状重构，直流分量因其代表平移信息而需舍去，高频部分和中频部分也不予采用；(3)将傅里叶描述子的低频成分(第1—5 项)所表示的形状信息加入到分类模型中，可以得到较高的分类精度，对于实验区的南京市 IKONOS 多光谱图像，其总体分类精度为 98.48%，Kappa 系数为 0.9714。

可见，利用傅里叶描述子的方法能够很好地表达遥感图像中地物轮廓的形状特征。下一步的研究可以继续探索这种形状特征表达方法在高分率遥感图像面向对象分类中的应用。

REFERENCES

Benz U C, Hofmann P, Willhauck G, Lingenfelder I and Heynen M. 2004. Multi-resolution object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, **58**: 239–258

Chen F S, Fu C M and Huang C L. 2003. Hand gesture recognition using a real—time tracking method and hidden markov models. *Image and Vision Computing*, **21**: 745–758

Chen Q X, Luo J C, Zhou C H, Zheng J, Lu X J and Shen Z F. 2004. Classification of Remotely sensed imagery using multi-features based approach. *Journal of Remote Sensing*, **8**(3): 239–245

Chen Y H, Feng T, Shi P J and Wang J F. 2006. Classification of remote sensing image based on object oriented and class rules. *Geomatics and Information Science of Wuhan University*, **31**(4): 316– 320

Cheng Q, Zhuang L J and Fu B. 2008. Gait recognition based on fourier descriptors and artificial neural network. *Journal of wuhan University of Technology*, **30**(1): 126–129

Cosgriff R L. 1960. Identification of shape. *Ohio State Univ. Res. Foundation, Columbus Rep*, 820–11

Duan W, Kuester F, Gaudiot J L and Hammami O. 2008. Automatic object and image alignment using Fourier Descriptors. *Image and Vision Computing*, **26**(9): 1196–1206

Ge Y, Guo X W and Wang L Q. 2005. The application of fourier descriptors to the recognition of alphabet gesture. *Computer Applications and Software*, **22**(6): 12–13

Gonzalez R C and Woods R E. 2002. Digital image processing (Second

Edition). Beijing: Publishing House of Electronics Industry, 656–657

Govindu V, Shekhar C and Chellapa R. 1998. Using geometric properties for correspondence-less image alignment. *Proceedings of the International Conference on Pattern Recognition*, **1**: 37–41

Huang X, Zhang L P and Li P X. 2007. Classification of high spatial resolution remotely sensed imagery based on the fusion of spectral and shape features. *Journal of Remote Sensing*, **11**(2): 193– 200

Liu Y, Teng X L and Liu C Q. 2005. Hand gesture recognition based on fourier descriptors with complex backgrounds. *Computer Simulation*, **22**(12): 158–161

Luo J C, Liang Y and Zhou C H. 1999. Scale space based hierarchical clustering method and its application to remotely sensed data classification. *Acta Geodaetica et Cartographica Sinica*, **28**(4): 319–324

Pierce L E, Ulaby F T and Sarabandi K. 1994. Knowledge-based classification of polarimetric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, **32** (5): 1081–1086

Rui Y, Huang T S, Ortega M and Mehrotra S. 1998. Relevance Feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Video Technology*

Wang T, Liu W Y, Sun J G and Zhang H J. 2002. Using fourier descriptors to recognize object's shape. *Journal of Computer and Development*, **39**(12): 1714–1719

Wei F M, Li X W, Gu X F, Yu T and Sun Y. 2008. Shape-based classification of "spectral similar" objects in remote sensing image processing. The 14th *National Academic Conference of Image and Graphics*, 466–470

Werff H M A and Meer F D. 2008. Shape-based classification of spectrally identical objects. *ISPRS Journal of Photogrammetry & Remote Sensing*, **63**: 251–258

Wong, W and Shih F Y. 2007. Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps. *Information Sciences*, **177**: 1878–1891

Yadav R B, Nishchal N K, Gupta A K and Rastogi V K. 2007. Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study. *Optics and Lasers in Engineering*, **45**: 695–708

Zahn C T and Roskies R Z. 1972. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, **21**(3): 269–281

Zhang X Y, Yang M H, Liu C J and Wang Z X. 2008. The application of the remote sensing classification method based on object-oriented in the second land use survey. *Remote Sensing Information*, **24**(5): 36–39

附中文参考文献

陈秋晓, 骆剑承, 周成虎, 郑江, 鲁学军, 沈占锋. 2004. 基于多特征的遥感影像分类方法. 遥感学报, **8**(3): 239–245

陈云浩, 冯通, 史培军, 王今飞. 2006. 基于面向对象和规则的遥感影像分类研究. 武汉大学学报(信息科学版), **31**(4): 316–320

程琼, 庄留杰, 付波. 2008. 基于傅里叶描述子和人工神经网络的步态识别. 武汉理工大学学报, **30**(1): 126–129

高丽, 杨树元, 李海强. 2007. 一种基于标记的分水岭图像分割新算法. 中国图像图形学报, **12**(6): 1025–1032

葛元, 郭兴伟, 王林泉. 2005. 傅里叶描述子在手势识别中的应用. 计算机应用与软件, **22**(6): 12–13

黄昕, 张良培, 李平湘. 2007. 融合形状和光谱的高空间分辨率遥感影像分类. 遥感学报, **11**(2): 193–200

刘寅, 滕晓龙, 刘重庆. 2005. 复杂背景下基于傅里叶描述子的手势识别. 计算机仿真, **22**(12): 158–161

骆剑承, 梁怡, 周成虎. 1999. 基于尺度空间的分层聚类方法及其在遥感影像分类中的应用. 测绘学报, **28**(4): 319–324

王涛, 刘文印, 孙家广, 张宏江. 2002. 傅里叶描述子识别物体的形状. 计算机研究与发展, **39**(12): 1714–1719

魏飞鸣, 李小文, 顾行发, 余涛, 孙源. 2008. 基于形状参数的遥感图像"同谱异物"目标区分. 第十四届全国图象图形学学术会议论文集, 466–470

张秀英, 杨敏华, 刘常娟, 王振兴. 2008. 面向对象遥感分类新技术在第二次土地调查中的应用. 遥感信息, **24**(5): 36–39