Text Region Adaptive Multi-temporal Stone Inscriptions Image Registration Framework

Yuan Cheng [1,2], Hanling Li [1], Jizhong Huang [1,2], Hongbin Yan [3]

[1] Institute for the Conservation of Cultural Heritage, School of Cultural Heritage and Information Management, Shanghai University, Shanghai, China, 200444 – lihanling@shu.edu.cn, chengyuan@shu.edu.cn, hjizhong@shu.edu.cn

[2] Key Laboratory of Silicate Cultural Relics Conservation (Shanghai University), Ministry of Education – chengyuan@shu.edu.cn, hjizhong@shu.edu.cn

[3]Yungang Research Institute, Datong, Shanxi, China, 037034 – 435135458@qq.com

Keywords： Image registration, Stone inscriptions Image, Match Points Enhancement.

Abstract

Multi-temporal comparative analysis of stone inscription images helps understand weathering processes and enables more targeted conservation measures, with accurate image registration serving as the prerequisite for such analysis. However, existing registration methods perform poorly on these images due to sparse and unevenly distributed feature points in text regions caused by natural weathering processes such as weathering and exfoliation. This study proposes a semantic-aware registration framework for multi-temporal stone inscription images, implementing a four-stage processing workflow: global initial registration using Scale-Invariant Feature Transform (SIFT) algorithm with geometric filtering; region partitioning through grid division and text region identification; sub-block enhancement applying SIFT-Accelerated KAZE (AKAZE) dual-feature fusion strategy for regions with insufficient matching points; and global fine registration integrating multi-source matching points with Random Sample Consensus (RANSAC) optimization. Experiments on stone inscriptions with varying degrees of weathering validate the effectiveness of this method: for well-preserved samples, the Mean Euclidean Error (MEE) decreased by 94.8% and structural similarity improved by 813.7%; for severely weathered samples, matching points in text regions increased 18-fold with an inlier ratio reaching 97.35%. This research represents the integration of text region semantic recognition with adaptive feature enhancement, providing reliable technical support for multi-temporal stone inscription analysis and cultural heritage conservation.

1. Introduction

Stone inscriptions, such as the Rosetta Stone, the Stone Drum Inscriptions and the Qin Dynasty Taishan Stone Inscriptions, preserve historical traditions and cultural memories across civilizations, representing one of the most important types of cultural heritage artifacts. However, numerous stone inscriptions have been exposed to environmental factors such as wind erosion, sand abrasion, water infiltration, and temperature fluctuations. Long-term exposure has led to surface exfoliation, gradually degrading inscribed features and ultimately resulting in the irreversible loss of carved text data.

Recent research has primarily focused on analyzing and evaluating the weathering degree and physical properties of stone heritage sites, employing non-destructive testing (NDT) methods such as impedance guns (Navacerrada et al., 2022), ultrasonic testing, and Schmidt hammer tests (Salvatici et al., 2020). However, there remains a gap in effective quantitative methods for evaluating the degradation of inscriptions on stone cultural heritage surfaces. Digital imaging technology offers a practical approach to address this challenge. While several studies utilize high-resolution imagery (Nasri & Huang, 2020) to document and analyze the surface conditions of stone inscriptions, it remains difficult to capture dynamic weathering processes from static images alone. Image registration, a technique for aligning multi-temporal images, provides a foundation for quantifying surface degradation indicators such as contrast variations and edge sharpness reduction. This approach can effectively integrate historical images with modern multi-temporal data, enabling the documentation of weathering patterns over time and providing valuable baseline data for systematic monitoring of stone inscriptions.

Image registration is a process of establishing spatial correspondence among images acquired from different sensors or at varying viewpoints and scales. This technique is applied in medical imaging (Chen et al., 2022; Leng et al., 2021), remote sensing (Hou et al., 2021; Ye et al., 2022) and various other fields. Image registration methods generally fall into two categories: region-based and feature-based methods (Gandhi and Panchal, 2014). Due to the reference points provided by character edges and stroke intersections, feature-based registration is particularly suitable for stone inscription analysis. In this method, factors that affect the registration accuracy include image quality (Ketcha et al., 2017), distribution of registration points (Bian et al., 2021), and differences in descriptors (Anzid et al., 2022). Research has shown (Tan et al., 2012) that an adequate number of well-distributed matching points can improve registration accuracy. However, for stone inscriptions, varying degrees of weathering on stones often result in texture ambiguity and local information loss, leading to fewer feature points in text regions. Therefore, obtaining sufficient quantity and even distribution of these points remains challenging.

This study proposes an adaptive registration framework for stone inscription images that prioritizes feature point distribution within core text regions while implementing an average-blocking strategy across the entire image to enhance spatial feature extraction. The framework integrates semantic identification of text regions with adaptive feature enhancement techniques, implementing a SIFT-AKAZE dual-feature fusion strategy optimized for stone inscription characteristics, which addresses the issue of insufficient matching points in the text region.

2. Method

The objective of this study is to develop a framework for adaptive region-enhanced registration based on stone inscriptions images. As shown in Figure 1, the workflow

primarily involves a four-stage treatment process: (1) Global initial registration: the SIFT algorithm extracts global feature points, followed by Fast Library for Approximate Nearest Neighbors (FLANN) matching and geometric distance filtering to retain high-confidence correspondences; (2) Regional segmentation and evaluation: the image is divided into an $N \times M$ grid, text areas are identified using text detection, and the distribution of matching points is assessed for each sub-block; (3) Match points enhancement: blocks with insufficient matches undergo image enhancement, then combining SIFT and AKAZE as dualfeature fusion strategy to increase the number of matches; (4) Global fine registration: all validated feature points are integrated, and the RANSAC algorithm is applied to optimize the transformation matrix, achieving precise alignment. This study employs a multi-dimensional validation framework to quantitatively evaluate registration performance and conduct comparative analysis with conventional methods.

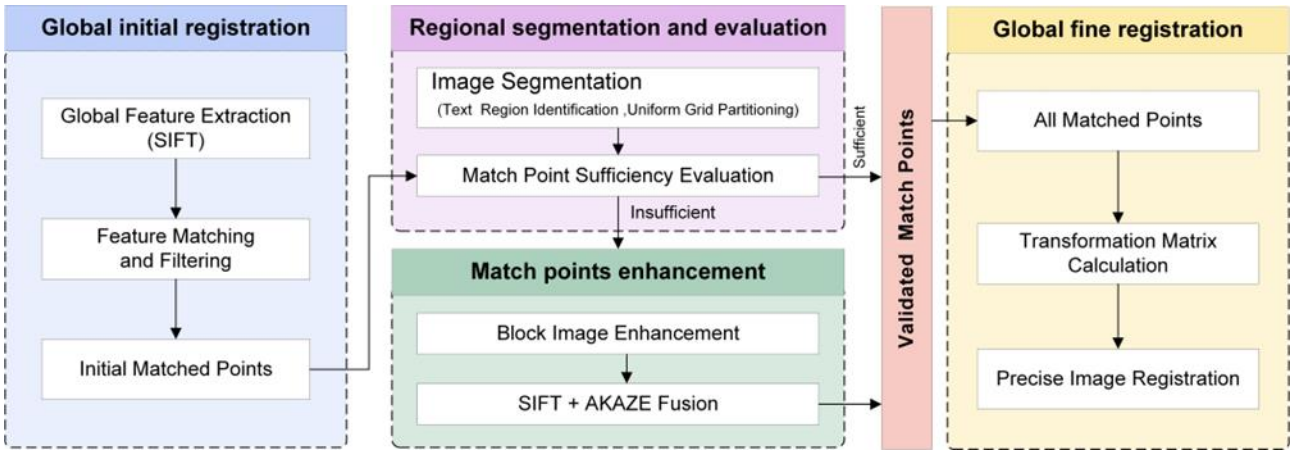


Figure 1. Method implementation process

2.1 Global Initial Registration

This method begins with the extraction of global feature points using the SIFT algorithm, followed by a two-step filtering mechanism to secure high-quality matches. The SIFT algorithm is particularly well-suited for feature extraction from weathered stone inscription images due to its inherent robustness to rotation, scale variations, and illumination changes. The matching process incorporates the following strategies:

Initially, Lowe's ratio test is applied for preliminary match screening. A threshold of 0.70, which is more stringent than the standard 0.75, is employed to enhance the accuracy of initial matches. This test effectively discards ambiguous matches exhibiting insufficient distinctiveness in the descriptor space by evaluating the ratio of distances between a feature point's nearest and second-nearest neighbors. Subsequently, to further eliminate mismatches, spatial consistency filtering based on geometric distance is performed, with the filtering threshold set to 1/20th of the image dimensions. This filtering mechanism, grounded in the principle of spatial projection consistency, utilizes the RANSAC algorithm to estimate a geometric transformation model. Outlier matches inconsistent with this global transformation are then identified and removed based on Euclidean distance evaluation.

This dual-validation approach, operating on both feature similarity and spatial geometric consistency, ensures the reliability of the initial match set, thereby providing a solid foundation for subsequent region partitioning and fine registration. For images of stone inscriptions, which often present with uneven weathering and localized damage, this strategy yields robust initial registration results.

2.2 Regional Segmentation and Evaluation

Surface weathering on stone inscription can lead to an uneven distribution of matching point pairs in certain regions, thereby compromising the global consistency of the homography transformation. To ensure a sufficient density of feature points across all local areas of an image, this study employs two complementary partitioning strategies: uniform grid partitioning and text region identification.

2.2.1 Uniform Grid Partitioning: This method initially divides the image into nine blocks, arranged in a 3×3 grid. The precise coordinates of each block are calculated based on the overall image dimensions. Such uniform partitioning ensures comprehensive coverage of the entire image area, facilitating an assessment of the global feature point distribution and the identification of regions with sparse features.

2.2.2 Text Region Identification: The acquisition of region blocks is approached through two methods, with strategies adapted to the varying degrees of weathering on the stele images:

For stone inscription images exhibiting minor weathering and clearly delineated text, the YOLOv11 model is utilized for automated text region localization. When applied to text detection tasks, especially after being trained on a domain-specific dataset, YOLOv11 demonstrates good stability in localizing text regions and reducing errors in severely weathered rock carving images with complex backgrounds, compared to traditional edge- or texture-based methods (e.g., Sobel operator) and other mainstream deep learning-based text detection algorithms (e.g., DBNet, PaddleOCR). A specialized dataset of rock carving text images was constructed, and the model was trained on this dataset to develop a deep learning model capable of automatically detecting text on such artifacts. Experimental results indicate that this model achieves higher detection accuracy on the rock carving text dataset than other leading text detection methods, thus providing a more reliable basis for selecting Regions of Interest (ROIs) for subsequent feature point extraction and image registration.

For stone inscriptions images characterized by severe weathering or indistinct text, where the accuracy of automated text detection is diminished, manual annotation is employed as a supplementary measure. To facilitate this, an image visualization and interactive annotation tool was developed using Python. This tool allows users to delineate rectangular bounding boxes on the image via mouse input; the coordinates of these boxes are then automatically exported for subsequent feature point extraction and registration. This semi-automated approach provides an effective alternative when fully automated detection methods are inadequate.

2.3 Match points enhancement

For each block, the method employs an adaptive matching point enhancement mechanism, augmenting feature points in sparse regions through a hierarchical feature extraction strategy. Enhancement is triggered by comparing the number of matching points within a block to a predefined threshold. This threshold is established through a heuristic analysis of feature distribution across various blocks. Notably, for small-sized images or those where text regions constitute a significant portion, the global partitioning check may be bypassed, and enhancement can be applied directly to text region blocks to improve algorithmic efficiency.

The match points enhancement process proceeds as follows: Initially, the number of matching points within the target region is assessed. If this count falls below the threshold, a three-stage enhancement strategy is executed:

(1) Image Enhancement: Preprocessing techniques, including image denoising, histogram equalization, and sharpening, are applied to enhance texture contrast and detail visibility within the block.

(2) Parameter-Adaptive Iteration: Key parameters of the SIFT feature detector, such as contrastThreshold, edgeThreshold, nOctaveLayers, and sigma, are adjusted. Concurrently, a matching parameter iteration mechanism is employed, incrementally increasing the Lowe's ratio threshold by 0.02 in each iteration. This progressively relaxes the matching criteria while ensuring matching reliability, up to a predefined safe upper limit of 0.8.

(3) Complementary Feature Extraction: If an adequate number of matching points is still not obtained after SIFT parameters reach their upper limits, the system automatically transitions to the AKAZE feature descriptor for complementary extraction. The AKAZE algorithm constructs its scale space using nonlinear diffusion filtering, which, compared to the Gaussian blurring used in SIFT, better preserves edge features and fine details in the image. For stone inscriptions images, particularly in regions where weathering has reduced contrast, AKAZE's nonlinear filtering mechanism enables the extraction of a greater number of salient feature points. Empirical evidence indicates that in severely weathered text regions with blurred edges, AKAZE yields a significantly higher density of feature points than SIFT, effectively supplementing areas with sparse features.

Should the required number of feature points still not be met after the full enhancement pipeline, the system flags the block as anomalous and logs the entire parameter adjustment trajectory. This log serves as a basis for potential manual intervention. Finally, all valid matching point pairs, aggregated from various blocks and yielded by different feature extraction algorithms, are consolidated. The RANSAC algorithm is then employed to compute the optimal homography matrix, facilitating high-precision geometric transformation and image alignment.

2.4 Registration performance evaluation

The assessment of registration quality employs a multi-metric fusion strategy, integrating three complementary metrics: Mean Euclidean Error (MEE), Structural Similarity Index (SSIM), and inlier ratio, to provide a comprehensive evaluation of registration accuracy.

MEE is calculated according to Equation (1). It serves to evaluate the geometric alignment accuracy between corresponding feature points after transformation. In this equation, (x_i, y_i) represents a feature point in the reference image, $(\hat{x}_i, \hat{y}_i)$ denotes the coordinates of the corresponding feature point from the sensed image after undergoing the estimated transformation, and N is the total number of matched point pairs. In image registration, a lower MEE value signifies a higher quality of registration.

$$MEE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \tag{1}$$

SSIM is computed as shown in Equation (2). It primarily serves to quantify the preservation of structural information during the image registration process. The SSIM value typically ranges from 0 to 1, where a value closer to 1 indicates greater similarity between the images.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{2}$$

In this equation, x and y represent corresponding image patches from the reference and target images, respectively; μ_x and μ_y are the mean intensities of patches x and y; σ_x^2 and σ_y^2 are their respective variances; σ_{xy} is their covariance; and C_1 and C_2 are stabilization constants included to prevent division by zero.

The inlier ratio refers to the proportion of matched point pairs within the entire set that are considered inliers (i.e., correctly matched points). This metric is used to assess the reliability of the matching process and is defined by Equation (3). Here, N_{inlier} is the number of matched point pairs identified as inliers, and N_{total} is the total number of initially matched point pairs. The inlier ratio ranges from 0 to 1, with a higher value indicating superior matching quality.

$$IR = \frac{N_{inlier}}{N_{total}} \tag{3}$$

3. Results and Discussions

3.1 Experimental setup

Experiments were conducted on a Windows 11 64-bit system equipped with an NVIDIA T600 Laptop GPU. The registration performance of the proposed algorithm was compared against traditional registration methods across diverse types of stone inscriptions texts.

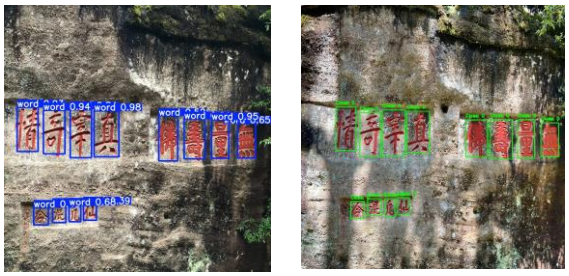
The training data for the automated text detection model was primarily sourced from photographs of stone inscription texts

collected from social media, encompassing major types of Chinese stone inscription texts. Text regions in each image were manually annotated, with the annotation format adhering to YOLO framework requirements. Each annotation comprised a class label and the corresponding bounding box coordinates. A total of 60 photographs were selected and divided into training and validation sets at a 4:1 ratio. The learning rate was set to 0.0001, the batch size is 8, the epoch is 1000. After training, the model achieved a mean Average Precision (mAP) of 80% on the validation set.

This study selected stone inscription images from Wuyi Mountain and Tianlong Mountain as experimental data to validate the applicability and stability of the proposed method on different types of stone inscription surfaces.Stone inscriptions from Wuyi Mountain, located in a Danxia landform region, were represented by images from 2018 and 2024 which showed no significant weathering changes. This dataset, sourced from social media, was primarily used to evaluate the method's matching performance on stone inscription texts exhibiting minimal weathering. Conversely, the Tianlong Mountain stone inscriptions are primarily situated on sandstone, a material prone to weathering. An exemplary case of significant weathering between 2012 and 2022 is the four-character inscription '揽胜留题' (Lǎn Shèng Liú Tí – "Enjoying the Scenery and Leaving an Inscription"), whose legibility notably deteriorated. This dataset was acquired through fieldwork by our research group and was principally employed to test the method's matching performance on stone inscriptions that have undergone substantial weathering changes. The experimental samples thus encompassed stone inscription images from diverse geological settings (Danxia landforms and sandstone formations) and exhibiting varying degrees of weathering (no significant change versus significant change). This provided a range of test scenarios and robust data support for investigating the matching performance for stone inscription texts, particularly in feature-sparse regions.

3.2 Results of Wuyishan stone inscriptions

The first set of experiments utilized stone inscription images from Wuyi Mountain as test data. The carvings in this region are relatively well-preserved and predominantly located in the central area of the images. Given the relative clarity of the target stone inscription, a pre-trained deep learning model was employed for automated text region localization firstly. Instances of overlapping detection boxes were observed in the detection results for the 2024 image. These were subsequently refined using the Non-Maximum Suppression (NMS) algorithm to filter and merge boxes, thereby ensuring the accuracy and uniqueness of the detected regions. The final text region localization results are presented in Figure 2.



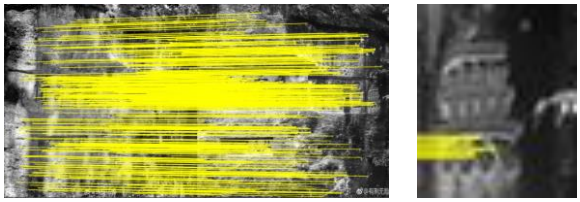
(a) 2018 Wuyishan stone inscriptions (b) 2024 Wuyishan stone inscriptions

Figure 2. Text area positioning results

Secondly, feature matching and evaluation were executed. Figures 3 and 4 illustrate a comparison of feature matching results between the traditional method and the proposed region-adaptive method, respectively. The results show that the proposed method not only maintains high matching accuracy, but also increases the number of valid matching points. Notably, The density of matching points within the text regions showed a marked improvement. For instance, in the region corresponding to the character '无' (Wú) (cf. Figures 3b and 4b), the traditional method yielded only 2 inlier matches, whereas the proposed method successfully extracted and matched 5 valid feature points, representing an increase of 150%.

(a) (b)

Figure 3. Registration result of SIFT+FLANN



(a) (b)

Figure 4. Registration result of optimised method

Finally, a quantitative evaluation analysis was conducted. Table 1 presents a detailed comparison of the traditional method and the proposed method across various evaluation metrics. The results demonstrate that the proposed method significantly outperforms the traditional method across all key indicator. The total number of matching points increased from 223 to 452, an increment of 102.7%, indicating that the algorithm successfully extracted a greater number of valid features in feature-sparse regions. The SSIM improved from 0.0337 to 0.3081, an increase of 813.7%, demonstrating a significant enhancement in the structural correspondence between the registered images. The inlier ratio experienced a slight increase from 93.27% to 94.47%, signifying that the algorithm maintained high matching precision despite a substantial increase in the number of matching points. Furthermore, the MEE decreased from 17.184 to 1.0109, a reduction of 94.11%.

Indicator	SIFT+FANLL	Ours
Total matching points number	223	452
SSIM	0.0337	0.3081
Inlier ratio	0.9327	0.9447
MEE	17.1843	1.0109

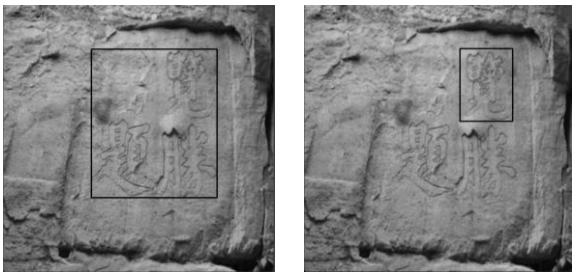
Table 1. Registration assessment result of Wuyishan stone inscriptions image

3.3 Results of Tianlongshan stone inscriptions

The second set of experiments was conducted using stone inscription images from Tianlong Mountain. The text regions on these carvings exhibit severe weathering, as exemplified by the character '览' (Lǎn), whose lower portion has nearly vanished

due to erosion. With other inscriptions similarly shallow, deep learning-based text detection methods struggle to accurately localize text regions. Therefore, for such severely weathered carvings, ROIs were defined through manual delineation. As illustrated in Figure 5, both a primary text region and representative sub-blocks were selected. These areas encompass typical weathering features and individual character units. The primary text region covers the core section of continuous text distribution within the image (approximately 150×175 pixels) and served as the baseline unit for feature enhancement. Individual character units (approximately 60×80 pixels) were used to validate the local enhancement efficacy of the multi-algorithm fusion.

Given the small image dimensions and severe background weathering, applying global partitioning for registration directly could introduce significant errors. Consequently, in this experimental set, feature point extraction and registration were performed exclusively on the manually annotated text regions. This approach was adopted to mitigate interference from background noise and thereby enhance the stability and accuracy of the matching process.



(a) (b)

Figure 5. ROIs regional positioning

Figure 6 depicts the matching results obtained using the traditional method. Only six matching pairs were identified within the text region, with some matches located in non-text areas, highlighting the limitations of traditional methods when applied to severely weathered carvings. Through the application of the regional enhancement strategy, the capability to detect keypoints within text regions was significantly improved, leading to a substantial increase in the density of matching points within these areas. As shown in Figure 7, the number of matching points in the primary text region increased to 114, thereby enhancing the stability and precision of the subsequent registration.

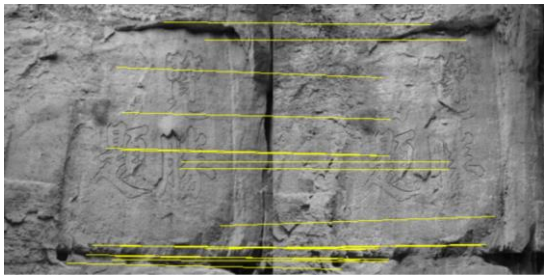


Figure 6. Registration result of SIFT+FLANN

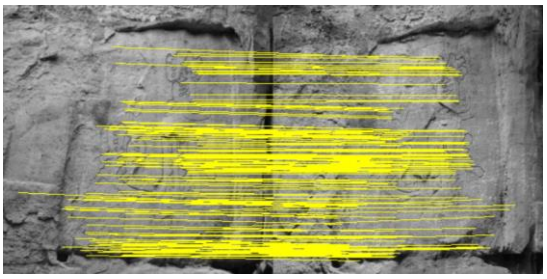


Figure 7. Registration result of optimised method

For feature-sparse regions, such as the '览' (Lǎn) character unit, the AKAZE algorithm was further employed for local feature enhancement. As illustrated in Figure 8, this algorithm successfully detected four sets of geometrically consistent matching points within the target region. These newly identified matching points, after coordinate correction, were fused with the initial match set and the regionally enhanced match set, ultimately forming a comprehensive, high-confidence distribution of matching points, as shown in Figure 9.

Figure 8. Registration result of '览' (Lǎn) character region

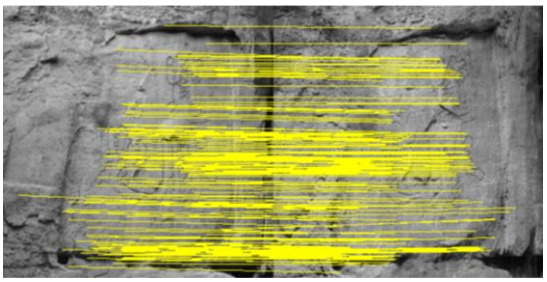


Figure 9. Final registration results

The indicators of the registration results are evaluated in Table 2. Within the text region, the number of matching points increased from 6 to 113, an increment of 1783%, which demonstrates a significant enhancement in the proposed algorithm's capability for feature extraction and matching in severely weathered areas. Concurrently, the total number of matching pairs across the entire image rose from 26 to 170, a 553% increase, reflecting a comprehensive improvement in global feature coverage. Registration robustness was also substantially improved, with the inlier ratio increasing from 76.92% to 97.35%, which represents an 88.8% reduction in the mismatch rate. This demonstrates that the optimized matching verification mechanism effectively identifies and discards outlier matches, significantly enhancing the system's resilience to interference. The MEE decreased from 15.37 to 14.51 pixels (a 5.6% reduction), reflecting an improvement in the spatial consistency of the matched points. The SSIM showed a marginal decrease of 0.5%, likely attributable to the fact that the more precise registration achieved by the proposed method exposed inherent content differences between the source images, namely, the actual weathering and exfoliation changes on the carving

surface. Despite the MEE not exhibiting a substantial decrease, an inlier ratio exceeding 97% is sufficient to ensure that the registration results are stable and usable.

Indicator	SIFT+FANLL	Ours
Matching points in ROI	6	114
Total matching points number	26	170
SSIM	0.4410	0.4358
Inlier ratio	0.7692	0.9735
MEE	15.3714	14.5107

Table 2. Registration assessment result of Tianlongshan stone inscriptions image

4. Conclusion

This study introduces a semantic-guided registration framework specifically for multi-temporal stone inscriptions images. Integrated the text region semantic recognition with region-adaptive feature enhancement, establishing a specialized registration mechanism tailored to the structural characteristics of stone inscriptions images. To address the challenges posed by the uneven distribution of feature points and the sparsity of features within text regions in stele images, the proposed method incorporates a four-stage processing pipeline: initial global registration, regional partitioning, sub-block enhancement, and refined global registration. At the same time, a feature point distribution assessment and augmentation mechanism is introduced, effectively resolving the persistent issue of feature point sparsity in text regions, a long-standing challenge in stone inscriptions image registration.

The results suggest the proposed method achieved better performance compared to existing approaches on both well-preserved and severely weathered stone inscriptions. For well-preserved stone inscriptions, the matching density was substantially increased while the MEE was reduced by 94.8%. In severely weathered samples, the proportion of valid match points increased by 23.6%, and registration accuracy improved by 5.6%, with only minor residual errors in localized areas attributable to weathering-induced deformation.

This method provides robust technical support for the precise registration of multi-temporal stone inscriptions images, thereby laying a technical foundation for subsequent image-based monitoring of weathering progression and assessment of conservation efficacy for stele artifacts.

Acknowledgements

This research was supported by Scientific and Technological Research Project of Cultural Relics of State Administration of Cultural Heritage (No. 2023ZCK014), Science and Technology Major Special Program Project of Shanxi Province (No. 202201150501024), Project of Key Laboratory of Silicate Cultural Relics Conservation (Shanghai University), Ministry of Education (No. SCRC2024KF06YQ, SCRC2024ZZ02ZD) and the Tencent Tanyuan Program 2024.

References

Anzid, H., Le Goic, G., Bekkari, A., Mansouri, A., Mammass, D., 2022: A new SURF-based algorithm for robust registration of multimodal images data. *The Visual Computer*. 39, 1667–1681 doi.org/10.1007/s00371-022-02435-z.

Bian, Y., Wang, M., Chu, Y., Liu, Z., Chen, J., Xia, Z., Fang, S., 2021: A cost function for the uncertainty of matching point distribution on image registration. *ISPRS International Journal of Geo-Information*, 10(7), 438. doi.org/10.3390/ijgi10070438.

Chen, J., Frey, E. C., He, Y., Segars, W. P., Li, Y., Du, Y., 2022: TransMorph: Transformer for unsupervised medical image registration. *Medical Image Analysis*, 82, 102615. doi.org/10.1016/j.media.2022.102615.

Gandhi, V. H., Panchal, S. R., 2014: Feature based image registration techniques: An introductory survey. *International Journal of Engineering Development and Research*, 2(1), 368–375.

Hou, X., Gao, Q., Wang, R., Luo, X., 2021: Satellite-borne optical remote sensing image registration based on point features. *Sensors*, 21(8), 2695. doi.org/10.3390/s21082695.

Ketcha, M. D., De Silva, T., Han, R., Uneri, A., Goerres, J., Jacobson, M. W., Vogt, S., Kleinszig, G., Siewerdsen, J. H., 2017: Effects of image quality on the fundamental limits of image registration accuracy. *IEEE Transactions on Medical Imaging*, 36(10), 1997–2009. doi.org/10.1109/TMI.2017.2725644.

Leng, C., Zhang, H., Cai, G., Chen, Z., Basu, A., 2021: Total variation constrained non-negative matrix factorization for medical image registration. *IEEE/CAA Journal of Automatica Sinica*, 8(5), 1025–1037. doi.org/10.1109/JAS.2021.1003979.

Nasri, A., Huang, X., 2020: A missing color area extraction approach from high-resolution statue images for cultural heritage documentation. *Scientific Reports*, 10(1), 21939. doi.org/10.1038/s41598-020-78254-w.

Navacerrada, M. A., Barbero-Barrera, M. M., Fort, R., De La Prida, D., Núñez, J. C., Gómez, T. S., 2022: Application of acoustic impedance gun to non-destructively monitor stone damage. *Construction and Building Materials*, 323, 126510. doi.org/10.1016/j.conbuildmat.2022.126510.

Salvatici, T., Calandra, S., Centauro, I., Pecchioni, E., Intrieri, E., Garzonio, C. A., 2020: Monitoring and evaluation of sandstone decay adopting non-destructive techniques: On-site application on building stones. *Heritage*, 3(4), 1287–1301. doi.org/10.3390/heritage3040071.

Tan, X., Sun, C., Sirault, X., Furbank, R., Pham, T. D., 2012: Feature correspondence with even distribution. *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*, 1–7. doi.org/10.1109/DICTA.2012.6411723.

Ye, Y., Tang, T., Zhu, B., Yang, C., Li, B., Hao, S., 2022: A multiscale framework with unsupervised learning for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–15. doi.org/10.1109/TGRS.2022.3167644.