

# A Precision Evaluation Method for Remote Sensing Data Sampling Based on Hexagon Discrete Grid

Yue Ma<sup>1,2</sup>, Guoqing Li<sup>1</sup>, Xiaochuang Yao<sup>3\*</sup>, Jin Ben<sup>4</sup>, Qianqian Cao<sup>1</sup>, Long Zhao<sup>1</sup>, Lianchong Zhang<sup>1</sup> and Rui Wang<sup>4</sup>

<sup>1</sup> Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; mayue@radi.ac.cn (Y.M.); ligq@radi.ac.cn (G.L.); caoqq@radi.ac.cn (Q.C.); zhaolong@aircas.ac.cn (L.Z.); zhanglc@aircas.ac.cn (L.Z.)

<sup>2</sup> School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049; mayue@radi.ac.cn (Y.M.)

<sup>3</sup> College of Land Science and Technology, China Agricultural University, Beijing 100083, China; yxc@cau.edu.cn

<sup>4</sup> Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052, China; benj@reis.ac.cn (J.B.); wr\_paper@126.com (R.W.)

\* Correspondence: yxc@cau.edu.cn; Tel.: +86-010-62-737-554

**Abstract:** With the rapid development of earth observation, satellite navigation, mobile communication and other technologies, the order of magnitude of the spatial data we acquire and accumulate is increasing, and higher requirements are put forward for the application and storage of spatial data. Under this circumstance, a new form of spatial data organization emerged-the global discrete grid. This form of data management can be used for the efficient storage and application of large-scale global spatial data, which is a digital multi-resolution the geo-reference model that helps to establish a new model of data association and fusion. It is expected to make up for the shortcomings in the organization, processing and application of current spatial data. There are different types of grid system according to the grid division form, including global discrete grids with equal latitude and longitude, global discrete grids with variable latitude and longitude, and global discrete grids based on regular polyhedrons. However, there is no accuracy evaluation index system for remote sensing images expressed on the global discrete grid to solve this problem. This paper is dedicated to finding a suitable way to express remote sensing data on discrete grids, and establishing a suitable accuracy evaluation system for modeling remote sensing data based on hexagonal grids to evaluate modeling accuracy. The results show that this accuracy evaluation method can evaluate and analyze remote sensing data based on hexagonal grids from multiple levels, and the comprehensive similarity coefficient of the images before and after conversion is greater than 98%, which further proves that the availability hexagonal grid-based remote sensing data of remote sensing images. And among the three sampling methods, the image obtained by the nearest interpolation sampling method has the highest correlation with the original image.

**Keywords:** Remote sensing; Global discrete grid; Accuracy evaluation; Hexagon grid

---

## 1. Introduction

With the rapid development of earth observation, satellite navigation, mobile communication and other technologies, with the construction and application of smart earth and smart cities, the explosive growth of the amount of data we obtain has updated from the current GB (Giga Byte) and TB (Tera Byte) level to EB (Exa Byte) or even ZB (Zetta Byte) level, which has brought challenges to the integration and utilization of spatial data[1,2]. However, the application and storage methods of spatial data have not been developed to the corresponding level, which means a large amount of data is just stored but not fully utilized. Researchers have proposed the concept of Global Discrete Grid System, which is a multi-resolution hierarchical structure that uses a specific method to subdivide the surface of the earth infinitely[3,4]. Compared with the traditional spatial data organization and application mode, the global discrete grid system has hierarchical and continuity, which can avoid the deformation caused by direct projection. It is suitable for solving large-scale issues, and supports efficient processing of multi-resolution in structure[5,6].

At present, various global discrete grids have emerged, including global discrete grids of equal latitude and longitude, global discrete grids of variable latitude and longitude, adaptive global discrete grids, and global discrete grids based on regular polyhedrons[7,8]. Because of 6-fold rotational symmetry, compared with the quadrilateral and triangular grids, the hexagonal grid has the advantages of equal circumference, equidistance, and uniform connectivity, which is suitable for processing tasks such as geographic spatial information processing such as DEM valley line extraction and river extraction[9,10]. The global discrete grid uses unique codes to indicate the membership of the unit, so as to further perform retrieval, addition and other operations[11,12]. By now, the global discrete grid has been initially applied in many fields such as global change, environmental monitoring and disaster analysis[13,14].

Furthermore, how to use appropriate methods to resample remote sensing data into hexagonal grids, accurately describe them, and use appropriate indicators to evaluate grid data is an urgent issue to be solved. So far, plenty of research teams have proposed accuracy evaluation indicators and systems for different global discrete grids[15,16]. These indicators and system can be used to evaluate the non-uniformity of grid cells, the irregularity of deformation distribution and the complexity of deformation with layer changes[17,18]. Therefore, the uncertainty of the global discrete grid model can be analyzed in spatial data management, spatial analysis and decision analysis[19,20]. However, in the process of transforming remote sensing data into global discrete grid data, there are very few evaluation method and evaluation indicators for expression accuracy. Thus ,there is currently no research result that can provide credibility that remote sensing data based on hexagonal grids can be used for later processing and application analysis[21].

This study analyzed the causes of errors and the theoretical accuracy of remote sensing image representation on a hexagonal grid. The error mainly comes from the error of the original image, which is transferred to the image of the hexagonal grid through the error[7,22,23]. There are some researches on the geometric optimization design and spatial measurement of the hexagonal integral discrete grid [24,25]. The checkpoint precision evaluation method can be used to extract the DEM valley line represented by a regular hexagonal grid, in which several geomorphic feature points in the study area and the edge are randomly selected as the detection points, and the hexagonal DEM and quadrilateral DEM established by the grid elevation error evaluation respectively data accuracy [9,26,27]. A series of geometric indicators when evaluating the properties of different grids, including the change rate of cellular area  $V_S$ , rate of change of cellular angle  $V_A$ , and change rate of cellular side length  $V_c$  [17,28,29].

The hexagonal grid to reorganize the rectangular pixels of original remote sensing images, and according to the application direction and basic characteristics of the hexagonal grid data, an appropriate accuracy evaluation system is established for the images on the hexagonal grid, which includes the following steps:

1). Remote sensing data modeling based on hexagonal grid

This study used several resampling methods including nearest neighbor pixel method, bilinear interpolation, bicubic interpolation to resample rectangular pixels, and compare these methods.

2). Evaluation of remote sensing image accuracy based on hexagonal grid

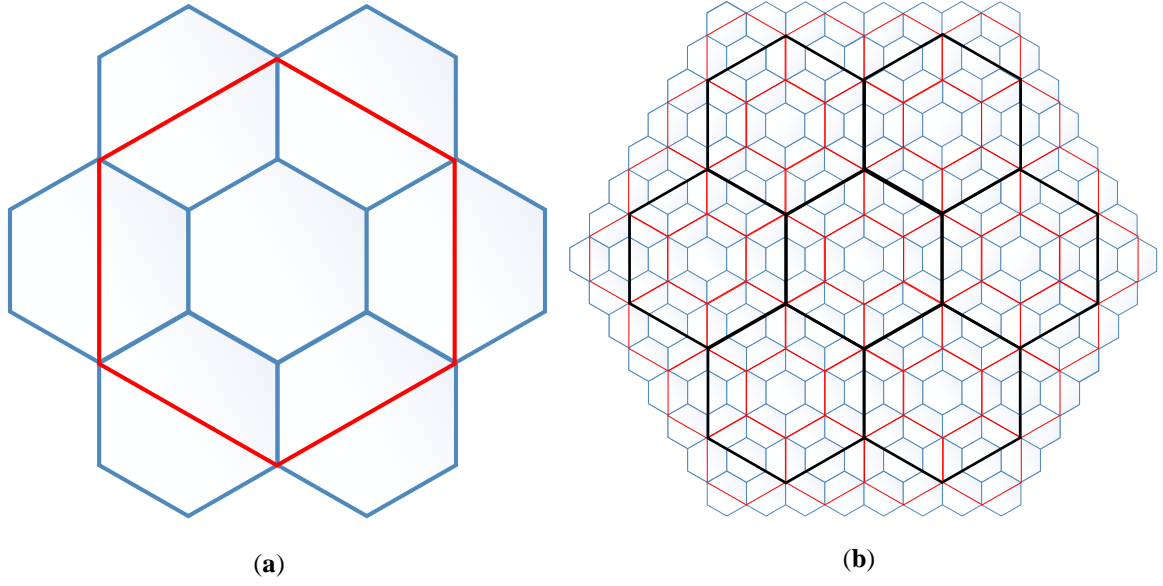
In order to prove the availability of the remote sensing image reorganized on the hexagonal grid, we select the appropriate accuracy evaluation index to evaluate the accuracy of the converted remote sensing data. According to the criteria for evaluating the accuracy of spatial data conversion: maintaining the conservation of composition information, maintaining the conservation of area information, and maintaining the conservation of regional spatial pattern and morphological information, this research proposes three types of evaluation indicators: basic evaluation indicators, evaluation indicators based on image features, and geometric evaluation index.

## 2. Materials and Methods

In the process of using a hexagonal grid to reorganize remote sensing images requires hexagonal grid pixels to represent rectangular pixels, the key step is image sampling. This study determines the level of the hexagonal grid to which the remote sensing data belongs based on the spatial resolution of the selected remote sensing data, and then sample the original remote sensing image into hexagonal grid pixels according to the corresponding relationship between the grid unit and the pixel unit.

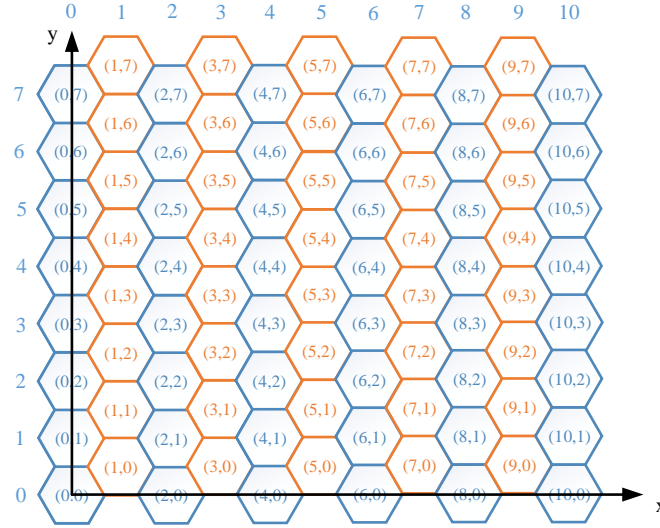
### 2.1. Conversion of original remote sensing image to hexagonal grid data

This study chose the planar aperture 4 hexagon grid system as the framework to transform the original remote sensing data. Because in this system the direction of the hexagon stays consistent and here is the hexagonal grid planar structure:



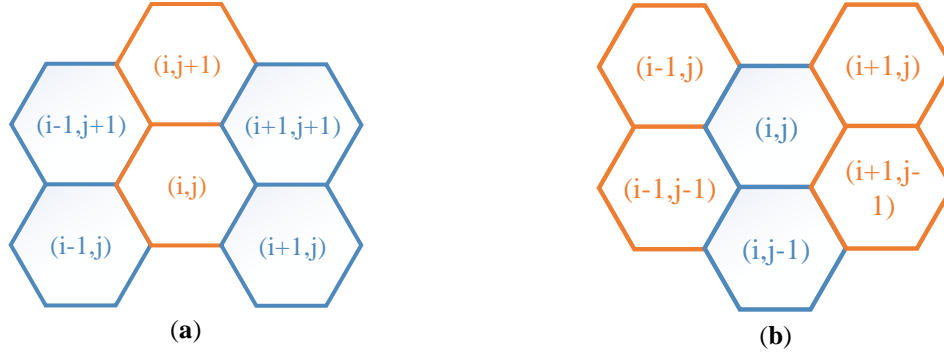
**Figure 1.** This is a figure of the hexagonal grid planar structure. (a) shows that the ratio of the area of two adjacent layers is 4; The ratio of the side lengths of the two layers is 2; (b) shows the unit structure between different layers, and all the unit directions are the same.

We used the concept of "point grid" to establish a planar hexagonal grid based on an orthogonal coordinate system. As shown in Figure 2, the number of rows and columns (i, j) where the grid is located is coded, and (0, 0) is the origin of the coordinates.



**Figure 2.** Planar hexagonal grid system based on orthogonal coordinate system

As shown in Figure 2, when a regular hexagon is tiled, the odd and even columns have a certain offset. If the side length of the hexagonal unit is  $l$ , then the offset between odd rows and even rows is  $\frac{\sqrt{3}}{2}l$ .



**Figure 3.** The misalignment relationship between odd and even rows, take the hexagonal unit  $(i, j)$  as an example. **(a)** Schematic diagram of grid structure in odd rows; **(b)** Schematic diagram of grid structure in even rows

According to the idea of “rasterization” and to simplify calculations, the attributes of the hexagonal grid cells and the calculation and analysis between grids are replaced by the corresponding grid center points, so that each grid in the grid structure is obtained. The geographic coordinate value of the center point of the grid is a key step in the application of the hexagonal grid structure to remote sensing data. According to the above analysis, the corresponding relationship between the planar hexagonal grid and geographic coordinates can be defined as:

$$\begin{cases} Y(i, j) = \begin{cases} y_0 + \sqrt{3}l \times j, & i \% 2 = 0 \\ y_0 + \sqrt{3}(j + \frac{1}{2})l, & i \% 2 = 1 \end{cases} \\ X(i, j) = y_0 1.5 \times l \end{cases} \quad (1)$$

Where  $Y(i, j)$  and  $X(i, j)$  are the geographical coordinates of the grid center point with grid coordinates of  $(i, j)$ ;  $l$  is the side length of the hexagonal grid cell.

The appropriate grid level needs to be selected for modeling and sampling. Since the size of the grid cannot be exactly the same as the size of each pixel unit of the remote sensing image, in actual processing, the level we choose is the grid area that is slightly smaller than the area represented by each pixel of the remote sensing image. For example, each pixel of the remote sensing image represents 4 square meters, the size of a single grid on the 22nd layer is about 3.86 square meters, and the size of a single grid on the 21st layer is about 15.26 square meters. According to this standard, Landsat 8 images with a resolution of 30m can be stored on the 18th layer, and images with a resolution of about 1km can be stored on the 13th layer. On the selected level, the correspondence between each grid and the raster data in the remote sensing image should be determined. Since the area of each grid of the grid that we selected to store a certain remote sensing image is smaller than a single pixel of the remote sensing image, each grid can store the pixel value of the grid with the largest area, so as to minimize the loss of the original remote sensing image data. The basic process is as follows:

It is the first step to determine the level of the hexagonal grid corresponding to the remote sensing image pixel according to the spatial resolution of the remote sensing image, and the corresponding discrete grid level should meet the requirement of  $n_1 \leq n \leq n_2$ . Then, based on the correspondence between geographic coordinates and hexagonal grid coding, a conversion model between remote sensing data and hexagonal grid data is established.

Hexagonal grids of different levels have corresponding areas of hexagonal elements. The best division area when dividing the grid is:

$$s = \frac{SR \times m}{MR} \quad (2)$$

Where  $SR$  is the area of remote sensing image,  $MR$  is the size of remote sensing image,  $m$  is preset size of remote sensing image after subdivision.

**Table 1** Grid levels corresponding to commonly used high, medium and low-resolution remote sensing images

Remote sensing data	Remote sensing data	Spatial resolution	Number of bands	Adjacent grid level	Corresponding grid level
	NOAA/AVHRR	1.1km	5	12,13	13

Low spatial resolution remote sensing data	Modis	250-1000m	36	13,14	13
	Spot Vegetation	1000m	4	12,13	13
	Landsat 8	30m	11	18,19	18
Medium spatial resolution remote sensing data	Spot 5	10m	6 Multispectral (4), Panchromatic (1), Shortwave Infrared (1)	19,20	19
	Terra/aster	15m	14	19,20	19
	Quick Bird	1m	5	22,21	22
High spatial resolution remote sensing data	Eros-b Satellite	0.41m	1	22,23	22
	Eros-a Satellite	1.9m	1	22,21	22

When converting a rectangular grid into a hexagonal grid, the spatial resolution of the hexagonal grid (the size of the cells) can be determined according to the spatial resolution of the rectangular grid. Generally, the resolution of a rectangular grid corresponds to the hexagonal grid between two adjacent layers, rather than the area of a certain layer of hexagonal units. In this way, it is necessary to determine which layer should be converted to and the six sides of the corresponding layer. The absolute value of the difference between the area of the rectangular element and the area of the hexagonal element should be the smallest.

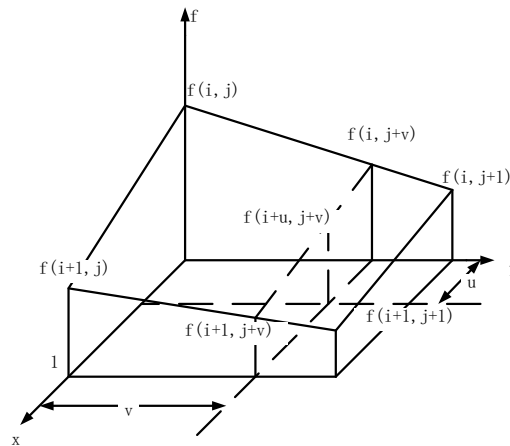
Then, on a two-dimensional plane, fill the specific study area with hexagonal grid, and resample the pixels of the remote sensing image with the following resampling methods to determine the attribute value of the hexagonal grid unit.

(1) The nearest neighbor interpolation

This method is aimed at a two-dimensional image, and takes the gray value of the nearest neighbor among the four neighboring pixels around the point to be sampled as the gray value of the point. Although this algorithm is simple to calculate, it only uses the gray value of the pixel that has the greatest impact on the sampling point (that is, the nearest) as the value of the point, and does not consider the influence (correlation) of other adjacent pixels, so resample the gray value of the latter image has obvious discontinuities, and the image distortion cannot be ignored.

(2) Bilinear interpolation

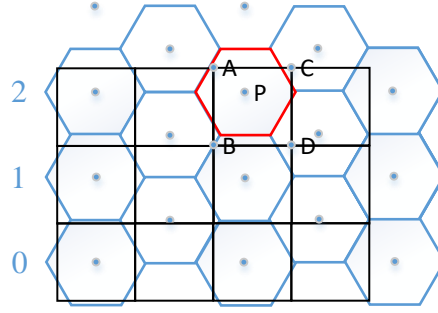
As an improvement to the nearest neighbor method, this method is using the gray values of the surrounding 4 neighboring points to perform linear interpolation in two directions to obtain the gray value of the point to be sampled. That is, the corresponding weight is determined according to the distance between the point to be sampled and the adjacent point to calculate the gray value of the point to be sampled.



**Figure 4.** Bilinear interpolation principle

Where X and Y coordinates represent the position of the pixel, and  $f(x, y)$  represents the DN value of the pixel. Its mathematical expression is:

$$f(i+u, j+v) = (i-u)(1-v)f(i, j) + (1-u)vf(i, j+1) + u(1-v)f(i+1, j) + uvf(i+1, j+1) \quad (3)$$



**Figure 5.** Schematic diagram of interpolation operation

According to Figure 5, the pixel value of P is determined by A, B, C, D by formula (3), and the attribute value of red-edged hexagon unit is determined by the value of P.

Compared with the nearest neighbor interpolation, the bilinear interpolation method takes the influence of the four direct neighbors around the sampling point on the sampling point into consideration, which basically overcomes the shortcomings of the former gray scale discontinuity, but the amount of calculation increased. However, because this method only considers the influence of the gray value of the four direct adjacent points, and does not consider the influence of the gray value change rate between each adjacent point, it has the nature of a low-pass filter to make the high-frequency components of the zoomed image with loss, the outline of the image becomes blurry. Compared with the original image, the image scaled by this method still has the problem of image quality degradation and accuracy reduction due to improper consideration of the calculation model.

### (3) Cubic convolution method

The cubic convolution method not only considers the influence of the gray value of the four direct neighboring points, but also the influence of the gray value change rate between the neighboring points. It uses 16 points to be sampled. The gray values of pixels in the larger surrounding neighborhood are interpolated three times. Three operations can get a zoom effect closer to high-resolution images, but it also leads to a sharp increase in the complexity of operations. This algorithm needs to select interpolation basis functions to fit the data.

The mathematical expression is as follows:

$$S(w) = \begin{cases} 1 - 2|w|^2 + |w|^3 & |w| < 1 \\ 4 - 8|w| + 5|w|^2 - |w|^3 & 1 \leq |w| \leq 2 \\ 0 & |w| > 2 \end{cases} \quad (4)$$

The cubic convolution interpolation formula is as follows:

$$f(i + u, j + v) = ABC \quad (5)$$

Where A, B, and C are all matrices, and their forms are as follows:

$$A = [S(1 + u)S(u)S(1 - u)S(2 - u)] \quad (6)$$

$$B = \begin{bmatrix} f(i - 1, j - 2) & f(i, j - 2) & f(i + 1, j - 2) & f(i + 2, j - 2) \\ f(i - 1, j - 1) & f(i, j - 1) & f(i + 1, j - 1) & f(i + 2, j - 1) \\ f(i - 1, j) & f(i, j) & f(i + 1, j) & f(i + 2, j) \\ f(i - 1, j + 1) & f(i, j + 1) & f(i + 1, j + 1) & f(i + 2, j + 1) \end{bmatrix} \quad (7)$$

$$C = [S(1 + v)S(v)S(1 - v)S(2 - v)]^T \quad (8)$$

Where  $f(i, j)$  represents the DN value of the pixel  $(i, j)$  in the original image.

The sampled grid data needs to be stored as the coordinates of the center of the hexagon, the direction, size, and attribute values of the hexagon. When this information is read, an image based on hexagonal grid can be reproduced based on the information.

## 2.2. Accuracy evaluation of hexagonal grid data

Since the original remote sensing images are also obtained by sampling and quantization, the image itself has errors, and after the image is interpolated, the error is transferred to the new sampled points through

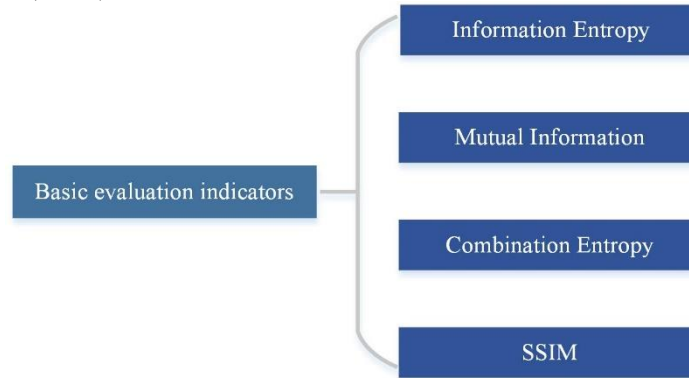
interpolation, so the error transfer needs to be quantified and the hexagonal-based remote sensing image needs to be compared with the original image[30,31] by calculating the loss of information and the error of geographic coordinates in the process of expression on the hexagonal grid. Therefore, it is necessary to establish an appropriate error evaluation model to evaluate the error of the converted data.

The relevant elements involved in the comprehensive evaluation of a certain thing constitute the evaluation element set. A series of indicators used to evaluate the thing constitute the evaluation index set, of which the weights is not the same. The evaluation index set is a mapping of the evaluation elements set[29,32], where there are multiple mapping index sets for an evaluation element set. There are currently two typical standards for the principle of establishing an index system: one is comprehensive, non-overlapping (or redundant), of which indicators are easy to obtain; the second is scientificity, rationality and applicability[13,33]. According to the three criteria for evaluating the accuracy of spatial data conversion: maintaining the conservation of composition information, maintaining the conservation of area information, and maintaining the conservation of regional spatial pattern and morphological information, this research proposes evaluation indicators from three perspectives: basic evaluation indicators, evaluation indicators based on image features, and geometric evaluation index[34-36].

### 2.2.1. Basic evaluation indicators

The basic evaluation index is a set of evaluation index established based on statistical information such as the mean, variance, and covariance between the original remote sensing image and the grid remote sensing image. These indicators are mainly used to evaluate the image composition information before and after conversion, with clear connotation, simple model and convenient evaluation. However, the basic evaluation indicators only focus on the information amount and composition type of the image, and they cannot reflect the image structure and image feature information.

The basic evaluation indicators include Information entropy, Mutual information, Combination entropy, Structural similarity index (SSIM).



**Figure 6.** The basic evaluation indicators

Information entropy reflects the information richness of the image, and calculating the information entropy of the image before and after conversion can reflect the loss of information during the conversion process[37]. If the amount of change in information entropy before and after the data conversion is small, it means there is not much information during the process. The calculation formula is as follows:

$$m_k = \frac{1}{N_k \times M_k} \sum_{j=1}^{N_k} \sum_{i=1}^{M_k} p[C_k(i, j)] \log_2 |C_k(i, j)| \quad (9)$$

The mutual information is an important concept in information theory. It can be used as a measurement of the correlation between two variables, or a measurement of the amount of information contained in another variable. Suppose there are two random variables  $P_X(x)$  and  $P_Y(y)$ , their marginal probability distributions are A and B, and their joint probability density is  $P_{XY}(x, y)$ . According to the related concepts of information theory, the mutual information between these two variables is:

$$M_{XY}(x, y) = \sum_{x, y} P_{xy}(x, y) \quad (10)$$

An image can be regarded as a two-dimensional random variable. The above concept can be easily extended to a two-dimensional space. For the two images A and B before and after conversion, the two images have the same gray level. Let the total gray level of the image be L,  $P_A(a)$  and  $P_B(b)$  are the probability densities of the images before and after conversion. The probability density can be obtained by dividing the histogram of the image by the total number of pixels in the image. Therefore, the mutual information of the images before and after conversion can be defined as:

$$M_{AB} = \sum_{f=0}^{L-1} \sum_{a=0}^{L-1} pF_A(f, a) \quad (11)$$

The mutual information is an objective indicator that reflects the abundance of information before and after data conversion. The larger its value, the more abundant information the hexagonal grid data can obtain from remote sensing data, and it can accurately evaluate the accuracy of the conversion. Joint entropy is also a very important parameter in information theory. It can be used to measure the correlation between two images and reflect the joint information. The joint entropy of images A and B before and after conversion can be defined as:

$$J_{AN} = - \sum_{a=0}^{t-1} \sum_{b=0}^{t-1} P_{AB}(a, b) \log_2 P_{AB}(a, b) \quad (12)$$

Where  $P_{AB}(a, b)$  represents the joint probability density of the converted image B and the original image A. Generally speaking, the larger the value of the joint entropy of the converted image and the original image, the richer the information contained in the image, so it can be used to evaluate the information richness of the image before and after the conversion.

The combination entropy of two images can be defined as:

$$H(A, B) = - \sum_{a,b} P_{AB}(a, b) \log_{P_{AB}}(a, b) \quad (13)$$

Where a represents the DN of image A, b represents the DN of image B,  $p_A(a)$  and  $p_B(b)$  represent the gray probability distribution density of A and B respectively, and  $p_{AB}(a, b)$  is the joint probability distribution of image A and B.

Structural similarity index establishes an image distortion evaluation model from the three factors of brightness, contrast and structure in terms of identifying things with human eyes. In this evaluation model, the mean is the estimated value of brightness, the standard deviation is the estimated value of image contrast, and the covariance can be used as a measure of the structural similarity of the image using the mean. This index can be used to compare the structural information of the image, thereby analyzing the distortion of the image, and objectively evaluating the image before and after conversion.

$$SSIM(A, B) = \frac{(2u_A u_B + c_1)(2\sigma_{AB} + c_2)}{(u_A^2 + u_B^2 + c_1)(\sigma_A^2 + \sigma_B^2 + c_2)} \quad (14)$$

Where  $u_A, u_B$  are the average of A, B;  $\sigma_{AB}$  is the covariance of image A and B;  $\sigma_A^2$  and  $\sigma_B^2$  are the variance of image A, B;  $c_1 = (k_1 L)^2, c_2 = (k_2 L)^2$  are constants to maintain stability, in which  $k_1 = 0.001, k_2 = 0.03$ .

### 2.2.2. Evaluation index based on image construction

The edge structure similarity (ESSIM) can reflect the edge information of the hexagonal grid image obtained from the source image during the conversion process. It is assumed that after the original image A is filtered by Sobel operator, the pixel  $f(i, j)$  gets the edge intensity  $g_A(i, j)$  and its direction  $a_A(i, j)$  are:

$$g_A(i, j) = \sqrt{S_A^x(i, j)^2 + S_A^y(i, j)^2} \quad (15)$$

$$a_A(i, j) = \tan^{-1} \left( \frac{S_A^y(i, j)^2}{S_A^x(i, j)^2} \right) \quad (16)$$

Where  $S_A^x(i, j)^2$  and  $S_A^y(i, j)^2$  are the horizontal and vertical amplitudes of pixels after Sobel respectively, the relative edge strength and direction values of the hexagonal grid image B and the original image A are:



$$G^{AB}(i, j) = \begin{cases} \frac{g_B(i, j)}{g_A(i, j)}, & \text{if } g_A(i, j) > g_B(i, j) \\ \frac{g_A(i, j)}{g_B(i, j)}, & \text{other wise} \end{cases} \quad (17)$$

$$A^{AB}(i, j) = \frac{\left| \alpha_A(i, j) - \alpha_B(i, j) \right| - \frac{\pi}{2}}{\frac{\pi}{2}} \quad (18)$$

Thus, we can get the edge retention of the converted image relative to the source image:

$$Q^{AB}(i, j) = Q_g^{AB}(i, j) Q_a^{AB}(i, j) = \Gamma_g \Gamma_a \left[ 1 + e^{K_g(G^{AB}(i, j) - \sigma_g)} \right]^{-1} \left[ 1 + e^{K_a(G^{AB}(i, j) - \sigma_a)} \right] \quad (19)$$

Where  $\Gamma_g, \Gamma_a, k_g, k_a, \sigma_g, \sigma_a$  are adjustable constant.

### 2.2.3. Normalization of evaluation indicators and calculation of weights

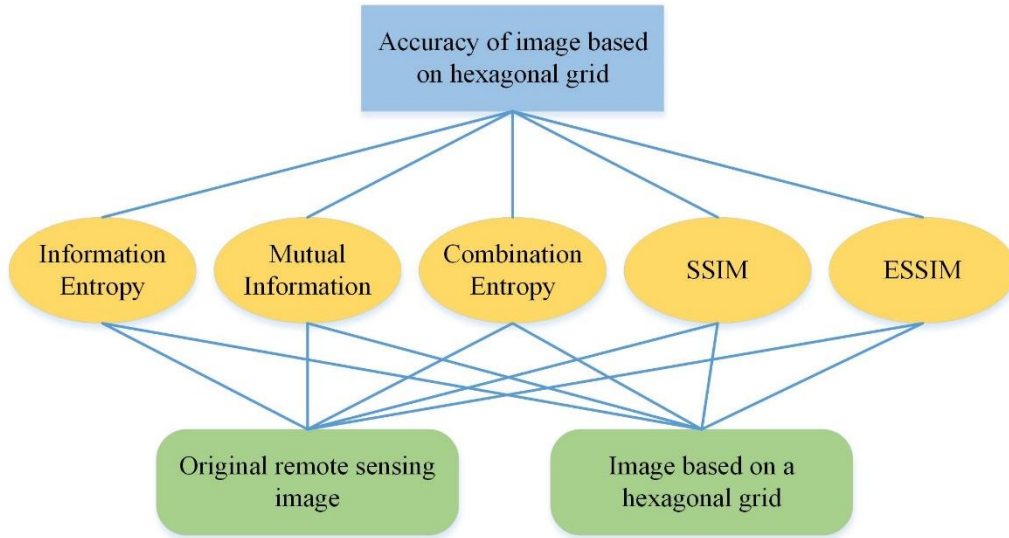
In order to simplify the calculation results, the index value calculated for the images before and after the conversion is calculated, with 1 as the standard. The closer the index is to 1, the higher the similarity of the images before and after the conversion. In order to simplify the basic evaluation index system, we use the analytic hierarchy process to determine the weights of the five major evaluation index systems.

The first layer: the target layer. The purpose of the model is to evaluate the information retention before and after conversion of the hexagonal grid image and the similarity to the original image. Therefore, the degree of similarity is regarded as the predetermined goal of decision-making.

The second layer: the indicator layer. Through literature research, this paper selects five indicators (Information Entropy, Mutual Information, Combination Entropy, SSIM and ESSIM) as indicators to quantify the preservation of image information before and after conversion, the similarity with the original image, and the structural similarity. These five indicators are intermediate evaluation standards.

The third layer: program layer. The objects we analyze are the original remote sensing image and the sampled remote sensing image based on the hexagonal grid, so the basic decision plan can be divided into the original remote sensing image and the hexagonal grid image.

Based on the above problem analysis, we established the following analytic hierarchy model:



**Figure 7.** The basic structure of AHP

And then, we use the analytic hierarchy process to determine the weight of the evaluation factor

(a). Construct a judgment matrix

We use the method of paired comparison and the 1-9 scale to construct the decision matrix  $A = (a_{ij})$

$$a_{ik} \times a_{kj} = a_{ij} \quad (20)$$

Where  $a_{ij}$  is set by standards of 1~9.

After analyzing the impact of the five major indicators on the overall goal, the following judgment matrix is constructed:

$$A = \begin{pmatrix} 1 & 1/2 & 4 & 3 & 3 \\ 2 & 1 & 7 & 5 & 5 \\ 1/4 & 1/7 & 1 & 1/2 & 1/2 \\ 1/3 & 1/5 & 2 & 1 & 1 \\ 1/3 & 1/5 & 3 & 1 & 1 \end{pmatrix} \quad (21)$$

According to the mutual influence of the five indicators, paired comparison matrixes are constructed as follows:

$$B_1 = \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix}, B_2 = \begin{pmatrix} 1 & 2 \\ 1/2 & 1 \end{pmatrix}, B_3 = \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix}, B_4 = \begin{pmatrix} 1 & 2 \\ 1/2 & 1 \end{pmatrix}, B_5 = \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix} \quad (22)$$

(b). Calculate eigenvalues and eigenvectors

MATLAB is used to find the maximum eigenvalue of matrix A  $\lambda_{max}$  and the corresponding matrix vector  $u = (u_1, u_2, \dots, u_n)^T$ , and the following formula to normalize u.

$$x_i = \frac{u_i}{\sum_{i=1}^n u_i} \quad (23)$$

(c). Consistency inspection

Calculate the consistency index CI:

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (24)$$

Calculate the consistency ratio CR:

$$CR = \frac{CI}{RI} \quad (25)$$

According to the calculation, we get  $CR=0.0145<0.1$ , which means that matrix has already passed the consistency check.

In terms of paired comparison matrixes  $B_1, B_2, B_3, B_4, B_5$ , we can find the weight vector of the hierarchical total sorting and perform the consistency test. Through calculation, it can be known that  $B_1, B_2, B_3, B_4, B_5$  passed the consistency test.

Then we calculated the total ranking weight and consistency test. According to the calculation, the weights of B on the overall goal are as follows:

**Table 2.** The Consistency index of five evaluation indexes

k	Information Entropy	Mutual Information	Combination Entropy	SSIM	ESSIM
CR	0.0048	0.0013	0.0025	0.0079	0.0014

$CR<0.1$ , Thus, total ranking passes consistency test.

**Table 3.** The weight of the objective options

Objective	Original remote sensing image	Hexagonal pixel image
CR	0.5179	0.4821

Thus, the weight of each index can be determined: Information Entropy: 0.21, Mutual Information: 0.18, Combination Entropy: 0.24, SSIM: 0.24 and ESSIM: 0.17.

#### 2.2.4. Geometric evaluation index

In addition, this study also calculated the change in the DN value of the pixel unit and in line length  $l_c$  before and after sampling.

In order to evaluate the credibility of the hexagonal grid image generated, this paper uses the checkpoint method. We randomly select 30 feature points in a scene image as the checkpoints, and use RMSE of pixel value to perform the accuracy evaluation remote sensing image based on the hexagonal grid. The formula is:

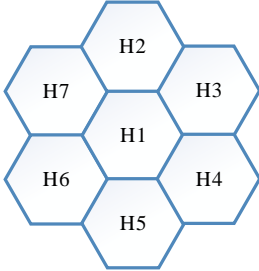
$$RMSE = \sqrt{\frac{\sum_{i=1}^n (H_i - S_i)^2}{n}} \quad (26)$$

Where  $H_i (i = 1, 2, \dots, n)$  is the pixel value of the checkpoint on the original image;  $S_i$  is the pixel value of the remote sensing image based on the hexagonal grid; RMSE is the root mean square error of the pixel value.

In order to evaluate the retention of the geometric features of the image during the image conversion process, this paper performs a refinement operation on the local area of the Landsat 8 images before and after the conversion, and compares the differences between the two.

When refining the same target, the hexagonal refinement algorithm only needs six templates, and each template includes 7 units, while the square refinement requires eight templates, and each template requires 9 units.

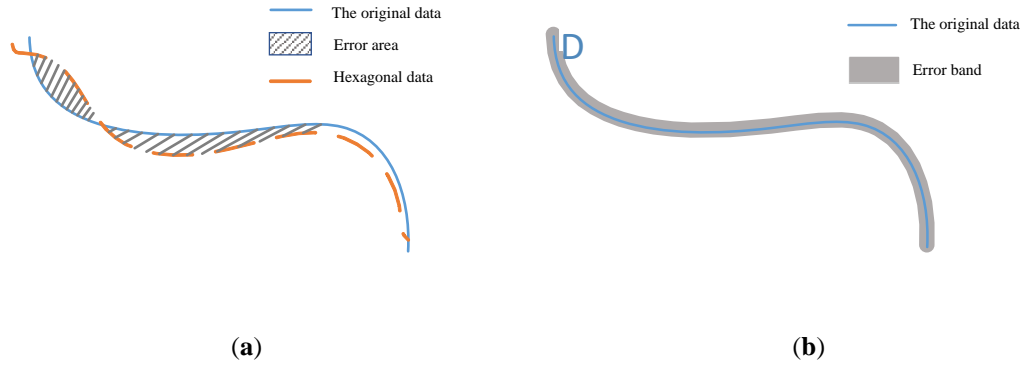
**Table 4.** Comparison of the refinement zero conditions of rectangular pixels and hexagonal pixels

<table border="1"> <tr> <td>P3</td><td>P2</td><td>P9</td></tr> <tr> <td>P4</td><td>P1</td><td>P8</td></tr> <tr> <td>P5</td><td>P6</td><td>P7</td></tr> </table>	P3	P2	P9	P4	P1	P8	P5	P6	P7	<p>Zero condition</p> <ol style="list-style-type: none"> <li><math>2 \leq NZ(P1) \leq 6;</math></li> <li><math>Z0(P1) = 1;</math></li> <li><math>P2 \times P4 \times P8 = 0;</math></li> <li><math>P2 \times P4 \times P6 \times P9 = 0</math></li> </ol>	 <p>Zero condition</p> <ol style="list-style-type: none"> <li><math>2 \leq NZ(H1) \leq 6;</math></li> <li><math>Z0(H1) = 1;</math></li> <li><math>H2 \times H4 \times H6 = 0;</math></li> <li><math>H3 \times H5 \times H7 = 0</math></li> </ol>
P3	P2	P9									
P4	P1	P8									
P5	P6	P7									
(a)		(b)									

In order to quantitatively analyze the results of the refinement of linear features, this paper uses the error band to indicate the degree of deviation between the original image refinement results and the hexagonal grid image refinement results.

$$D = \frac{\Delta s}{L} \quad (27)$$

Where  $\Delta s$  is the error area between the linear features extracted by the two image formats, and this parameter is the number of grid cells between the two linear features multiplied by the cell area.  $L$  is the average length of linear elements extracted by the two image formats, which is determined by the number of grid cells through which the linear elements pass.  $D$  is the average error width under this length.



**Figure 8.** Schematic diagram of error band calculation

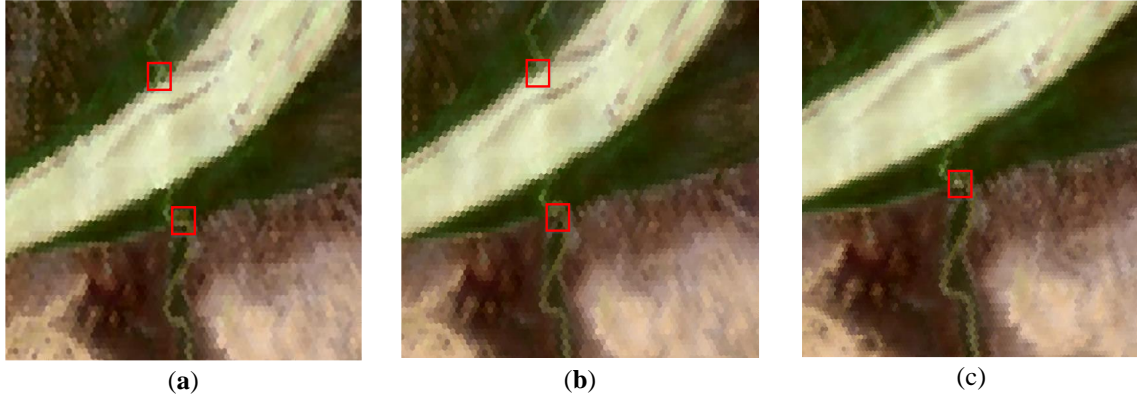
This parameter can quantify the degree of agreement between the linear elements extracted from the original image and the hexagonal grid image.

### 3. Results

#### 3.1. Remote sensing data modeling based on hexagonal grid

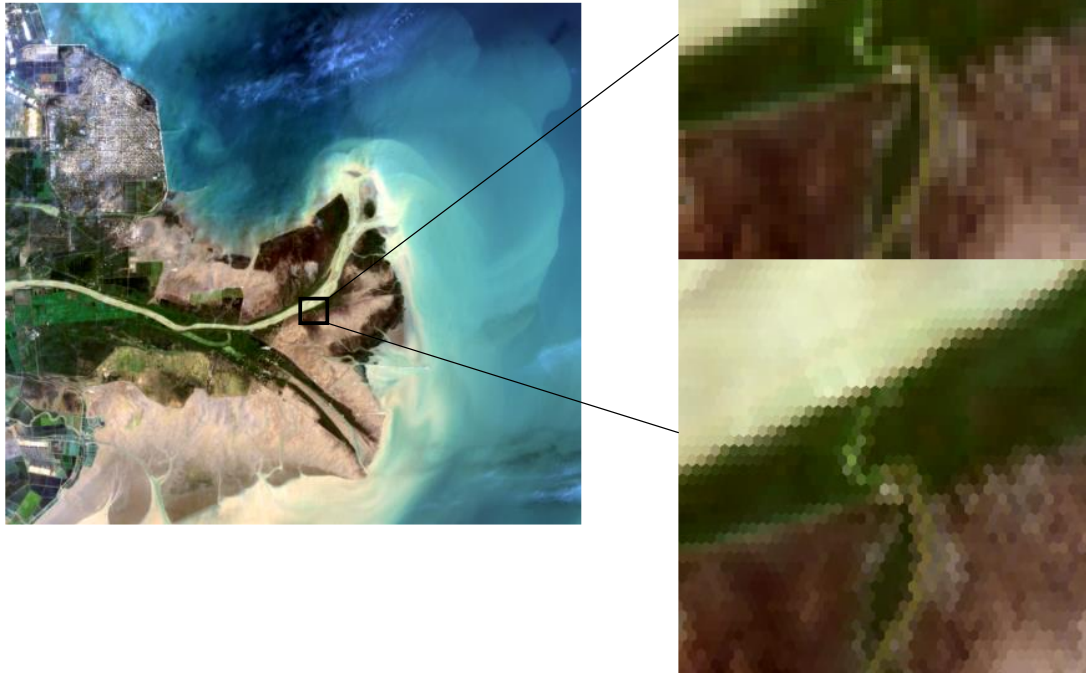
The nearest neighbor interpolation, bilinear interpolation, and cubic convolution interpolation methods are used to convert remote sensing images of different resolutions (high, medium, and low resolution respectively, using GF-1, Landsat 8 images as examples) to data based on hexagonal grid.

The conversion of remote sensing data to hexagonal grid data includes two stages: the first stage is the construction of the regional hexagonal grid; the second stage is the remote sensing data sampling based on the hexagonal grid. First, it is necessary to determine the four-corner coordinates of a scene of remote sensing data, and determine the scope of grid construction according to the four-corner coordinates. Secondly, the unit size of the hexagonal grid is determined according to the spatial resolution of the remote sensing image. For example, the spatial resolution of Landsat 8 is 30 meters, then according to formula (1) and Table 1, it can be determined that the grid level corresponding to Landsat 8 is 18 layers, and the area of the hexagonal unit of the grid on the 18th layer is 989.657 square meters. After determining the scope of grid construction, grid level, we can fill in the hexagonal grid of the corresponding level within the four corner coordinates of the remote sensing image.



**Figure 9.** The hexagonal grid images converted from Landsat 8 images in the same area using different resampling methods. (a) is generated by nearest neighbor interpolation, (b) is generated by bilinear interpolation and (c) is generated by bicubic interpolation.

Using different resampling methods to convert remote sensing images in the same area into hexagonal grid images has different effects. Figure 6 shows the interpolation results of Landsat 8 images in the same area (Yellow River Estuary) using different interpolation methods. It can be clearly seen that, compared with nearest neighbor interpolation, bilinear interpolation and bicubic interpolation perform pixel value operations, and the image is smoother. The pixel value of the nearest interpolation has not been calculated, so the original characteristics of the remote sensing image can be better maintained, and the result of linear interpolation is smoother, for example, the boundary of the river channel does not appear jagged. The result of cubic interpolation can reflect more information and image features, but the calculation of a single pixel involves 16 adjacent pixels, which requires more work.



**Figure 10.** The figure shows the expression of Landsat 8 image based on hexagonal grid and rectangular pixels



**Figure 11.** The figure shows the expression of GF-1 image based on hexagonal grid and rectangular pixel

Figure 10 and figure11 show the results of the conversion of Landsat 8 (Yellow River Estuary Region) and GF-1 (Hainan Region) images using bilinear interpolation.





**Figure 12.** The figure shows the expression of Landsat 8 image based on hexagonal grid on different Level. (a) is on level 18, (b) is on level 19

The side length ratio of the hexagons between adjacent grid levels is 1:2. As shown in the figure 12, the grid unit side length ratio of the 19th level grid and the 18th level grid is 1:2, the 19th level Grid image can show more information and image details.

### 3.2. Accuracy evaluation of different interpolation methods

#### 3.2.1. Basic evaluation indicators

This study calculated the basic evaluation indicators of Landsat 8 and GF-1 remote sensing image before and after conversion.

**Table 5** Evaluation results based on basic evaluation indexes and image structure evaluation indexes

Image	Interpolation Method	Information Entropy	Mutual Information	Combination Entropy	SSIM	ESSIM	Comprehensive Index
Landsat 8	Nearest neighbor interpolation	0.989	1.028	1.011	0.987	0.965	0.9968
	Bilinear interpolation	0.978	1.042	1.014	0.981	0.964	0.9963
	Bicubic interpolation	0.984	1.047	1.017	1.011	0.958	1.0042
GF-1	Nearest neighbor interpolation	0.973	1.021	1.019	0.978	0.969	0.9930
	Bilinear interpolation	0.972	1.043	1.017	1.008	0.964	1.0014
	Bicubic interpolation	0.977	1.028	1.011	0.974	0.960	0.9908

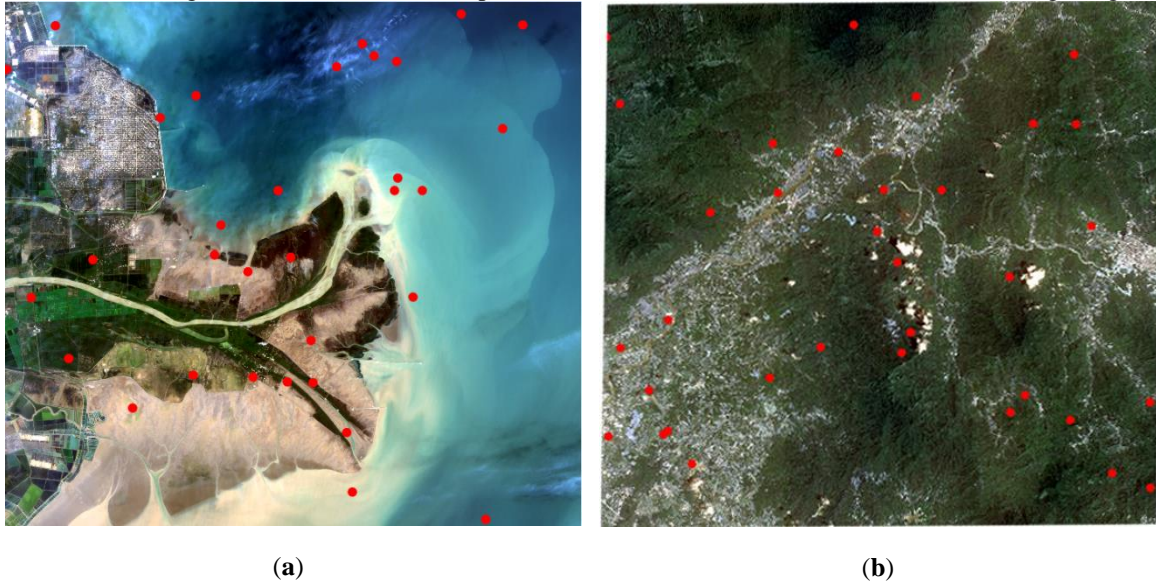
The calculation results are shown in Table 5. Comparing the values evaluated by comprehensive indicators before and after conversion of the images of the Landsat 8 and GF-1, it can be seen that the comprehensive evaluation result of the hexagonal image obtained using the nearest neighbor interpolation is closest to 1, Which shows that the hexagonal grid image obtained by nearest neighbor interpolation maintains the most information

and structural features compared with the original image. The main reason is that in the process of using the nearest neighbor interpolation, the attribute value of the hexagonal unit has not undergone further calculations, so the characteristics and attributes of the original remote sensing data can be retained to the greatest extent.

### 3.2.2. Evaluation index based on image feature

#### (1) Accuracy evaluation for points

The following are the results of random points selection of Landsat 8 and GF-1 remote sensing images:



**Figure 13** Checkpoints randomly selected on remote sensing images

The above is a schematic diagram of randomly selected check points for a scene image to calculate. It can be seen that selecting thirty checkpoints for a scene image can basically cover different types of features, so that degree of retention of the point feature of the hexagonal grid image can be calculated for the original image.

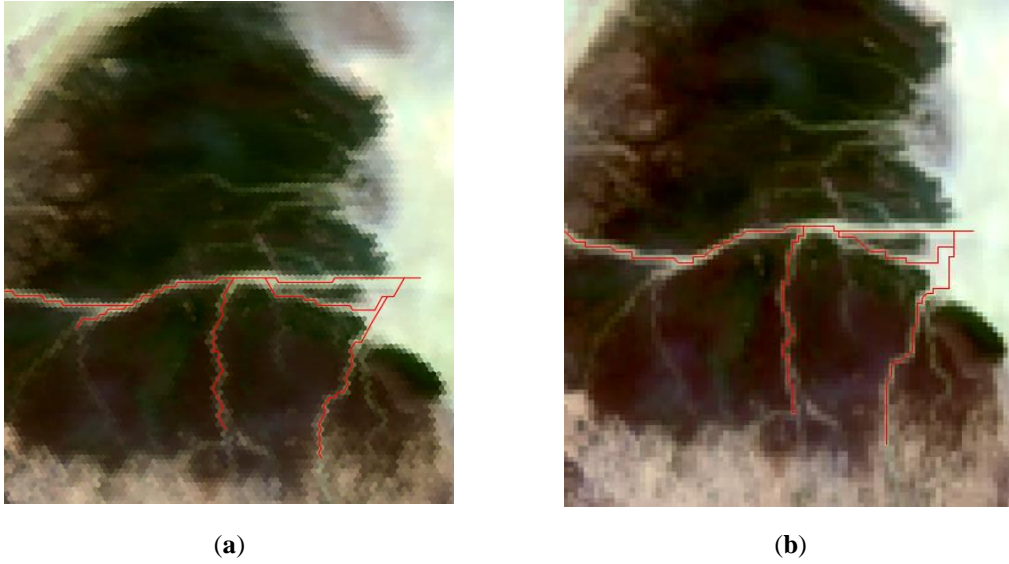
**Table 6.** The RMSE of hexagon grided images from different data resources

Image	Interpolation method	RMSE of grid pixel value
Landsat 8	Nearest neighbor interpolation	6.3
	Bilinear interpolation	7.1
	Bicubic interpolation	8.3
GF-1	Nearest neighbor interpolation	11.3
	Bilinear interpolation	10.5
	Bicubic interpolation	12.4

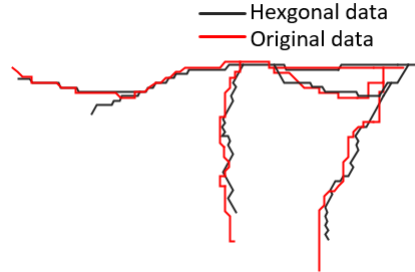
The resolution of GF-1 image and Landsat 8 image is not the same. The spatial resolution of Landsat 8 OLI land imager is 30 meters, and the spatial resolution of GF-1 PMS image is 16m. As shown in Table 2, We calculated the RMSE of the pixel value of the checkpoint on images obtained by different sampling methods. Compared with bilinear interpolation and bicubic interpolation, the RMSE of the nearest neighbor interpolate on is the smallest. The reason is that the process of establishing a hexagonal grid is regular, and the nearest neighbor interpolation directly takes the pixel value of the nearest pixel in Euclidean geometry and assigns it to the target grid unit without linear calculation. Therefore, the nearest neighbor interpolation does not have a smooth process, which makes it easier to save the characteristic pixel points and land reflection information of the original pixels.

#### (2) Accuracy evaluation for the line

As shown in the figure below, the sensitivity of the hexagonal grid image to angle changes is better than that of the original image, and it is better than the quadrilateral image in terms of maintaining traits, and can well maintain the curved characteristics of linear features. Hexagonal have consistent direction and better angular resolution than quadrilateral elements. The quadrilateral has two directions, the distance between these two directions is 1 and  $\sqrt{2}$ .



**Figure 14.** The picture shows the original image and the hexagonal grid image for linear element extraction and the result of the refinement operation; (a) is the Landsat 8 hexagonal grid image of Level 18; (b) is the original remote sensing image of Landsat 8.



**Figure 15.** Comparison of linear feature extraction results between Landsat 8 original image and hexagonal grid image.

According to equation (27), it can be calculated that the width of the error band is 22.3m. The error band between the linear elements extracted from the 19-layer grid image and the current elements extracted from the original image is 17.6m. This shows that the higher the grid image level, the better the image can maintain the geometric characteristics of the original image.

#### 4. Discussion

The hexagonal grid and the original remote sensing data are two different form of data management and data storage, and there are big differences in pixel units. This study tends to prove that the remote sensing data based on the hexagonal grid has practical applications. Availability requires quantitative indicators to prove the similarity between the hexagonal grid data and the original data. This paper used basic evaluation indicators to prove that the amount of information in the hexagonal grid data is within a controllable range compared to the original remote sensing data. There is inevitable data loss and smoothing, but the amount of lost data is within an acceptable range. This study compared different resampling methods, including nearest neighbor interpolation, linear interpolation and cubic interpolation. The weights calculated in this study are not the only indicator weights. When being used in different application scenarios, other appropriate judgment matrix can



be selected according to the focus of remote sensing applications, so as to determine the weights of different indicators accordingly.

The calculation results of the quantization index show that the image obtained by the nearest interpolation is the closest to the original image and retains more information. In addition, based on the differences between hexagons and quadrilaterals, this paper proposed a new algorithm for thinning hexagonal grid images, and further compares the features of linear features obtained by the thinning algorithm. This also shows that the hexagonal grid image is available in remote sensing geometry applications. For example, large area yield estimation land use classification, river extraction, etc. based on hexagonal grid image. In addition, according to the feature that the level of the hexagonal grid can be infinitely divided, the grid can be used to store unstructured emerging spatial big data, including point, line, human activity, luminous etc.

## 5. Conclusions

This research brought up and implemented the conversion method from the original remote sensing image to the hexagonal discrete grid image, and is committed to finding a suitable way to express the remote sensing data on the discrete grid, and establish a suitable accuracy evaluation system to evaluate the accuracy of modeling aims at the remote sensing data based on the hexagonal grid. This proved its availability and provides support for the further use of remote sensing data based on hexagonal grids. The converted hexagonal data can basically maintain the basic and geometric characteristics of the original data. However, due to the geometrical characteristics of the hexagon compared to the quadrilateral, the direction and angle of the hexagonal data are more sensitive to line elements, which means it can keep the direction better. Among the three resampling methods, the image obtained by the nearest neighbor interpolation has the highest similarity to the original image. The main reason is that the pixel value in the nearest interpolation has not been calculated, and the geographic coordinates and the hexagonal grid coordinates are different. Because the correspondence is regular, so the nearest neighbor interpolation can better retain the characteristics of the original image.

**Author Contributions:** Conceptualization, Y.M. and G.L.; methodology, X.Y., Z.L.; software, Y.M. and X.Y.; validation, Y.M. and Y.X.; formal analysis, Y.M.; investigation, J.B.; resources, Y.M.; data curation, Q.C., L.Z.; writing—original draft preparation, Y.M.; writing—review and editing, Y.M. and X.Y.; visualization, Y.M. and R.W.; supervision, X.Y., G.L.; project administration, X.Y.; funding acquisition, G.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the Strategic Priority Research Program of Chinese Academy of Sciences [grant number XDA19020201].

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Guo, H. Big Earth data: A new frontier in Earth and information sciences. *Big Earth Data* **2017**, *1*, 4-20.
2. Yao, X.; Li, G.; Xia, J.; Jin, B.; Zhu, D. Enabling the Big Earth Observation Data via Cloud Computing and DGGS: Opportunities and Challenges. *Remote Sensing* **2019**, *12*, 62.
3. Nan, L.; Cheng, C.; Ma, H.; Yang, Y. Global discrete grid systems analysis and comparison. In Proceedings of Geoscience & Remote Sensing Symposium.
4. Sahr, K.; White, D.; Kimerling, A.J. Discrete Global Grid System. *Cartography & Geographic Information Science* **2003**, *30*, 121-134.
5. Sahr, K.; White, D.; Kimerling, A.J. Geodesic Discrete Global Grid Systems. *American Cartographer*.
6. Yao, X.; Li, G. Big spatial vector data management: a review. *Big Earth Data* **2018**, *2*, 108-129, doi:10.1080/20964471.2018.1432115.
7. Sirdeshmukh, N. Utilizing a Discrete Global Grid System For Handling Point Clouds With Varying Locations, Times, and Levels of Detail. **2018**.
8. Goodchild, M.F.; Kimerling, A.J.E. Discrete Global Grids: A Web Book. **2002**.
9. Wang, L.; Ai, T.; Shen, Y.; Li, J. The isotropic organization of DEM structure and extraction of valley lines using hexagonal grid. *Transactions in Gis* **2020**.

10. Zhou, J.; Ben, J.; Wang, R.; Zheng, M.; Yao, X.; Du, L. A novel method of determining the optimal polyhedral orientation for discrete global grid systems applicable to regional-scale areas of interest. *International Journal of Digital Earth* **2020**, *13*, 1553-1569, doi:10.1080/17538947.2020.1748127.
11. Her, I. Geometric transformations on the hexagonal grid. *IEEE Transactions on Image Processing* **2002**, *4*, 1213-1222.
12. Xia, J.; Yang, C.; Li, Q. Building a spatiotemporal index for Earth Observation Big Data. *International Journal of Applied Earth Observation & Geoinformation* **2018**, *73*, 245-252.
13. Devendran, A.A.; Lakshmanan, G. A Review On Accuracy and Uncertainty of Spatial Data and Analyses with special reference to Urban and Hydrological Modelling. In Proceedings of Isprs Annals of the Photogrammetry, Remote Sensing & Spatial Information Sciences, li-8, Isprs Technical Commission VIII Symposium, 09 – 12 December, Hyderabad, India.
14. Freeman, L.K. WebGeochem: DGGS Geochemical Sample Analysis Search Engine. **2005**.
15. Wenshuai, Y.U.; Tong, X.; Jin, B.; Xie, J. The Accuracy Control in the Process of Vector Line Data Drawing in the Hexagon Discrete Global Grid System. *Journal of Geo-Information Science* **2015**, *6*, 70-76.
16. Gregory, M.J.; Kimerling, A.J.; White, D.; Sahr, K. A comparison of intercell metrics on discrete global grid systems. *Computers Environment & Urban Systems* **2008**, *32*, 188-203.
17. White, D.; Kimerling, A.J.; Sahr, K.; Song, L. Comparing area and shape distortion on polyhedral-based recursive partitions of the sphere. *International Journal of Geographical Information Systems* **1998**, *12*, 805-827.
18. Zhang, B.; Yuan, Z.Y.; Zhao, X.S.; Zhang, X.J.; Technology. An Geometry Deformation Evaluation Index of the Spherical Discrete Grid Based on the Fuzzy Similarity. *Geography and Geo-Information Science* **2015**.
19. Sahr, K. Hexagonal discrete global grid systems for geospatial computing. *Archives of Photogrammetry* **2011**, *22*.
20. Chandio, I.A.; Matori, A.N.B.; Wanyusof, K.B.; Talpur, M.A.H.; Lawal, D.U. GIS-based analytic hierarchy process as a multicriteria decision analysis instrument: A review. *Arabian Journal of Geosciences* **2013**, *6*.
21. Ye, S.; Liu, D.; Yao, X.; Tang, H.; Xiong, Q.; Zhuo, W.; Du, Z.; Huang, J.; Su, W.; Shen, S., et al. RDCRMG: A raster dataset clean & reconstitution multi-grid architecture for remote sensing monitoring of vegetation dryness. *Remote Sensing* **2018**, *10*, doi:10.3390/rs10091376.
22. Luo, J.; Zhang, W.; Su, J.; Xiang, F. Hexagonal Convolutional Neural Networks for Hexagonal Grids. *IEEE Access* **2019**, *7*, 142738-142749, doi:10.1109/ACCESS.2019.2944766.
23. Purss, M.B.J.; Gibb, R.; Samavati, F.; Peterson, P.; Jin, B. The OGC® Discrete Global Grid System core standard: A framework for rapid geospatial integration. In Proceedings of 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
24. Jin, B.; Chong, T, X.; Wang, L.; Wen, B.G. Key Technologies of Geospatial Data Management Based on Discrete Global Grid System. *Journal of Geomatics Science & Technology* **2010**.
25. Tong, X. The Principles and Methods of Discrete Global Grid Systems for Geospatial Information Subdivision Organization. *Acta Geodaetica Et Cartographica Sinica* **2011**.
26. Ai, T.; Li, J. A DEM generalization by minor valley branch detection and grid filling. *Isprs Journal of Photogrammetry & Remote Sensing* **2010**, *65*, 198-207.
27. wang, L.; Ai, T. The valley extraction based on the hexagonal grid-based DEM. *Journal of Geodesy and Geoinformation Science* **2019**, *048*, 780-790.

28. Tong, X.; Jin, B.; Qing, Z.; Zhang, Y. The hexagonal discrete global grid system appropriate for remote sensing spatial data. *Proceedings of SPIE - The International Society for Optical Engineering* **2009**, 7146.
29. Zhou, M.; Chen, J.; Gong, J. A pole-oriented discrete global grid system. *Computers & Geosciences* **2013**.
30. Tao, M.; Wen, Y.; Peng, G.X.; Su, L.; Wang, Z. The Study on Error Analysis of Discretization Area in Discrete Global Grid System. *Procedia Environmental Sciences* **2011**, 10, 0-1128.
31. Wang, H. A Study of Evaluation Methods on Performance of the Multi-Source Remote Sensing Image Fusion. *Computer Engineering & Applications* **2003**.
32. Zhao, S.H.; Feng, X.Z.; Kang, G.D.; Elnazir, R. MULTI-SOURCE REMOTE SENSING IMAGE FUSION BASED ON SUPPORT VECTOR MACHINE. *Chinese Geographical Science* **2002**, 12, 244-248.
33. Fan, S.; Zao, C.; Zou, K. A Review on the Accuracy Analysis of Spatial Scaling Data. *Progress in Geography* **2013**, 31, 1574-1582.
34. Feng, X.; Qin, K.; Cui, W.; Chen, Y.; Li, X. Multiscale description and recognition of target shape in high-resolution remote sensing images. *Journal of Remote Sensing* **2014**, 18, 90-104.
35. Tulloch; Ayesha, I.T.; Sutcliffe; Patricia; Naujokaitis-Lewis; Ilona; Tingley; Reid; Brotons; Lluís. Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation* **2016**.
36. XU Zhiying, H.Y., LIU Yue, YAN Yan. A Review on the Accuracy Analysis of Spatial Scaling Data. *PROGRESS IN GEOGRAPHY* **2012**, 31, 1574-1582, doi:10.11820/dlkxjz.2012.12.002.
37. Hong, W.; Cheng, C.; Xiao, Z. An hierarchical information entropy model for coverage estimation of coastal areas based on an adaptive Discrete Global Grid System. In Proceedings of International Conference on Geoinformatics.