

MEASURES OF INFORMATION IN REMOTE SENSING IMAGERY AND AREA-CLASS MAPS

Zongjian Lin^a, Yan Chen^b, Jingxiong Zhang^b, Bing Deng^b

^aChinese Academy of Surveying and Mapping, 16 Beitaping Road, Beijing 100039

^bSchool of Remote Sensing Information Engineering of Geomatic Engineering, Wuhan University, 129 Luoyu Road, Wuhan 43007

Commission II, WG II/7

KEY WORDS: Measures of Information, Remote Sensing Imagery, Area-class Maps, Uncertainty, Correlation, Equivocation

ABSTRACT:

This paper discusses quantitative evaluation of information in remote sensing and GIS data. The formula for calculating information contents of remote sensing imagery are developed, taking into account of signal-noise-ratio (SNR), inter-pixel and cross-band correlations, while information amounts in GIS data are calculated by referring to geometric position, attributes, and the equivocation in both. Tests were carried out to compute and analyze information contents in remote sensing imagery and area-class maps produced by four different methods. It was found that amounts of information in resultant area-class maps are consistent with classifiers' performances, and are decreased in comparison with the source imagery. Further research will be directed towards the relationships between information contents and uncertainty.

1. INTRODUCTION

With the advancements of remote sensing, it is possible to acquire large quantity of imagery at increasing spatial, spectral, and temporal resolution. However, it remains an issue as how to utilize the information to the fullest extent. By calculating the amount of information contained in remote sensor images, it becomes rational to perform integrated analysis of and information extraction from remote sensing data and GIS data.

At present, the amount of information is generally calculated on the basis of the famous Shannon entropy in information theory. Information theory has been proven to be very important and found wide applications in different fields. In fact, the amount of information in remote sensing image is associated with some factors, such as degrees of gray quantification, noise, geometrical distortion, correlation of neighbouring pixels, correlation of wave bands, and etc. The quantitative measurement of a categorical map also needs to consider the attributes, position and the equivocation of the both.

In this paper, with full consideration of factors involved, the information of remote sensing imagery and classification maps are discussed in detail.

2. METHODS

2.1 Uncertainty and information amounts

At present, the amount of information is generally calculated on the basis of the famous Shannon entropy in information theory. Shannon pointed that there is certain relationship between the amount of information and the degree on which the uncertainty of source information has been eliminated. The more uncertainty is eliminated, the more information is available. So the amount of information is also called as information entropy

or entropy. Information entropy is a quantitative measure of the information transmitted by the image (Du-Yih Tsai, 2007).

Suppose object X exists in n different states with different probabilities of occurring $p_1, p_2, \dots, p_i, \dots, p_n$. Uncertainty in X can be computed by the formula of entropy as follows:

$$E = \sum_{i=1}^n p_i \log \frac{1}{p_i} = - \sum_{i=1}^n p_i \log p_i \quad (1)$$

If the base 2 is used the resulting units may be called binary digits, or more briefly bits.

All the information can be acquired by observation of this object in all states, that is to say, all uncertainty of this object is eliminated. If the object is observed in k ($k < n$) states, then the information obtained can be represented as following formula:

$$H = - \sum_{i=1}^k p_i \log p_i \quad (2)$$

This remaining uncertainty of the object is:

$$U = E - H \quad (3)$$

Considering the existence of noise, the content of information is:

$$H = H_{(\xi)} - H_{(\eta)} \tag{4}$$

where $H_{(\xi)}$ is the entropy of signal, $H_{(\eta)}$ is the entropy of noise.

2.2 Information amounts of remote sensing imagery

The digital image of remote sensing is the discrete source with memory. It is difficult to express the memorial source. So we suppose that signals are correlative with the previous m signals, but independent of others. And the Markov source has the characteristics. The digital image is a Markov information source. However, information amount of remote sensing image involves in some factors besides correlation of neighboring pixels. Other factors such as degrees of gray quantification, noise, geometrical distortion should be considered.

2.2.1 The average information amount of single-pixel greyscale image

Remote sensing image is made up of pixels with different greyscale. Information entropy of remote sensing image has relation to change degree of grey. Therefore the amount of information of remote sensing image should be taken into the entropy, noise equivocation, and correlation.

Some experiments prove that both the signal and the noise of remote sensing image are belong to Gaussian distribution, suppose their variance are δ_{ξ}^2 and δ_{η}^2 . Therefore, noise equivocation of single-pixel can be expressed as (Lin,1988):

$$H_{(g)} = H_{(\xi)} - H_{(\eta)} = \ln \sqrt{2\pi e \delta_{\xi}^2} - \ln \sqrt{2\pi e \delta_{\eta}^2} = \ln(\delta_{\xi} / \delta_{\eta}) \tag{5}$$

It shows that the key of noise equivocation lies in signal-noise-ratio of single-pixel.

As mentioned above, the digital remote sensing image is a Markov information source, and it has the statistical features of 1st order Markov process. The covariance matrix is:

$$C = \delta^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & & \\ \rho^2 & \rho & 1 & & \\ \vdots & & & & \vdots \\ \rho^{n-1} & & & \dots & 1 \end{bmatrix} \tag{6}$$

where δ^2 is the variance of signal; ρ is the autocorrelation coefficient between neighboring pixels.

Figure 1 shows the correlation between the neighboring pixels according to the characteristics of 1st order Markov. In Figure.

1, “1” denotes central pixel, ρ is the correlation coefficient between the neighboring pixels and the central pixel when $\tau = 1$, and ρ^2 is the correlation coefficient between the neighboring pixels under $\tau = 2$ and the central pixel. According to information theory, the mutual information shared between two neighboring pixels is (xie, 1985)

$$I(\tau, \tau + 1) = -\ln(1 - \rho) \tag{7}$$

If we add t pixels to the central pixel, then the information added becomes

$$H_{(t)} = t \ln(\delta_{\xi} / \delta_{\eta}) + t \ln(1 - \rho) \tag{8}$$

It is worth noting that the first term of expression (8) is positive, and its second term is negative. Therefore we considered that the average information of each pixel was decreased $\ln(1 - \rho)$ comparing with single pixel on account of the correlation of neighboring pixels.

ρ^2	ρ^2	ρ^2	ρ^2	ρ^2
ρ^2	ρ	ρ	ρ	ρ^2
ρ^2	ρ	1	ρ	ρ^2
ρ^2	ρ	ρ	ρ	ρ^2
ρ^2	ρ^2	ρ^2	ρ^2	ρ^2

Figure 1 Correlations of Neighboring Pixels

So the average information of single-pixel greyscale image can be expressed as:

$$H' = H - H_0 - H_1 \tag{9}$$

Where H is the entropy of image
 H_0 is the noise equivocation
 H_1 is the mutual information shared between two neighboring pixels.

If the number of pixels is $m \times n$, the total information of single-band is:

$$H_f = m \times n \times H \tag{10}$$

2.2.2 Information in multi-spectral remote sensing imagery

The mutual correlation between bands determines the mutual information shared between them. Assuming that the synthetic

color images is composed by the 3th, 4th and 5th bands, and the total information of the images at the 3th, 4th and 5th bands are respectively H_3 , H_4 and H_5 , then considering their correlation, the total information of the three bands is

$$H = H_3 + H_4(1 - \rho_1) + H_5(1 - \rho_2)(1 - \rho_3) \quad (11)$$

Where ρ_1 is the correlation between the 3th band and the 4th band; ρ_2 is the correlation between the 3th band and the 5th band; ρ_3 the correlation between the 4th band and the 5th band. So the amount of information of multi-spectral remote sensing image can be calculated.

2.2.3 The influence of geometric distortion

Geometric distortion of remote sensing images results in the deviation of pixel position from its position of the corresponding ground targets. If the distortion offset is τ , displacement equivocation will be expressed as:

$$H_\tau = \ln \sqrt{4\pi e \delta^2 (1 - \rho^\tau)} \quad (12)$$

This formula shows that, the smaller the correlation coefficient is or the bigger the grayscale deviation is, the poorer the tolerance on geometric distortion is. So, the geometric correction of high-resolution images is very important. The geometric correlation is helpful to reducing the equivocation H_τ , and then to increase the amount of information.

2.3 Information amounts of maps

2.3.1 Positional data

The amount of geometric figures involves in the information of position and attribution (Lin 2006).

As is shown in figure 2, a square is divided into $2^k \times 2^k$ grids. If a point is located on a grid of this square, the information to describing this point is $2k$ bit. In other words, the uncertainty of location is the logarithmic function of the spatial resolution. If this point is painted a colour which is one of 2^c colours with the same probability, then the amount of information obtained is C bit. Assuming that each colour represents an attribute of the geography object, and there are n geography attributes according to the probability in the expression (1), then attributes information of this square is:

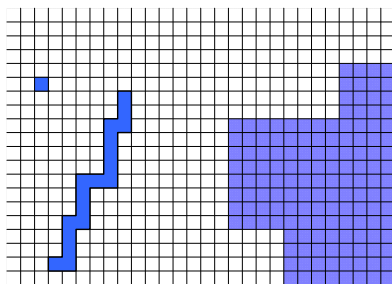


Figure2 Point, line and area

$$H = -\sum_{i=1}^n p_i \log_2 p_i \quad (13)$$

So the information of the point is $C_1 + 2k$. Given the key points determined area contour are s , information amount of this area is $C_2 + 2ks$. And the line has l key points. Thus the information of all figures is $2ks + 2k + 2kl + C_1 + C_2 + C_3$.

2.3.2 Attribute equivocation

The quantitative measurement of a map also needs to take attribute equivocation into account. The equivocation of attribute describes the existing uncertainty after the attribute is known. Assuming each attribute of the object is represented by a colour, there are n categories of attribute and the probability of these attributes are $p_1, p_2, \dots, p_i, \dots, p_n$. Then the quantitative measurement of attribute information in the image can be expressed as follow:

$$H = -\sum_{i=1}^n p_i \lg p_i \quad (14)$$

If the accuracy ratio of n categories attribution expressed by image is q_1, q_2, \dots, q_n , the amount of information and the equivocation of the image can be calculated by following formula:

$$H_{(q_n)} = -\sum_{i=1}^n q_i p_i \lg p_i \quad (15)$$

$$V = H_{(r_n)} = -\sum_{i=1}^n (1 - q_i) p_i \lg p_i \quad (16)$$

2.3.3 Position equivocation

Map includes the error $(\Delta x, \Delta y)$ from surveying and mapping process. Because the distribution of the most of measurement errors are close to normal distributions, therefore, position equivocation of map is:

$$H(\Delta x) = H(\Delta y) = \ln \sqrt{2\pi \delta^2} \quad (15)$$

$$H(\Delta p) = H(\Delta x) + H(\Delta y) = 2 \ln \sqrt{2\pi \delta^2} = 2(1.42 + \ln \delta) \quad (16)$$

Where δ is the standard deviation from map coordinate; $H(\Delta x)$ and $H(\Delta y)$ are coordinate equivocation; $H(\Delta p)$ is position equivocation.

So the overall information of figure 2 taken into the equivocation is $[2k - 2(1.42 + \ln \delta)][l + s + l] + C_1 + C_2 + C_3 - V_1 - V_2 - V_3$.

3. EXPERIMENT AND ANALYSIS

The study region is an area of central Montana, USA. It is situated at about $46^{\circ}29' \sim 48^{\circ}23'$ north latitude, $107^{\circ}52' \sim 111^{\circ}08'$ east longitude, over a highly rugged terrain almost exclusively covered by vegetation. The number of vegetation type is seven. They are low cover grasslands, moderate/ high cover grasslands, mixed broadleaf / cottonwood forest, mixed conifer forest, limber pine/ ponderosa pine, douglas-fir/ lodgepole pine, and xeric shrublands/rock. The size of the image is 1000×1000 .

Information amount of source imagery and its result maps produced by four different classification methods are compared using the different calculation method. The classification map is regarded as remote sensing image, which is one calculation method. And another is that classification map is regarded as geographical information system data. In this article, classification methods include mahalanobis distance, minimum distance, spectral angle mapper and parallelepiped. The source Landsat TM image consists of the 3th, 4th and 5th bands after feature selection. Two results are compared by table 2 and table 3. The source image and classification maps are shown in figure 3.

Table 1 and table 2 show the single pixel's information of the image and the classification images respectively. In table 1 ρ is the auto-correlation coefficient of image; $H_{(a)}$ is the information amount of certain single band; ρ' is the correlation coefficient of two bands. So the total information is 3666800 bit, and the total data quantity is 24000000 bit. In table 2 the total information amount of classification map adopt mahalanobis distance is 1388100 bit. Compared with the results of table 1 and table 2, we can draw a conclusion that the information amount reduces after classification. This is consistent with the fact because a lot of information lost after classification. In table 3 the another calculation method is used. The key of this method is to determine the number of key points.

Firstly, the noises in the classification images are filtered out by post-treating. Secondly, the boundaries of all the regions assigned by specific class labels which are sequentially judged pixel by pixel. In order to figure out whether the current pixel pertaining to the boundary or not, each pixel is checked by using a 3×3 window whose centre is located at the current pixel. In other words, if there exists any different class pixel in the window, then the central pixel will be set as boundary or background. Thirdly, each boundary of the specific class region is tracked into vector line, and Douglas - Puke vector method is used to remove redundant points. The remaining points are the key points.

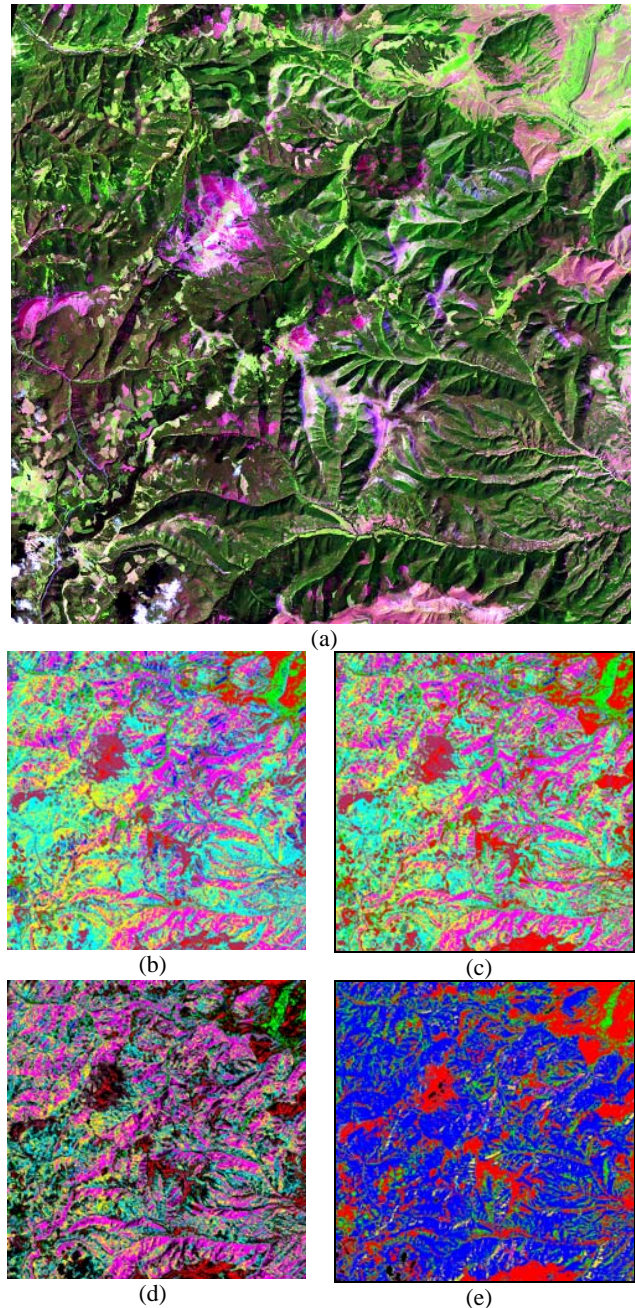


Figure 3 (a) source image (TM image)
 (b) the categorical image based on the mahalanobis distance
 (c) the categorical image based on the minimum distance
 (d) the categorical image based on the spectral angle mapper
 (e) the categorical image based on the parallelepiped

Table 3 shows the total information employed this method. Compared the results in table 2 and 3, we can draw the conclusion that the total information using this method is greater than that of other method. When we are interested in different objects, information amount varies. Specifically, area-class information is the object of interest in the first method, while boundary information is more concerned in the second.

Image (a)	Entrop _y	ρ	$H_{(a)}$	ρ'
Band 3	4.8453	0.9306	2.1774	0.56663
Band 4	5.8870	0.9369	3.1240	0.70313
Band 5	6.6624	0.9389	3.8672	0.88197

(Single pixel's information is 3.6668 bit, data quantity is 24 bit)

Table 1 Single pixel's information of source image

Classification methods	Entropy	ρ	Single's information
Mahalanobis distance	2.5213	0.6780	1.3881
Minimum distance	2.5288	0.7201	1.2555
Spectral angle mapper	2.5099	0.6312	1.5124
Parallelepiped	1.7367	0.7131	0.4881

Table 2 Single pixel's information of classification images

Classification methods	Classification Accuracy(%)	Information
Mahalanobis distance	45.591	3530600
Minimum distance	44.7368	3506500
Spectral Angle mapper	28.5178	2640800
Parallelepiped	19.137	2339700

Table 3 The total information of classification images

Classification methods	attributes information	position information
Mahalanobis distance	7.3870	3530600
Minimum distance	7.6311	3506500
Spectral angle mapper	3.9542	2640800
Parallelepiped	5.7883	2339700

Table 4 The attributes and position information

We come to some conclusions from tables 1, 2, 3, and 4 as follows:

- (1) The amount of information of an image is related to a number of factors as discussed previously, not equal to data quantity of the image. Generally, data quantity is much greater than information amount.
- (2) Comparing the results in table 2 and 3, as for the same image, different classification methods will result in

different information amounts of maps. The higher the accuracy of classification is, the greater the corresponding information amount is, if the classification image is regarded as GIS data.

- (3) From table 3, it can be seen that the information amount of images is mainly determined by position information as there is obvious complexity of class distributions in the map. At the same time, the number of classes of a map is usually not too many. If the size of map is rather large, attribution information may be neglected. The corresponding position information is bigger for the images with higher classification accuracy.
- (4) Comparing the information between source image and classification maps, we can find that it decrease in information amount, which is coincident with the facts, because in course of classification some information is lost.
- (5) The uncertainty of position and attribute can be measured by uniform information theory.

To calculate and analyze the information quantity of remote sensing imagery and area-class maps is a very important but complex problem. In this paper, we analyze the factors which affect the calculation of information amount, such as degrees of gray quantification, noise, geometrical distortion, correlation of neighbouring pixels, correlation of wave bands, and so on. A series of formulas to calculate information amount of remote sensing imagery are derived. At the same time, information amount of the classification image is calculated using two methods. The result shows that information amount will be different with different image/map contents or purposes. And we can establish a uniform criterion for measuring information amount of remote sensing imagery and map.

REFERENCES

- Du-Yih Tsai, Yongbum Lee, Eri Matsuyama, 2007. Information Entropy Measure For Evaluation of Image Quality. Journal of Digital Imaging, 0(0), pp.1-10.
- LIN Zongjian, ZHANG Yonghong, 2006. Measurement of Information and Uncertainty of Remote Sensing and GIS Data, Geomatics and Information Science of Wuhan University, 31(7), pp. 569-572.

XIE Zhongjie, 1985. Probability Theory. Posts & telecom press.

ACKNOWLEDGEMENTS

The research is partially supported by a "973 Program" grant (2007CB714402-5), and a grant received from the Key Lab in Geospatial Information Engineering, Chinese Academy of Surveying and Mapping, State Bureau of Surveying and Mapping (2007020).

