Color-Enhanced Local Feature Fusion for FPFH-Based Point Cloud Registration

Zezhi Ding[1], Jiaqi Li[1], Xincong Yang[1,*] , Yangfan Wu[1], Ruinan Tan[1]

[1]School of Intelligent Civil and Ocean Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, 518055, PR China

Abstract –

To address the limitations of the traditional Fast Point Feature Histogram (FPFH) in heterogeneous building point cloud registration—over-reliance on geometric information, neglect of color cues, and susceptibility to lighting variations—this paper proposes a color-enhanced local feature fusion framework. First, the RGB color space is transformed to HSV, with histogram equalization applied to the Hue channel to improve illumination robustness and color discriminability. Then, normalized FPFH geometric features and PCFH color features are fused into a 44-dimensional joint vector. Experimental validation on real shopping mall indoor point clouds shows that the proposed method performs well in complex scenarios: RRE and RMSE are reduced by up to 74.9% and 74.6% under different initial poses; robust registration is achieved in asymmetric point-count scenarios where FPFH fails; and RMSE remains 27.73% lower than FPFH even with only 500 sampling points. An outdoor campus road experiment further verifies generalizability, with RRE, RTE, and RMSE reduced by 66.2%, 63.4%, and 59.8%, respectively. Overall, integrating compact color features introduces negligible computational overhead while improving registration accuracy and robustness.

Keywords –

Point cloud, Registration, Feature Fusion

1 Introduction

The application of point cloud technology in civil engineering has attracted widespread attention in recent years, serving as a key driver for the industry's digital transformation. Traditional surveying and mapping methods in this field are constrained by high labor costs, low efficiency, limited handling of complex scenarios, poor adaptability to dynamic environments, and time-consuming post-processing. In contrast, point cloud technology enables non-contact, high-precision capture of spatial information for engineering structures through high-density 3D data acquisition. Its advantages in efficiency, digitalization, 3D visualization, and adaptability to complex scenarios offer innovative solutions for full lifecycle management of engineering projects [1-4].

Point cloud registration, which estimates the rigid transformation (i.e., 3D rotation and translation) between source and target point clouds, is a fundamental challenge in applications such as 3D building reconstruction [5-7]. Among traditional methods, many rely on geometric feature descriptors, with the Fast Point Feature Histogram (FPFH) [8] being a common choice. By statistically analyzing angular distributions between query points and neighbors, FPFH demonstrates robustness to mild noise. However, when applied to heterogeneous building point clouds, FPFH has inherent limitations: its exclusive reliance on geometry overlooks rich color cues in RGB data. Furthermore, local geometric neighborhoods can suffer significant distortions when scanned from different viewpoints, reducing feature similarity.

As a complementary data type, color information is expected to enhance local feature discriminability. However, raw RGB values are sensitive to lighting variations, leading to suboptimal results in direct applications. For example, walls scanned on sunny days appear brighter than those captured on cloudy days, causing color discrepancies at the same physical locations [9]. Naively combining color features with FPFH fails to achieve reliable fusion; instead, targeted color enhancement is necessary to improve lighting invariance and feature distinctiveness.

To address these challenges, this paper presents a novel FPFH-based registration framework integrating color-enhanced local feature fusion. The core principle is to synergize lighting-invariant color features with geometric descriptors, leveraging their complementary strengths to improve matching accuracy and robustness for heterogeneous data. The contributions are summarized as follows:

(1) A tailored color enhancement strategy for building

point clouds: By converting RGB to HSV color space and applying histogram equalization to the Hue channel, lighting interference is mitigated, and color discriminability is enhanced.

(2) Fusion of local features by integrating color-enhanced attributes with FPFH descriptors. Through normalization and dimension balancing, both types of features effectively facilitate matching.

(3) Experimental validation on real-world point cloud datasets demonstrates superior registration accuracy and robustness compared to traditional FPFH methods.

2 Related work

Point cloud registration methods are generally classified into coarse and fine registration. Coarse registration estimates an initial transformation matrix between point clouds with significant pose differences. RANSAC [10] is a classic method that estimates the transformation by randomly sampling point pairs and evaluating their consistency. To improve robustness, 4PCS [11] selects coplanar point sets to reduce the influence of outliers, but it suffers from high computational cost on large-scale point clouds. Recently, optimal transport-based methods such as Sinkhorn algorithms [12-14] have been introduced to enhance correspondence confidence and reduce false matches. However, these methods still rely on discriminative initial features, highlighting the importance of high-quality descriptors.

For fine registration, the Iterative Closest Point (ICP) algorithm [15] and its variants [16] are widely used. These methods iteratively minimize the distance between corresponding points to achieve precise alignment, but they are highly sensitive to the initial pose. Poor initial alignment may lead to convergence to a local optimum rather than the global solution. Therefore, reliable coarse registration remains a critical prerequisite for successful point cloud registration.

Geometric feature descriptors form the foundation of point cloud registration by encoding local structural information. Early descriptors such as Shape Context [17] and Spin Images [18] laid the groundwork for 3D feature extraction but often involve high computational complexity. As an efficient variant of PFH, FPFH reduces point-pair calculations while maintaining strong geometric discriminability, making it a benchmark in geometry-based registration. Nevertheless, FPFH is sensitive to viewpoint changes and point-density variations in heterogeneous scenarios. More robust descriptors have therefore been developed. For example, SHOT [19] improves viewpoint invariance through structured neighborhood sectors but has high dimensionality, while ISS [20] detects key points based on geometric saliency but still requires auxiliary descriptors for matching.

As a complement to geometric information, color features have been increasingly used to enhance the discriminative capability of point cloud descriptors. Early studies directly incorporated RGB values to improve nearest-neighbor correspondence or pose estimation [21-24], but raw color data are highly sensitive to illumination changes, resulting in unstable matching. To improve illumination invariance, researchers have explored color space transformations—such as converting from RGB to HSV or Lab color spaces—which separate chromatic information from brightness, thereby mitigating the effects of variable lighting [25,26]. Inspired by these approaches, this work fully leverages the color information embedded in point cloud data and adopts the HSV color space, following established practices. This enables effective registration without compromising the integrity of spatial structure.

To summarize, existing registration methods face challenges in heterogeneous building environments: limited discriminability of geometric features alone, vulnerability of raw color features to illumination variations, and suboptimal performance of naive feature fusion. To address these issues, this paper proposes a color-enhanced local feature fusion framework. Compared with ColorPCR, which relies on multi-stage color–geometry fusion, and SHOT-color, which extends high-dimensional descriptors, the proposed method constructs a lightweight Hue-based color histogram and efficiently fuses it with FPFH while retaining low computational cost.

3 Methodology

3.1 Problem Statement

Given two color point clouds $\mathcal{P} = \{p_i \in \mathbb{R}^3 \mid i = 1, \dots, N\}$ with $\mathcal{P}_c = \{p_{ci} \in \mathbb{R}^3 \mid i = 1, \dots, N\}$ and $\mathcal{Q} = \{q_i \in \mathbb{R}^3 \mid i = 1, \dots, M\}$ with $\mathcal{Q}_c = \{q_{ci} \in \mathbb{R}^3 \mid i = 1, \dots, M\}$, in which, $\mathcal{P}_c$ and $\mathcal{Q}_c$ denote the color of $\mathcal{P}$ and $\mathcal{Q}$ respectively, our goal is to estimate a rigid transformation $\boldsymbol{T} = \{\boldsymbol{R}, \boldsymbol{t}\}$ which aligns the two point clouds, comprising a 3D rotation $\boldsymbol{R} \in SO(3)$ and a 3D translation $\boldsymbol{t} \in \mathbb{R}^3$. With the estimated transformation, we can align the two color point clouds. To achieve the most precise alignment, we need to solve the following optimization problem:

$$\min_{\boldsymbol{R}, \boldsymbol{t}} \sum_{(p_i^*, q_i^*) \in C^*} \|R \cdot p_i^* + t - q_i^*\|_2^2 \tag{1}$$

3.2 Overall Algorithm Design

We improve coarse registration based on FPFH by incorporating color information into the FPFH descriptors. The resulting combined features demonstrate robust performance even in environments with limited geometric features or surface details, and show strong resilience to noise and variations in point cloud density.

The main procedure is as follows (the pipeline is illustrated in **Figure 1**.):

(1) Downsample the input source and target point clouds;

(2) Estimate surface normals for both point clouds using normal estimation methods;

(3) Compute FPFH features for each point in both point clouds;

(4) Convert the RGB values to the HSV color space and extract color descriptors using only the Hue (H) channel;

(5) Apply L2 normalization separately to the geometric (FPFH) and color descriptors, then concatenate them into a unified feature vector;

(6) Perform registration using the RANSAC algorithm.

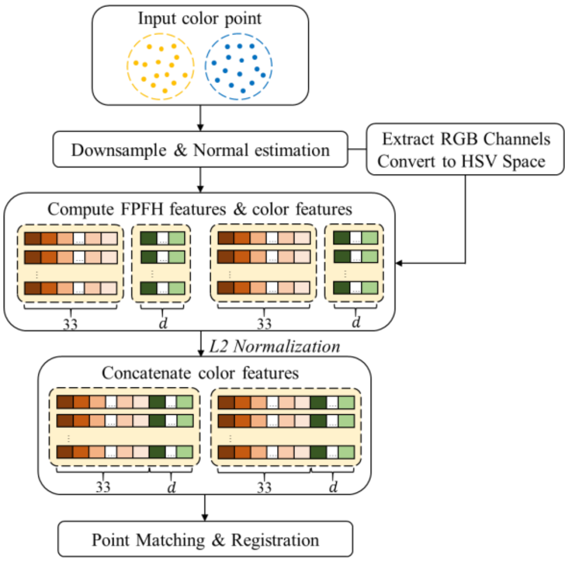
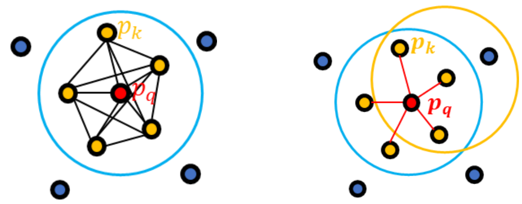


Figure 1. Pipeline of the proposed method

3.3 Computation of Fast Point Feature Histogram

The core idea of FPFH lies in first computing the Simplified Point Feature Histogram (SPFH) for each point in the k-neighborhood of the query point, and then weighting all the SPFH features via a formula to form the final FPFH feature. While FPFH reduces the computational complexity of the algorithm to O(nk), it still retains most of the discriminative properties of Point Feature Histograms (PFH). The computation process of FPFH is as follows:

(1) Only the three feature elements ($< \alpha, \phi, \theta >$) between each query point p_q and its neighborhood points are computed, which differs from PFH: while PFH calculates the feature elements for all combinations of neighborhood points, this step only computes those between the query point and its neighboring points. As shown in **Figure 2**, **Figure 2(a)** illustrates the feature computation process of PFH—i.e., the feature values for all combinations of neighborhood points (all lines in the figure, including but not limited to the lines between p_q and p_k, while **Figure 2(b)** depicts the computation scope of FPFH, which only requires calculating the feature elements between the query point p_q and its immediate neighboring points p_k (the red lines in **Figure 2(b)**). It can be seen that this design reduces the computational complexity, and this histogram is termed the Simplified Point Feature Histogram (SPFH).



(a) Spatial coverage of PFH (b) Spatial coverage of FPFH

Figure 2. The spatial coverage of FPFH and PFH

(2) A modified statistical method for k-neighborhood point pairs is proposed, comprising two components: point pairs between the query point and its k surrounding points; point pairs between each neighborhood point and its own k surrounding points. For the second component, the weight is usually set to the reciprocal of the distance, such that spatially closer neighborhood points make a greater contribution, as shown in **Eq. (2)**:

$$FPFH(p_q) = SPFH(p_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{w_k}\cdot SPFH(p_k) \tag{2}$$

3.4 Computation of Point Color Feature Histogram

We transform the RGB values of the point cloud data into the HSV color space and extract color features using only the Hue (H) channel. First, for each point, a d-dimensional Histogram of Hue Differences (HHD) is calculated between the query point p_q and its neighboring points p_k. Next, for each point again, the Point Color Feature Histogram (PCFH) is computed based on the HHD calculated in the first step, using **Eq. (3)-(4)**, where the weight is determined by the distance between the neighboring point and the query point.

$$HHD(p_q) = Histogram(|h_q - h_k|, d) \tag{3}$$

$$PCFH(p_q) = HHD(p_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{w_k}\cdot HHD(p_k) \quad (4)$$

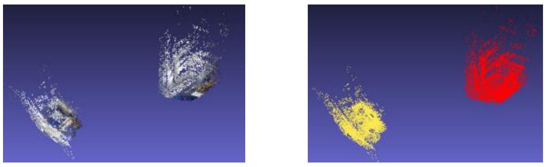
After obtaining the PCFH descriptor, we first apply L2 normalization to it, and then concatenate it with the FPFH descriptor to form a 44-dimensional feature vector (where the dimension of PCFH is fixed at 11 in subsequent experimental designs). The purpose of normalizing before concatenation is to facilitate better collaboration between the geometric and color descriptors.

4 Experiments and Results

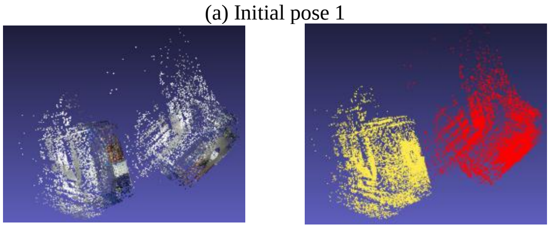
4.1 Experiment Design

The point clouds used in this study were collected from a large shopping mall using the Leica BLK360 imaging scanner. The source and target point clouds were obtained from two scans at the same location. Since the scanner does not provide initial pose information, the ground truth transformation matrix in Eq. (5)–(7) was generated by applying an artificial transformation to the target point cloud, ensuring an exact reference. Five experiments were designed: (1) registering two 10,000-point clouds under three initial poses; (2) registering asymmetric point counts of 12,000 and 4,000 points; (3) testing reduced densities of 4,000, 2,000, 1,000, and 500 points; (4) comparing the computation time of FPFH and FPFH + PCFH; and (5) testing outdoor generalization on a campus road segment with two 10,000-point clouds.

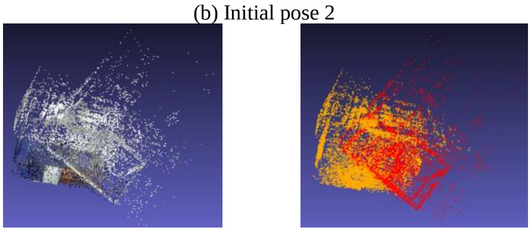
Experiment 1: The three initial poses are detailed in Table 1. This experiment evaluated the registration accuracy of the Color-enhanced FPFH algorithm under diverse initial pose conditions. Figure 3 illustrates the downsampled point clouds and their initial relative poses, with the source cloud colored red and the target orange. The RANSAC algorithm was set to 600,000 iterations, with an inlier threshold of 0.2 m, and the results were compared with FPFH in Section 4.2.



(a) Initial pose 1



(b) Initial pose 2



(c) Initial pose 3

Figure 3. Three different initial poses in Experiment 1

Table 1 Three different initial poses in Experiment 1

	Initial pose 1	Initial pose 2	Initial pose 3
Roll/°	55.08	60.63	33.68
Pitch/°	-9.61	-84.00	41.03
Yaw/°	-55.52	97.76	17.52
X-axis translation distance/m	-20.097	8.376	-3.523
Y-axis translation distance/m	20.474	-7.014	-0.243
Z-axis translation distance/m	24.060	8.380	1.931

Experiment 2: The source and target point clouds were downsampled to 12,000 and 4,000 points, respectively, to amplify point-count discrepancy and reduce the proportion of feature-similar points. The initial pose is shown in Figure 4 and Table 2. The RANSAC settings were the same as in Experiment 1, and the results were compared with FPFH in Section 4.2.

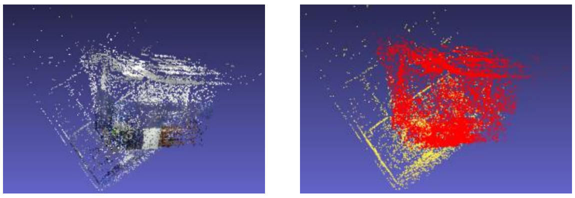


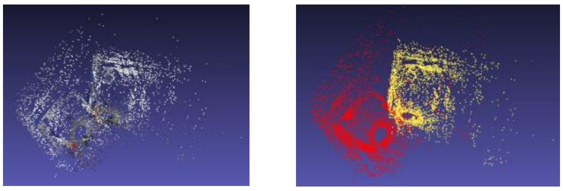
Figure 4. The initial pose in Experiment 2

Table 2 The initial pose in Experiment 2

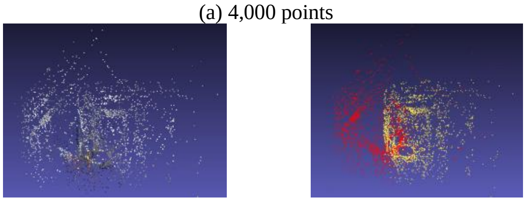
	Initial pose
Roll/°	-2.38
Pitch/°	-68.74
Yaw/°	30.71
X-axis translation distance/m	3.362
Y-axis translation distance/m	3.933
Z-axis translation distance/m	-3.000

Experiment 3: Four downsampling point counts were tested: 4,000, 2,000, 1,000, and 500. The objective was to evaluate the accuracy and robustness of the Color-enhanced FPFH algorithm under continuously decreasing sampling densities. The point clouds and initial poses are shown in Figure 5 and

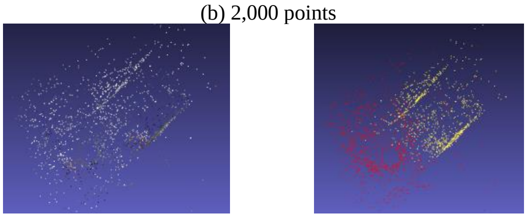
Table 3.



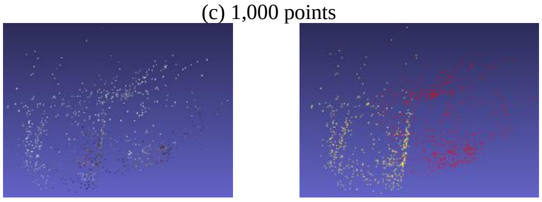
(a) 4,000 points



(b) 2,000 points



(c) 1,000 points



(d) 500 points

Figure 5. The four downsampling settings in Experiment 3

Table 3 The initial poses in Experiment 3

	Initial pose (4k)	Initial pose (2k)	Initial pose (1k)	Initial pose (0.5k)
Roll/°	55.97	-39.97	-11.23	-28.28
Pitch/°	14.39	35.47	-58.75	-61.12
Yaw/°	-34.77	-18.12	-22.35	-19.60
X-axis translation distance/m	4.479	0.666	4.966	-3.183
Y-axis translation distance/m	2.380	0.087	4.855	4.865
Z-axis translation distance/m	1.579	-2.754	-4.406	1.291

Experiment 4: To quantify the computation time of the fused descriptor versus the FPFH, we designed a comparative experiment. To ensure a fair comparison, both methods utilized the same parameters.

Experiment 5: The initial pose is detailed in Table 4. This experiment evaluated outdoor generalization using a campus road segment point cloud. Figure 6 shows the source and target point clouds and their initial relative poses, with the source colored red and the target orange.

Table 4 The initial pose in Experiment 5

	Initial pose
Roll/°	52.47
Pitch/°	23.51
Yaw/°	−56.22
X-axis translation distance/m	−97.50
Y-axis translation distance/m	81.86
Z-axis translation distance/m	81.90

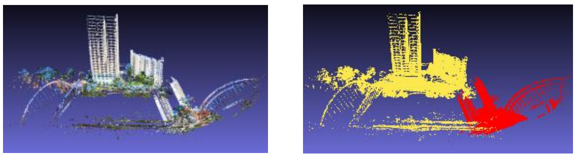


Figure 6. The initial pose in Experiment 5

In this paper, normal estimation employs a hybrid neighborhood search strategy, with a search radius of 2 meters and a maximum of 30 nearest neighbors. Similarly, the FPFH feature extraction algorithm utilizes a hybrid search strategy, setting the FPFH search radius to 5 meters and the maximum number of neighbors to 100.

The custom PCFH descriptor (dimension = 11) is constructed based on the Hue (H) channel. Its neighborhood search adopts a spherical neighborhood with a radius of 2.0 m (skipping points with fewer than 3 neighbors). The Hue difference calculation accounts for the 360° cyclic nature of the H channel by taking the minimum angular difference (0–180°) and normalizing it to the [0,1] range. Histogram binning uses 11 equal-width bins covering the full [0,1] range.

During feature matching, the nearest neighbor in the target point cloud is first identified for each source point in the concatenated 44-dimensional feature space. Mutual nearest neighbor filtering retains bidirectional correspondences, and incorrect matches are further filtered using edge length verification (threshold 0.9) and a correspondence distance threshold of 0.2 m. The valid correspondences are then used for RANSAC, configured with 600,000 iterations, confidence of 0.999, three sample points, and the same 0.2 m distance threshold.

4.2 Discussion

To further quantify the registration results, the Relative Rotation Error (RRE), Relative Translation Error (RTE), and Root Mean Square Error (RMSE) for different algorithms were calculated. The formulas for RRE, RTE and RMSE are as follows:

$$RRE = \arccos\left(\frac{tr(R_{gt}^T R_{est}) - 1}{2}\right) \times \frac{180}{\pi} \quad (5)$$

$$RTE = \left\| t_{gt} - t_{est} \right\|_2 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N}\sum\left\|(T_{gt} \cdot P) - (T_{est} \cdot P)\right\|_2^2} \qquad (7)$$

where R_{gt} and t_{gt} denote the ground truth rotation and translation, respectively; R_{est} and t_{est} denote the estimated rotation and translation; T_{gt} denote the ground truth transformation matrix; T_{est} denotes the estimated transformation matrix, and P denotes the source point cloud.

Under controlled conditions where all other experimental parameters were fixed, Tables 5-9 summarize the registration results. As shown in Table 5, under different initial poses, the Color-enhanced FPFH-based algorithm consistently outperformed standard FPFH. The most significant improvement occurred under Initial Pose 2, where RRE, RTE, and RMSE improved by 74.9%, 67.4%, and 74.6%, respectively. Although the RTE under Initial Pose 1 slightly increased from 0.0638 m to 0.104 m, this value remains within the 0.2 m RANSAC correspondence threshold and does not compromise overall stability. In Experiment 2 (Table 6), standard FPFH essentially failed under significantly different point counts, while the proposed method achieved low errors of 2.100° (RRE), 0.048 m (RTE), and 0.2 m (RMSE), indicating better matching robustness under asymmetric density. In Experiment 3 (Table 7), the proposed method maintained high accuracy at 4,000, 2,000, and 1,000 points, and even with only 500 points still improved RRE, RTE, and RMSE by 26.2%, 30.82%, and 27.73%, respectively. Table 8 shows that adding the 11-dimensional color feature causes almost no increase in computation time. Table 9 further demonstrates outdoor generalization, with RRE, RTE, and RMSE reduced by 66.2%, 63.4%, and 59.8%, respectively.

Table 5 Comparison of registration results (Experiment 1)

	FPFH-based algorithm			Color-enhanced FPFH-based algorithm		
Initial pose	1	2	3	1	2	3
RRE/°	2.904	4.084	1.313	0.824	1.025	0.635
RTE/m	0.0638	0.279	0.099	0.104	0.091	0.070
RMSE/m	0.234	0.379	0.169	0.127	0.096	0.108

Table 6 Comparison of registration results (Experiment 2)

	FPFH-based algorithm	Color-enhanced FPFH-based algorithm
RRE/°	172.510	2.100
RTE/m	2.947	0.048
RMSE/m	8.550	0.200

Table 7 Comparison of registration results (Experiment 3)

	FPFH-based algorithm			
Point Count	4k	2k	1k	0.5k
RRE/°	1.440	39.795	10.109	29.795
RTE/m	0.066	1.063	0.201	2.164
RMSE/m	0.142	3.012	0.818	2.517
	Color-enhanced FPFH-based algorithm			
Point Count	4k	2k	1k	0.5k
RRE/°	1.633	2.108	2.987	21.990
RTE/m	0.065	0.045	0.261	1.497
RMSE/m	0.147	0.185	0.329	1.819

Table 8 The processing times in Experiment 4

Point Count	4k	2k	1k	0.5k
FPFH	0.0034s	0.0025s	0.0010s	0.0006s
FPFH + PCFH	0.00036s	0.0027s	0.0011s	0.0009s

Table 9 Comparison of registration results (Experiment 5)

	FPFH-based algorithm	Color-enhanced FPFH-based algorithm
RRE/°	0.281	0.095
RTE/m	0.254	0.093
RMSE/m	0.241	0.097

5 Conclusions

This paper proposes a color-enhanced local feature fusion framework to address the limitations of FPFH-based registration in heterogeneous building point clouds, including over-reliance on geometry, neglect of color cues, and sensitivity to lighting interference. The method transforms RGB to HSV, applies histogram equalization to the Hue channel, and fuses normalized FPFH and PCFH into a 44-dimensional joint descriptor. Experiments on real shopping mall point clouds demonstrate strong performance under different initial poses, asymmetric point counts, and sparse sampling: RRE and RMSE are reduced by up to 74.9% and 74.6%; robust registration is achieved when FPFH fails under asymmetric point counts; and at 500 points, RMSE remains 27.73% lower than FPFH. The computation-time comparison shows negligible additional overhead, and the outdoor experiment confirms promising generalization ability.

Despite the promising results, this study has limitations. The impact of PCFH dimensionality on registration performance has not been fully explored; the value of 11 was chosen following the common bin setting of PFH and FPFH. Future work will optimize the color feature dimensionality and weighting strategy through sensitivity analysis or adaptive learning. Moreover, this study only uses the Hue channel, while the influence of Saturation and Value channels should be investigated in more complex scenarios.

6 Acknowledgement

This research is supported by Key Technologies R&D Program (Grant No. 2024YFC3808602), Shenzhen Science and Technology Programs (Grant No. KJZD20230923114310021, Grant No. JCYJ20241202123722029).

References

[1] H. Yue, Q. Wang, H. Zhao, N. Zeng, and Y. Tan. Deep learning applications for point clouds in the construction industry. Automation in Construction, 168:105769, 2024.

[2] Q. Wang and M. K. Kim. Applications of 3D point cloud data in the construction industry: A fifteen-year review from 2004 to 2018. Advanced engineering informatics, 39:306–319, 2019.

[3] L. J. Sanchez-Aparicio, F. L. del Blanco-Garcia, D. Mencias-Carrizosa, P. Villanueva-Llaurado, J. R. Aira-Zunzunegui, D. Sanz-Arauz, and J. Garcia-Gago. Detection of damage in heritage constructions based on 3D point clouds. A systematic review. Journal of Building Engineering, 77:107440, 2023.

[4] P. Stałowska, C. Suchocki, and M. Rutkowska. Crack detection in building walls based on geometric and radiometric point cloud information. Automation in Construction, 134:104065, 2022.

[5] M. R. Sanzana, T. Maul, J. Y. Wong, M. O. M. Abdulrazic, and C. C. Yip. Application of deep learning in facility management and maintenance for heating, ventilation, and air conditioning. Automation in Construction, 141:104445, 2022.

[6] C. Zhang and D. Arditi. Advanced progress control of infrastructure construction projects using terrestrial laser scanning technology. Infrastructures, 5(10):83, 2020.

[7] R. Maalek, D. D. Lichti, and J. Y. Ruwanpura. Automatic recognition of common structural elements from point clouds for automated progress monitoring and dimensional quality control in reinforced concrete construction. Remote Sensing, 11(9):1102, 2019.

[8] R. B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (FPFH) for 3D registration. In 2009 IEEE International Conference on Robotics and Automation, pages 3212–3217. IEEE, 2009.

[9] J. Mu, L. Bie, S. Du, and Y. Gao. ColorPCR: Color point cloud registration with multi-stage geometric-color fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 21061–21070, 2024.

[10] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381–395, 1981.

[11] D. Aiger, N. J. Mitra, and D. Cohen-Or. 4-points congruent sets for robust pairwise surface registration. In ACM SIGGRAPH 2008 papers, pages 1–10, 2008.

[12] R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. The Annals of Mathematical Statistics, 35(2):876–879, 1964.

[13] R. Sinkhorn and P. Knopp. Concerning nonnegative matrices and doubly stochastic matrices. Pacific Journal of Mathematics, 21(2):343–348, 1967.

[14] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. Advances in Neural Information Processing Systems, 26, 2013.

[15] P. J. Besl and N. D. McKay. Method for registration of 3-D shapes. In Sensor Fusion IV: Control Paradigms and Data Structures, volume 1611, pages 586–606. SPIE, 1992.

[16] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, pages 145–152. IEEE, 2001.

[17] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(4):509–522, 2002.

[18] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(5):433–449, 2002.

[19] F. Tombari, S. Salti, and L. Di Stefano. Unique signatures of histograms for local surface description. In European Conference on Computer Vision, pages 356–369. Springer Berlin Heidelberg, 2010.

[20] Y. Zhong. Intrinsic shape signatures: A shape descriptor for 3D object recognition. In 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, pages 689–696. IEEE, 2009.

[21] J. Park, Q. Y. Zhou, and V. Koltun. Colored point cloud registration revisited. In Proceedings of the IEEE International Conference on Computer Vision, pages 143–152, 2017.

[22] M. Korn, M. Holzkothen, and J. Pauli. Color supported generalized-ICP. In 2014 International Conference on Computer Vision Theory and Applications, VISAPP, volume 3, pages 592–599. IEEE, 2014.

[23] J. Servos and S. L. Waslander. Multi-Channel Generalized-ICP: A robust framework for multi-

channel scan registration. Robotics and Autonomous Systems, 87:247–257, 2017.

[24] J. H. Joung, K. H. An, J. W. Kang, M. J. Chung, and W. Yu. 3D environment reconstruction using modified color ICP algorithm by fusion of a camera and a 3D laser range finder. In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 3082–3088. IEEE, 2009.

[25] J. Mu, L. Bie, S. Du, and Y. Gao. ColorPCR: Color point cloud registration with multi-stage geometric-color fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 21061–21070, 2024.

[26] H. Men, B. Gebre, and K. Pochiraju. Color point cloud registration with 4D ICP algorithm. In 2011 IEEE International Conference on Robotics and Automation, pages 1511–1516. IEEE, 2011.