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-C57518.2022.00079

Chinese Character Creation Algorithm: Evaluation of the Effectiveness of Topological Similarity of Remote Sensing Images Considering the Description of Sparse Geographic Structure

Ziyang Weng[1,*], Shuhao Wang[1], Weixin Yan[2] and Anna Li[3]
[1]School of Information Management, Wuhan University, Wuhan, China
[2]Shenzhen Museum of Contemporary Art and Urban Planning, Shenzhen, China
[3]College of Professional Studies, Northeastern University, Boston, MA, USA
weng_ziyang@whu.edu.cn, wangshuhao@whu.edu.cn, wilson_yan841@163.com, anna.li6811@yahoo.com
*corresponding author

***Abstract*—This study proposes a method for evaluating the effectiveness topological projection similarity between potential semantic analysis and stochastic remote sensing images based on sparse geographic structural descriptions (word building laws), which is applicable to general field of air and space humanities research. Inspired by unsupervised text clustering, the sparse geographic structure description word-making set is first established, the labeled text to be processed is converted into a bag-of-words model, and the TF-IDF algorithm is used to assign weights to it to obtain the weight vector set. Then the weight vector set is processed using the LSA algorithm to get the LSA index base, and then the topological mechanism of 2D image and 3D data model in the previous study is used for the weight vector set to realize the labeling of geo-feature pose, and the random projection algorithm is used to process it to get the RP index base. Finally, the corpus to be computed is processed by TF-IDF and then compared with the contents of the index library using LSA algorithm and RP algorithm respectively to obtain the text similarity, and the spatial location similarity effectiveness evaluation is realized through the spatial migration of labels.**

Keywords- aerospace humanities；word building; semantic analysis; geographic evidence-based analysis

I. Introduction

The method of taking into account the description of sparse geographic structures proposed in this paper can be understood as a kind of recommendation tags based on fuzzy concepts. The precise definition of geographic concepts is considered as one of the effective solutions to the problem of vocabulary in the scientific tagging system of social space and natural space, which is a high-quality tag that is continuously grown and compounded by the disciplinary system to the researchers according to certain computational rules. In the perspective of artificial intelligence research field, according to their sources can be divided into user labeled recommendation tags and machine extracted recommendation tags. These two different sources of recommendation tags both play an important role in guiding the selection of labeling perspectives and topic discovery in the user labeling process, which not only facilitates users to add tags, but also improves the semantic stability in the labeling system and improves the lexical problems such as semantic ambiguity of tags, over-personalization of tags and proliferation of meaningless tags.

However, most of the current research on recommendation tags involves the optimization of recommendation algorithms, and few scholars have focused on how recommendation tags change annotation habits and the tagging quality. It is important to study the impact of recommendation labels on users' labeling behaviors to better improve the system functionality. Especially in the environment of cross-disciplinary research, the role of recommendation labels on user labeling behavior is more complicated due to the diversity and disciplinary differences of recommendation label representation language and user labeling language.

II. Motivation

This study draws on latent semantic analysis (LSA), an unsupervised approach that migrates research logic originally applied to textual topic analysis to the field of image interpretation [1]. It is characterized by discovering topic-based semantic relationships between text and words through matrix decomposition. Latent semantic analysis is a non-probabilistic topic analysis method. The text set is represented as a word-text matrix and the matrix is decomposed by singular values to obtain the topic vector space and the representation of the text in the topic vector space. This study attempts to establish the basis of research on the description of sparse geographic structures (character creation laws), to extract, to label and to identify geospatial feature gestures, and to couple the ontological features of spatial joints with the word vector space. The basic idea is that given a geo-feature coordinate, a vector is used to represent each dimension of the "location semantic" vector in the coordinate corresponding to the word formation table, and its value is the frequency or weight of the geo-feature mapping in the spatial data.

By choosing the path of " character - phrase - word - character creation standard - social observation - natural perception", we propose for the first time that the observation of the structure of sparse geographical locations in the ecology. In the process of cultural evolution, precise descriptions of geomorphology, water bodies, and soil are constantly coalesced. By establishing an experimental method of directed word lists, the effect of recommended labels and their word-building norms on the spatial interpretation with

2693-9371/22/$31.00 ©2022 IEEE
DOI 10.1109/QRS-C57518.2022.00079

images in the image interpretation research environment is explored, which on the one hand helps to uncover the mechanism of the influence of recommended labels on research behavior, so as to feed back to the recommendation algorithm to optimize the recommendation results. On the other hand, it can decipher the scientific components of ancient people's observation of geographic environment, so as to improve the positive evaluation of the effectiveness of remote sensing image topological similarity calculation under the care of social geography and image psychology research [2].

III. Background

In the evolution of various civilizations, the laws of character creation and pattern recognition in today's academic research form a connection between two poles. Usually, pattern recognition refers to the process of processing and analyzing various forms of information (numerical, textual and logical relationships) that characterize things or phenomena in order to describe, identify, classify and explain them, and is an important part of information science and artificial intelligence. Pattern recognition is also often referred to as pattern classification. From the viewpoint of the nature of the processing problem and the method of solving it, pattern recognition is divided into two types of classification: supervised classification and unsupervised classification [3-6].

The character creation patterns covered in this paper can also be divided into two forms: abstract and concrete. The former, which is coalesced by abstract consciousness such as consciousness, thought, and speech, belongs to the category of concept recognition research and is another research branch of artificial intelligence. The character creation pattern recognition referred to in this paper is mainly to identify and classify the concrete patterns of images, scenes, pictures, photographs, words, symbols and other objects such as geography, geology, landforms, mountain shapes, peaks, valleys, water flows, water systems, water networks, and roads. Pattern recognition research is mainly focused on two aspects: one is the study of how living organisms (including human beings) perceive objects, which belongs to the scope of cognitive science; the second is the theory and method of how to achieve pattern recognition by computer under a given task. The computer is applied to identify and classify a set of events or processes [7]. The identified events or processes can be concrete objects such as words, sounds, images, etc., or abstract objects such as states and degrees. These objects are distinguished from information in digital form and are called pattern information. Pattern recognition is related to statistics, psychology, linguistics, computer science, biology, cybernetics, etc. It is cross-relative to the study of artificial intelligence and image processing.

IV. Methods

A. *Construct the word-making word list described by the scarce geographic structure*

Vectorize the tag to get the word tag vector set V_1, and use the TF-IDF algorithm [8] on it to get the tag weight vector set V_2, as follows:

1) Definition D_1 is a pre-detection of all Kangxi dictionary characters containing "mountain", "earth", "water" and "stone". Due to the unique Chinese standard of character formation and meaning, this type of dataset contains the complete observation and refinement of natural geopolitical features in various ancient social and geographical environments .
   $D1 =$
   $\{id_1,\ title_1,\ paragraph_1, image_1,\ url_1,\ tag_1\}$,
   where $id_1, title_1, paragraph_1, image_1, url_1, tag_1$ denote the number, title, paragraph, image link, document link and word tag, respectively.
2) By using split method on tag_1 we get $T_1 = \{wi_1, wi_2, \dots, wi_n\}$, wi_A is the Ath word tag set of the word-making word dataset, where the variable $A \in [1, n]$.
3) Obtain the dictionary $Dict_1$ by using the Dictionary method on T_1.
4) Saving the dictionary Dict1 locally.
5) The set of lexical label vectors $V_1 = \{vi_1, vi_2, \dots, vi_n\}$ is obtained by using the Doc2Bow method on T_1, vi_A is the Ath lexical label vector of the lexical label vector set V_1, where the variable $A \in [1, n]$.
6) The set of word label weight vectors $V_2 = \{vj_1, vj_2, \dots, vj_n\}$ is obtained by performing the TF-IDF method on V_1, vj_A is the Ath lexical label weight vector of the lexical label weight vector set V_2, where the variable $A \in [1, n]$.
7) Save label weight vector set V_2 locally.

B. *Using the LSA algorithm for V_2 to obtain the LSA model M_1 and the index library I_1.*

The specific steps are as follows:

1) Load the tag weight vector set V_3 from local, $V_3 = \{vk_1, vk_2, \dots, vk_n\}$, vk_B is the Bth weight vector of the set of lexical label weight vectors V_3, where $B \in [1, n]$.
2) Loading the dictionary $Dict_2$ locally.
3) Defining $id2word = Dict_2$ and the number of topics $num\_topics = 300$.
4) Obtain model M_1 by training V_3 using the LSA method, passing in the parameters $id2word$ and num_topics.
5) Processing of V_3 by model M_1 to obtain packaging corpus C_1.
6) Create an index base for C_1. to obtain index base I_1.
7) Save model M1 and index library I_1.

C. *Use the random projection algorithm for V_2 to obtain the RP model M_2 and index library I_2, and topologize with the remote sensing DEM data mode [9]l, as follows:*

1) Load the label weight vector set V_4 from local, $V_4 = \{vl_1, vl_2, \dots, vl_n\}$, vlC is the Cth weight vector of the set V4 of lexical label weight vectors, where $C \in [1, n]$.
2) 3.2 Define the number of topics $num\_topics = 500$.

3) 3.3 Obtain model M_2 by training V_4 using the RP method, passing in the parameter num_topics.
4) 3.4 Processing of V_4 by model M_2 to obtain packaging corpus C_2.
5) 3.5 Create an index base for C_2 to obtain index base I_2.
6) 3.6 Save the model M_2 and index library I_2.
7) 3.7 In the course of the preliminary research, we know that the location discrimination experiment process of the landscape map (2D image spatial interpretation) can realize the node location extraction and implant the location information, and by calling the index library I_2, the index label annotation of the image location can be realized and the topology with the remote sensing DEM data model.

D. *The corpus to be processed is processed using TF-IDF with LSA and RP to obtain the final recommendation set and evaluate the frequency or weight of the correlation components in it, and form the conclusion.*

1) Define D_2 as the test set for word-making entries, $D_2 = \{id_2, title_2, paragraph_2, image_2, url_2, tag_2\}$,where id_2, $title_2$, $paragraph_2$, $image_2$, url_2, tag_2 denote number, title, paragraph, image link, document link and word tag, respectively.
2) Taking $title_2$ as input, we obtain $T_2 = \{wj_1, wj_2, \ldots., wj_n\}$ by using split method on tag2, wj_D is the Dth word tag set of the encyclopedia word dataset, where the variable $D \in [1, n]$.
3) The set of lexical label vectors $V_5 = \{vm_1, vm_2, \ldots, vm_n\}$ is obtained by using the Doc2Bow method on T2. ,vmn}, vmE is the Eth lexical label vector of the set V_5 of lexical label vectors, where the variable $E \in [1, n]$.
4) The set of word label weight vectors $V_6 = \{vo_1, vo_2, \ldots, vo_n\}$ is obtained by performing the TF-IDF method on V_5. vo_F is the Fth lexical label weight vector of the lexical label weight vector set V6, where the variable $F \in [1, n]$.
5) Define the variable $k = 1$ as the loop variable to be used to traverse V_6.
6) Define the sets R_1 , R_2 and R_3 , $R_1 = \{simi_1, simi_2, \ldots, simi_n\}$, $R_2 = \{simj_1, simj_2, \ldots, simj_n\}$, R_3 is the empty set, $simi_G$ and $simj_G$ denote the Gth similarity set in R_1 and R_2 respectively, and $simi_G$ and $simj_G$ have the initial value of empty, where $G \in [1, n]$.
7) Importing LSA model M3 and random projection model M_4 , importing LSA index library I_3 and random projection index library I_4
8) If $k <= n$ then go to step 4.9), otherwise go to step 4.14).
9) vec_{1k} is obtained by wrapping vok with the LSA method, and vec_{1k} is obtained by wrapping vo_k with the random projection method vec_{2k}.
10) The similarity set with the elements in I_3 to vec_{1k} is obtained by retrieving the index library I_3 for vec_{1k}, using the cosine similarity [10] calculation and depositing it in simik the set of similarities with the elements in I_3 and vec_{1k} and store them in simik, and the set of similarities with the elements in I_4 and vec_{2k} by retrieving the index library I_4 for vec_{2k} and store them in $simj_k$ using the cosine similarity calculation.
11) Summing the corresponding elements of $simi_k$ and $simj_k$ and taking the average to obtain simlk.
12) Insertion of simlk into R_3.
13) $k = k + 1$, go to step 4.8); 4.14) Take the eight elements with the highest similarity in each set of R_3 and deposit them in the result set R_4, each element in $R_4$is the recommendation set, evaluate the correlation components of the frequency or weight in it, and form the conclusion.

V. EXPERIMENTAL EVALUATION

A. *Constructing data sets*

Chinese characters have been linked for thousands of years, from the earliest perceptions of natural landscapes, comprehending the laws of mountains and water bodies, and based on this, combining behavior and abstraction, forming a unique overall logic of expressing feelings, recording observations, abstract descriptions, and forming a complex combination of character creation laws. From the initial microscopic description of regional topography and living areas, it gradually expands to the logic of character formation for the entire mountain series, mountain ranges, and landscapes, forming complex semantic labels based on the character "mountain", collecting the display forms of up-and-down composition, left-right composition, and enveloping composition, and expanding to nearly 363 independent Chinese characters, 3630 features were obtained.

Among them, there are progressive structures depicting records of independent mountains, combined mountains (mountain systems), and extensive mountains (mountain ranges), and there also exist various types of mountain coordinate systems from the top to the bottom of mountain tops, ridges, valleys, slopes, mountains, slopes, water banks, rivers, etc., which complete records of various geopolitical features; and create adjective characters with realistic features, completing precise granularity descriptions and subjective feelings records. The data analysis of the character

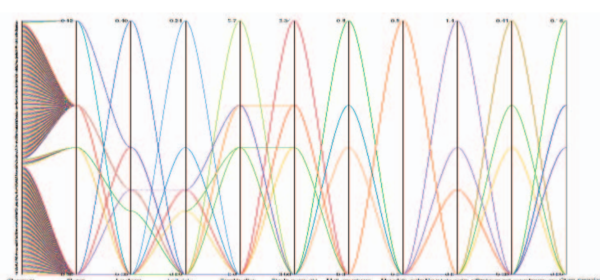

Figure 1. Assigning weights to 3630 feature labels using TF-IDF algorithm

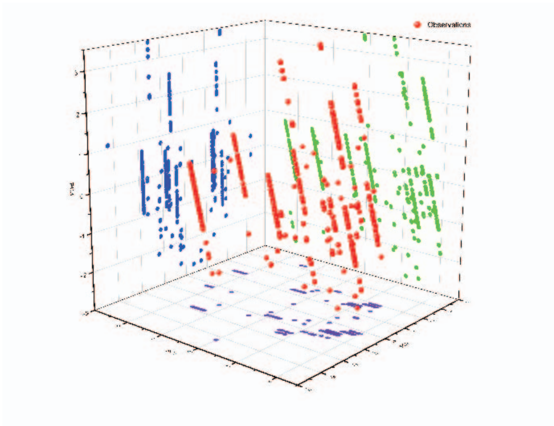

Figure 2. Reduced dimensional analysis of the set of tags in the word bags

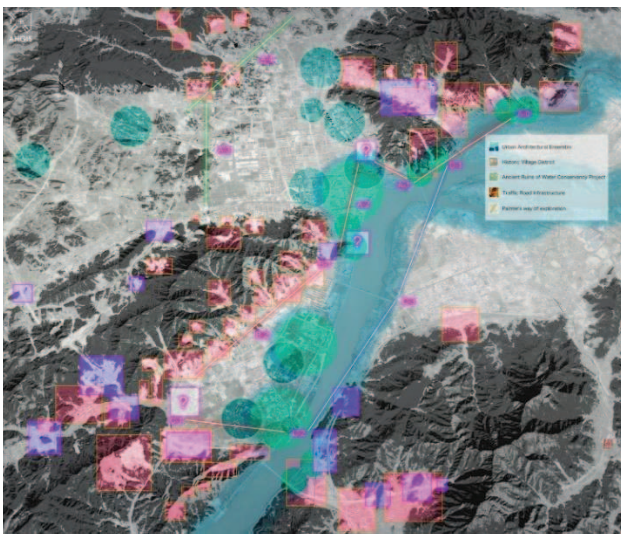

Figure 3. Exclusion of mountain destruction and building coverage of character creation law recognition

creation law can effectively realize the accurate image space annotation. What's more, because of the abstract meaning, implicit features and semantic complexity throughout the civilization evolution of nearly 8,000 years, this kind of data ontology has the possibility to transform linguistic location labels into calculation basis in the process of rich image interpretation.

B. *Observation experiment*

The compound geo-structural features reflected in the word-building laws interacted with the precise needs of the ancient people in the process of observing the landscape, and the need for more subtle, macroscopic, and specific records gave rise to more aggregated, more accurate, and more complex independent single words. Therefore, the use of the decomposition of the word creation law label, get rid of the multi-foot domain restrictions that rely solely on spatial location data, the spatial labeling of specific mountain from modern remote sensing, satellite viewpoint to the ancient exploration and observation perspective, reasonable and effective to achieve the experimental process of the observation trajectory through time and space through the labeling of word creation.

Take the above figure 1 as an example, all the light red areas belong to the top areas of mountains such as mountain tops, peaks and ranges, where valleys, roads, mountain streams and river valleys can be labeled by editing with associated semantic word-making and achieving point and area labeling; while some areas near rivers and located at the foot of mountains can be labeled by word-making through the exploration behavior of looking up at the mountains on the ground; green areas such as rivers and wetlands lacking mountain interference features The tagging set can not be used in this category.

Through tagging, the location of macro mountains, specific peaks, significant ridges, typical rivers, and specific streams and valleys can be clearly recorded, and the data

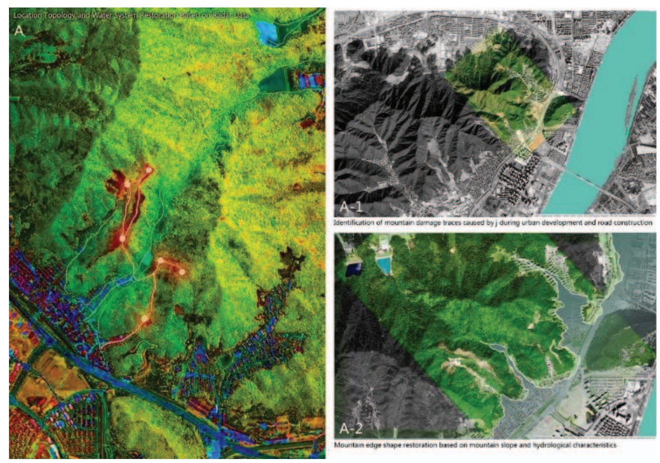

Figure 4. Semantic tagging based on the law of character creation: remote sensing image tagging of the region of Jiuzhu Peak, China

supply with calculation possibilities can be formed by converting the location with the coordinate system. This kind of data contains the observer's and recorder's geophysical feature cognition and the basis of scientific and technological concerns of the era, and has the possibility of extremely strong spatial humanistic calculation.

C. *Evaluation of topological efficiency between sparse geographical features in flat character images and remote sensing images*

1) *Hypothesis:* Based on the comparison of 2D image classification coordinates (mountain peak coordinates, hydrological coordinates and road network coordinates) with 3D remote sensing spatial classification coordinates, the location topological effectiveness of 2D Chinese character image scenario space and remote sensing positions can be improved.

2) *Experiment Setup*:

   a) *Data preparation:* partial slices of "Dwelling in the Fuchun Mountains (Useless Master - Leftover Mountain Scroll)", high-definition digital images of "Nine Pearl Peaks Cui Tu" with resolution above 600 dpi; 5M digital elevation model of

Fuyang area, using 3dMAX to obtain image maps with similar perspective to the 2D images in the given model of Fuyang area.

b) *Contour extraction:* use ArcGIS to draw contours and merge the projection maps of each viewpoint; with the intervention of experts in the field of landscape imagery, use ArcGIS to outline the contours of the mountains, hydrology and road network in the ancient painting, outline the contours of the mountains, hydrology and road network in the DEM, determine the peak, ridge and valley feature points in the ancient painting and DEM, and save them as shp files.

c) *Format conversion:* Use ENVI to convert shp to jpg format image.

3) *Experimental procedure:* Peak and hydrological contours were extracted for four groups of samples respectively. The two types of contour coordinates are used as feature regions and the similarity is calculated separately. Python 3.10 programming is used to call the function to realize the calculation of cosine similarity and Jaccard similarity (as above, the average of 1000 calculations is taken). The function is called to realize the calculation of perceptual hash similarity and histogram similarity.

4) *Results:* The experimental results show that the topology of remote sensing images and word-making laws through the classification coordinates of scarce geographical features has better results.

Table 1. Matching results of 2D image and 3D spatial classification coordinate mapping

Coordinates	Cosine Similarity	Jaccard Similarity	Perceptive Hash Similarity	RGB Histogram Similarity
Fuyang Mountain	0.9340	0.8828	0.9727	0.6635
Fuyang Points	0.8814	0.1911	0.9973	0.8161
Fuyang Road	0.9271	0.5215	0.9473	0.6055
Fuyang River	0.8071	0.9695	0.7832	0.6242
Jiuzhu Mountain	0.7876	0.9457	0.9424	0.6952
Jiuzhu Points	0.7396	0.1756	0.9980	0.7524
Jiuzhu Road	0.8766	0.9806	0.9219	0.7048
Jiuzhu River	0.7996	0.9924	0.7598	0.6910

VI. DISCUSSION

Latent semantic analysis in images can be used for image knowledge acquisition and verification of the authenticity and validity of spatial information in them. By using statistical methods to analyze a large number of text sets, the potential semantic structure between words can be extracted, and this potential semantic structure can be used to represent words and texts to achieve the purpose of eliminating the correlation between words and simplifying the text vector to achieve dimensionality reduction. And using multiple dimensionality reduction methods to reduce the vector space can effectively improve the accuracy and reliability of the results.

Usually, random projection is done with a simple and effective dimensional approximate subtraction method, which is different from some epistasis-based feature extraction methods. Feature extraction is completely independent of the original sample data set and does not cause significant distortions in the data. The dimensionality reduction still retains the important feature information contained in the original high-dimensional data, and does not require matrix decomposition to derive the transformation matrix, which can greatly improve the ability to process data in real time. In this paper, the underlying logic of random projection is migrated to the DEM model in the GIS environment, and the uniqueness of specific continuous geographic landscape patterns, linear observation survey trajectory features, and spatial-temporal autocorrelation assembled from typical geographic feature pose data can realize constraints to rapidly improve the topological effectiveness of remote sensing images of target images and change the limitations of traditional methods that rely on subjective speculation of domain experts.

VII. RELATED WORK

In this study, we further explore the impact of interference caused during algorithm preparation, experimentation, and evaluation after different domain studies are transformed into a collaborative recommendation mechanism. The experimental results show that in the process of providing interdisciplinary word lists by domain experts for labeling and discrimination, it can be found that the experimental model in the traditional GIS domain is less likely to directly understand the labels and achieve computation. Along with the increase in the number of added labels, the deficiency of cross-domain unfamiliarity with each other in the case of continuous iterations of the algorithm causes a greater cost in the experimental process. Eventually, after the preferential selection, the coupling of multiple classes of algorithms can effectively solve the research problem, leading to an increase in the effectiveness of understanding the labeling, but there are still limitations of slow decoding speed and bias in the decoding results.

In the matching process, multi-category geographic elements and character creation rules achieve effective mutual proof. It can be found that the recommended tags require more familiarity with their representational languages in the early stage, which shows certain advantages in helping researchers' semantic understanding: reducing the difficulty of labeling, improving the speed and experience of labeling.

The above indicates that reasonable research path planning, scientific research clues and interdisciplinary professional guidance are the differences from the previous single-disciplinary language annotation environment. Further analysis revealed that this finding can be explained more specifically by drawing on the Anchoring Effect [11-13] in the field of psychology. The Anchoring Effect suggests that when a researcher makes a decision, important information that is already available to the researcher is unconsciously perceived by the researcher as a criterion that acts as an "anchor" that governs the researcher's decision. The anchor effect suggests that as long as the anchor is noticed, it will have some influence on the researcher's decision, regardless of whether the data has actual reference utility.

Previous studies have found that when researchers produce more knowledge consistent with the "anchor", or even show stronger anchoring effects, they feel more negative emotions and affect the output of valid conclusions, which leads to resistance at the beginning of the experiment due to slow decoding speed and sparse distribution of annotations. In this study, the cross-domain study played the role of "anchoring", making the researcher aware of the "anchor" in the labeling process, and analyzing its rules and avoiding its effects through the "anchoring" learning behavior. The cross-disciplinary study acted as an "anchor", making the researcher aware of the "anchor" during the labeling process, and analyzing its rules and avoiding its effects through the "anchoring" learning behavior.

VIII. Conclusion

This experiment fully realizes the reuse of natural language processing logic in cross-research fields, especially involving potential semantic analysis in the scenario space based on the image interpretation process, and effectively introduces the text similarity calculation method and calculation results of random projection to multidimensional, multi-source flow, multi-temporal and multiphase research. In traditional text recommendation algorithms, researchers choose to perform similarity computation by ordinary vector space models and so on [14]. By discovering semantic associations between texts, we combine latent semantic analysis and random projection to provide a reliable text similarity calculation method for relevant systems.

1) For the first time, it is proposed to realize the identification and labeling of image data mimesis through the construction of map archives and the development of assisted retrieval functions, to form the mapping and differentiated labeling of the principal components and location structure of image scenarios, and to standardize the paradigm of spatio-temporal composite image perception and computation in the field of spatial humanities.
2) Constructing a meta-learning framework for landscape images, defining the division between natural description semantics and structural description semantics of image scenarios, and providing new semantic labels for the organization and invocation of image resources.
3) We constructed the Chinese categorical geographic location and traffic road network dataset to achieve effective topology and complementation with contemporary mapping and remote sensing geographic location data, gradually improve the scarce location information in the calculation of map scenarios, greatly enrich the spatial semantic narrative granularity of landscape scenarios, and initially solve the problem of misinterpretation caused by scarce information in this kind of research.

Empirical research is an important path to read, analyze, and evaluate the mapping effectiveness of geographic information in landscape images. The empirical evaluation of several research cases verifies the feasibility of the path to reveal the value of landscape images based on situational spatial mapping scenario extrapolation (the above three aspects of work) proposed in this paper. It provides a multi-dimensional and multi-level perspective on how to initiate value research on the political implication, academic support, philosophical thinking, general knowledge representation, and effectiveness communication of landscape images.

Acknowledgment

Thanks to Yan Weixin, the Director, and the Art Technology Lab of the Shenzhen Museum of Contemporary Art and Urban Planning for helping us with art consultation. 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] Dumais, Susan T. "Latent semantic analysis." Annu. Rev. Inf. Sci. Technol. 38.1 (2004): 188-230.

[2] Sun, Xian, et al. "Research progress on few-shot learning for remote sensing image interpretation." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021): 2387-2402.

[3] Parrilla-Gutierrez, Juan Manuel, et al. "A programmable chemical computer with memory and pattern recognition." Nature communications 11.1 (2020): 1-8.

[4] Li, Dongcheng, W. Eric Wong, et al. “Detection and Mitigation of Label-Flipping Attacks in Federated Learning Systems with KPCA and K-Means.” 2021 8th International Conference on Dependable Systems and Their Applications (DSA) (2021): 551–559.

[5] Mahamunkar, Geetanjali S., Arvind W. Kiwelekar, et al. “Deep learning model for black spot classification.” International Journal of Performability Engineering 18.3 (2022): 222–230.

[6] Li, Dongcheng, W. Eric Wong, et al. “Automatic test case generation using many-objective search and principal component analysis.” IEEE Access 10 (2022): 85518–85529.

[7] Zhang, Xuelong. "Research on data mining algorithm based on pattern recognition." International Journal of Pattern Recognition and Artificial Intelligence 34.06 (2020): 2059015.

[8] Qaiser, Shahzad, and Ramsha Ali. "Text mining: use of TF-IDF to examine the relevance of words to documents." International Journal of Computer Applications 181.1 (2018): 25-29.

[9] Polidori, Laurent, and Mhamad El Hage. "Digital elevation model quality assessment methods: A critical review." Remote sensing 12.21 (2020): 3522.

[10] Gunawan, Dani, C. A. Sembiring, and Mohammad Andri Budiman. "The implementation of cosine similarity to calculate text relevance between two documents." Journal of physics: conference series. Vol. 978. No. 1. IOP Publishing, 2018.

[11] Gonnerman, Laura M. "A linguistic analysis of word morphology." Morphological Processing and Literacy Development. Routledge, 2018. 3-15.

[12] Huang, Jiacheng, and Yong Peng. "Simulation of Life Game Based on Cellular Automata." Journal of Computer and Communications 9.1 (2021): 44-58.

[13] Röseler, Lukas. Anchoring Effects: Resolving the Contradictions of Personality Moderator Research. Diss. Otto-Friedrich-Universität Bamberg, Fakultät Humanwissenschaften, 2021.

[14] Chu, Shu-Chuan, et al. "Optimizing ontology alignment in vector space." Journal of Internet Technology 21.1 (2020): 15-22.