AN APPROACH FOR EVALUATING THE INFORMATION CONTENT OF REMOTE SENSING IMAGES

Shimin Fang [1], Xiaoguang Zhou[1,*]

[1]Department Geo-Informatics, Central South University, Changsha 410083, China -zxgcsu@foxmail.com

Commission IV, ICWG IV/III, WG IV/4

KEY WORDS: Remote sensing images, Information entropy, Neighborhood information, Scatterplot matrix, Geometrical mapping model

ABSTRACT:

Due to being affected by the rapid development of open science and the increasing popularity of mobile devices (e.g., smartphones), remote sensing data as frequently used data sources are broadly applied to our daily life. At the same time, remote sensing data collection also presents a trend of popularization. To improve the utilization efficiency and availability of the obtained diversified remote sensing data, we propose a novel evaluation method based on information theory and scatterplot mapping model, i.e., geometrical mapping entropy (GME). The goal is to construct a unified model of measurement to be much more effectively and accurately evaluate the information content and quality of remotely sensed imagery. Different experimental data are used to verify the performance of the proposed method, i.e., a group of the dataset that contains different four types of images; the other group of image data contains the images with different modalities and different imaging times (2016-05, 2017-08, 2018-04, and 2018-06). Experimental results indicate that the proposed approach can better characterize the spectrum features and spatial structural features contained in images and visual perception information. Additionally, it can also reflect the difference in the quality of different modality images, especially the effect for the images that contain clouds or poor lighting conditions, is better.

1. INTRODUCTION

In recent years, with the rapid development of remote sensing technology and the establishment of global earth observation system, the obtaining ability of various remotely sensed data is increasingly strong (Verde et al., 2018). As the development of open science and the increasing popularity of mobile devices (e.g., smartphones), remote sensing data as an important data sources are widely applied in human life. Additionally, the remote sensing data collection also shows a certain trend of popularization (Shan., 2017). That means that end users can access and use various remotely sensed data to achieve different operations and apply in much more actual applications too. In the process of image processing, how to quickly and accurately evaluate the amount of information of images has become one of the critical factors of improving the utilization ratio of image data, reducing data cost and meeting in some degree the timeliness of image processing (Datcu et al., 2007; Quartullinet al., 2013; Xia et al., 2017).

The related information measurement method of remote sensing images have been developed through the physical sciences, information theory and computer vision, and also plays a significant important role in lots of remote sensing processing and applications (Finkelstein et al., 2009; Kong et al., 2019). Image entropy as one frequently-used quantitative criterion is used to evaluate the amount of information contained in images. It could reflect the difference in the correlative attribute characteristics of different sources of remote sensing data, to some degree (Tsai et al., 2008; Hu et al., 2015). Additionally,, as the quantity and categories of remote sensing images that are applied to various application areas gradually increases, the image entropy is also regarded as a very valid means to improve the utilization efficiency and availability of obtained diversified remote sensing data (Ma et al., 2017).

Nevertheless, the traditional image entropy, similar to the global histogram of the image, merely reflects the global statistical information from the quantitative angle, but cannot effectively describe more and a deeper level of significant information contained in images, e.g., the information on the local spatial distribution, organizational structure, geometric features and visual perception (Cushman 2016; Li et al., 2016).

In remote sensing field, some improved methods or models of calculating the information content of images were proposed by related scholars over the past decades. Some scholars have proposed the Markov-based measurement methods, such as the first-order markov model employed in Lin et al. (2006) (Lin et al., 2006), and the computational models using quadrilateral markov random field developed by Razlighi et al. (2009) (Razlighi et al., 2009) and Liu. (2016) (Liu., 2016). Zhang et al. (2015) (Zhang et al., 2015) developed a measurement method of combining information theory with geostatistics. Furthermore, some scholars research on the influence of resolution (e.g., spatial resolution, radiometric resolution) on the amount of information contained in remote sensing images (Narayanan et al., 2002, Verde et al., 2018), and other researchers discuss the relationship between information content of images and signal-noise-ratio (SNR). Recently, the configurational entropy model based on Boltzmann entropy and its improved model is constructed to characterize the spatial structural information in geoscience data (Gao et al., 2017, Cushman., 2018; Gao et al., 2019).

Remarkably, compared with traditional image entropy, the reliability, effectiveness, and accuracy of the measurement approaches above mentioned are ameliorated to some degree. To further characterize much more attribute feature information and the difference in the quality of images, we present a neighborhood mapping-based solution to evaluate the geometrical mapping entropy of remotely sensed images (Sun et al., 2006; Li et al., 2016).

The rest of this paper is organized as follows. Section 2 reviews traditional image entropy and then present a novel approach for combining the idea of scatterplot mapping with information theory. In Section 3, we validate the feasibility and

effectiveness of the method (i.e., geometrical mapping entropy) proposed, do further discussion. Section 4 provides the conclusion.

2. METHODOLOGY

2.1 Image Entropy

As we all known, remotely sensed data as one of frequently-used information data is widely applied to various fields in human life. Hence, in the process of image processing, the image entropy model on the basis of information theory is usually regarded as a valid evaluation index of disorder degree in remotely sense imagery, which can also reflect in some degree information richness and texture complexity of images (Shannon 1948; Verde et al., 2018). The formula of image entropy is defined as.

$$H(I) = -\sum_{i=0}^{255} p_i \log p_i \qquad (1)$$

where i and p_i represent the gray levels in images and the corresponding probability of grey levels, respectively.

In the process of image processing, image entropy can be used to characterize the difference in the information content contained in images and image quality. It can also reduce the feature dimensionality and computational complexity of image data. However, the traditional image entropy only describes part of two-dimensional feature information of images from the perspective of quantity but doesn't reflect related spatial feature information in images from the qualitative and subjective perception angle. That is, it lacks the capability to characterize the spatial structure validly and cannot eliminate the influence of symmetry of entropy on measurement results. As a result, we propose an approach for evaluating the attribute information and quality of images, i.e., geometrical mapping entropy (GME). This model can better characterize much more attribute feature information contained in images, to accurately and validly reflect the differences in the information content and quality of multimodality images. Meanwhile, it can eliminate the influence of symmetry of entropy on measurement results.

2.2 A Strategy for Evaluating Geometrical Mapping Entropy of Images

To better evaluate the information content contained in remote sensing images, on the basis of information theory, we present a new geometrical mapping entropy model (GME) by incorporating spatial neighborhood information and the idea of scatterplot mapping into modelling processing (Touchette et al., 1985; Itti et al., 2001; Elmqvist et al., 2008; Quweider., 2012). The descriptions of the GME method are as follows:

1) First, let us construct the two-dimensional grayscale array (i, j). Thereinto i is the grayscale value of the pixel, and j is the corresponding average grayscale value of its neighboring pixels. The grayscale array (i, j) can validly reflect the information on the local neighborhood distribution and pixel position. The formula is as follows.

$$(i, j) \begin{cases} i, (i \in [0, 255]) \\ j = \dfrac{1}{N} \sum_{s=1}^{N} i^s, j \in [0, 255], s = 1, 2, ..., N \end{cases} \qquad (2)$$

where i_s is the grayscale value of the s-th neighborhood pixel of the target pixel with grayscale value i, N is the true number of pixels in the effective neighborhood of the target pixel. The average value j is calculated by the ratio of the sum of the practical neighboring pixels of the target pixel to N. Additionally, the grayscale array (i, j) is mapped into the planar mapping matrix by the two-dimensional scatter plot model, and $n_{(i,j)}$ is the statistical value of (i, j) in the planar mapping matrix.

2) The next step is to construct the three-dimensional array $(i, j, n_{(i,j)})$, which consists of the grayscale value i of the pixel, the average grayscale value j of its neighboring pixels, and the statistical value $n_{(i,j)}$ of the grayscale array (i, j). The $(i, j, n_{(i,j)})$ can be mapped into the geometrical mapping space by the three-dimensional scatterplot model, and $n_{(i,j,n_{(i,j)})}$ is the statistical value of the three-dimensional array $(i, j, n_{(i,j)})$ in the geometrical mapping space.

3) Third, the single level geometrical mapping entropy (GME) is developed by incorporating the mapping ideas of the scatterplot matrices and the statistical value $n_{(i,j,n_{(i,j)})}$ into modeling processing. The probability $p_{(i,j,n_{(i,j)})}$ of $n_{(i,j,n_{(i,j)})}$ in the geometrical mapping space can be calculated by:

$$p_{(i,j,n_{(i,j)})} = \frac{n_{(i,j,n_{(i,j)})}}{N}, (N = \sum_{n_{(i,j)}=0}^{\max(n_{(i,j)})} \sum_{i=0,j=0}^{255} n_{(i,j,n_{(i,j)})}) \qquad (3)$$

where $n_{(i,j,n_{(i,j)})}$ is the statistical value of $(i, j, n_{(i,j)})$, and N is the total number of $n_{(i,j,n_{(i,j)})}$ with different values.

Definition 1 The geometrical mapping entropy (GME) of the image is defined as:

$$H_{GME} = -\sum_{n_{(i,j)}=0}^{\max(n_{(i,j)})} \sum_{i=0,j=0}^{255} p_{(i,j,n_{(i,j)})} \log p_{(i,j,n_{(i,j)})} \qquad (4)$$

The model can validly characterize the comprehensive feature information of images, e.g., spectrum features, spatial structural features, and visual perception. In the meantime, it also reflects the difference in the quality of images. Additionally, the GME method eliminates the adverse influence of traditional image entropy symmetry on the results of the information measurement.

The general framework of this approach is described in Figure 1.

3. RESULTS

3.1 Study Area and Data

In this study, we use the two groups of different datasets to verify the feasibility and validity of the proposed method (see below).
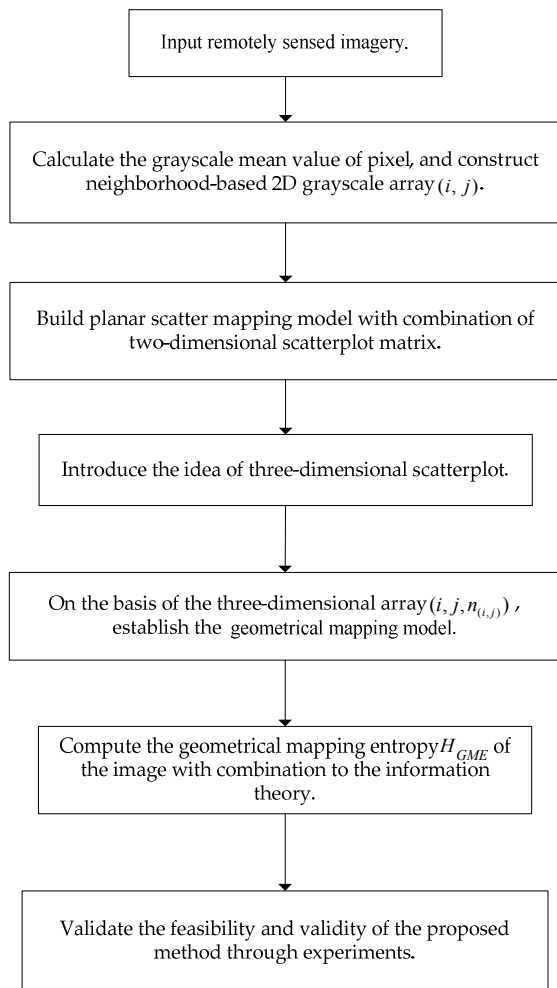
Figure 1. The general framework of the GME model.

We testify the performance of the proposed approach using these experimental images (see below). Figure 2a denotes a mixed reservoir area that contains different kinds of land cover (e.g., reservoir, town region, farmland, trees, and roadway). Figure 2b is the cultivated land that has certain distribution regularity; Figure 2c describes a water area; Figure2d describes a certain mountain area. Additionally, Figure 2a1-2d1 depicts the planar mapping distribution of different images, respectively; Figure 2a2-2d2 characterizes the corresponding results of the geometrical mapping model, respectively.

➤ The first group of the dataset 1 contains the following four types of image data:

(1) A 0.5 m resolution image of a reservoir area located in the Zhengzhou region, obtained from the Digital Globe platform in 2018;
(2) An image of farmland obtained from the UC Merced Land Use Dataset with USGS National Map Urban Area Imagery in 2010 with 0.3 m resolution;
(3) A UAV image of a local area in the district of the lower and middle reaches of the Yellow River in 2015;
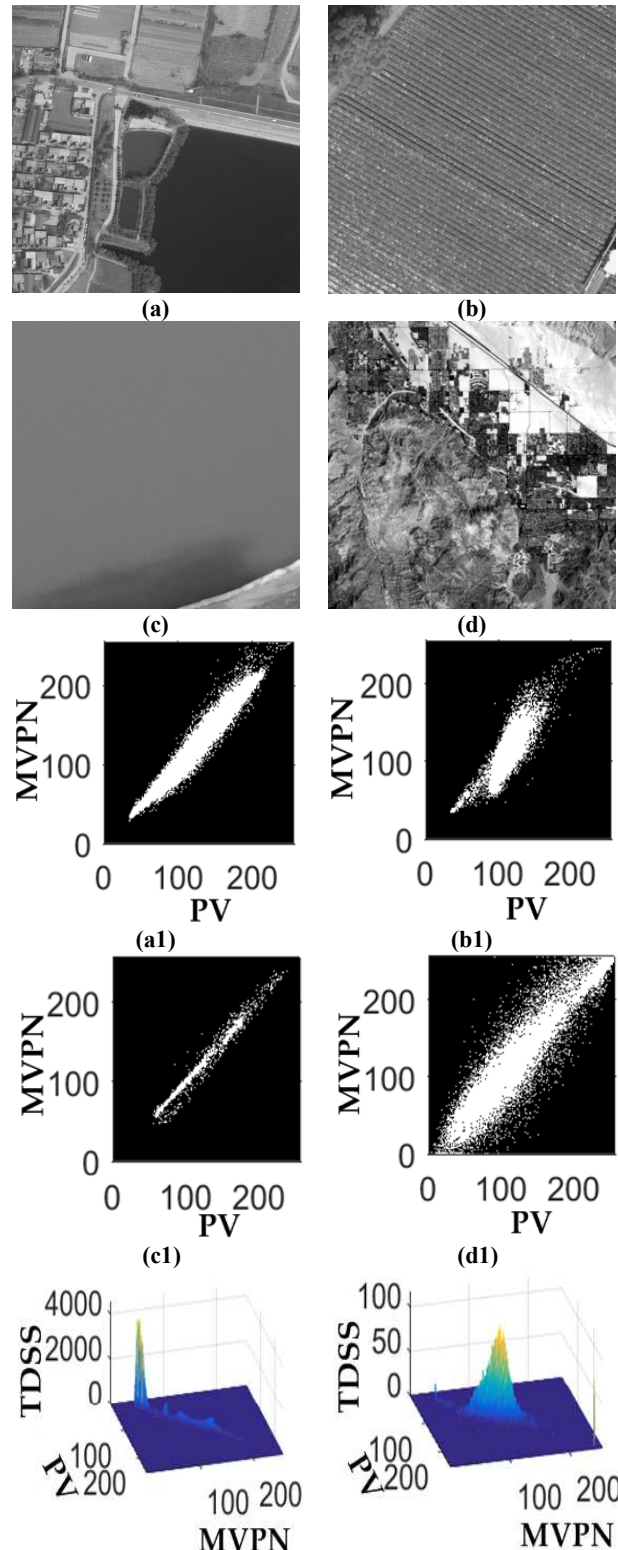(4) Landsat TM image of a mountainous region provided by NASA;

➤ The second group of the dataset 2 is as follows：

(1) A 1 m resolution, remote sensing images with the different modalities (e.g., different imaging conditions, various imaging quality) in the same urban region of Changsha with different times (i.e., 2016-05, 2017-08, 2018-04, and 2018-06);

3.2 Experiments

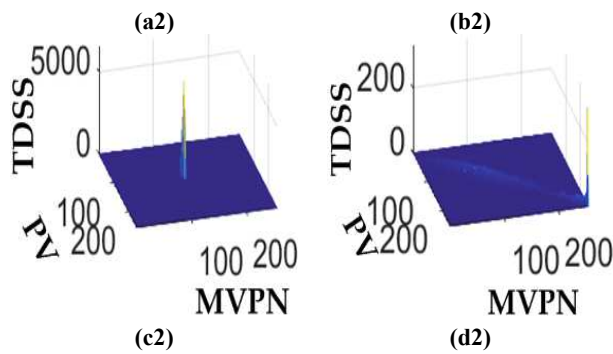### 3.2.1 Analysis of the Results on the Dataset 1

Figure 2. The plane and geometrical mapping models of experimental images with different types: (a-d) experimental images; (a1-d1) planar scatterplot mapping model; (a2-d2) geometrical space mapping model.

Firstly, let us compare and analyze the mapping results of different experimental images based on the mapping strategy constructed (see Figure 2a1-2d1 and Figure 2a2-2d2). As one can see, the scatterplots of two categories of mapping model proposed can relatively well reflect spatial distribution trends and differences of information richness of experimental data when different types of images in this section are mapped into unified geometrical space. Meanwhile, there is also consistent with the visual perception results of the human eye system.

Secondly, we can evaluate the information content of remote sensing images using different approaches from the quantitative perspective (see Table 1).

Table 1. Comparative analysis of different methods on the dataset 1

Methods	Figure 2a	Figure 2b	Figure 2c	Figure 2d
$H(I)$	7.193	6.432	4.425	7.827
H_{GME}	0.514	0.436	0.195	0.832

The experimental results indicate that the measurement model constructed, similar to the traditional image entropy, can effectively evaluate image information. That is, the proposed method in this paper is feasible and valid. The grayscale distribution and spatial distribution of water area are all comparatively centralized, as demonstrated in Figure 2c1 and Figure 2c2, i.e., the dispersion degree and uniformity of its brightness values are very lower. It means that the uncertainty involved in this image is smaller; thus, the amount of information of the image is less.

Additionally, we further compare the mountain image with the farmland image and reservoir image. It is seen that the image of the mountain area contains a more complex spatial structure; its visual perception is more disordered. Identical conclusions are obtained in the corresponding planar and geometrical mapping results of images. This is consistent with the size relationship of the two kinds of metric values of these images in Table 1. The uncertain information contained in Figure 2d is the largest.

3.2.2 Analysis of the Results on the Dataset 2

As shown in Figure 3, the lighting conditions of Figure 3a and Figure 3b are poor lighting condition and slightly poor lighting condition, respectively; Figure 3c represents the image

with thin clouds; and Figure 3d is the image of good quality. Compared with the objects in the same regions of Figure 3c-3d, the ground objects covered by the red circle in Figure 3a-3b have already changed, i.e., the object type described is mainly changed from woodland to bare land. It can be seen that there was a relative reduction in the texture complexity and diversity of the spatial structure of the images above, and the uncertainty information contained in images also decreases.
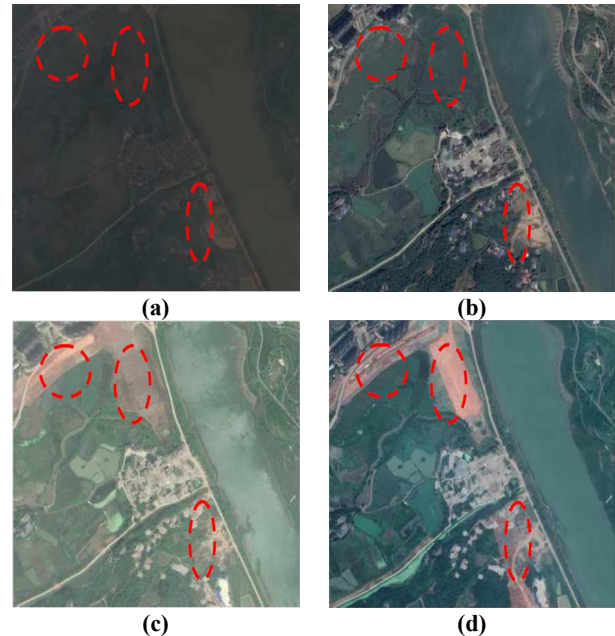

Figure 3. Experimental images with different imaging conditions, e.g., different light conditions, clouds.

In Table 2, the metric results of Experimental images (Figure 3a-3d) obtained by the geometrical mapping entropy model and traditional image entropy are listed.

Table 2. Comparative analysis of different methods on the dataset 2

Methods	Figure 3a	Figure 3b	Figure 3c	Figure 3d
$H(I)$	4.518	6.235	6.422	6.648
H_{GME}	0.074	0.536	0.407	0.462

We can discover that Figure 3a is the image with a bad lighting condition. It leads to that there has scarcely any useful information in this image. That is, its quality is poor, which is consistent with the GME value of Figure 3a, i.e., 0.074, which approaches zero. Furthermore, it can be seen from Tab.1 that, the distribution tendency of traditional image entropy for this group of experimental images in Figure 3 is overall on the rising trend. This conclusion, however, deviates from the real judging results of human visual perception.

Additionally, Figure 3c affecting by the thin clouds, especially compared with Figure 3b and Figure 3d, the quality of which should be relatively worse. It can be observed from Table1that, the GME value of Figure 3c is less than the corresponding metric results of the images in Figure 3b and Figure 3d, which also matches the subjective judging results.

However, the result of traditional image entropy has no reflecting the same correct conclusion effectively and accurately.

Compared with the same areas in Figure 3b, the ground objects of local areas covered by the red circle in Figure 3d are changed from woodland to bare land. It means that the information richness and spatial structural complexity of Figure 3d affected by object changes should be less than the corresponding feature information of Figure 3b. The identical conclusion is supported by the GME values of Figure 3b and Figure 3d.

Based on the analysis above, experimental results validate the feasibility and validity of GME model. In addition, it can be concluded that the geometrical mapping entropy, in comparison with traditional image entropy, can better evaluate and distinguish the difference in the information content of images with different imaging quality, especially for remotely sensed images affected by different imaging conditions, such as different light conditions and clouds.

4. CONCLUSION

In this study, we propose a geometrical mapping entropy model by incorporating the neighborhood information and the idea of scatterplot mapping into the modelling process. Experimental results show that the method proposed in this paper is feasible and valid, which can overcome the existed shortcomings of traditional image entropy. Furthermore, it can distinguish the influence of the difference in the images with different imaging conditions on measurement results.

Future work will focus on a few directions below. First, the proposed measurement method can be further improved. Second, it is necessary to explore the related attribute characteristics of geometrical mapping entropy deeply. Third, it is a meaningful thing to make a comparison with other improved methods such as the configurational entropy proposed by Gao et al. (2017). Fourth, we could attempt to investigate the respective applicable application fields of different measurement models by contrasting and analyzing the experimental effect of different methods in image processing.

ACKNOWLEDGEMENTS (OPTIONAL)

REFERENCES

Cushman, S. (2018). "Calculation of configurational entropy in complex landscapes." Entropy **20**(4): 298.

Cushman, S. A. (2016). "Calculating the configurational entropy of a landscape mosaic." Landscape ecology **31**(3): 481-489.

Datcu, M., S. d'Elia, R. L. King and L. Bruzzone (2007). "Introduction to the special section on image information mining for earth observation data." IEEE Transactions on Geoscience and Remote Sensing **45**(4): 795-798.

Elmqvist, N., P. Dragicevic and J.-D. Fekete (2008). "Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation." IEEE transactions on Visualization and Computer Graphics **14**(6): 1539-1148.

Finkelstein, L. (2009). "Widely-defined measurement–An analysis of challenges." Measurement **42**(9): 1270-1277.

Gao, P., J. Wang, H. Zhang and Z. Li (2019). "Boltzmann entropy-based unsupervised band selection for hyperspectral image classification." IEEE Geoscience and Remote Sensing Letters **16**(3): 462-466.

Gao, P., H. Zhang and Z. Li (2017). "A hierarchy-based solution to calculate the configurational entropy of landscape gradients." Landscape ecology **32**(6): 1133-1146.

Hu, L., Z. He, J. Liu and C. Zheng (2015). "Method for measuring the information content of terrain from digital elevation models." Entropy **17**(10): 7021-7051.

Itti, L. and C. Koch (2001). "Computational modelling of visual attention." Nature reviews neuroscience **2**(3): 194-203.

Kong, L., H. Pan, X. Li, S. Ma, Q. Xu and K. Zhou (2019). "An Information Entropy-Based Modeling Method for the Measurement System." Entropy **21**(7): 691.

Li, Z., Q. Liu and P. Gao (2016). "Entropy-based Cartographic Communication Models: Evolution from Special to General Cartographic Information Theory." Acta Geodaetica et Cartographic Sinica **45**(7): 757-767.

Lin, Z. and Y. Zhang (2006). "Measurement of information and uncertainty of remote sensing and GIS data." Geomatics and Information Science of Wuhan University **31**(7): 569-572.

Liu, F. (2016). "Research on analysis and information measurement of remote sensing images." Wuhan University.

Ma, C., X. Wei, C. Fu, J. Liu and L. Wei (2017). "A Content-Based Remote Sensing Image Change Information Retrieval Model." ISPRS International Journal of Geo-Information **6**(10): 310.

Narayanan, R. M., M. Desetty and S. Reichenbach (2002). "Effect of spatial resolution on information content characterization in remote sensing imagery based on classification accuracy." International Journal of Remote Sensing **23**(3): 537-553.

Quartulli, M. and I. G. Olaizola (2013). "A review of EO image information mining." ISPRS journal of Photogrammetry and Remote Sensing **75**: 11-28.

Quweider, M. (2012). "Spatial entropy-based cost function for gray and color Image segmentation with dynamic optimal partitioning." International Journal of Computer Science and Network Security (IJCSNS) **12**(4):

64.

Razlighi, Q. R., N. Kehtarnavaz and A. Nosratinia (2009). "Computation of image spatial entropy using quadrilateral Markov random field." IEEE transactions on image processing **18**(12): 2629-2639.

Shan, J. (2017). "Remote sensinf : from trained professional to general public." Acta Geodaetica et Cartographic Sinica **46**(10): 1434-1446.

Shannon, C. E. (1948). "A mathematical theory of communication." Bell system technical journal **27**(3): 379-423.

Sun, J., X. Zhang, J. Cui and L. Zhou (2006). "Image retrieval based on color distribution entropy." Pattern Recognition Letters **27**(10): 1122-1126.

Touchette, P. E., R. F. MacDonald and S. N. Langer (1985). "A scatter plot for identifying stimulus control of problem behavior." Journal of applied behavior analysis **18**(4): 343-351.

Tsai, D.-Y., Y. Lee and E. Matsuyama (2008). "Information entropy measure for evaluation of image quality." Journal of digital imaging **21**(3): 338-347.

Verde, N., G. Mallinis, M. Tsakiri-Strati, C. Georgiadis and P. Patias (2018). "Assessment of radiometric resolution impact on remote sensing data classification accuracy." Remote Sensing **10**(8): 1267.

Xia, G. S., X. Y. Tong, F. Hu, Y. Zhong, M. Datcu and L. Zhang (2017). "Exploiting Deep Features for Remote Sensing Image Retrieval: A Systematic Investigation."

Zhang, Y. and J. Zhang (2015). "Measure of Information Content of Remotely Sensed Images Accounting for Spatial Correlation." Acta Geodaetica et Cartographic Sinica **44**(10): 1117-1124.