IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 6, 2019, accepted December 25, 2019, date of publication January 6, 2020, date of current version January 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2963987

Color Point Cloud Registration Based on Supervoxel Correspondence

YANG YANG[1], (Member, IEEE), WEILE CHEN[2], MUYI WANG[1], DEXING ZHONG[3], AND SHAOYI DU[1], (Member, IEEE)

[1]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[3]Shenzhen Research School, Xi'an Jiaotong University, Shenzhen 518057, China

Corresponding authors: Dexing Zhong (bell@xjtu.edu.cn) and Shaoyi Du (dushaoyi@gmail.com)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFA0700800, and in part by the project of Shenzhen Technology Plan under Grant JCYJ20170816100724089.

ABSTRACT With the development of RGBD sensors, the high-quality color point cloud can be obtained expediently. In this paper, we propose a novel registration method for 3D color point clouds from different views, which is a critical issue in many applications. Different from traditional feature-based methods, we design a hybrid feature representation with color moments of the point, which could be applied naturally for any color point cloud. And these features are extracted from point clouds based on the supervoxel segmentation. By jointly conducting these features for similarity measure, a weight parameter is dynamically adapted between the color and the spatial information. The registration algorithm is under a classic iterative framework for building the correspondence and estimating transformation parameters. In addition, we provide a mutual correspondence matching condition with hybrid features to build some more robust relationships for estimating transformation parameters. Experimental results demonstrate that our method can effectively reduce the number of point data for registration and achieve good matching results even in a poor initial condition.

INDEX TERMS Color point cloud registration, hybrid feature, mutual correspondence matching.

I. INTRODUCTION

Point cloud registration is a classical problem in the fields of computer vision, computer graphics, and robotics. It has a wide application in scene reconstruction [1], 3D printing [2], medical image analysis [3], 3D object recognition [4], and so on. The goal of point cloud registration is to find an optimal spatial transformation to align two given 3D point clouds. For solving this problem, the classic algorithm is the Iterative Closest Point (ICP) algorithm [5]. It iteratively establishes the correspondence by distance similarity between points, and estimates the spatial transformation parameters according to the corresponding points. In reality, point clouds are captured in non-ideal conditions. Some researchers [6], [7] improved the ICP algorithm to convert the registration problem into a geometric alignment issue for the point clouds with noise or partial overlapping data.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaoqing Pan.

With the development of RGBD sensors, high-quality color point cloud can be obtained expediently. Color information provides a great help for registration especially when the geometry alignment process is confused by some symmetry structures. One direct way to use the color feature is to add the color vector to its coordinate vector for each point. Then the correspondence is built by combining the similarity measure of color feature. However, this easy way ignores the context information. The correspondence is interfered by noise. More robust way is to extract some local appearance features, such as Scale Invariant Feature Transform (SIFT) feature [8], inspired by the feature extraction method used in image processing. These features are useful for the point cloud with obvious local invariant features, whose registration results are depending on a few of robust feature points.

In this paper, we aim at aligning two color point clouds from different views, which is useful especially for the hand-held devices. Since the initial geometric aligning is bad, we prefer to add the help of color information. In feature level, we design a hybrid feature representation with color moments. The color moments express the color distribution of a local region in high-level. By jointly conducting the features for similarity measure, we dynamically adjust the weight between color and spatial information. Accordingly, the alignment metric is built by measuring these hybrid features. Then, the registration method is under the classic ICP framework. In addition, we propose a mutual correspondence matching condition based on the hybrid features. By filtering out the ambiguous feature relationships, the transformation estimation could be more accurate for robust registration. This paper is an extension version of [9]. Comparing with the previous work, more improvements are provided in the pre-processing for the large-scale data to improve the efficiency. We firstly conduct a supervoxel segmentation for point clouds. Then the down-sampling is used to remain the center point in each supervoxel and calculate the color moments in the supervoxel region. In this way, the point clouds become very sparse, but the local color and spatial features are still retained.

To sum up, the following contributions are presented in this paper.

1. To reduce the amount of the large-scale data, the point clouds are down-sampled by the supervoxel segmentation. The hybrid feature is designed by combining the spatial vector with color moments in each supervoxel.

2. A new registration objective function is proposed for the color point clouds by the hybrid similarity measure.

3. A mutual correspondence matching based on the hybrid features is suggested to select the accurate correspondences between features.

II. RELATED WORK

According to the features utilized in registration, we roughly separate the methods into two classes: the method based on the geometric alignment, and the method combined with color information.

A. REGISTRATION BY GEOMETRIC ALIGNMENT

Some researchers improved the ICP algorithm for the specific geometry structures. Chen and Medioni [10] and Rusinkiewicz and Levoy [11] proposed the point-to-plane ICP methods by minimizing the distance from each point to the nearest plane. Similar method using the plane-to-plane strategy was presented in [12]. By introducing a planar structure of point clouds, these algorithms have high accuracy for indoor scenes. To align two point clouds with low overlapping parts, Chetverikov *et al.* [6] proposed the trimmed ICP (TrICP) algorithm, which introduced an overlapping ratio into the objective function to trim outliers. It had great robustness and high precision for point clouds with only partly overlapping shapes. Then, Philllips *et al.* [7] proposed the fractional ICP algorithm (FICP) to improve the speed of TrICP algorithm. To deal with outliers, a robust ICP method was improved by adding a constraint of the overlapping ratio [13]. Bariya *et al.* [14] used a set of scale-invariant local shape descriptors according to object structures, and applied them into an automatic registration for multiple objects. Ahmed *et al.* [15] extracted some implicit quadric surfaces from the point cloud, and defined a group of virtual interesting points to establish correspondences. Rusu et al. proposed Point Feature Histograms (PFH) [16], [17] as a multi-dimensional feature, which described the local geometry around a single point in point cloud. Then, they improved the speed of the PFH and used it for the point cloud registration [18]. However, the traditional ICP method and these geometric-based approaches established the correspondence fully depending on the shape of point cloud. So that the geometric matching ability affects the registration result.

Recently, some learning based point cloud registration methods have been proposed. Xie *et al.* [19] learned a binary spectral shape descriptor with the deep neural network for 3D shape correspondence. Deng *et al.* [20] proposed the PPFNet with global context aware local featurts for point matching. Then they also presented a rotation invariant local descriptors by unsupervised learning [21]. Wang and Solomon [22] introduced a point cloud embedding network with an attention-based module. However, these methods are focus on the pure geometric features. They ignore the important color features.

B. REGISTRATION WITH COLOR INFORMATION

In recent years, color-based methods have received widespread attentions. By introducing the color information, feature matching has improved the accuracy of building the correspondence, especially when the geometric information is not sufficient. In [23], Johnson et al. introduced a distance function that combines the color information to search the nearest neighbor point. Furthermore, Men *et al.* [24] and Korn *et al.* [25] constructed the joint point-color space using Hue value from Hue-Saturation-Lightness (HSL) color space and L, a, b from CIELab color space, respectively. They presented the metric to find the correspondence according to their respective spaces. Godin *et al.* [26] used the color attribution as a constraint to judge the similarity. Danelljan *et al.* [27] used a Gaussian mixture model to represent color features. Park *et al.* [28] established correspondences in the physical three-dimensional space, and defined a joint optimization objective that integrated both geometric and photometric terms. Besides, some studies extracted more advanced color features to assist registration. For example, Joung *et al.* [29] used SIFT to estimate the initial alignment. Chu and Nie [30] and Zheng *et al.* [31] used SIFT features to establish the correspondence between point clouds. Ji *et al.* [32] proposed a probabilistic ICP registration algorithm using Oriented FAST and Rotated BRIEF (ORB) color features.

Our method also belongs to the second class by combining the spatial geometric features with color information. However, our work is different from the above methods. On the one hand, considering the large-scale registration problem, we use a sparse geometric representation for point cloud by sampling the center points of its supervoxels. As the supervoxel is a local clustering result by neighbors, it can represent

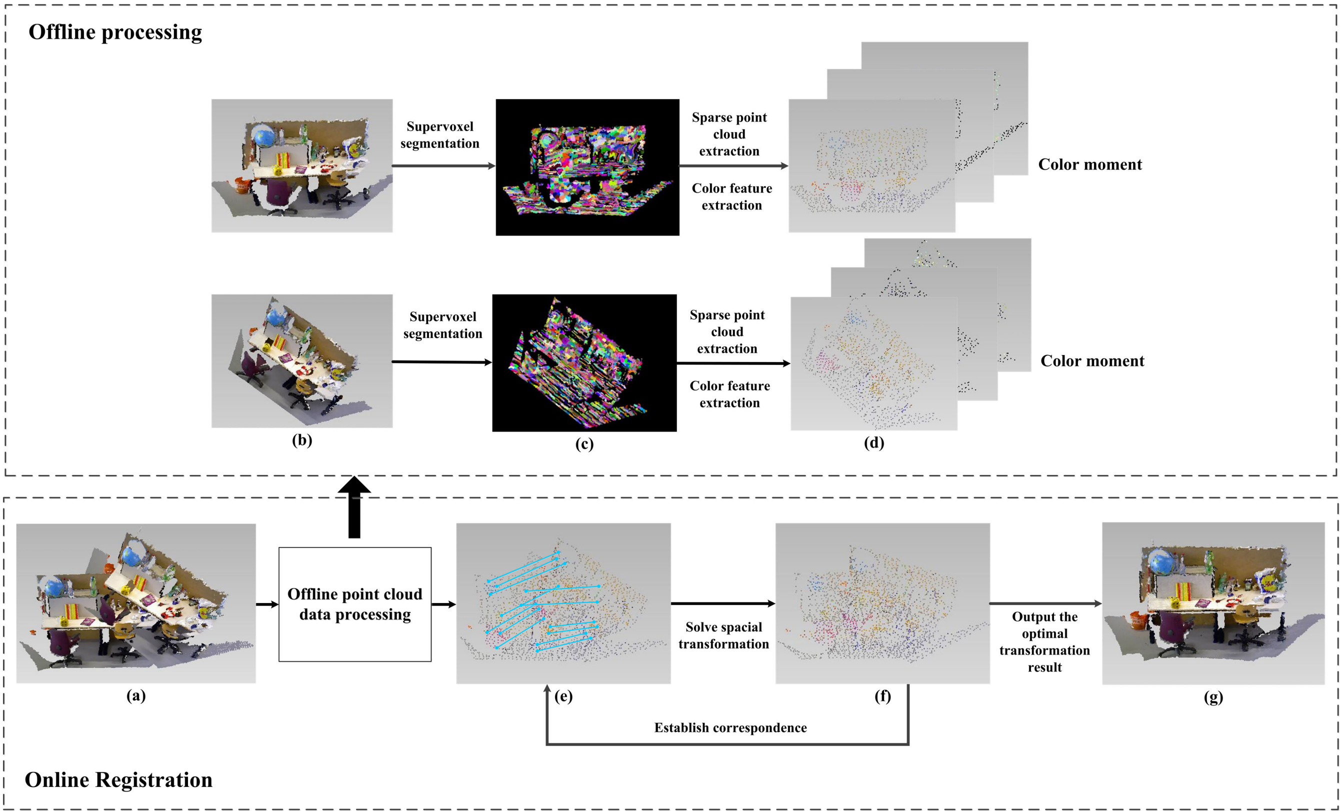


FIGURE 1. System overview. (a) Original point clouds. (b) Model point cloud and data point cloud. (c) Supervoxel segmentation results. Different colors represent different supervoxels. (d) Three orders color moment features. (e) The correspondence features. (f) The spatial position solved by the correspondence. (g) The final registration result.

the spatial distribution of the original point cloud. On the other hand, the color feature is extracted in a more general way as color moments in each supervoxel with context information. And the weight parameter is dynamically adapted between the color and the spatial information. To achieve accurate results, the corresponding feature pairs are built based on the bidirectional measures. Only when the mutual correspondences are satisfied, they can be used for the registration task. Other points will be neglected. So our method has the applicability for partial overlapping data. In conclusion, the color and point information we used are both common features of the color point cloud. There is no requirement about the point cloud structure or other local shape features. So our method can be used universally. The experimental results will demonstrate that our method has robustness for different initial positions.

III. COLOR POINT CLOUD REGISTRATION

The system overview is shown in Figure 1. Since a large amount of data in the point cloud brings a lot of computational burden for the following registration work. Downsampling is helpful to accelerate the registration process. But how to choose the sampling point will impact our registration result. We hope the sampling point could stand a local area. And it is known in a supervoxel that the location and color for all points are close. So to sample the center point of a supervoxel is the best way. The proposed method is firstly to segment the point clouds into several supervoxels. Then some sparse point clouds are sampled and their features are extracted by combining with color moments. The main registration algorithm is an iterative process as the ICP. Especially, a set of correspondence is built up by the mutual correspondence matching. And the transformation parameters are estimated accordingly.

A. SUPERVOXEL SEGMENTATION

We segment the point cloud into several local supervoxels for processing. The clustering method [32] is applied for generating supervoxels that takes into account of the color and geometric features of the point cloud. Segmentation results are consistent with the human's visual perception. We briefly introduce the segmentation method as following. More details can be found in [32].

Each 3D voxel is clustered according to a 39 dimensional vectors, given as:

$$F = [x, y, z, L, a, b, FPFH_{1,\ldots,33}], \tag{1}$$

where x, y, z are spatial coordinates, L, a, b are color values in CIELab space, and $FPFH_{1,\ldots,33}$ are the 33 elements of local geometrical feature proposed in [16] extracted by Fast Point Feature Histograms.

The distance measure D for clustering is:

$$D = \sqrt{\frac{\lambda D_c^2}{m^2} + \frac{\mu D_s^2}{3R_{seed}^2} + \epsilon D_{Hik}^2}, \quad (2)$$

where D_s is the spatial distance, D_c is the color distance calculated in the CIELab space, normalized by a constant m. D_{Hik} is the FPFH space distance calculated using the Histogram Intersection Kernel [30]. R_{seed} is the resolution of the seed voxel in space, which determines the number of initial supervoxel. λ, μ and ϵ are weight parameters of each distance measurement.

The main steps for segmenting a point cloud are as follows:

Step 1. The 3D space is divided into several regular spatial voxels according to the predefined resolution. A number of seed voxels is selected for initializing the centers of supervoxels. A small radius is used to filter out the noisy seed whose area does not contains sufficient voxels.

Step 2. The clustering is performed by a local K-Means clustering method. More specifically, the distance from the seed center to each candidate voxel is measured. The nearest seed is the supervoxel which voxel belongs. Then, these supervoxel centers are updated by their belonging voxels. Finally, this step is repeated until the cluster centers are stable.

B. HYBRID FEATURE REPRESENTATION WITH COLOR MOMENTS

After segmenting the point cloud into supervoxels, we present the spatial information of point cloud by a set of central points of each supervoxel. The color information is extracted to further enrich the feature presentation for these central points. Especially, three-order color moments are utilized, which can independently represent the color distribution characteristics of each supervoxel.

For all points contained in one supervoxel, we calculate their first three-order center moments on RGB color channels of these points. We define the color value of i-th points of u-th supervoxels at the k-th ($k = 3$) color channel as c_{iuk}. And the number of points in the u-th supervoxel is defined as N_u. The formulas for calculating the color moments are as follows:

$$E_k = \frac{1}{N_u}\sum_{i=1}^{N_u} c_{iuk}, \quad (3)$$

$$\sigma_k = (\frac{1}{N_u}\sum_{i=1}^{N_u}(c_{iuk} - E_k)^2)^{\frac{1}{2}}, \quad (4)$$

$$s_k = (\frac{1}{N_u}\sum_{i=1}^{N_u}(c_{iuk} - E_k)^3)^{\frac{1}{3}}, \quad (5)$$

where E_k is the color mean of the supervoxel. σ_k is the standard deviation which represents the differences between the color distribution and mean value. s_k is the skewness which denotes the degree of asymmetry which is related to the mean value.

According to Eqs. (3)-(5), we can acquire a nine-dimensional color moment eigenvector as follows:

$$\vec{M}_{p_u} = (E_r, \sigma_r, s_r, E_g, \sigma_g, s_g, E_b, \sigma_b, s_b),$$

which represents the statistical distribution of the color in a single supervoxel region.

In that case, for a color point cloud, both the spatial feature and local color distribution features are obtained by combining the color moment vector and the spatial coordinates of center points. We define a hybrid feature point cloud extracted from the original point cloud P as P_{SV}. For each point p_u in P_{SV}, combining its spatial coordinate $\vec{l}_{p_u} = (x, y, z)$ and its color moment eigenvector $\vec{M}_{p_u}$, we give the hybrid feature vector $\vec{p}_u$ as:

$$\vec{p}_u = (x, y, z, E_r, \sigma_r, s_r, E_g, \sigma_g, s_g, E_b, \sigma_b, s_b).$$

Compared to the original data, the number of the above hybrid features is much smaller. However, it still remains the main geometric features and color feature of the point cloud. In addition, compared with the original point clouds, since the spatial position of the feature point cloud does not change, our registration problem can be solved through two feature point clouds P_{SV} and Q_{SV}. In the following sections, we will see the advantages of these sparse hybrid features in the point-to-point registration algorithm.

C. OBJECTIVE FUNCTION

Based on the above point cloud features, the goal for registration is to find the spatial transformation $(\mathbf{R}, \vec{t})$ to align two sets of feature point clouds P_{SV} and Q_{SV}, where $\mathbf{R}$ is the rotation matrix and $\vec{t}$ is the translation vector in 3D space. So we propose a novel objective function as follows:

$$\begin{aligned} &\min_{\mathbf{R},\vec{t}} \sum_{i=1}^{N_{c(u,v)}} \|\mathbf{R}\vec{l}_{p_{u_i}} + \vec{t} - \vec{l}_{q_{v_i}}\|_2^2 + \omega\|\vec{M}_{p_{u_i}} - \vec{M}_{q_{v_i}}\|_2^2, \\ &s.t. \quad \mathbf{R}^T\mathbf{R} = I_n,\ det(\mathbf{R}) = 1. \end{aligned} \quad (6)$$

where $c(u, v)$ is the corresponding point pair set, p_{u_i} and q_{v_i} are the corresponding point pairs belong to $c(u, v)$, $N_{c(u,v)}$ is the number of $c(u, v)$. $\vec{l}_{p_{u_i}}$ and $\vec{l}_{q_{v_i}}$ is the spatial coordinates of the point p_{u_i} and q_{v_i}, ω is the weight of color feature, $\vec{M}_{p_{u_i}}$ and $\vec{M}_{q_{v_i}}$ are the color moment feature vectors.

There two sets of unknown parameters in this objective function: the correspondence and the transformation parameter. It can be solved under the ICP framework. In the following, we will present the details about how to build the correspondence including the similarity measure with color moments and the parameter estimation method by corresponding features.

D. SIMILARITY MEASURE WITH COLOR MOMENTS

To build the correspondence between two feature sets, the similarity measure should be defined according to the hybrid feature. The spatial coordinates represent the geometric feature of the point cloud, and its color moment vector represents the local color distribution characteristics of the point cloud. Therefore, giving the appropriate weights to balance them, we will find the corresponding points between the data point cloud and the model point cloud. The formula of similarity measure is defined as:

$$d(p_u, q_v) = \sqrt{d_s(p_u, q_v) + \omega d_c(p_u, q_v)}, \tag{7}$$

where

$$\begin{aligned} d_s(p_u, q_v) &= (\frac{x_p}{P_{range}} - \frac{x_q}{Q_{range}})^2 + (\frac{y_p}{P_{range}} \\ &\quad - \frac{y_q}{Q_{range}})^2 + (\frac{z_p}{P_{range}} - \frac{z_q}{Q_{range}})^2, \\ d_c(p_u, q_v) &= (E_{pr} - E_{qr})^2 + (\sigma_{pr} - \sigma_{qr})^2 + (s_{pr} - s_{qr})^2 \\ &\quad + (E_{pg} - E_{qg})^2 + (\sigma_{pg} - \sigma_{qg})^2 + (s_{pg} - s_{qg})^2 \\ &\quad + (E_{pb} - E_{qb})^2 + (\sigma_{pb} - \sigma_{qb})^2 + (s_{pb} - s_{qb})^2, \end{aligned}$$

where P_{range} and Q_{range} are the maximum acquisition ranges of point clouds P and Q. ω is the weight of color feature same as in Eq. (6). The value of ω is calculated by the following dynamic adapting method.

The main purpose of the dynamic color weight ω is to make the spatial and color features properly combining and perform suitable a similarity measurement according to different point cloud positions. The basic idea is: when the spatial positions of two point clouds are different greatly, it is more accurate to use the color feature rather than the spatial feature. And when the point clouds are roughly aligned, it is more suit to use spatial features to match the geometric details of the point clouds.

We define $sc(p_u, Q_{SV}) = q_v$ as the spatial correspondence, which means that in the model point cloud Q_{SV}, the point q_v is the nearest neighbor of p_u by Euclidean distance measure. We define $K(q_v)$ to represent the amount of point $p_u \in P_{SV}$ whose spatial correspondence point in model point cloud Q_{SV} is q_v:

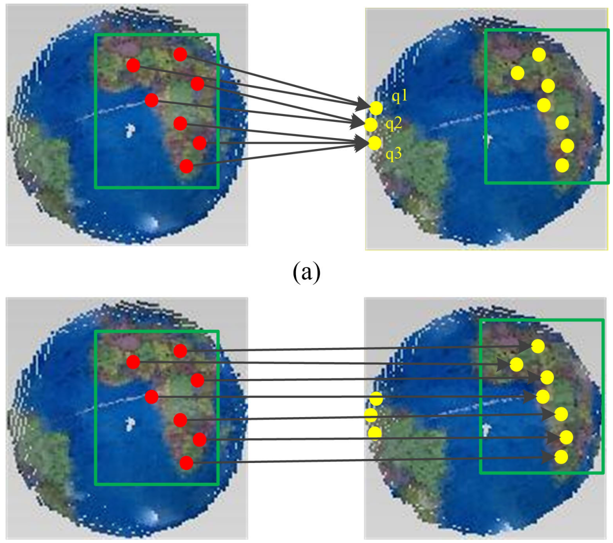
$$K(q_v) = |\{p_u \in P_{SV} : sc(p_u, Q_{SV}) = q_v, u = 1, \ldots, N_{P_{SV}}\}|. \tag{8}$$

Then the weight ω is calculated as follows:

$$\omega = \frac{1}{N_{P_{SV}}} \sum_{i=1}^{mN_{Q_{SV}}} \underset{v \in 1,\ldots,N_{Q_{SV}}}{sort} (K_i(q_v)), \tag{9}$$

where the $sort(\mathbf{.})$ is a descending sorting function, m is a constant term whose value is set to 0.01, which means that only the top one percent points K value in the point cloud Q_{SV} is calculated.

Figure 2 presents an example to explain the weight setting in Eq. (9). There are two point clouds of the 'Globe' object from different views. p_u are red points on the data point cloud, and q_v are yellow points on the model point cloud. The black line in Figure 2 (a) connects two sets of points with the shortest Euclidean distance. We can see that $K(q_1) = 2$, $K(q_2) = 2$, $K(q_3) = 3$ and the remaining points $K(q_{others}) = 0$. In this case, the spatial corresponding points are concentrated in a small part q_1, q_2, q_3, and their K values are much larger than the rest of them. So ω has a larger value according to Eq. (9). So the color feature has a higher weight in the similarity measure. Figure 2 (b) shows the correspondences established by hybrid features. If two point clouds have been roughly aligned, the spatial correspondence of their points is closer to one-to-one relationship. So the ω is relatively small. The distance mainly depends on the geometric features to find the appropriate correspondence.



(a)

(b)

FIGURE 2. Schematic diagram for calculating color weights. (a) Spatial correspondence established by Euclidean distance between data point cloud (left) and model point cloud (right). (b) The correspondence established by hybrid feature between data point cloud (left) and model point cloud (right).

E. BUILDING MUTUAL CORRESPONDENCE

Usually, the correspondence is built on the nearest neighbor searching from one point to another point set. However, as shown in Figure 2(a), when two point clouds are not roughly aligned, the unidirectional search may cause many points concentrate on a small part of points. But only a small number of feature pairs are correct in such correspondence. To filter out these interference features, we provide the bidirectional mutual correspondence based on the hybrid feature. It means only when two points are mutual nearest neighbors in each set, their correspondence will be believed to remain.

The bidirectional mutual similarity measure is to establish the correspondence. We use $c(u, v)$ to represent the bidirectional correspondence between data point cloud and model point cloud. $c_{DM}(u)$ represents the corresponding point of data point p_u in the model point cloud Q_{SV}. $c_{MD}(v)$ is the opposite. $c_{DM}(u)$ and $c_{MD}(v)$ can be defined as:

$$c_{DM}(u) = \arg\min_{v} d(p_u, q_v), \quad v = 1, \ldots, N_{Q_{SV}}, \tag{10}$$

$$c_{MD}(v) = \arg\min_{u} d(q_v, p_u), \quad u = 1, \ldots, N_{P_{SV}}, \tag{11}$$

where $d(p_u, q_v)$ and $d(q_v, p_u)$ are the measures of hybrid feature distance defined in Eq. (7).

So the condition for point pairs p_u and q_v belonging to the bidirectional mutual correspondence is $c(u, v) = c_{DM}(u) \wedge c_{MD}(v)$, which means they are corresponding to each other in $c_{DM}(u)$ and $c_{MD}(v)$, respectively.

F. REGISTRATION PARAMETER ESTIMATION

After the mutual correspondence $c(u, v)$ has been established, the object of solving the registration problem is simplified to optimize the spatial transformation between corresponding feature pairs. For the corresponding feature pairs p_{u_i} and q_{v_i} in $c(u, v)$, $\vec{l}_{p_{u_i}} = (x, y, z)$ and $\vec{l}_{q_{v_i}} = (x, y, z)$ are their spatial coordinates. Supposed that there is a spatial transformation as:

$$\vec{l}_{p_{u_i}} = \mathbf{R}\vec{l}_{q_{v_i}} + \vec{t}. \tag{12}$$

The optimal rotation and translation transformation between point clouds are:

$$(\mathbf{R}, \vec{t}) = \underset{\mathbf{R}^T\mathbf{R}=I_n, det(\mathbf{R})=1}{\arg\min} \sum_{i=1}^{N_{c(u,v)}} \|\mathbf{R}\vec{l}_{p_{u_i}} + \vec{t} - \vec{l}_{q_{v_i}}\|_2^2. \tag{13}$$

To estimate the transformation, we firstly calculate the centroids of the corresponding point sets P_{SV} and Q_{SV} in Eq. (14). Then, we align two centroids by Eq. (15).

$$\overline{l_p} = \frac{1}{N_{c(u,v)}} \sum_{i=1}^{N_{c(u,v)}} \vec{l}_{p_{u_i}}, \quad \overline{l_q} = \frac{1}{N_{c(u,v)}} \sum_{i=1}^{N_{c(u,v)}} \vec{l}_{q_{v_i}}, \tag{14}$$

$$\vec{l}'_{p_{u_i}} = \vec{l}_{q_{v_i}} - \overline{l_p}, \quad \vec{l}'_{q_{v_i}} = \vec{l}_{q_{v_i}} - \overline{l_q}. \tag{15}$$

where $N_{c(u,v)}$ are the numbers of point pairs in the bidirectional correspondence $c(u, v)$. After the coordinate transformation, an orthonormal transformation matrix of associated points can be constructed by Eq. (16).

$$H = \begin{bmatrix} \sum_{i=1}^{N_{c(u,v)}} x_p x_q, & \sum_{i=1}^{N_{c(u,v)}} x_p y_q, & \sum_{i=1}^{N_{c(u,v)}} x_p z_q \\ \sum_{i=1}^{N_{c(u,v)}} y_p x_q, & \sum_{i=1}^{N_{c(u,v)}} y_p y_q, & \sum_{i=1}^{N_{c(u,v)}} y_p z_q \\ \sum_{i=1}^{N_{c(u,v)}} z_p x_q, & \sum_{i=1}^{N_{c(u,v)}} z_p y_q, & \sum_{i=1}^{N_{c(u,v)}} z_p z_q. \end{bmatrix} \tag{16}$$

Singular value decomposition is performed for the $\mathbf{H}$ matrix to determine the rotation matrix $\mathbf{R}$, as:

$$\mathbf{H} = \mathbf{U}\Lambda\mathbf{V}^T, \tag{17}$$

$$\mathbf{R} = \mathbf{V}\mathbf{U}^T. \tag{18}$$

Then the translation vector $\vec{t}$ is calculated as:

$$\vec{t} = \overline{l_q}^T - \mathbf{R}\overline{l_p}^T. \tag{19}$$

To sum up, the whole registration processes for color point clouds are as follows:

Step 1. Segment the point clouds P and Q to get the supervoxel sets of point clouds.

Step 2. Calculate the center point coordinate and local color moment feature of each supervoxel to obtain sparse feature point clouds P_{SV} and Q_{SV}.

Step 3. Calculate the correspondence between P_{SV} and Q_{SV} with the $(k-1)^{th}$ rigid transformation $(\mathbf{R}_{k-1}, \vec{t}_{k-1})$ to find out the bidirectional reliable correspondence $c_k(u, v)$ by Eqs. (10) and (11).

Step 4. Solve the new rigid transformation $(\mathbf{R}_k, \vec{t}_k)$ according to the current correspondence $c_k(u, v)$ by Eqs. (18) and (19).

Repeat steps 3-4 until the registration error is small enough or reaches the maximum number of iterations.

TABLE 1. Description of TUM data used in the experiment.

File name	Fragment duration (seconds)	# of Extracted frames	Average # of points per point clouds
freiburg1_desk	14.8	24	229579
freiburg2_desk	25	24	179332
freiburg3_office	6.6	24	233067
freiburg1_teddy	5.7	24	235533

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the algorithm in dealing with color point cloud registration with partial overlapping and poor initial positions. Two different indoor datasets are used for testing. One is the TUM dataset [35]. The other is the Indoor Lidar-RGBD Scan Dataset [28]. Both datasets are captured in indoors by RGBD acquisition devices, with some continuous synchronized RGB image sequences and depth image sequences. Then, the testing data is generated by mapping the depth image to the corresponding RGB image based on the known camera parameters. In the TUM dataset, four sequences are selected from different indoor scenes. We use continuous data for a period of time in each scene, and uniformly select 24 sets of point clouds as experimental data in each sequence. Since the movement rates of devices are different, these point clouds are sampled by different time lengths to ensure enough identical contents are contained. And the source and statistics data of the TUM experiment data are shown in Table 1. In the Indoor Lidar-RGBD Scan Dataset, we randomly select adjacent point clouds from different scenarios. Some original point cloud data used in the experiment is shown in Figure 4(a).

The algorithms for comparing include the traditional ICP algorithm [26], the ICP-4D method [24] which uses the Hue value in HSL color space to assist registration, the improved trimmed ICP algorithm (ITrICP) [13], our previous method ICP-CM in the conference version [9], which used color moment features based on ITrICP, and the state-of-the-art CPCRR color registration algorithm [28]. Same as the proposed method, these methods are all based on the point feature. The ICP and ITrICP use the spatial feature. And other methods are based on the spatial and color features. In our experiments, we set 0.1 meter as the scale of supervoxel for indoor scences empirically.

The testing hardware environment is: Intel Core i5-6400 2.7GHz CPU and 8GB RAM. The running software for the

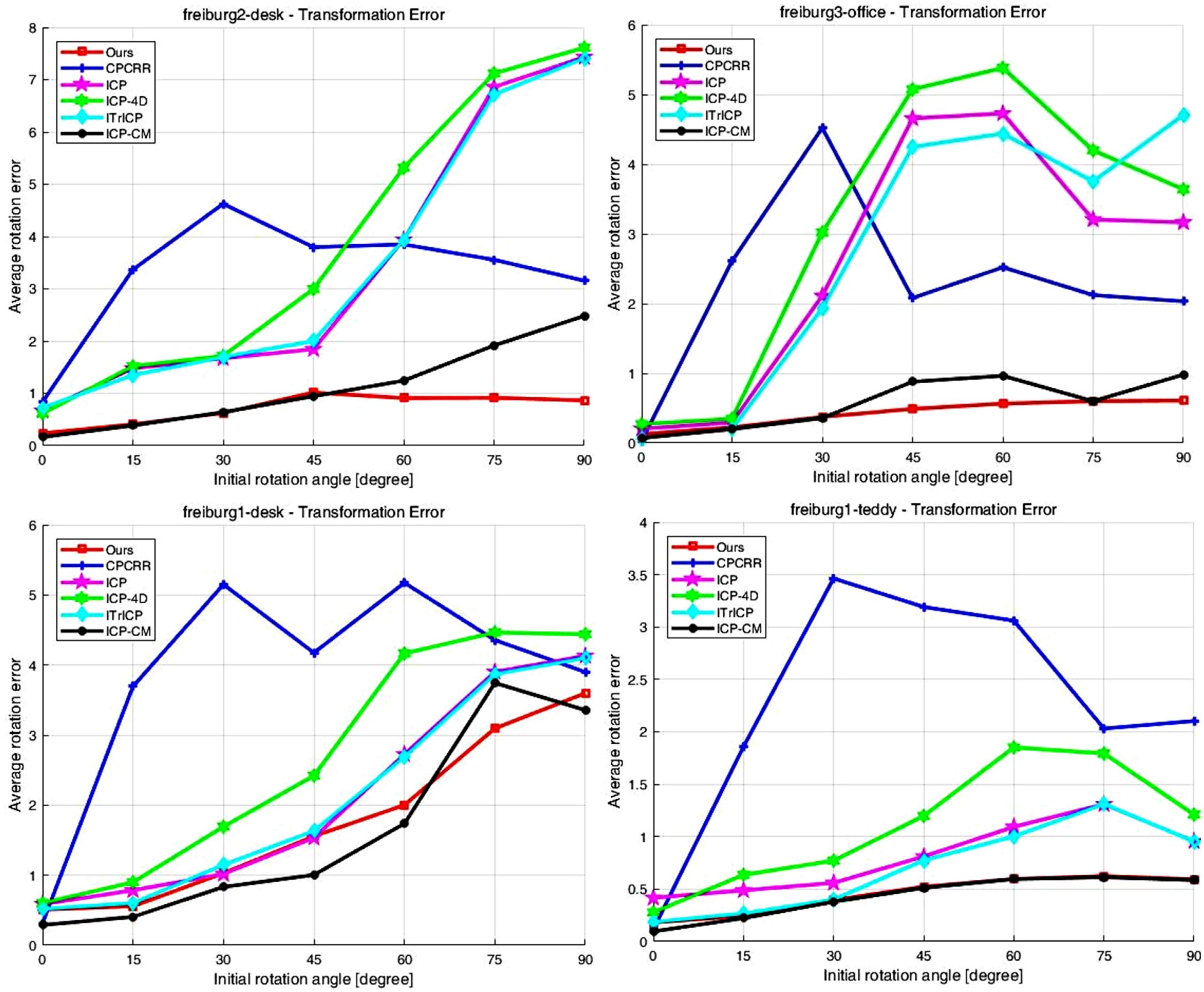


FIGURE 3. Average transformation error of experiment on TUM dataset. The title of each figure is the data name used in this group of experiments. And each group of experiments used 24 consecutive point cloud frames and performed 23 times registration experiments on all adjacent frames.

ICP, ICP-4D, ITrICP, and ICP-CM is Matlab2017A, and the environment of CPCRR is Python3.5. The running environment of our algorithm for point cloud segmentation is C++, and the registration process uses Matlab2017A.

A. ALGORITHM ROBUSTNESS AND EFFECTIVENESS TESTING

In order to evaluate the effectiveness of the algorithm, we set different initial transformation angles to simulate the situation when handheld device moves. Some specific rotation matrix in 3D space are conducted on one of two point clouds. And the above registration algorithms are performed to align them from these relative initial positions.

In the first dataset, we perform the registration experiment on all adjacent point clouds in each group. The effectiveness is evaluated by comparing with the ground truth, which is provided by the TUM dataset. And the registed error is defined as:

$$\varepsilon_{\mathbf{R}} = \|\mathbf{R} - \mathbf{R}_G\|^2, \tag{20}$$

where $\mathbf{R}$ is the rotation angle of the registration result. And $\mathbf{R}_G$ is the ground truth.

These registration results are shown in Figure 3. Here, seven sets of experiments are tested with different initial rotation angles from 0 to 90 degrees. Original point clouds can be roughly registered when the angle is zero. When the initial transformation angle increases, two point clouds become more difficult to regist. In Figure 3, we can see that as the initial angle increases, the registration errors for most algorithms increases significantly. But our method presents more accurate transformation results than other compared algorithms. The reason for the big errors in ICP and ITrICP algorithms is that: they both use the unidirectional Euclidean distance to find the corresponding points. When the initial positions between two point clouds are far, the Euclidean distance can only find the point pairs with closest positions between two point clouds. These correspondences are not correct and lead to misaligning of the final result. In the ICP-4D algorithm, the spatial coordinates and the Hue color

TABLE 2. Performance comparison of our algorithm with different strategies.

File name	Unidirectional strategies & No color features		Unidirectional strategies & Using color features		Bidirectional strategies & No color features		Bidirectional strategies & Using color features	
	$\mathbf{R}$ Error	$\vec{t}$ Error	$\mathbf{R}$ Error	$\vec{t}$ Error	$\mathbf{R}$ Error	$\vec{t}$ Error	$\mathbf{R}$ Error	$\vec{t}$ Error
freiburg1_desk	1.9463	0.6339	1.3458	0.6080	2.8179	0.5554	**1.2165**	**0.5473**
freiburg2_desk	0.6222	0.4549	2.7149	1.1527	3.2496	0.8152	**0.4111**	**0.2997**
freiburg3_office	2.2476	1.3447	0.4349	0.3536	2.9031	0.8152	**0.1940**	**0.2334**
freiburg1_teddy	1.0034	0.5550	0.3551	0.1343	2.5187	0.5316	**0.3333**	**0.1182**

TABLE 3. Algorithm average running time on the TUM dataset.

Algorithm	Average amount of experiment data points	Average running time (seconds) at different initial rotation angles (degrees)						
		0	15	30	45	60	75	90
ICP	223978 (Original data)	141.88	227.45	293.73	365.79	421.69	473.28	439.95
ICP-4D	11101 (Downsampled data)	38.08	52.30	58.71	57.60	54.87	57.63	57.47
ITrICP	223978 (Original data)	344.48	439.58	583.43	719.19	855.81	973.96	955.31
ITrICP-CM	11101 (Downsampled data)	131.05	139.59	125.03	121.58	119.62	118.01	117.31
Our (segmentation)	223978 (Original data)	105.25	105.25	105.25	105.25	105.25	105.25	105.25
Our (registration)	847 (segmented data)	13.23	5.57	18.69	20.59	20.17	19.77	19.17
Our (total)	223978 (Original data)	118.48	110.82	123.94	125.84	125.42	125.02	124.42

values of points are combined to find the corresponding point pairs. But the identification of the point information is still insufficient to find the accurate correspondence when the initial position is poor. The CPCRR algorithm defines the continuous color function and depth function for point clouds as the objective. But its optimization result is easily trapped into a local minimum. Therefore, the registration result of the CPCRR method is greatly disturbed by the initial position of the point clouds. We can see that the results by ICP-CM are close to the algorithm proposed in this paper. In fact, the ICP-CM can be seen as a special case of the proposed method with the smallest data cell in segmentation. But we will compare the improved efficiency by our algorithm in the next section. For color point cloud data, two kinds of features for these methods are based on color and geometric information. In that case, when the color is relatively homogenize, these performances become similar as other geometric based method. If the geometric structure of the data is homogenize too. All methods will have big errors, such as the ''freiburg1-desk'' shown in Figure 3.

In the second dataset, we set different initial transformation angles in the same way as before. Figure 4 (a) shows the relative position of two point clouds before registration. In this experiment, we mainly compare the registration results with the classic registration method ICP and the state-of-the-art color registration method CPCRR. As we can see from the Figure 4, our registration results are superior to other methods. This is mainly due to the color moments features can build accurate corresponding point pairs when the initial position is not good. Thus it leads the algorithm to reach to the optimal result. Since the capture device is moved, two point clouds for registration are partial overlapping. But our method can register them correctly. This experiment also verify that our algorithm can be applied in various indoor scenes.

Furthermore, in order to verify the effectiveness of our proposed hybrid features and bidirectional measurement, we also compared our algorithms under different strategies. Four different combinations of strategies are compared in the experiments, including unidirectional and bidirectional correspondences with and without color features. As shown in Table 2, we test different strategies on the TUM dataset with different initial transformation angles. By comparing the rotation and translation errors, we can conclude that the hybrid features with color information have better distinctiveness for corresponding points. And the measurement of bidirectional mutual correspondence can obtain more precise correspondence between point clouds. But comparing with the unidirectional search, the processing will cost more time. So we apple the downsampling based on supervoxel segmentation to improve the efficiency.

(a) (b) (c) (d)

FIGURE 4. Comparison of registration results using Indoor Lidar-RGBD Scan Dataset. (a) Original point cloud data. (b) Registration result of ICP algorithm. (c) Registration result of CPCRR algorithm. (d) Registration result of our algorithm.

B. ALGORITHM EFFICIENCY TESTING

The ICP-based algorithm has the common drawback in the computational speed: the computational time-consuming increases as the amount of data increases. The main reason is that the ICP algorithm needs to recalculate the correspondence in each iteration. When the number of points is tremendous, it is very time consuming to find the correspondence of each point. Our previously ICP-CM algorithm improves the accuracy of the registration result by color moment features, but it is still need to calculate the correspondence of all points in each iteration.

The statistical time about the algorithm efficiency analysis is shown in Table 3. We can see the efficiency of the proposed method. In order to solve the efficiency problem, our core solution is to reduce the amount of data to be processed during the iteration. We divide the proposed algorithm into two stages: the first stage is the supervoxel segmentation of point cloud, and the second stage is the feature extraction and registration process by sparse point features. Since the extraction of sparse point cloud is completed in the first stage, the proposed algorithm reduces the amount of data to be processed in the registration process. The other advantage is that supervoxel segmentation can be used as part of offline data processing because it does not rely on any registration algorithm. In that case, the point cloud segmentation and feature extraction do not need repeated computation in the algorithm, which reduces the computational cost in the iterative process.

V. CONCLUSION

This paper presents a novel registration method for color point clouds which extracts sparse feature points by supervoxel segmentation to reduce the burden of large scale registration task. A similarity measure is proposed for feature matching by dynamically combining spatial information and color information. Furthermore, we construct the objective function with hybrid features under a mutual bidirectional correspondence strategy. Experiments demonstrate that our method has better performance in dealing with the problem of partical overlapping data and poor initial position in real-world scenes.

REFERENCES

[1] S. Choi, Q.-Y. Zhou, and V. Koltun, "Robust reconstruction of indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5556–5565.

[2] Q. Wang, P. Yang, L. Huang, Q. Wei, and Y. Liang, "An adaptive slicing thickness adjustment method based on cloud point in 3D printing," in *Proc. 13th Int. Conf. Embedded Softw. Syst. (ICESS)*, Aug. 2016, pp. 108–1137.

[3] J. Barabas, B. Babusiak, M. Gala, R. Radil, and M. Capka, "Analysis 3D reconstruction and anatomical feature extraction from medical images," in *Proc. Int. Conf. Biomed. Eng. Biotechnol.*, 2012, pp. 731–735.

[4] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 998–1005.

[5] P. Besl and N. D. Mckay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[6] D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek, "The trimmed iterative closest point algorithm," in *Proc. Object Recognit. Supported User Interact. Service Robots*, vol. 3, Jun. 2003, pp. 545–548.

[7] J. M. Phillips, L. Ran, and C. Tomasi, "Outlier robust ICP for minimizing fractional RMSD," in *Proc. Int. Conf. 3D Dig. Imag. Mod.*, 2007, pp. 427–434.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[9] W. Chen, Y. Yang, D. Fan, Y. Yang, Z. Chen, and Q. Kou, "Registration of color point cloud by combining with color moments information," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2018, pp. 2098–2104.

[10] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 3, Apr. 1991, pp. 2724–2729.

[11] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. Int. Conf. 3-D Digit. Imag. Modeling*, May/Jun. 2001, pp. 145–152.

[12] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," *Robot., Sci. Syst.*, vol. 2, no. 4, p. 435, 2009.

[13] S. Du, "Robust iterative closest point algorithm for registration of point sets with outliers," *Opt. Eng.*, vol. 50, no. 8, Aug. 2011, Art. no. 087001.

[14] P. Bariya, J. Novatnack, G. Schwartz, and K. Nishino, "3D geometric scale variability in range images: features and descriptors," *Int. J. Comput. Vis.*, vol. 99, no. 2, pp. 232–255, Sep. 2012.

[15] M. T. Ahmed, J. A. Marshall, and M. Greenspan, "Point cloud registration with virtual interest points from implicit quadric surface intersections," in *Proc. Int. Conf. 3D Vis.*, 2017, pp. 649–657.

[16] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, "Learning informative point classes for the acquisition of object model maps," in *Proc. Int. Conf. Control, Automat., Robot. Vis.*, 2008, pp. 643–650.

[17] R. B. Rusu, N. Blodow, Z. C. Marton, and M. Beetz, "Aligning point cloud views using persistent feature histograms," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3384–3391.

[18] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2009, pp. 3212–3217.

[19] J. Xie, M. Wang, and Y. Fang, "Learned binary spectral shape descriptor for 3D shape correspondence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 3309–3317.

[20] H. Deng, T. Birdal, and S. Ilic, "PPFNet: Global context aware local features of rotation invariant 3D point matching," in *Proc. IEEE Conf. Comput. Vis. Pattern. Recognit.*, Jun. 2018, pp. 195–205.

[21] H. Deng, T. Birdal, and S. Ilic, "PPF-FoldNet: Unsupervised learning of rotation invariant 3D local descriptors," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 620–638.

[22] Y. Wang and J. Solomon, "Deep closest point: Learning representation for point cloud registration," in *Proc. IEEE Int. Conf. Comput. Vis.*, May 2019. [Online]. Available: https://arxiv.org/abs/1905.03304v1

[23] A. E. Johnson and S. B. Kang, "Registration and integration of textured 3D data," *Image Vis. Comput.*, vol. 17, no. 2, pp. 135–147, Feb. 1999.

[24] H. Men, B. Gebre, and K. Pochiraju, "Color point cloud registration with 4D ICP algorithm," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2011, pp. 1511–1516.

[25] M. Korn, M. Holzkothen, and J. Pauli, "Color supported generalized-ICP," in *Proc. Int. Conf. Comput. Vis. Theory Appl.*, 2014, pp. 592–599.

[26] G. Godin, D. Laurendeau, and R. Bergevin, "A method for the registration of attributed range images," in *Proc. Int. Conf. 3-D Digit. Imag. Modeling*, 2001, pp. 179–186.

[27] M. Danelljan, G. Meneghetti, F. S. Khan, and M. Felsberg, "A probabilistic framework for color-based point set registration," in *Proc. IEEE Conf. Comput. Vis. Pattern. Recognit.*, Jun. 2016, pp. 1818–1826.

[28] J. Park, Q. Y. Zhou, and V. Koltun, "Colored point cloud registration revisited," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 143–152.

[29] 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 *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 3082–3088.

[30] J. Chu and C. Nie, "Multi-view point clouds registration and stitching based on SIFT feature," in *Proc. 3rd Int. Conf. Comput. Res. Develop.*, 2011, pp. 274–278.

[31] Z. Zheng, Y. Li, and W. Jun, "LiDAR point cloud registration based on improved ICP method and SIFT feature," in *Proc. IEEE Int. Conf. Prog. Inform. Comput.*, Dec. 2015, pp. 588–592.

[32] Z. Ji, F. Zhou, X. Tian, R. Jiang, and Y. Chen, "Probabilistic 3D ICP algorithm based on ORB feature," in *Proc. IEEE 3rd Int. Conf. Inf. Sci. Technol.*, Mar. 2013, pp. 300–304.

[33] J. Papon, A. Abramov, M. Schoeler, and F. Worgotteer, "Voxel cloud connectivity segmentation—Supervoxels for point clouds," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2027–2034.
[34] A. Barla, F. Odone, and A. Verri, "Histogram intersection kernel for image classification," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2003, p. III-513.
[35] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 573–580.

YANG YANG received the B.E. degree in information engineering, in 2005, and the Double Ph.D. degree in pattern recognition and intelligent system from Xi'an Jiaotong University, China, and the Systems Innovation Engineering from Tokushima University, Japan, in 2011. She is currently an Associate Professor with the School of Electronic and Information Engineering, Xi'an Jiaotong University. Her research interests include image processing, multimedia, and machine learning.

WEILE CHEN received the B.E. degree in software engineering from Shanxi University, China, in 2016. He is currently pursuing the master's degree in software engineering with the School of Software Engineering, Xi'an Jiaotong University, China. His research interests include image processing and template matching.

MUYI WANG received the B.E. degree in electrical engineering and automation from the Changchun Institute of Technology, China, in 2006, and the M.E. degree in power electronics and power drives from Northwestern Polytechnical University, China, in 2009. She is currently pursuing the Ph.D. degree in control science and engineering with Xi'an Jiaotong University, China. Her current research interests include machine vision, pattern recognition, and big data processing.

DEXING ZHONG received the B.E. and Ph.D. degrees from Xi'an Jiaotong University, in 2005 and 2010, respectively. He was a Visiting Scholar with the University of Illinois at Urbana–Champaign, USA. He is currently an Associate Professor with the School of Electronic and Information Engineering, Xi'an Jiaotong University, China, where he is also a Researcher with the Shenzhen Research School. His main research interests include biometrics and computer vision.

SHAOYI DU received the Double bachelor's degree in computational mathematics and in computer science, in 2002, the M.S. degree in applied mathematics, in 2005, and the Ph.D. degree in pattern recognition and intelligence system from Xi'an Jiaotong University, China, in 2009. He is currently a Professor with Xi'an Jiaotong University. His research interests include computer vision, machine learning, and pattern recognition.