2022 IEEE 22nd International Conference on Software Quality, Reliability, and Security Companion (QRS-C) | 979-8-3503-1991-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/QRS-C57518.2022.00101

Mapping Method between 2D Landscape Image and 3D Spatial Data based on Adversarial Relative Depth Constraint Network

Shuhao Wang[1], Zhijun Lin[2,*], Zhimo Weng[3], and Anna Li[4]
[1]School of Information Management, Wuhan University, Wuhan, China
[2]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China
[3]China-Korea Institute of New Media, Zhongnan University of Economics and Law, Wuhan, China
[4]College of Professional Studies, Northeastern University, Boston, MA, USA
wangshuhao@whu.edu.cn, Linzhijun@whu.edu.cn, 202121250239@stu.zuel.edu.cn, anna.li6811@yahoo.com
*corresponding author

Abstract—**This paper proposes a mapping method of 2D image and 3D spatial data based on the adversarial relative depth constraint network. The steps are as follows: 1) Input pixel coordinates of key nodes of 2D landscape image, and conduct normalization preprocessing; 2) Input two-dimensional pixel coordinates into the depth prediction network and output the depth values of key nodes; 3) Using depth values and two-dimensional pixel coordinates to reconstruct three-dimensional coordinates of key nodes; 4) Input DEM data to the discriminator of the generated adversarial network to calculate the authenticity error value, and use the relative depth information between the attitude characteristics of mountain and hydrology and the corresponding key nodes of the image to calculate the relative depth error; 5) Add the authenticity error and relative depth error calculated above to get the total error, and feed back to the depth prediction network to get a more accurate mapping evaluation, so as to realize mapping discovery. The problems solved in this paper include: lack of characteristic pose data in the traditional geo-evidence-based process of 2D landscape images; The results of the generative adversarial network method do not conform to the relative depth relationship of feature points in 3D spatial data.**

Keywords- Adversarial learning; constrained networks; spatial data; DEM; position superposition

I. INTRODUCTION

Chinese landscape painting, which has lasted for thousands of years, is a unique category of Chinese cultural resources. The scientific interpretation of landscape images is an urgent requirement for the continuation of Chinese civilization, and a concrete manifestation of technology-enabled humanities research. The traditional category of landscape painting research is mainly based on mutual evidence, comparison and discrimination between art archives and painting recording techniques, which stays in the textual research limitation. What cannot be ignored is that the landscape records originated in China, in addition to the beauty of art, also contain a variety of concrete geographical characteristics of mountains, hydrology, plants and other information in various historical stages of China. In order to build up cultural confidence, build mutual learning among civilizations, and construct the scientific discourse system of Chinese art research, it is urgent to carry out higher level research on the civilization landscape recorded in landscape images.

The technical path to perform evidence-based landscape painting is to process a single landscape image and achieve prediction by analyzing multiple consecutive complex structural features. In deep learning-based landscape feature estimation [1], the deep network is converted into a series of heat maps representing probabilities, where the most probable points are key coordinate points. Each key coordinate point represents a node in the mountain survey trajectory. The use of the adversarial network can enhance the ability of deep neural networks to grasp the hydrological features of mountains in images. In the field of image interpretation intelligence, identifying the spatial information in traditional landscape paintings is an important part of predicting the location and inferring the behavior of the trajectory. The purpose of the mountain hydrological feature mapping experiment is to locate the boundary and position of geo-edge information in the image using relevant algorithms. The technique of locating mountain hydrological features from images allows the machine to locate specific geo-environmental locations more accurately and to determine the movement trajectory and behavior habits of the observing recorder [2].

II. MOTIVATION

The pinnacle of Chinese landscape painting, "Dwelling in the Fuchun Mountains", is selected as a case study. We take the author Huang Gongwang's observation and recording behaviors of the mountains and rivers during the whole creation process as the core of the analysis, and focus on the evolutionary factors of the survey perspective participation intensity, cartographic efficiency changes, and the fusion of the survey and the process of merging with the subjective painting orientation during the image generation stage. The effective integration of data, images and associated space is realized, which is transformed into effective image data resources, meta-learning framework construction and spatial calculation evaluation of image scenario information, attempting to discover the precise location of continuous mountain bodies recorded by creation and restoring the complete trajectory of the painter's creation sketching through spatial analysis.

2693-9371/22/$31.00 ©2022 IEEE
DOI 10.1109/QRS-C57518.2022.00101

Before discussing the study of situational space in Chinese landscape painting, we must objectively understand the scientific value of objective observation and subjective recording of geographical environment in ancient Chinese painting. At the same time, it is necessary to break the wrong evaluation system that directly adopts the western evaluation perspective or forcibly matches in the GIS environment. Therefore, the main research goals are as follows: 1) Establish the label specification of geographical features implanted in map scene space [3]; 2) Realize the looming space position marking in landscape images; 3) Through the intervention of the concept of picture scene, the spatial data in the target image can be activated and the spatial topological efficiency can be calculated under fuzzy logic.

III. Background

Since the cross-disciplinary research paradigm has not yet been uniformly defined in the academic community, in order to facilitate the reader's understanding, this paper summarizes the background proposed in the study as follows: 2D landscape scenario refers specifically to the total set of time dimensions, spatial scenes and activity relations constituted by the presence of people in them in ancient Chinese landscape images. The data environment, in which people's presence and activities are present, extends the image information ontology to the ubiquitous domains. The involvement of multiple disciplines has triggered the study of image scenarios, and the extensive participation of researchers urgently requires the fusion of image scenario-embodied data with intertwined reality studies, so that any image scenario (whether real or virtual) can be effectively transformed into linked domains with diversified data characteristics, thus realizing the activated computation and instantly triggered experience of image scenario perception.

And to study 2D landscape images, it is necessary to first construct the landscape map situational study association relations, solve the literature collection of painting theory, geography, prescriptions, and notes in the historical generations, realize the text extraction and arrangement, and solve the problem of cognitive biases and historical geographic virtual location data for annotation matching and collaborative conversion; then, the landscape image meta-learning framework is proposed so that the ontological association knowledge of the situational space can be effectively converted to have diversified Data features are linked so that the situational location-aware annotation with the intervention of domain experts triggers the activation of instant computation; finally, it reveals how the landscape image meta-learning framework takes the existence of objective geospatial features from the perspective of ancient cartographic observation of French logic, with the route traces in the process of sketching by the painter traveling the road as the link and the remote sensing elevation data as the reference. For the above information to achieve matching with DEM data, it must appear in the form of measurable data or mimetic state.

The image data mimicry proposed in this paper introduces the concept of ecology and consists of the mimicry image, the simulated image and the perceptual error (subjective reality) together. There is a certain degree of homogeneity and simultaneity among the three, but they do not appear absolutely at the same spatial-temporal location. Part of the image mimicry consists of only the mimic and the simulated object. On the one hand, the environment of image mimicry is not a mirror reflection of the real environment, not a record of the "real" objective environment, and more or less deviates from the real environment; on the other hand, the image mimicry is not completely separated from the real environment, but is based on it.

By combining the complex behavior generation adversarial network [4], scarcity detection network [5] and DEM spatial data set [6], the label information is transformed into activation function, so that the matrix data can be nonlinear processed. Meanwhile, the cosine dynamic learning rate attenuation strategy is used to control gradient updating to reduce overfitting. The self-focused adversarial network is used to replace the common generation adversarial network to extract the key nodes, achieve the spatial comparison similarity evidence-based composite results, and verify the effectiveness through the evaluation mechanism [7].

IV. Methods

A method for mapping 2D landscape image and 3D spatial data based on an adversarial relative depth constrained network includes the following steps:

Step A: Input 2D pixel coordinates of 2D landscape image nodes and normalize pre-processing.

Step B: Input the geographic survey coordinate nodes tagged by domain expert identification and the normalized pre-processed 2D pixel coordinates into the depth prediction network to output the depth values of mountain hydrological nodes.

Step C: Reconstructing the 3D coordinates of the nodes using the depth values of the nodes with 2D pixel coordinates to obtain the reconstructed 2D image-2.5-dimensional view of the geoid feature pose.

Step D: Inputting the DEM data model, selecting the matching viewpoint and then corresponding the 3D coordinates of the marked nodes to obtain the reconstructed 3D spatial data-2.5-dimensional viewpoint of the sparse geoid feature pose.

Step E: To the discriminator of the generative adversarial network for the authenticity error calculation, and the relative depth error calculation using the relative depth information between the reconstructed geoid feature poses and the corresponding nodes of the images.

Step F: The authenticity error calculated by the discriminator of the generative adversarial network is summed with the relative depth error to obtain the total error and fed back to the depth prediction network, which constrains the depth prediction network to predict a more accurate depth value and thus reconstructs to obtain a more accurate mapping evaluation, then enables mapping discovery.

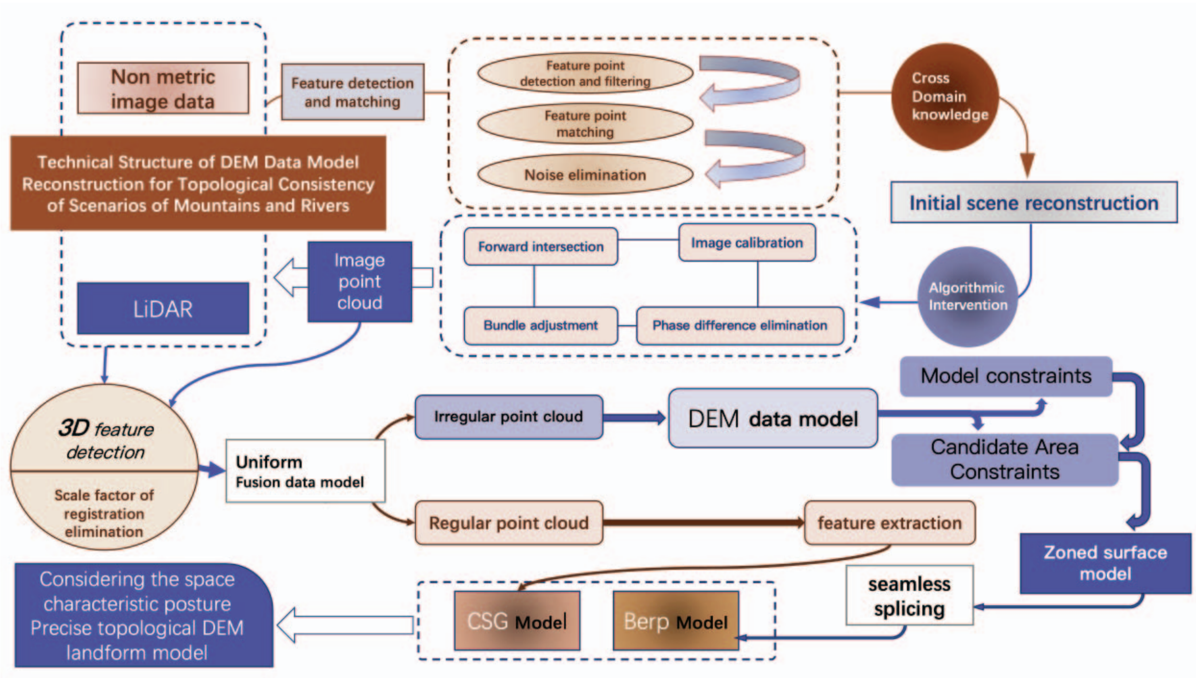

Figure 1. DEM Data Model Reconstruction Technical Structure

1) In step A, for each mountain area, the 2D pixel coordinates of each detection coordinate point (including the peaks, ridges, valleys, cliffs, etc. of the first category; the intersections of road networks, water sources, habitat areas, vegetation areas, etc. of the second category) are subtracted from the mean value of the 2D pixel coordinates in the mountain discrimination curve, and then divided by the standard deviation of the 2D pixel coordinates in the mountain discrimination curve to obtain the normalized pre-processed 2D pixel coordinates.
2) In step B, the normalized pre-processed two-dimensional pixel coordinates of each node obtained in the previous step are inputted to a depth prediction network consisted of three modules for predicting the depth values of the mountain feature nodes, which includes the following steps [8][9]:
   a) Input the normalized pre-processed two-dimensional pixel coordinates into the feature extraction module to extract features, which consists of a fully connected layer containing 1024 neurons and a linearly rectified activation function layer.
   b) The features extracted by the feature extraction module are inputted to the scarcity network module for feature learning, which consists of two scarcity blocks, each scarcity block inputs the output value of the previous layer of the neural network to a layer of fully connected layer containing 1024 neurons and a layer of linear rectification activation function layer to output preliminary feature values, followed by inputting the preliminary feature values to a fully connected layer containing 1024 neurons to output further eigenvalues, then the further eigenvalues are added to the input values input to the scarcity block, and finally the summed eigenvalues are inputted to a linear rectification activation function layer to output the scarcity block eigenvalues to the next layer of the neural network.
   c) The output features of the scarcity network module are inputted to the depth-value regression module, which consists of a fully connected layer containing the corresponding number of neurons, and the depth-value regression module inputs the output features of the scarcity network module and outputs the depth values of the corresponding number of nodes.
3) In step C, the depth values of the corresponding number of joints with 2D pixel coordinates are used to reconstruct the 3D coordinates of the joints as follows:

Suppose the two-dimensional pixel coordinates of a joint (a class of coordinates) point of the mountain are (u, v), where u is the lateral coordinate of the joint in the image and v is the longitudinal coordinate of the joint in the image; suppose the depth value of the joint obtained in the previous step prediction is H and the focal length corresponding to the image is f. The three-dimensional coordinates of the joint are $(\frac{uH}{f}, \frac{vH}{f}, H)$; the three-dimensional coordinates of each joint are reconstructed to be able to reconstruct the three-dimensional coordinates of the mountain contour pose, and the three-dimensional coordinates of the corresponding number of joints form the fuzzy three-dimensional structural feature pose.

4) In step D, the reconstructed 2.5-dimensional geoid feature pose is inputted to the discriminator of the generative adversarial network for authenticity error calculation, while the relative depth error calculation is performed using the relative depth information between the 3D DEM data model pose and each node corresponding to the image, the details are as follows:
   a) The 3D structural feature pose obtained from the previous reconstruction step is treated as a false sample, the existing acquired image pose data is treated as a true sample, input them into the discriminator of the generative adversarial network, which can make the 3D DEM data model pose conform to the data distribution of the existing acquired image pose, so as to obtain a more reasonable coupled mapping pose; the discriminator of the generative adversarial network consists of upper and lower the discriminator of the generative adversarial network consists of a two-layer fully-connected feature extraction module and a fully-connected true-false prediction module; first, the 3D DEM data model pose samples are inputted to the two-layer fully-connected feature extraction module for feature extraction, then the features extracted from the two-layer fully-connected feature extraction module are spliced to obtain the merged features, and the merged features are inputted to a fully-connected true-false prediction module for determining the true-false of the samples. The combined features are inputted to a full-connected true-false prediction module to determine the truth of the sample, and the decision value is outputted to calculate the 3D DEM data model pose truth error by using the loss function of generative adversarial network; among them, the upper fully connected feature extraction module and the lower fully connected feature extraction module have the same structure, both of which are composed of a feature extraction module in a deep prediction network and a scarcity network module composed of a scarcity block, and the full-connected true-false prediction module is composed of an algorithm group consisting of coupled computations.
   b) Using the relative depth information between the 3D DEM data model pose and the image pose data corresponding to each joint point for relative depth error calculation [10], with the human eye observation and marker traces of the image by the domain experts, we are able to obtain the relative depth information between each joint point of the geoid data in the image, using the form of matrix to store the relative depth relationship information between the joint points, specifically: from the image observation, the assuming that the i-th joint of the mountain is closer to the observation position than the j-th joint, the element value $r(i,j)$ in the i-th row and j-th column of the matrix is 1; the i-th joint is farther away from the observation position than the j-th joint, then $r(i,j)$ is -1; the difference between the i-th joint and the j-th joint is within the set range, then $r(i,j)$ is 0; where both i and j are values taken in the interval $[1,16]$, r is a matrix used to store the relative depth information between the joint points, and $r(i,j)$ is the value of the element in the i-th row and j-th column of the matrix, which is used to represent the relative depth relationship between the i-th joint point and the j-th joint point; the obtained matrix of relative depth information is used to calculate the relative depth error between each pair of joint points in the 3D-DEM data model pose in step C), specifically:

$$L_{i,j}\begin{cases}\max\left(0, r(i,j)\left(H_i - H_j\right)\right), & |r(i,j)| = 1\\ \left(H_i - H_j\right)^2, & r(i,j) = 0\end{cases} \tag{1}$$

Where, $L_{i,j}$ denotes the relative depth error value of the point pair formed by the i-th node and the j-th node in the mountain feature pose; $r(i,j)$ denotes the relative depth relationship between the i-th node and the j-th node, taking the values $\{1,-1,0\}$; $|r(i,j)|$ denotes the absolute value of r(i,j); Hi and Hj denote the depth values of the i-th node and the j-th node obtained in the depth prediction network respectively; finally, the relative depth error between each pair of nodes is mapped to calculate the sum of the relative depth errors formed by the two pairs of mountain hydrological nodes as follows:

$$L_{rank} = \sum_{(i,j)\in B} L_{i,j} \tag{2}$$

Where L_{rank} is the sum of the relative depth errors formed by the two pairs of mountain hydrological joints, (i,j) is the point pair formed by the i-th joint and the j-th joint in the mountain hydrology, and B is the set formed by the two pairs between the mountain hydrological joints; the sum of the relative depth errors formed by the two pairs of the calculated joints is expressed as the relative depth error of the characteristic pose of the mountain profile.

5) In step E, the authenticity error calculated by the discriminator of the generative adversarial network and

the relative depth error are added to obtain the total error of the reconstructed mountain contour feature pose in terms of authenticity and relative depth, and the error is fed back to the depth prediction network through the backward gradient descent propagation of the neural network to update the parameters in the depth prediction network, so that the neural network can learn the authenticity of the mountain contour feature pose and the relative depth information between the corresponding joints of the image, predict more accurate joint depth, and reconstruct more accurate 2D images and 3D spatial data mapping efficiency.

V. Evaluation

A. Experiment 1: Landscape spatial perception feature data extraction

1) *Hypothesis formulation:* In the process of 2D landscape spatial perception data extraction, the spatial perceptual features of landscape scenarios perceived by subjects with normal vision can be searched for by eye movement trajectories and thermal distribution images from eye movement experiments.
2) *Experimental procedure:*
   a) *Experimental apparatus:* The eye-movement experiment in this paper uses Tobii Pro X3-120 research tracker, which is a portable eye-tracking module in the ErgoLAB human-computer environment synchronization platform with a sampling rate of 120 Hz. The eye-tracking can be recovered in real time when the subjects achieve out-of-tracking range or blink, and the data loss rate is reduced to a minimum.
   b) *Subjects:* 15 university students with good corrected vision were selected, instructed to be familiar with the use of the eye-tracking device, described the experimental procedure for the subjects, and briefly described the experimental content as finding mountain peaks, hydrology and road network features in the landscape images.
   c) *Experimental materials*: 4 HD digital image slices of "Dwelling in the Fuchun Mountains (Wuyongshi - Shengshan Scroll)"; 4 HD digital image slices of remote sensing DEM images with similar angles to the peaks of the ancient painting selected under expert intervention. All images were adjusted to 1920×1080 size.
   d) *Experimental operation:* Subjects were subjected to the eye movement test in turn, and the results were recorded at the end of the experiment. The landscape image perceptual heat maps were obtained.
3) *Experimental results:* After visualizing the eye-movement experimental data, the eye-trajectory images and eye-motion thermal images were superimposed, and the morphological contour delineation of the scenario space of the landscape map was completed with the intervention of experts in the field of landscape painting. The experimental results are basically consistent with the original hypothesis.

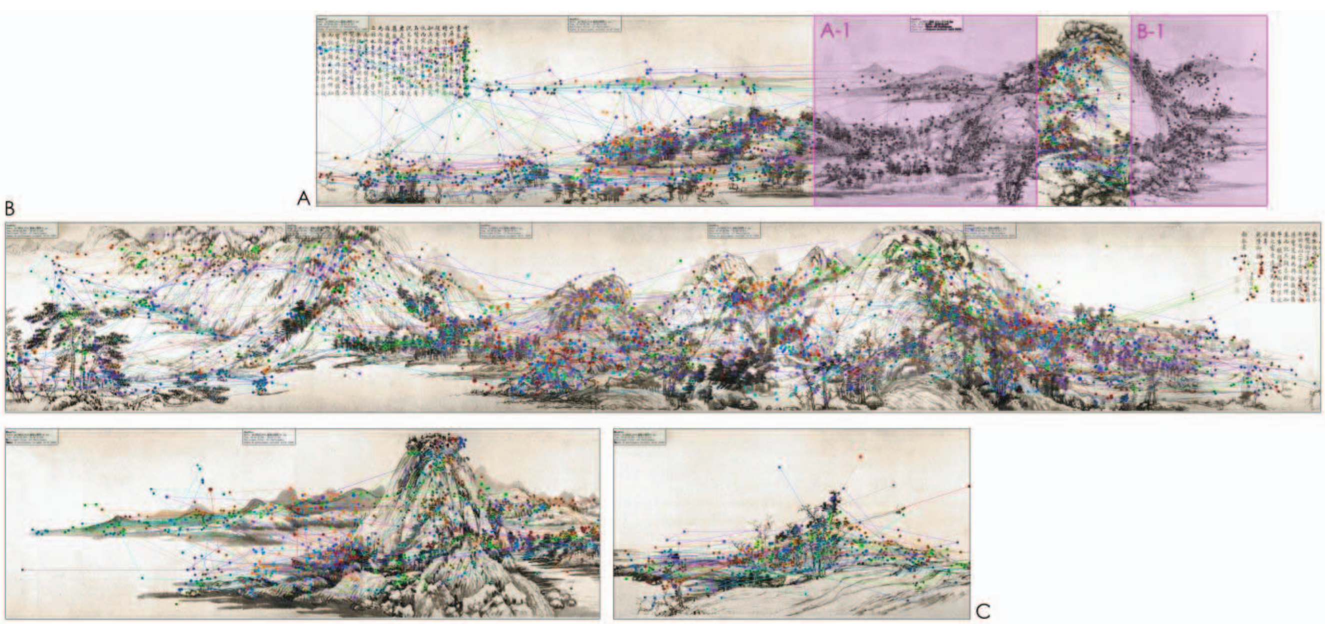

Figure 2. (spatial feature pose) Full-field image eye-tracking detection of areas containing digital restoratio

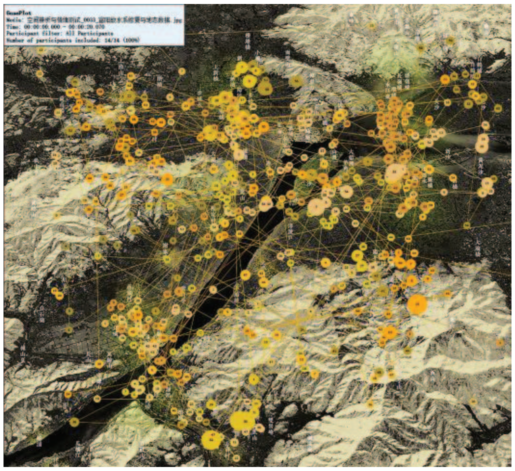

Figure 3. Perception of ancient water system restoration and geological features in Fuyang under knowledge annotation

B. Experiment 2: Overall matching of 2D landscape image and remote sensing DEM placement

1) *Hypothesis formulation:* The morphological contours of ancient painted mountain peaks extracted by semi-supervised method in 2-dimensional-2.5-dimensional view have high similarity with those of remotely sensed DEM mountain peaks. The predefined evaluation indexes are realized by the mapping method of 2D landscape image-3D spatial data based on adversarial relative depth constraint network.

2) *Evaluation metrics:*

a) *Hamming distance:* Hamming distance [11] is commonly used in the field of data transmission error control coding to indicate the number of different characters in the corresponding positions of two (same length) strings. The Hamming distance is the value of the number of strings that can be heterogeneous and counted as 1. The Hamming distance between two samples x,y is usually denoted by d(x,y).

$$d(x, y) = \sum x[i] \oplus y[i] \qquad (3)$$

where $i = 0,1, \ldots, n-1$, x, y are all n-bit encodings, and $\oplus$ denotes a heterogeneous or.

b) *Perceptual Hash Algorithm:* Perceptual Hash Algorithm (PHA) is a hashing algorithm to calculate the similarity of images [12]. The algorithm calculates the similarity by reducing the size, simplifying the color, calculating and reducing the DCT (Discrete Cosine Transform), calculating the average value, calculating the hash value, comparing the Hamming distance and other steps, independent of the height and width, brightness and color of the picture. The algorithm can better remove noise (such as inscriptions and small tears in ancient paintings) and judge the overall shape of the graphics from the perspective of subjective visual perception with high accuracy. When using the perceptual hashing algorithm to calculate the picture similarity, it is generally considered that if the Hamming distance $d(x, y) < 10$, the two are similar pictures.

3) *Experimental procedure:*

a) *Data processing:* Put the ancient painting peak morphological outline and remote sensing DEM corresponding peak morphological outline pictures under Photoshop, adjust them to the same size, remove the noise such as lifting and fine breakage, convert the pictures into grayscale images and save them as JPG format.

b) *Programming calculation:* Using Python 3.10 programming, we read in the ancient paintings and remote sensing DEM morphological images in 4 groups to realize the perceptual hash similarity calculation.

4) *Experimental results:* The experimental data are shown in the following table. Through calculation, it can be found that the Hamming distance of group 1, group 2 and group 4 does not exceed 10, which means that the similarity of these three groups of pictures is very high and can be basically confirmed as the same location, and the accuracy of the ancient painting is better. The reason for the lower similarity of group 3 is that the remote sensing DEM data has a farther perspective, but the average similarity of the four groups is higher, which is basically consistent with experimental hypothesis 2. Through the experiment, we can basically confirm that the group of peaks painted in "The Dwelling in Fuchun Mountain (Wuyongshi - Shengshan Scroll)" are the corresponding peaks in Fuyang area.

Table1. Perceptual hash algorithm similarity between ancient paintings and the overall morphology of DEM peaks

Comparison\ Group	Slice 1	Slice 2	Slice 3	Slice 4	Average
Perceptual Hash Similarity	0.8594	0.9063	0.7500	0.8750	0.8477
Hamming Distance	9	6	16	8	9.75

C. Experiment 3: Fine matching of mountain pose features

1) *Hypothesis formulation:* The features of 2.5-dimensional paleographs and the classified coordinates of peaks, mountains, hydrology, and road networks in DEM images have high matching similarity.

2) *Evaluation indicators:*

a) *Cosine similarity:* Cosine similarity measures the similarity between two vectors by measuring the cosine of the angle between them and is applicable to vector comparisons in any dimension [13]. The similarity between images can be measured by transforming the set of image features in the sample into vectors in a high-dimensional space and calculating the cosine of the angle of the inner product space of these vectors representing each image feature. The cosine value between two vectors can be found by the Euclidean dot product formula.

$$\mathbf{a} \cdot \mathbf{b} = \| \mathbf{a} \| \| \mathbf{b} \| \cos\theta \tag{4}$$

Given two attribute vectors A and B, the remaining chord similarity θ is given by the dot product and the vector length, as shown in Equation (2). where A_i,B_i represent each component of vectors A and B, respectively.

$$\cos(\theta) = \frac{A \cdot B}{\| A \| \| B \|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}} \tag{5}$$

b) *Jaccard similarity:* Jaccard similarity is the calculation of the proportion of the number of elements of the intersection of two sets A and B in the concatenation of A and B. It is also called the Jaccard coefficient of these two sets, and is denoted by the symbol J(A,B). The Jaccard similarity coefficient can measure the similarity of two sets of image sample coordinates [14].

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{6}$$

3) *Experimental procedure:*

a) *Data processing:* Combining the spatial feature perception results of the landscape generated by the eye-movement experiment in Experiment 1 and the similar perspective of remote sensing DEM confirmed in Experiment 2, the contour maps of mountain peaks, mountain bodies and hydrological coordinates were calibrated for the four groups of ancient paintings and remote sensing DEM images respectively with the intervention of domain experts. The contours were adjusted to a finer degree using Photoshop, the contour color was adjusted to black, and the contour maps of mountain coordinates and hydrological coordinates were exported and saved in JPG format, respectively.

b) *Programming calculation:*

- The four sets of mountain coordinate samples and hydrological coordinate samples are calculated separately for similarity.
- Using Python 3.10 programming, the functions are called to realize the calculation of cosine similarity and Jacquard similarity. Program to read in the images of DEM mountain contour and ancient painting mountain contour respectively, the size of the two images are similar. Obtain the corresponding Y coordinates on the two images under the same X coordinate and store them separately, then calculate the similarity of Y coordinates. It should be noted that in the calculation process, as the number of Y coordinates fetched from the two images may be different, a random discard is used to make the two Y coordinate sets the same. To reduce the calculation error, the calculation results are averaged over 1000 calculations. The function is called to implement the perceptual hash similarity calculation.

4) *Experimental results:*The results of Experiment 3 are shown in the following table. From the calculation results, it can be seen that all 34 indexes reached 0.75 or above, except for 2 indexes of group 3 which did not reach the standard due to the distant view angle. The group comparison of classification coordinates can significantly improve the matching effectiveness between landscape image space and remote sensing DEM space.

Table2. Coordinate matching results of ancient paintings and remote sensing DEM image classification

Group	Slice	Cosine Sim	Jaccard Sim	Perceptual Hash	Group Average
G1	Mountain1	0.8305	0.7576	0.9863	0.8729
	Water1	0.8917	0.7667	0.9492	
	Both1	0.8292	0.9289	0.9160	
G2	Mountain2	0.8190	0.8723	0.9971	0.9013
	Water2	0.9609	0.7930	0.9727	
	Both2	0.8627	0.8739	0.9600	
G3	Mountain3	0.7153	0.8512	0.9775	0.8597
	Water3	0.9280	0.7393	0.9287	
	Both3	0.7536	0.9468	0.8965	
G4	Mountain4	0.8283	0.9338	0.9961	0.9063
	Water4	0.9570	0.7513	0.9629	
	Both4	0.8492	0.9367	0.9414	

VI. Discussions

In the process of interpreting 2D landscape images, emphasis is placed on the transmission characteristics of spatial information (geographic data and coordinate system) rather than time (social environment and technological development stage). Among them, the diversity of data originating from ontological knowledge fusion production, reasoning, and computation and the temporal and spatial changes of its generation process are established under the formation of image ontological metadata and knowledge graph structure [15], and the introduction of anthropological

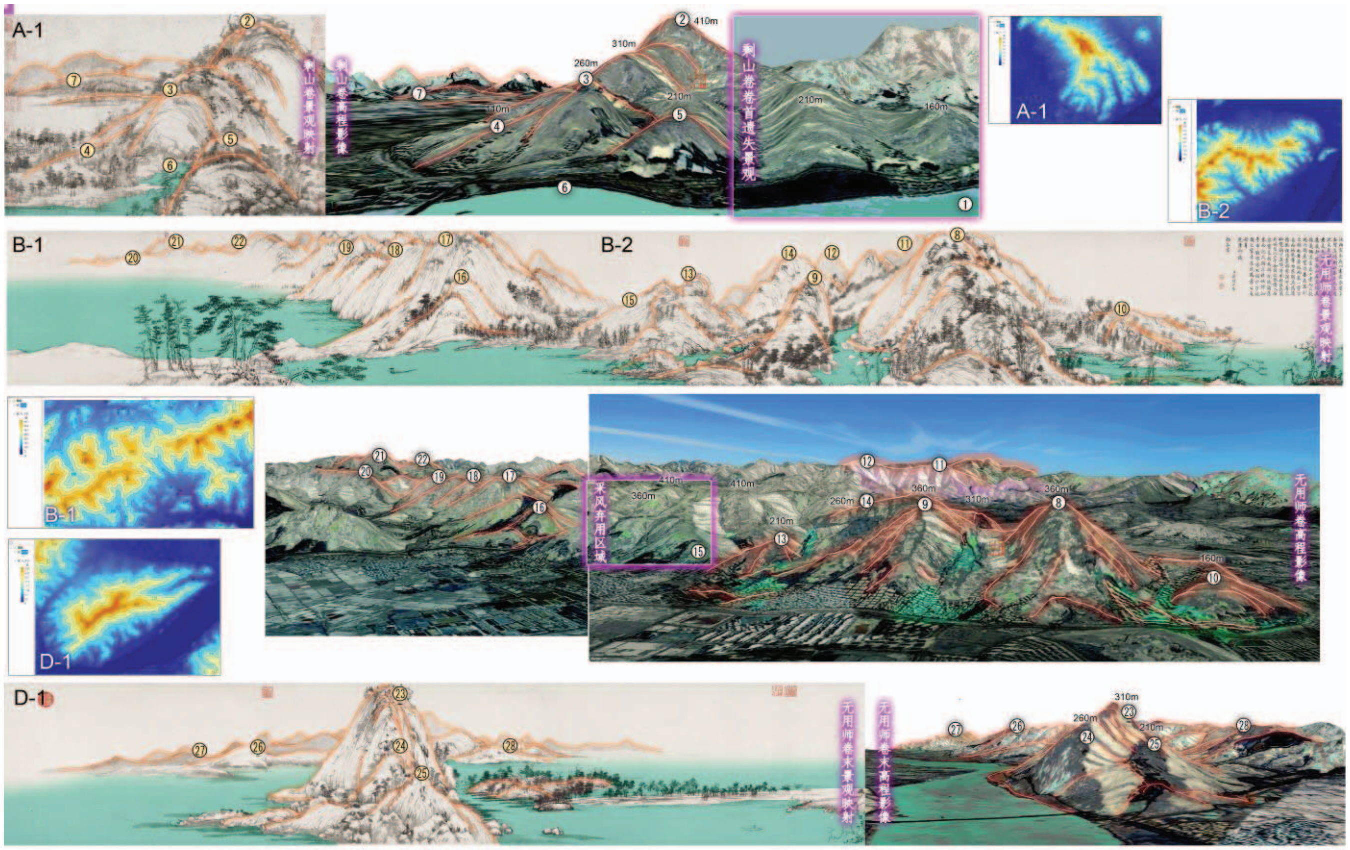

Figure 4. Elevation model mapping based on spatial tagging of graph scenarios

perspectives on the culture under investigation called "emic" and "etic The "emic" and "etic" perspectives, i.e., the "native perspective" and the "other perspective," are introduced. In the process of transformation with the "objective image" that exists in the image and the "self-position" that is driven by the researcher's goal, the behavioral analysis of the individual's "etic perspective" is derived with the objective "etic perspective" of the society at the time. The scientific analysis and description of the objective "other perspective" of the civilizational mode of existence provides a more "outsider's" perspective and "vicarious" humanized record of interpretation possibilities for landscape painting research. In other words, the evolution laws of ancient Chinese geographical observation and cartographic norms do not depend on the subjective consciousness of cartographers, observers and researchers, and can neither be created in isolation nor eliminated artificially. As long as the analytical conditions can be found, the rule, germinal effect can be found. Firstly, in the current process of re-cognition, because the enabling law is hidden in the image resources, only by exerting the subjective initiative can the phenomenon be penetrated and the law be grasped. Secondly, in the past generation stage, the cartographers also relied on their subjective initiative to change the conditions on which the law works according to the purpose of practice and the situation, so as to guide the law to work, resulting in different specific ways of image resources presentation. It is necessary to change the dimension of time and space stage and start from the specific scene to accurately realize its value mining. Thirdly, in the stage of fusion research, the technical support of natural cross-disciplinary enablement is not only direct comparison calculation, but the realization of causality, interpretation and prediction deduction.

The conclusion of the study of landscape painting and landscape scenario space should satisfy the dialectical principle of both opposition and unity in order to be falsifiable as a demarcation criterion; only when the construction based on "empirical foundation" is completed can the scientific objectivity and subjective certainty of the value of the conclusion be guaranteed.

VII. Conclusion

The validation through several cases shows that a mapping method based on adversarial relative depth constraint network for 2D images and 3D spatial data proposed in this paper can obtain more accurate spatial feature pose mapping evaluation, thus realizing mapping discovery. The problems of the lack of feature information data and the non-conformity between the results of the generative adversarial network method and the relative depth relationship between nodes in 3D spatial data in the traditional 2D landscape image geo-evidence process are solved.

The landscape image meta-learning framework can effectively improve the discovery of associated knowledge. In addition, this paper also draws the following important conclusion: to realize the study of landscape images with the intervention of five core elements, including political implication, academic support, philosophical thinking, general knowledge representation, and validity dissemination, it is necessary to develop a new understanding and new revelation with the evaluation indexes of cultural similarity, political and economic entity manifestation, independent scientific and technological prestige support, and geo-geographical environment.

ACKNOWLEDGMENT

We are grateful to Prof. Deren Li, academician of Chinese Academy of Sciences and Chinese Academy of Engineering, for guiding our research path, Prof. Renyi Liu for providing arithmetic support and data resources, and Prof. Jiansong Li, Prof. Kun Qin and Prof. Xudong Lai for their valuable suggestions. This paper was completed under the support of Major Project of the National Social Science Foundation of China (No.21&ZD332) "The Collection, Arrangement and Research of Ancient Chinese Agricultural Books".

REFERENCES

[1] Kuppala, Kavitha, Sandhya Banda, and Thirumala Rao Barige. "An overview of deep learning methods for image registration with focus on feature-based approaches." International Journal of Image and Data Fusion 11.2 (2020): 113-135.

[2] Murray, Brian, and Lokukaluge Prasad Perera. "An AIS-based deep learning framework for regional ship behavior prediction." Reliability Engineering & System Safety 215 (2021): 107819.

[3] Gu, G., et al. "Image hierarchical classification based on semantic tag generation and biased order structure." Journal of Software 31.02 (2020): 531-543. doi:10.13328/j.cnki.jos.005630.

[4] Jin, Lianchao et al. “Generative Adversarial Network Technologies and Applications in Computer Vision.” Computational intelligence and neuroscience vol. 2020 1459107. 1 Aug. 2020, doi:10.1155/2020/1459107

[5] Li, Chu-Wei, Zhang, Zhi-Long, and Yang, Wei-Ping. "Combining Boolean graph and grayscale sparsity for small target saliency detection." (2020).

[6] Xu, Junlong, et al. "Automatic extraction of lineaments based on wavelet edge detection and aided tracking by hillshade." Advances in Space Research 65.1 (2020): 506-517.

[7] Gao, Hongmin, et al. "A hyperspectral image classification method based on multi-discriminator generative adversarial networks." Sensors 19.15 (2019): 3269.

[8] Yang, Ke, et al. "Convolutional neural network multilayer feature union for remote sensing image retrieval." Surveying and Mapping Science 44.7 (2019): 9-15.

[9] Du, Jin, et al. "Low-power expression recognition based on improved deep residual networks." Computer Science 45.9 (2018): 303-307.

[10] Zhang Huan, et al. "Accuracy assessment of remote sensing elevation data in alpine mountains based on airborne LiDAR." Remote Sensing Technology and Applications 36.6 (2022): 1311-1320.

[11] Yuan, Li, et al. "Central similarity quantization for efficient image and video retrieval." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020.

[12] Song, Bo, et al. "Image alignment algorithm combining fast feature extraction and perceptual hashing." Computer Engineering and Applications 54.7 (2018): 206-212.

[13] Zhou, Lina, Yin Xiao, and Wen Chen. "Imaging through turbid media with vague concentrations based on cosine similarity and convolutional neural network." IEEE Photonics Journal 11.4 (2019): 1-15.

[14] Chen, H.-Chang, et al. "Candidate marker information-aware biased marker learning algorithm." Journal of Electronics and Information 41.10 (2019): 2516-2524.

[15] Xia Cuijuan. "Knowledge integration of cultural memory resources: from heterogeneous resource metadata application outline to integrated ontology design." Library and Intelligence Knowledge 38.1 (2021): 53-65.