

Guided Sampling for Multistructure Data via Neighborhood Consensus and Residual Sorting

Taotao Lai^{ID}, Yizhang Liu^{ID}, Jie Chang, Lifang Wei^{ID}, Zuoyong Li^{ID},
and Hamido Fujita, *Life Senior Member, IEEE*

Abstract—Robust model fitting is a critical technique for artificial intelligence. The performance of most robust model fitting techniques heavily depends on the use of sampling algorithms. In this paper, we propose an efficient guided sampling algorithm for multi-structure data by using the neighborhood consensus and the residual sorting. Specifically, a Neighborhood Consensus based Strategy (NCS) is first proposed to select the first datum (i.e., seed datum) of a minimal subset, and then a Residual Sorting based Strategy (RSS) samples the rest data of the minimal subset based on the seed datum. This strategy effectively combines the benefits of neighborhood consensus and residual sorting, where neighborhood consensus can judge whether a selected data point is an inlier, and residual sorting encourages this strategy to select data points from the same structure of the first selected data point. Moreover, to achieve better fitting performance, the Markov Chain Monte Carlo process is used to combine NCS with the random selection strategy to select the seed datum, and an appropriate size is set to the initial block of randomly sampled hypotheses for RSS. Experimental results on three vision tasks (e.g., two-view motion segmentation and 3D motion segmentation) demonstrate that the proposed algorithm achieves superior performance to several state-of-the-art sampling algorithms.

Manuscript received 19 July 2022; revised 28 October 2022 and 5 December 2022; accepted 17 December 2022. Date of publication 19 December 2022; date of current version 3 July 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62172197, Grant 61972187, and Grant 62171130; in part by the Natural Science Foundation of Fujian Province and Anhui Province under Grant 2020J01825, Grant 2020J02024, and Grant 2108085MF205; in part by the Fuzhou Science and Technology Project under Grant 2022-R-001; and in part by the National Key Research and Development Program of China under Grant 2022YFC3302200. This article was recommended by Associate Editor V. Stankovic. (*Corresponding authors:* Lifang Wei; Zuoyong Li.)

Taotao Lai and Zuoyong Li are with the Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, College of Computer and Control Engineering, Minjiang University, Fuzhou 350121, China (e-mail: laitaotao@gmail.com; fzulzytdq@126.com).

Yizhang Liu is with the School of Software Engineering, Tongji University, Shanghai 200092, China (e-mail: lyz8023lyp@gmail.com).

Jie Chang is with the Department of Medical Information, Wannan Medical College, Wuhu 240001, China (e-mail: cfuture@mail.ustc.edu.cn).

Lifang Wei is with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China (e-mail: weilifang1981@163.com).

Hamido Fujita is with the Faculty of Information Technology, HUTECH University, Ho Chi Minh City 700000, Vietnam, also with the Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada, Granada 18182, Spain, and also with the Regional Research Center, Iwate Prefectural University, Takizawa, Iwate 020-0193, Japan (e-mail: hfujita@hutech.edu.vn; HFujita-799@acm.org).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCSVT.2022.3230736>.

Digital Object Identifier 10.1109/TCSVT.2022.3230736

Index Terms—Robust model fitting, guided sampling, neighborhood consensus, residual sorting, multi-structure data.

I. INTRODUCTION

ROBUST model fitting is a fundamental task in artificial intelligence, and has been employed in various artificial intelligence applications, such as predictive transform video coding [1], visual augmented reality [2], 3D rigid registration [3], facial landmark detection [4] and plane detection [5]. The aim of robust model fitting is to fit geometric model hypotheses among data with outliers and noise extracted from input images, and further recover meaningful structures, such as shapes of objects and moving objects.

The successes of most robust model fitting techniques such as [6], [7], [8], [9], and [10] are depended on accurate hypotheses generated by their sampling algorithms. Most sampling algorithms try their best to sample clean minimal subsets to generate more accurate hypotheses. A clean minimal subset contains the minimum number of inliers from the same model instance (i.e., structure) for estimating the parameters of a geometric model. There is a uniform random sampling algorithm which is still used by many robust model fitting methods such as [7] and [11] and, also, there are guided sampling algorithms. However, as the dimension of the geometric model increases, the probability of sampling a clean minimal subset by using a uniform random sampling algorithm decreases exponentially.

To reduce the disadvantage of the uniform random sampling algorithm, many guided sampling algorithms have been proposed. In the uniform random sampling algorithm, each data point in input data is assigned an equal sampling probability value. In contrast, in the guided sampling algorithms, each data point is given a distinct sampling probability value by using various auxiliary information (e.g., spatial proximity information and residual sorting information). The guided sampling algorithms can be roughly grouped into four types: matching score-based, spatial proximity-based, greedy search-based and residual sorting-based algorithms. The residual sorting-based sampling algorithm is one of the most promising guided sampling algorithms, because this kind of algorithms innately encode the probability of two inliers belonging to the same structure [12].

However, computing sampling probabilities from residual sorting is relatively expensive, and most existing residual

sorting-based sampling algorithms (e.g., [12], [13], [14]) require computing sampling probabilities of all input data ($p - 1$) times to sample a minimal subset including p data points. Recently, we propose an improved algorithm (i.e., AGS [15]), which only computes one time sampling probabilities for sampling a minimal subset. Specifically, AGS first randomly chooses a data point from input data as most existing residual sorting-based sampling algorithms (e.g., [12], [13]) do, and then selects significant data by using both information theory principles and sampling probabilities computed from residual sorting, and finally samples minimal subsets from the chosen significant data. Nevertheless, if the first datum chosen by these sampling algorithms is an outlier, then the sampled data subset will be invalid [14]. Meanwhile, the higher outlier rate the data contains, the more likely an outlier the first randomly chosen datum will be. Thus, these algorithms cannot effectively handle multi-structure data with a high outlier rate. To reduce the drawback, a mode seeking technique is used in [14] to select the first datum, but it is computationally expensive.

To overcome the above drawbacks, this paper proposes an efficient sampling algorithm, called Neighborhood Consensus and Residual Sorting (NCRS). The first aim of the proposed NCRS is to efficiently select inliers for the first datum (i.e., seed datum) of a minimal subset. Different from ModeSamp using a computationally expensive mode seeking technique, a neighborhood consensus strategy is proposed to judge whether a selected data point is an inlier, by computing the neighborhood consensus information of the selected data point. If the selected data point is an inlier, the data point will be selected as the seed datum. Otherwise, the neighborhood consensus strategy is repeatedly performed until an inlier is selected or the maximum repetitions are reached. If no single inlier is found after the iterative conduct of the neighborhood consensus strategy, a data point is randomly selected as the seed datum. In the Section III-C, we will show that the probability of no single inlier found is relatively low.

Moreover, intuitively, sampling more clean minimal subsets by using the neighborhood consensus strategy may lead to better performance. However, this intuition is not true for the residual sorting-based sampling algorithms. The detailed analyses will be given in Section IV-B. Thus, the proposed NCRS combines the neighborhood consensus strategy with the uniform random sampling strategy by using the Markov Chain Monte Carlo (MCMC) process for further improving the performance of the proposed algorithm, where the uniform random sampling strategy selects the seed datum randomly. Furthermore, in Sections II and III-D, we will discuss the influence of the initial block on the performance of the proposed sampling algorithm. We will set an appropriate size to the initial block for better performance.

The main contributions of this paper are summarized as follows. First, by using the neighborhood consensus information, a novel sampling strategy is proposed to select inliers as the seed data for efficiently sampling clean data subsets, i.e., the proposed sampling strategy effectively combines the benefits of residual sorting and neighborhood consensus information for data sampling. Second, the MCMC process is used to

combine the proposed neighborhood consensus strategy with the uniform random sampling strategy for better performance. Moreover, we set a relatively large size to the initial block of randomly sampled hypotheses to further improve the performance of the proposed sampling algorithm.

The rest of this article is organized as follows. In Section II, we review the related work. In Section III, we present the proposed sampling algorithm. In Section IV, we show experimental results on three computer vision applications by using the proposed sampling algorithm. Finally, we give conclusions in Section V.

II. RELATED WORK

In this paper, we mainly study data sampling. Thus, we only review guided sampling algorithms as follows.

A. Matching Score-Based Algorithms

This group sampling algorithms such as [5], [10], and [23] assume that the data points with larger matching score values are more likely to be inliers. These algorithms can effectively sample clean minimal subsets for data with single structure, while for multiple structure data, these algorithms are prone to sample minimal subsets, whose data points are from different structures [14]. The reason is that inliers from different structures are with larger matching score values. Thus, these algorithms are difficult to sample clean minimal subsets for multi-structure data.

B. Spatial Proximity-Based Algorithms

This group sampling algorithms such as [19], [20], [21], and [22] assume that the neighbors of inliers are likely to be inliers. These algorithms sample minimal subsets from neighbour data points. However, the assumption is not true when the data has high outlier ratio [12]. In this case, the neighbors of inliers are more likely to be outliers. Thus, these algorithms are difficult to sample clean minimal subsets for data with a high outlier ratio.

C. Greedy Search-Based Algorithms

This group sampling algorithms such as [10], [5], and [23] use a greedy search for fast generating promising hypotheses. However, these algorithms have a key input parameter, i.e., the minimal acceptable size of a structure. If the given minimal acceptable size is close to the true value of the structure, these sampling algorithms may fast generate accurate hypotheses, and vice versa. The authors of [5] and [23] need to give this parameter for each structure manually. The method in [10] tries to give the same value to all structures. Nevertheless, in practice, the true structure sizes in data are unknown in advance. This leads to difficultly specifying the parameter.

D. Residual Sorting-Based Algorithms

The authors of [12] find that the probabilities computed from residual sorting can be effectively used to judge whether two data points are from the same structure. Because Multi-GS [12] can relatively fast sample clean minimal subsets,

several improved versions of Multi-GS have been proposed. Multi-GS-Offset [24] tries to sample minimal subsets with large spans to generate more accurate hypotheses. We find that, to sample a minimal subset containing p data points, the two methods need to compute sampling probabilities for all input data ($p-1$) times. To reduce the computational cost, our previous work [13] uses local constraints to only compute sampling probabilities of part of rather than all input data. However, RHG [13] still need to compute sampling probabilities ($p-1$) times for sampling a minimal subset. To further reduce the computational cost, RCMSA [9] first uses a random cluster model to cluster input data by using probabilities computed from residual sorting, then generates a hypothesis by using the data points of one of clusters. In our another previous work [15], information theory principles are used to choose significant data, and minimal subsets are sampled from the chosen significant data. AGS [15] only computes one time sampling probabilities for sampling a minimal subset.

The sampling algorithms (i.e., Multi-GS, Multi-GS-Offset, RHG and AGS) select the first datum (i.e., seed datum) of a data subset randomly. If the seed datum is an outlier, then the sampled data subset will be invalid. In contrast, RCMSA assumes that inliers from the same structure are prone to stay together, which may be not true in practice [14]. To reduce the drawback of the above methods, ModeSamp [14] uses information obtained from the generated hypotheses by the mode seeking technique to select the seed datum. However, the mode seeking step becomes the most computationally expensive part as the experimental results in [14] show. Thus, this reduces its practicality.

In addition, all the residual sorting-based sampling algorithms need to sample an initial block hypotheses by using the uniform random sampling algorithm to guide their residual sorting-based sampling strategy. In the whole sampling process, these algorithms re-sort their residuals for computing sampling probabilities when each new block including the initial block of hypotheses is sampled. All these algorithms set the size of the initial block the same as that of the block for simplicity. However, this is suboptimal.

III. THE PROPOSED SAMPLING ALGORITHM

The proposed sampling algorithm is based on our previous proposed AGS sampling algorithm [15]. In this section, we first review AGS [15] in Section III-A. Then, we introduce the neighborhood consensus in Section III-B. After that, we present the proposed NCRS sampling strategy in Section III-C. Finally, we describe the complete sampling algorithm in Section III-D.

A. Our Previous Proposed AGS

Suppose that the input data X has N data points and a set of M model hypotheses Θ has been generated from the input data, where $X = \{x_1, x_2, \dots, x_N\}$ and $\Theta = \{\theta_1, \theta_2, \dots, \theta_M\}$. Let the residual between the i th data point x_i and the j th hypothesis denote as r_i^j , then the residual vector r_i between x_i and Θ can be denoted as

$$r_i = [r_i^1, r_i^2, \dots, r_i^M]. \quad (1)$$

Sorting the elements in the residual vector r_i in a non-descending order can obtain the residual index vector

$$\kappa_i = [\kappa_i^1, \kappa_i^2, \dots, \kappa_i^M], \quad (2)$$

where κ_i ranks the preference of the data point x_i to the M hypotheses, i.e., the higher a hypothesis ranks, the more likely that x_i will be an inlier of the hypothesis.

The authors of [12] find that two inliers from the same structure share more common residual indices of the top of their residual vectors. If x_i is selected, the inlier probability of x_j is computed as

$$f_{i,j} = \frac{1}{h} |\kappa_i^{1:h} \cap \kappa_j^{1:h}|, \quad (3)$$

where the vector $\kappa_i^{1:h}$ is the first h elements of κ_i , and $|\kappa_i^{1:h} \cap \kappa_j^{1:h}|$ is the number of common elements between κ_i and κ_j . h is a window size, and $1 \leq h \leq M$. For the input data, our previous proposed AGS [15] samples a minimal subset by the following four steps: it first computes inlier probability vector by Eq. (3). By using information theory principles, AGS then selects the significant data based on the inlier probability vector. After that, AGS [15] computes sampling probabilities by using both inlier probabilities and matching scores for the significant data. Finally, AGS samples a minimal subset from the significant data by using the computed sampling probabilities.

In addition, in most residual sorting-based algorithms, residual sorting is time-consuming because these algorithms re-sort all the residuals after generating each new block hypotheses. We find that the computation of the inlier probability only uses the first h rather than all elements of residual vectors. Thus, AGS [15] only sorts part of the top ranked elements of residual vectors for improving the performance of the residual sorting step.

B. The Neighborhood Consensus

Although AGS [15] can relatively efficiently generate promising hypotheses, it still selects the seed data for minimal subsets randomly. However, for the input data with high outlier rate, AGS tends to sample non-clean minimal subsets. In this section, we show that how to use neighborhood consensus information for efficiently selecting inliers as seed data of minimal subsets.

In the use of spatial information (i.e., neighborhood consensus information uses spatial information of two images), the differences between the proposed method and the existing methods (e.g., [21], [22], [25]) are as follows. First, different from [21], [22], and [25] directly using spatial information to sample data subsets, the proposed method uses spatial information only to determine whether a selected data point is an inlier. In other words, spatial information is only used to select the first datum of a data subset, and the rest data of the data subset are selected by using residual sorting. In summary, the main aim of using spatial information is to select an inlier as the first datum of a minimal subset to improve the performance of the residual sorting-based sampling algorithm. Second, to compute the neighbors of data points, the authors

of [21] and [25] used a Fast Approximated Nearest Neighbors technique [26]. Torr et al. [22] only used one of two views. The proposed method computes neighbors of a data point in two views respectively as in [22], and then computes an inlier probability by using Eq. (5) and the neighbors of a data point.

We further denote a match as $\mathbf{x}_i = \{\mathbf{u}_i, \mathbf{v}_i\}$, where multiple view images can be regarded as multiple two-view images. Feature points \mathbf{u}_i and \mathbf{v}_i are extracted from the first and the second images, respectively. The neighborhood consensus means that if \mathbf{u}_i and \mathbf{v}_i share more common neighborhoods of the top of their sorted neighborhood index vectors, then \mathbf{x}_i is more likely to be an inlier [27], where the sorted neighborhood index vectors are obtained by sorting the elements in the neighborhood vectors in a non-descending order as Eq. (2). The probability of \mathbf{x}_i to be an inlier is computed as follows.

First, we compute the first K nearest neighbors for \mathbf{u}_i and \mathbf{v}_i by using Euclidean distance, denoted as $\mathcal{N}_{\mathbf{u}_i}^K$ and $\mathcal{N}_{\mathbf{v}_i}^K$, respectively. K is a given threshold. Then the inlier probability of \mathbf{x}_i is computed as

$$\hat{p}_i^K = \frac{1}{K} |\mathcal{N}_{\mathbf{u}_i}^K \cap \mathcal{N}_{\mathbf{v}_i}^K|. \quad (4)$$

Different from Eq. (3), which is used to compute the inlier probability of \mathbf{x}_j related to the selected \mathbf{x}_i based on the generated hypotheses, Eq. (4) is used to compute the inlier probability of \mathbf{x}_i by using neighbors of its two feature points (i.e., \mathbf{u}_i and \mathbf{v}_i).

As in [27], because the distribution of the putative matches is not uniform, multi-scale neighborhood consensus information is used to compute inlier probability for better performance as

$$p_i = \frac{1}{\alpha} \sum_{j=1}^{\alpha} \hat{p}_i^{K_j}, \quad (5)$$

where $\mathcal{K} = \{K_j\}_{j=1}^{\alpha}$ is the multiple scales. If the inlier probability p_i of \mathbf{x}_i is larger than a threshold, then \mathbf{x}_i is regarded as an inlier. Otherwise, \mathbf{x}_i is regarded as an outlier.

We note that the neighborhood consensus information has been widely used in feature matching such as [27], [28], and [29]. Feature matching aims to cluster the input data into two clusters (i.e., inlier and outlier clusters), whereas robust model fitting aims to cluster the input data to multiple clusters (i.e., multiple model instances and one outlier) by estimating geometric models. Thus, robust model fitting is a more challenging task than feature matching.

We also note that the neighborhood consensus information has been used in recently proposed robust model fitting technique MCF [30]. MCF not only uses the neighborhood consensus information to select the seed datum, but also uses it to search the neighbors of the seed datum and then samples the rest data of the minimal subset from the neighbors. However, MCF may sample inliers of two model instances from a minimal subset resulting in poor performance, when inliers of two model instances are close, even in contact. To overcome the drawback, the proposed algorithm samples the rest data of a minimal subset by using sampling weights computed from residual sorting.

C. The Proposed NCRS Sampling Strategy

The proposed NCRS sampling strategy mainly consists of two steps: select the seed datum and sample the minimal subset by using AGS [15] based on the seed datum. The detailed descriptions are shown in Lines 6 to 18 of Algorithm 1 (i.e., the proposed sampling algorithm). To increase the probability of sampling the clean minimal subset, we introduce the neighborhood consensus information to select the inlier as the seed datum. Intuitively, the more clean subsets sampled by the proposed NCRS sampling strategy, the better performance of the proposed sampling algorithm will be. However, this intuition is not true. The corresponding experimental results and analyses are shown in Section IV-B. To achieve better fitting performance, NCRS uses the MCMC process to combine the neighborhood consensus information with the random selection to select the seed datum.

Algorithm 1 The Proposed Sampling Algorithm

Input: input data X , number of samples M , number of initial hypotheses L , parameter β for MCMC, maximum number of selecting seed datum G and batch size b .

Output: the generated hypothesis set Θ .

```

1: Let  $\Theta = null$ .
2: for  $\tau = 1$  to  $M$  do
3:   if  $\tau \leq L$  then
4:     Sample a minimal subset  $S_\tau$  by the uniform random sampling algorithm.
5:   else
6:     if  $\text{rand}(0, 1) < \beta$  then
7:       for  $i = 1$  to  $G$  do
8:         Select a data point  $\mathbf{x}_{s_i}$  from input data randomly.
9:         Determine whether  $\mathbf{x}_{s_i}$  is an inlier by Eq. (5).
10:        if  $\mathbf{x}_{s_i}$  is an inlier then
11:          Let  $\mathbf{x}_{s_i}$  be the seed datum and break.
12:        end if
13:      end for
14:    end if
15:    if the seed datum is not found then
16:      Select a data point as the seed datum randomly.
17:    end if
18:    Sample a minimal subset  $S_\tau$  by using AGS and the seed datum.
19:  end if
20:  Generate a model hypothesis  $\theta_\tau$  using  $S_\tau$ .
21:  if  $\tau \geq b$  and  $\text{mod}(\tau, b) = 0$  then
22:    Update residual indices for Eq. (3).
23:  end if
24:  Let  $\Theta = \Theta \cup \{\theta_\tau\}$ .
25: end for
```

On the one hand, the neighborhood consensus is used to select the seed datum with a β probability. The selection process contains two steps. First, NCRS randomly selects a data point from input data. Second, NCRS judges whether the selected data point is an inlier by using Eq. (5). The first and the second steps are repeated until either an inlier is found or the number of repetitions is larger than the specified one

(i.e., the parameter G). If an inlier is found after performing the first and the second steps, the inlier is regarded as the seed datum. Given an appropriate G value, the probability of no single inlier found is low after performing the first and the second steps G times. For example, if the inlier rate is 0.2 and the G value is 20, then the probability of no single inlier found will be $(1 - 0.2)^{20} \approx 0.01$. On the other hand, the random selection is used to select the seed datum with a $(1 - \beta)$ probability. In addition, failure to find the inlier point may still occur after performing the neighborhood consensus information G times. In this case, the random selection is used to select the seed datum. For simplicity, the two cases using the random selection to select the seed datum are combined in our implementation (please see Lines 15 to 17 of Algorithm 1). The influence of β and G on the proposed sampling algorithm is evaluated in Section IV-B.

After the seed datum is selected, a minimal subset is sampled by using AGS [15] and the seed datum as below: (1) Compute the inlier probabilities of input data based on the generated hypotheses and the seed datum by using Eq. (3). (2) Select the significant data based on the inlier probabilities. (3) Compute sampling probabilities by using both the inlier probabilities and the corresponding matching scores. (4) Sample a minimal subset from the significant data by using the computed sampling probabilities.

In [31], an MCMC process is designed to sample data subsets from inliers and sample data subsets from outliers based on a greedy search-based sampling algorithm. In contrast, the proposed method uses the MCMC process to combine the neighborhood consensus-based strategy with the random selection strategy for selecting the first datum rather than all data of a minimal subset. To the best of our knowledge, the proposed method is the first one using the MCMC process to select the first datum of a minimal subset for the residual sorting-based sampling algorithm. Because of effectively using the MCMC process to select the first datum for the residual sorting-based sampling algorithm, the proposed method achieves superior performance to the method proposed in [31] (experimental results please see Table I).

D. The Proposed Sampling Algorithm

The complete proposed sampling algorithm is summarized in Algorithm 1, which mainly consists of the uniform random sampling strategy and the proposed NCRS sampling strategy. The main aim of the uniform random sampling strategy is to sample initial hypotheses to compute inlier probabilities for the NCRS sampling strategy. Different from the existing residual sorting-based sampling algorithms, which aim to sample as many clean minimal subsets as possible and thus sample a small batch of initial minimal subsets by using the uniform random sampling algorithm, the proposed sampling algorithm aims to achieve as good fitting performance as possible. We find that sampling a large batch of initial minimal subsets by using the uniform random sampling algorithm will improve the fitting performance. In contrast, the aim of the NCRS sampling strategy is to sample accurate hypotheses based on the generated hypotheses. As the existing residual

sorting-based sampling algorithms, we update the residual indices for the inlier probability computation after generating a new batch of hypotheses (Lines 21 to 23).

IV. EXPERIMENTS

In this section, we first introduce three datasets and an evaluation metric in Section IV-A. Then, we evaluate the influence of both different components and parameters on the performance of the proposed sampling algorithm in Section IV-B. After that, we compare the proposed method with several state-of-the-art model fitting methods on three vision tasks (i.e., two-view motion segmentation, image registration and 3D motion segmentation) in Sections IV-C, IV-D and IV-E, respectively, where the proposed method uses the proposed sampling algorithm and the existing model fitting frameworks. Finally, we give the discussion about what will happen when the assumptions made in the proposed method are violated in Section IV-F. We run all experiments on a desktop computer with a 3.4 GHz i7 CPU on the Windows 10.

A. Datasets and Evaluation Metric

We use image pairs from *SNU* [32] and the popular benchmark *AdelaideRMF* [33] for the two-view motion segmentation. We also use Printed Circuit Board (PCB) images from [30] to further evaluate the performance of the proposed method. In addition, we use video sequences from *KT3DMoSeg* [11] and the popular benchmark *Hopkins 155* [34] for the 3D motion segmentation. As in [35] and [36], we use the Clustering Error (CE) to evaluate the fitting performance, and CE is computed as

$$CE = \frac{\text{number of incorrectly clustered data points}}{\text{number of input data points}}. \quad (6)$$

The lower value of CE means the better fitting performance.

B. Influence of Parameters and Components

In this section, we use the image pairs from *AdelaideRMF* [33] for the two-view motion segmentation to evaluate the influence of the components and the parameters on the performance of the propose method. We report the mean and the standard deviation of clustering errors over 200 repetitions.

For two-view motion segmentation, we use the framework of CBS [10] due to its relative efficiency. We first evaluate the influence of the parameter β on the performance of the propose method. Fig. 1 shows the results. $\beta = 0$ means that the proposed NCRS is not used. In contrast, $\beta = 1$ means that the proposed method totally uses NCRS for the seed datum selection. Intuitively, the more inliers selected by the proposed NCRS, the better performance of the proposed method is. However, this is not true in fact. The reasons behind this are as follows. (1) As shown in Eq. (3), the computation of the inlier probability only uses the first h residual indices. In most of residual-sorting based sampling methods (e.g., [12], [15]), h is fixed to one-tenth of the number of the generated hypotheses. It is difficult to ensure that the first h residual indices of two inliers from the same structure have many common elements when many clean minimal subsets are

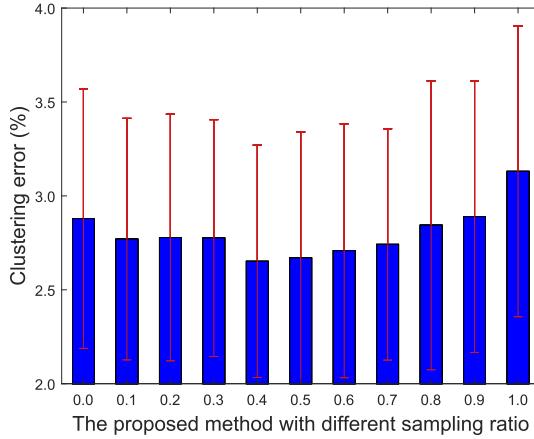


Fig. 1. Influence of the parameter β on the performance of the proposed method.

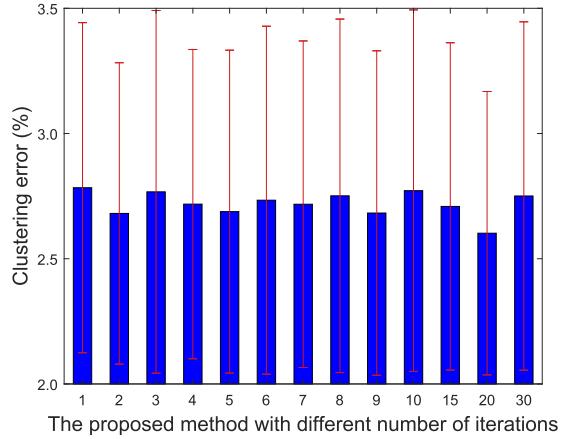


Fig. 2. Influence of the proposed method with different number of iterations.

TABLE I
THE CLUSTERING ERRORS (%) ACHIEVED BY THE TWELVE METHODS
FOR THE TWO-VIEW MOTION SEGMENTATION ON
THE *AdelaideRMF* DATASET

Method	PEARL	T-linkage	RCMSA	MSMH	Prog-X	MCNC
Mean	29.54	13.39	8.05	7.41	10.73	7.01
std	14.80	7.20	4.65	2.98	8.73	4.22
Method	MLink	CBS	HRMP	MCF	AGS	Ours
Mean	8.59	4.65	7.37	4.01	2.80	2.38
std	4.67	2.68	6.90	4.64	1.27	1.03

sampled. (2) The inlier rates of different structures in an input data may vary greatly, and it is also true in different input data. Thus, it is hard to set an adaptive h for different input data or its structures.

In Table I, the results of AGS seem to be the results of Fig. 1 when $\beta = 0$. However, the results of the proposed method are different from Fig. 1 when $\beta = 0.4$. The reason is that in Fig. 1, the number of initial hypotheses sampling by the uniform random sampling algorithm is the same as that of AGS. In contrast, in Table I, the proposed method samples a relatively large number of initial hypotheses based on the evaluation in Fig. 3, and thus achieves better performance.

Both the proposed method and our previous proposed AGS method [15] achieve low mean clustering errors on the *AdelaideRMF* dataset for the two-view motion segmentation. Thus, it is hard to significantly improve the performance of AGS on this dataset. On the relatively challenging dataset (i.e., the *SNU* dataset), from Table II, we can see that the proposed method achieves a much lower mean clustering error than AGS (3.21 vs. 8.39). In addition, from Fig. 1, we can see that the mean clustering error first decreases and then increases generally as the value of β increases. Although the standard deviation seems random, both the mean clustering error and the standard deviation are low when $\beta = 0.4$.

Fig. 2 shows that the influence of the proposed method with different number of iterations (the parameter G in Line 7 of Algorithm 1) on the performance. The experimental results demonstrate that the proposed method is tolerant to

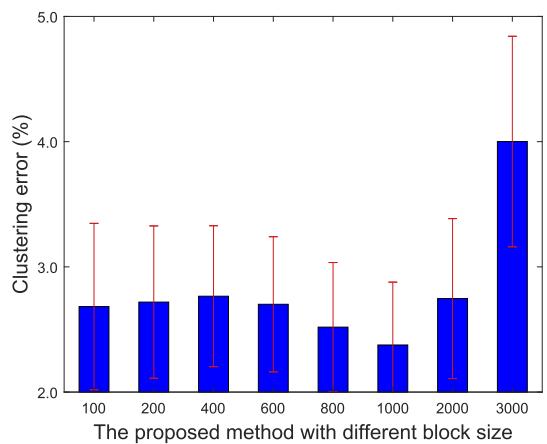


Fig. 3. Influence of different block size on the proposed method.

the parameter G . The main reason is the proposed method can sample accurate hypotheses in the given time. We also evaluate the influence of the proposed method with different initial block sizes in Fig. 3. The proposed method achieves the best performance when the block size is set to 1000. When the block size is set to 2000 and 3000, the performance is bad because the proposed method mainly and almost totally uses the uniform random sampling strategy in these two cases respectively. We finally evaluate the influence of the proposed method with different sampling time in Fig. 4. The mean and the standard deviation of the clustering errors obtained by the proposed method decrease generally as the value of sampling time increases, especially for the standard deviation of the clustering errors.

In addition, the neighbourhood consensus threshold \mathcal{K} in Eq. (5) has been discussed in detail in [27]. We empirically set the default values for the multiple scale threshold $\mathcal{K} = [4, 6, 8]$ as [27]. The window size h in Eq. (3) is set to $0.1 * M$ as in [11] and [15], where N is the number of the generated model hypotheses.

C. Two-View Motion Segmentation

The two-view motion segmentation aims to segment correspondences belonging to the same moving object in two

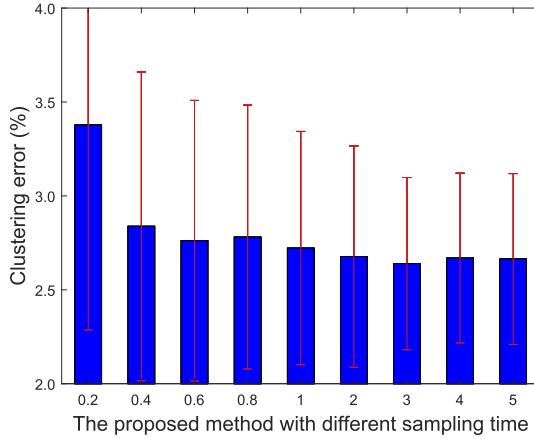


Fig. 4. Influence of different sampling time (in seconds) on the performance of the proposed method.

views [10]. Assume that N correspondences $\{\mathbf{x}_i\}_{i=1}^N$ of two input images include multiple moving objects, where $\mathbf{x}_i = \{\mathbf{u}_i, \mathbf{v}_i\}$. A fundamental matrix $\mathbf{F} \in \mathbb{R}^{3 \times 3}$ can be used to model correspondences belonging to the same moving object by $(\mathbf{v}_i)^T \mathbf{F} \mathbf{u}_i = 0$. The two-view motion segmentation has been widely employed in various vision applications such as visual localization, camera calibration and 3D reconstruction.

We compare the proposed method with eleven model fitting methods, containing PEARL [37], MCNC [31], T-linkage [38], RCMSA [9], MSMH [8], Prog-X [39], MLink [40], CBS [10], HRMP [41], MCF [30], and AGS [15]. The reasons of choosing the above competing methods are as follows: (1) RCMSA and AGS are residual sorting-based sampling algorithms. The proposed sampling algorithm is based on AGS. (2) We use the overall framework of CBS for two-view motion segmentation. (3) MLink, HRMP, MSMH and MCF are most recently proposed model fitting methods. Both the proposed method and MCF use the neighborhood consensus information for their data sampling. (4) The other competing methods (i.e., PEARL, MCNC, T-linkage and Prog-X) are popularly compared in recent literature for fitting data with multiple structures. The source code of CBS is provided by its authors. The results of PEARL, Prog-X and MLink are cited from [40]. The results of T-linkage and MSMH are cited from [8]. The results of MCNC are cited from [31]. The results of RCMSA, HRMP and MCF are cited from [30]. CBS, AGS and the proposed method are performed 200 times. The results obtained by twelve competing methods are shown in Table I and Fig. 5.

From Table I, we can see that the proposed method achieves the lowest mean clustering error and the lowest standard deviation of the clustering errors compared with the other eleven competing methods. Our previous proposed AGS achieves the second lowest mean clustering error and the second lowest standard deviation of the clustering errors. In other words, the proposed method reduces 15% the mean clustering error obtained by our previous proposed AGS (i.e., 2.38% vs. 2.80%). MCF and CBS also achieve good performance. The mean clustering errors obtained by the two methods are less than 5%, but are significantly larger than that obtained by the proposed method.

TABLE II
THE CLUSTERING ERRORS (%) ACHIEVED BY THE SEVEN METHODS FOR AFFINE MATRIX ESTIMATION ON THE SNU DATASET

Method	RCMSA	MSMH	HOMF	HRMP	MCF	AGS	Ours
Mean	11.60	15.31	13.02	12.76	4.57	8.39	3.21
std	3.21	4.13	4.29	7.45	2.70	3.70	1.72

RCMSA, MSMH, MLink, MCNC and HRMP achieve relatively low clustering errors, and the mean clustering errors obtained by the four methods are larger than 5% but lower than 10%. The mean clustering error obtained by Prog-X is slightly larger than 10%. The performance obtained by T-linkage is worse than Prog-X, and the mean clustering error obtained by T-linkage is larger than 13%. PEARL achieves the worst performance, and the mean clustering errors obtained by the two methods are larger than 20%.

Finally, Fig. 5 provides a qualitative results obtained by the proposed method. From the Fig. 5, we can see that the proposed method accurately segment eight example image pairs of the *AdelaideRMF* dataset.

To further show the effectiveness of the proposed method, we evaluate the proposed method on the *SNU* dataset. In this experiment, we estimate the affine matrix as [30]. AGS and the proposed method are performed 200 times. The results of RCMSA, MSMH, HOMF, HRMP and MCF are cited from [30]. The experimental results obtained by seven methods are shown in Table II.

From Table II, we can observe that the proposed method achieves the lowest mean clustering error and the lowest standard deviation of the clustering errors compared with other six competing methods. MCF (AGS) achieves the second (third) lowest mean clustering error. The reasons of AGS achieving the large mean clustering error are as follows. (1) Matches extracted from many image pairs in the *SNU* dataset have a large proportion outliers and multiple structures. For example, the outlier ratio in the *Mickeys* image pair is 73.75% and this image pair has 3 structures. The mean clustering error obtained by AGS is 21.34% on this image pair. (2) As stated before, AGS is prone to sample invalid data subset when the outlier ratio is high. In addition, the other four competing methods (i.e., RCMSA, MSMH, HOMF and HRMP) achieve relatively bad performance, and the mean clustering errors obtained by the four methods are larger than 10%.

D. Registration of Printed Circuit Board Images

Printed Circuit Boards (PCBs) are widely used in computers, televisions and other electronic equipments. The defect inspection is an important step in the PCB production process. The automatic optical inspection is one of most common means of the defect inspection of PCBs [42], where image registration plays a critical role in the automatic optical inspection. However, feature matches obtained from PCB image pairs probably contain a large proportion of outliers because PCBs often contain a large number of repetitive elements (see Fig. 6).

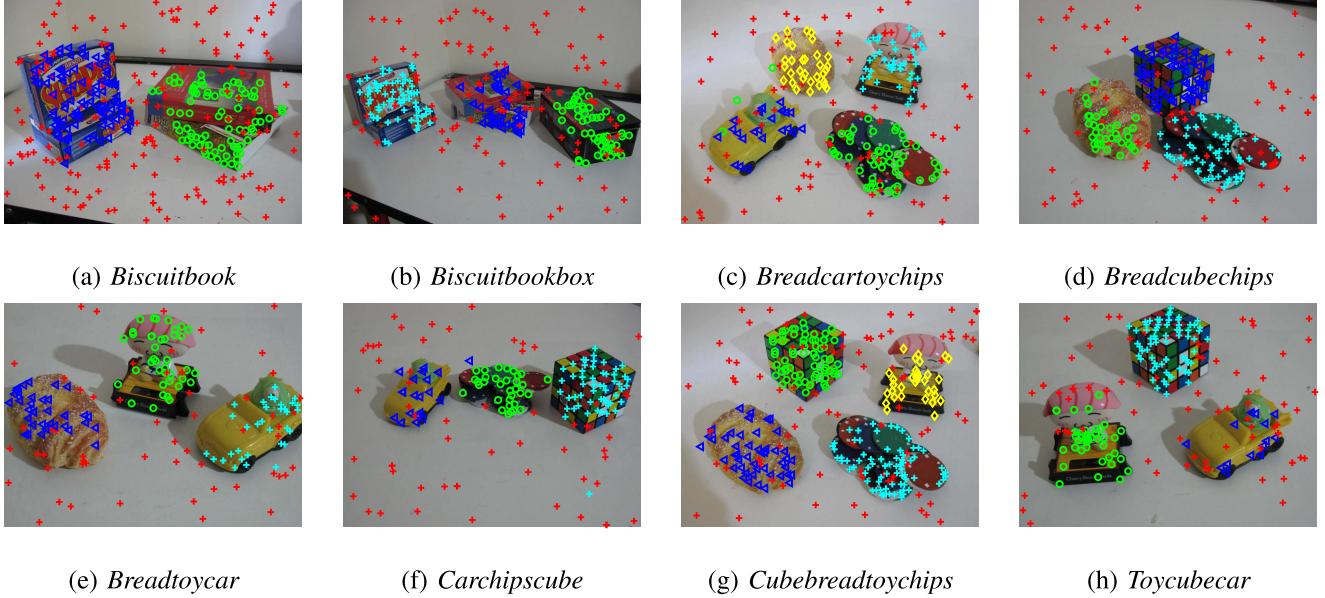


Fig. 5. Eight successful two-view segmentation examples obtained by the proposed method (only showing the first view).

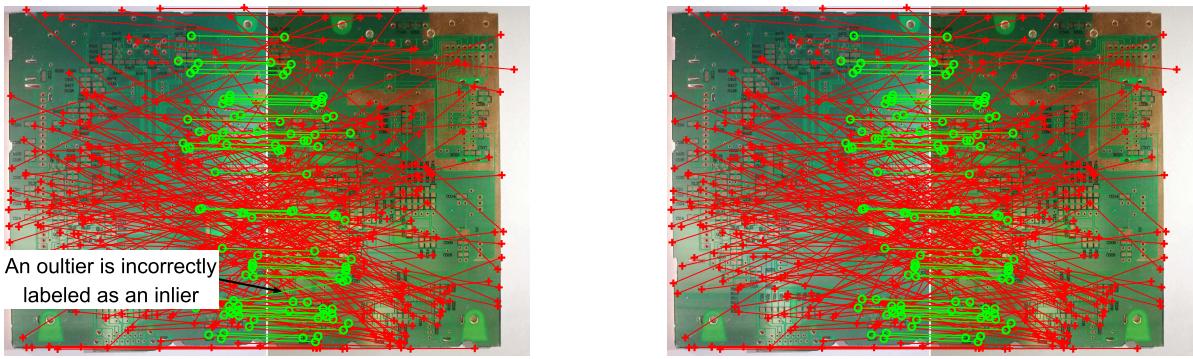


Fig. 6. Segmentation results obtained by MCF (left) and the proposed method (right). Inliers and outliers are marked with green circles and red crosses, respectively.

Thus, the defect inspection of PCBs needs to use a robust model fitting method for image registration.

In this experiment, we compare the proposed method with the recently proposed and better-performing method MCF. To clearly show the detailed scenes in the PCB image pair, we only select 10% of the totally 2514 feature matches. The results obtained by MCF and the proposed method are shown in Figs. 6 and 7. Fig. 6 shows that the proposed method accurately segments inliers from outliers. In contrast, MCF cannot accurately segment inliers from outliers, and an example is shown in Fig. 6. Based on the obtained inliers, we achieve the registration results. Fig. 7 shows the proposed method achieves satisfactory registration result, whereas the result achieved by MCF contains multiple “ghosting” areas highlighted with the red ellipses, in which the words are not clear.

E. 3D Motion Segmentation

The aim of the 3D motion segmentation is the same as that of the two-view motion segmentation. The main difference between the two tasks is: The video sequences used for the

TABLE III
THE CLUSTERING ERRORS (%) ACHIEVED BY THE TWELVE METHODS ON THE *KT3DMoSeg* DATASET

Method	LSA	SSC	MSMC	ALC	Affine	Homo
Mean	38.30	33.88	27.74	24.31	15.76	11.45
Median	38.58	33.54	35.80	19.04	11.52	7.14
Method	KerAdd	CoReg	Subset	Fund	<i>Ours_H</i>	<i>Ours_F</i>
Mean	8.31	7.92	8.08	13.92	10.13	7.80
Median	1.02	0.75	0.71	5.09	3.59	2.63

3D motion segmentation contain multiple frames while the image pairs used for the two-view motion segmentation only contain two views. In the 3D motion segmentation, given N feature points $\{\mathbf{t}_m^f \in \mathbb{R}^2\}_{m=1, \dots, N}^{f=1, \dots, F}$ tracked through F frames in a video sequence, the aim is to segment the N feature points into multiple motions (i.e., moving objects). The 3D motion segmentation has been widely used as an important pre-processing step for many vision applications such as visual surveillance, object tracking and action recognition.

We compare the proposed method with ten motion segmentation methods, containing LSA [43], SSC [35], MSMC [44],

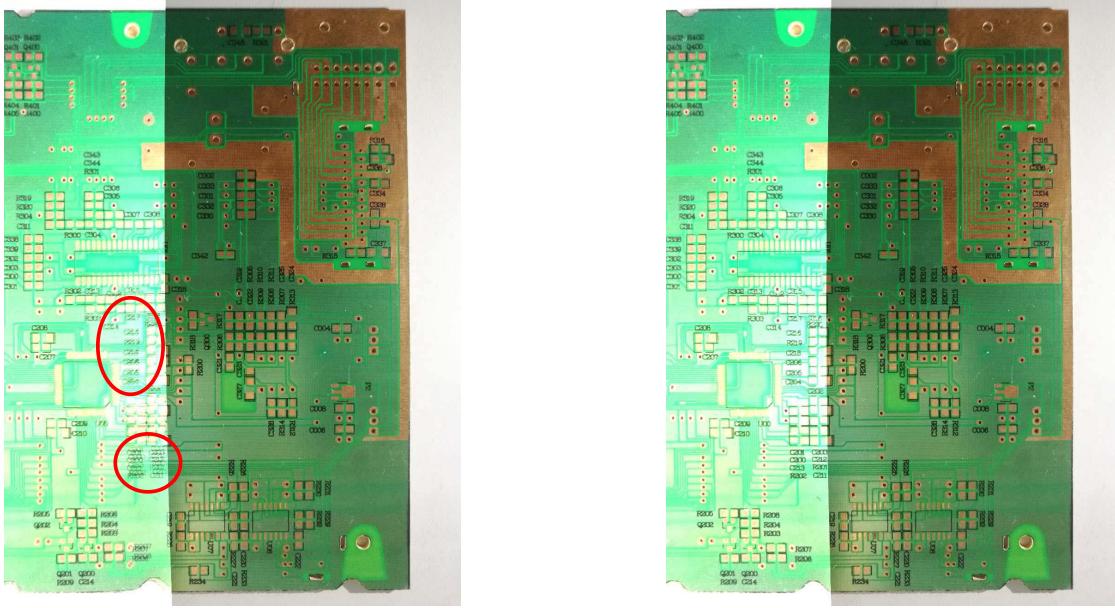


Fig. 7. Registration results obtained by MCF (left) and the proposed method (right), respectively.

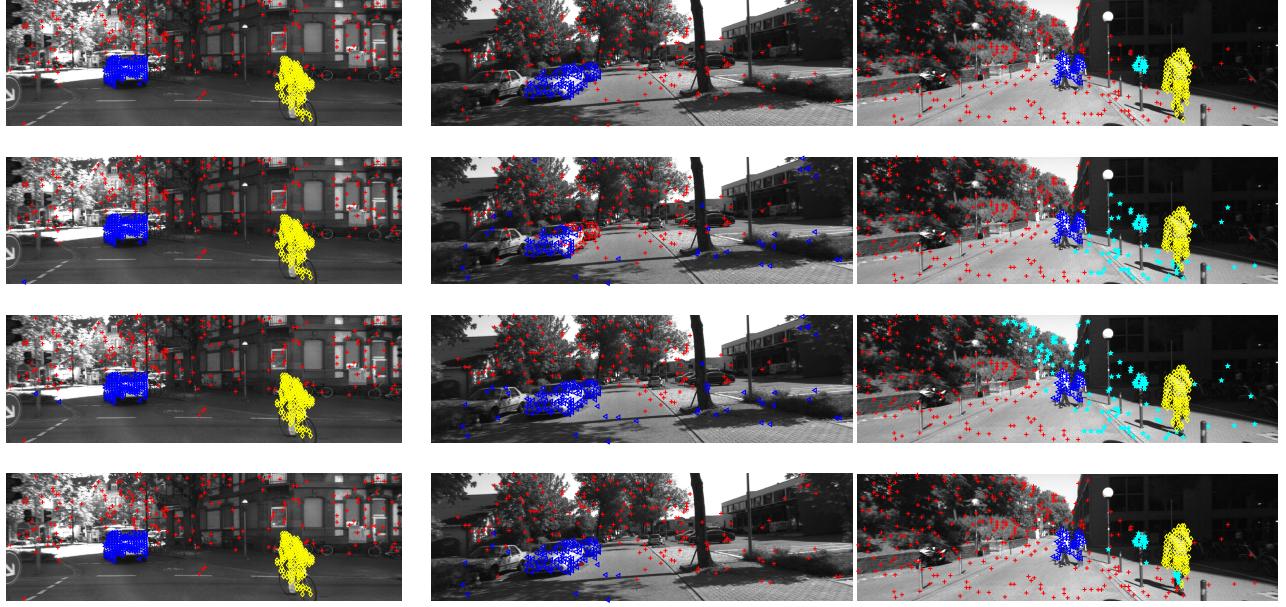
ALC [45], Affine, Homo, KerAdd, CoReg, *Subset* and Fund [11] on the *KT3DMoSeg* dataset. Among these competing methods, LSA is a local subspace affinity-based method. SSC is a sparse subspace clustering method. ALC is an information-theory method. All other competing methods (i.e., MSMC, Affine, Homo, KerAdd, CoReg, *Subset* and Fund) are model fitting-based methods. Affine, Homo, KerAdd, CoReg, *Subset* and Fund are introduced in [11]. Affine, Homo and Fund only use single motion model, and they use affine, homography and fundamental matrix, respectively. In contrast, KerAdd, CoReg and *Subset* use multiple motion models. In other words, three motion models (i.e., affine, homography and fundamental matrix) are used in the three methods. They use different techniques to exploit three types of motion model information to achieve better segmentation performance, and they use kernel addition, co-regularization and constrained clustering, respectively. The experimental results of the ten competing motion segmentation methods are cited from [11]. We use the framework of Fund [11] for the proposed method, where we use two motion models (i.e., homography and fundamental matrix) and denote them as *Ours_H* and *Ours_F*. As in [11], each consecutive frame pair (defined as the f th and the $(f + 1)$ th frames) is selected as two views, i.e., the proposed method will be performed $(F - 1)$ times to sample data subsets for a video sequence with F frames. Table III shows the experimental results achieved by the twelve motion segmentation methods.

Table III shows that the proposed *Ours_F* achieves the lowest mean clustering error compared with all the other eleven competing methods. Although the gap between the proposed *Ours_F* and the best of the other ten competing methods is small (i.e., 7.80% vs. 7.92%), the gap between the proposed *Ours_F* and Fund is significantly large (i.e., 7.80% vs. 13.92%). We emphasize that the same as Fund, the proposed *Ours_F* only uses one motion model (i.e., the fundamental matrix). As the second best method CoReg, the proposed method can also fuse three motion models together to

achieve better segmentation performance. However, our main aim is only to demonstrate the effectiveness of the proposed guided sampling algorithm in this paper. The above results demonstrate that.

For other competing methods, *Subset* and KerAdd achieve the similar results as CoReg. The mean clustering errors obtained by the three methods are about 8.00%. Homo, Fund and Affine achieve relatively worse performance due to the use of only one motion model. The mean clustering errors obtained by the three methods range from 11.45 to 15.16. Although MSMC is also a model fitting based method, the mean clustering error obtained by it is large because it cannot effectively use the motion information of all frames of a video sequence. The non-model fitting methods cannot effectively handle practical problems such as strong perspectives and object occlusions, and thus the mean clustering errors obtained by these methods (i.e., LSA, MSMC, SSC and ALC) are larger than 20%.

The performance obtained by the proposed method using homography model is worse than that obtained by the proposed method using fundamental matrix model. The reasons are as follows: (1) As stated in the introduction, as the dimension of the geometric model increases, the probability of sampling a clean minimal subset by using the uniform random sampling algorithm decreases exponentially. In [11], the authors use the uniform random sampling for data sampling. Thus, by using the uniform random sampling, sampling accurate hypotheses for homography model is easier than sampling accurate hypotheses for fundamental matrix model, and as expected, the performance obtained by the method Homo using homography model is better than that obtained by the method Fund using fundamental matrix model. (2) The experimental results in [12] and [15] show that as the dimension of the geometric model increases, the probability of sampling a clean minimal subset by using guided sampling algorithms still decreases. By using the same fundamental matrix model, the performance obtained by the proposed



(a) Seq005_Clip01

(b) Seq009_Clip01

(c) Seq038_Clip02

Fig. 8. 3D motion segmentation examples (only showing the one view). The first row shows the ground truth segmentations. The second to the fourth rows show the results obtained by Fund, SubSet and the proposed *OursF*, respectively.

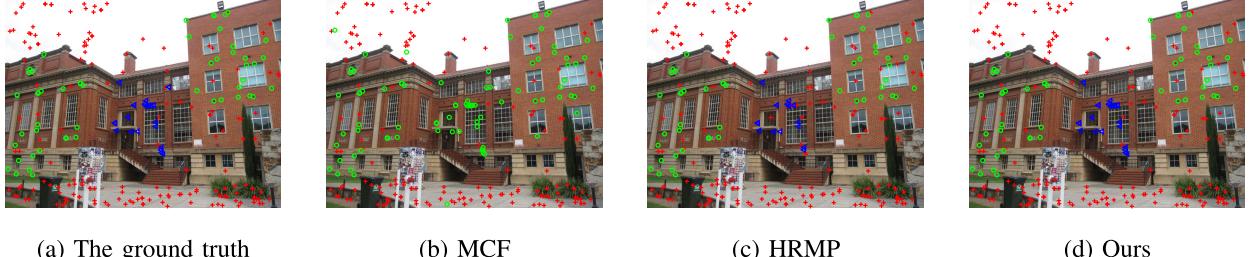


Fig. 9. Qualitative results on the *Barrsmith* image pair.

method is much better than that obtained by Fund [11]. This demonstrates that the proposed sampling algorithm generates significantly more accurate hypotheses than the uniform random sampling. However, by using homography model, the performance of the proposed becomes worse. This is because the proposed method by using homography model generates more accurate hypotheses than the proposed method by using fundamental matrix model, but the value of the parameter h in Eq. (3) for computing affinity matrix for clustering is still set to the same value as [11]. It should be set a relatively large value to h for better performance. However, it is hard to set an adaptive value for h , and this is out of the scope of this paper.

The final segmentation results obtained by the proposed method on three video sequences from the *KT3DMoSeg* dataset are shown in Fig. 8. Fig. 8 shows the proposed method achieves better results than the state-of-the-art *SubSet* method.

To further strengthen this work, we perform more evaluations on the *Hopkins 155* dataset. We compare the proposed method with eleven motion segmentation methods, including LSA, SSC, MSMC, ALC, MCNC [31], Affine, Homo, KerAdd, CoReg, *SubSet* and Fund. In this experiment, we only use one motion models (i.e., fundamental matrix) for the proposed method. The results of MCNC are cited from [31],

TABLE IV
THE CLUSTERING ERRORS (%) ACHIEVED BY THE TWELVE METHODS
ON THE *Hopkins 155* DATASET. “—” MEANS NOT REPORTED

Method	LSA	SSC	MSMC	ALC	MCNC	Affine
Mean	4.86	2.18	4.33	3.56	2.34	0.59
Median	0.89	0.00	0.00	0.50	-	-
Method	Homo	KerAdd	CoReg	SubSet	Fund	Ours
Mean	0.71	0.36	0.46	0.31	1.79	0.85
Median	-	-	-	-	-	0.00

and the results of LSA, SSC, ALC are cited from [35]. The other seven competing methods are cited from [11].

Table IV shows that the proposed method achieves comparable results with the state-of-the-art methods. We emphasize that the mean clustering error obtained by the proposed method is about only half of that obtained by Fund, where the proposed method and Fund use the same motion model and framework, but use different sampling algorithms. This further demonstrates the effectiveness of the proposed sampling algorithm.

F. Discussion

As stated before, different from most of spatial proximity-based algorithms [21], [22], [25] sampling all data of

a minimal subset from neighbour data points, the proposed method only uses a neighborhood consensus information to judge whether a selected data point is an inlier, and tries to select an inlier as the first datum of a minimal subset. The proposed method repeatedly performs the following two steps (please see Lines 7 to 13 of Algorithm 1) G times: (1) Select a match $\mathbf{x}_i = \{\mathbf{u}_i, \mathbf{v}_i\}$ from input data randomly; (2) Judge whether the selected match is an inlier; (3) If \mathbf{x}_i is an inlier, then let \mathbf{x}_i be the seed datum and break. When the spatial proximity assumption is violated, \mathbf{x}_i will be discarded probably because \mathbf{u}_i and \mathbf{v}_i share fewer or no common neighborhoods. Even when the spatial proximity assumption is violated and \mathbf{x}_i is selected as the seed datum with a certain smaller probability, but the proposed method still achieves good results. Experimental results show that the proposed method successfully segments input data because it repeatedly performs the sampling process M times rather than only one time (please see Line 2 of Algorithm 1), and it only needs to sample appropriate data points to generate an accurate hypothesis one time. For example, the mean segmentation errors obtained by the proposed method are 2.58% and 1.14% on the image pairs of *Toycubecar* and *Barrsmith* in Adelaide RMF, where the *Toycubecar* image pair has a small structure (i.e., its *car* structure only has 7% inliers) and the *Barrsmith* image pair has segregated structures. Qualitative results please see Figs. 5(h) and 9.

V. CONCLUSION

In this paper, we propose a novel guided sampling algorithm by using the neighborhood consensus and the residual sorting. Based on the neighborhood consensus, we propose a sampling strategy to choose inliers as the seed data for increasing the probabilities of sampling clean data subsets. Moreover, we use the MCMC process to combine the above proposed sampling strategy with the uniform random sampling strategy for better fitting performance. Finally, we set a relatively large size to the initial block of the randomly generated hypotheses to further improve the performance of the proposed algorithm. We evaluate the proposed algorithm on the tasks of two-view motion segmentation, image registration and 3D motion segmentation. The promising results show the efficiency of the proposed algorithm for guided sampling.

REFERENCES

- [1] V. Sanchez, "Rate control for predictive transform screen content video coding based on RANSAC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 11, pp. 4422–4438, Nov. 2021.
- [2] F. Tang, Y. Wu, X. Hou, and H. Ling, "3D mapping and 6D pose computation for real time augmented reality on cylindrical objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 9, pp. 2887–2899, Sep. 2020.
- [3] J. Yang, Z. Huang, S. Quan, Q. Zhang, Y. Zhang, and Z. Cao, "Toward efficient and robust metrics for RANSAC hypotheses and 3D rigid registration," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 2, pp. 893–906, Feb. 2022.
- [4] M. Jeong, B. C. Ko, S. Kwak, and J.-Y. Nam, "Driver facial landmark detection in real driving situations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 10, pp. 2753–2767, Oct. 2018.
- [5] T. Lai, H. Fujita, C. Yang, Q. Li, and R. Chen, "Robust model fitting based on greedy search and specified inlier threshold," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 7956–7966, Oct. 2019.
- [6] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [7] H. Isack and Y. Boykov, "Energy based multi-model fitting & matching for 3D reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1146–1153.
- [8] H. Wang, G. Xiao, Y. Yan, and D. Suter, "Searching for representative modes on hypergraphs for robust geometric model fitting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 3, pp. 697–711, Mar. 2018.
- [9] T. T. Pham, T.-J. Chin, J. Yu, and D. Suter, "The random cluster model for robust geometric fitting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 8, pp. 1658–1671, Aug. 2014.
- [10] R. Tennakoon, A. Sadri, R. Hoseinnezhad, and A. Bab-Hadiashar, "Effective sampling: Fast segmentation using robust geometric model fitting," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4182–4194, Sep. 2018.
- [11] X. Xu, L.-F. Cheong, and Z. Li, "3D rigid motion segmentation with mixed and unknown number of models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 1–16, Jan. 2021.
- [12] T.-J. Chin, J. Yu, and D. Suter, "Accelerated hypothesis generation for multistructure data via preference analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 625–638, Apr. 2012.
- [13] T. Lai, H. Wang, Y. Yan, D.-H. Wang, and G. Xiao, "Rapid hypothesis generation by combining residual sorting with local constraints," *Multimedia Tools Appl.*, vol. 75, no. 12, pp. 7445–7464, Jun. 2016.
- [14] H. S. Wong, T.-J. Chin, J. Yu, and D. Suter, "Mode seeking over permutations for rapid geometric model fitting," *Pattern Recognit.*, vol. 46, no. 1, pp. 257–271, Jan. 2013.
- [15] T. Lai, H. Wang, Y. Yan, T.-J. Chin, J. Zheng, and B. Li, "Accelerated guided sampling for multistructure model fitting," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4530–4543, Oct. 2020.
- [16] O. Chum and J. Matas, "Matching with PROSAC—Progressive sample consensus," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 220–226.
- [17] V. Fragoso and M. Turk, "SWIGS: A swift guided sampling method," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2770–2777.
- [18] E. Brachmann and C. Rother, "Neural-guided RANSAC: Learning where to sample model hypotheses," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4322–4331.
- [19] Y. Kanazawa and H. Kawakami, "Detection of planar regions with uncalibrated stereo using distributions of feature points," in *Proc. Brit. Mach. Vis. Conf.*, 2004, pp. 247–256.
- [20] K. Ni, H. Jin, and F. Dellaert, "GroupSAC: Efficient consensus in the presence of groupings," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2193–2200.
- [21] D. Barath, J. Noskova, M. Ivashechkin, and J. Matas, "MAGSAC++, a fast, reliable and accurate robust estimator," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1304–1312.
- [22] D. R. Myatt, P. H. S. Torr, S. J. Nasuto, J. M. Bishop, and R. Craddock, "NAPSAC: High noise, high dimensional robust estimation—It's in the bag," in *Proc. Brit. Mach. Vis. Conf.*, 2002, pp. 458–467.
- [23] T. Lai et al., "Efficient robust model fitting for multistructure data using global greedy search," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3294–3306, Jul. 2020.
- [24] Q. H. Tran, T.-J. Chin, W. Chojnacki, and D. Suter, "Sampling minimal subsets with large spans for robust estimation," *Int. J. Comput. Vis.*, vol. 106, no. 1, pp. 93–112, Jan. 2014.
- [25] D. Barath and J. Matas, "Graph-cut RANSAC," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6733–6741.
- [26] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *Proc. Int. Conf. Comput. Vis. Theory Appl.*, 2009, pp. 331–340.
- [27] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, "Locality preserving matching," *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, 2019.
- [28] Y. Liu, B. N. Zhao, S. Zhao, and L. Zhang, "Progressive motion coherence for remote sensing image matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022.
- [29] Y. Liu et al., "Robust feature matching via advanced neighborhood topology consensus," *Neurocomputing*, vol. 421, pp. 273–284, Jan. 2021.
- [30] H. Guo et al., "Motion consistency guided robust geometric model fitting with severe outliers," *IEEE Trans. Ind. Electron.*, vol. 69, no. 4, pp. 4065–4075, Apr. 2022.

- [31] A. Sadri, R. Tennakoon, R. HosseiniNezhad, and A. Bab-Hadiashar, "Robust visual data segmentation: Sampling from distribution of model parameters," *Comput. Vis. Image Understand.*, vol. 174, pp. 82–94, Sep. 2018.
- [32] M. Cho, Y. M. Shin, and K. M. Lee, "Co-recognition of image pairs by data-driven Monte Carlo image exploration," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2008, pp. 144–157.
- [33] H. S. Wong, T.-J. Chin, J. Yu, and D. Suter, "Dynamic and hierarchical multi-structure geometric model fitting," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 1044–1051.
- [34] R. Tron and R. Vidal, "A benchmark for the comparison of 3-D motion segmentation algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [35] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.
- [36] T. Lai, H. Wang, Y. Yan, T.-J. Chin, and W.-L. Zhao, "Motion segmentation via a sparsity constraint," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 973–983, Apr. 2017.
- [37] A. Delong, A. Osokin, H. N. Isack, and Y. Boykov, "Fast approximate energy minimization with label costs," *Int. J. Comput. Vis.*, vol. 96, no. 1, pp. 1–27, 2012.
- [38] L. Magri and A. Fusiello, "T-linkage: A continuous relaxation of J-linkage for multi-model fitting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3954–3961.
- [39] D. Barath and J. Matas, "Progressive-X: Efficient, anytime, multi-model fitting algorithm," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2019, pp. 3780–3788.
- [40] L. Magri, F. Leveni, and G. Boracchi, "MultiLink: Multi-class structure recovery via agglomerative clustering and model selection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 1853–1862.
- [41] S. Lin, X. Wang, G. Xiao, Y. Yan, and H. Wang, "Hierarchical representation via message propagation for robust model fitting," *IEEE Trans. Ind. Electron.*, vol. 68, no. 9, pp. 8582–8592, Aug. 2020.
- [42] M. H. Annaby, Y. M. Fouad, and M. A. Rushdi, "Improved normalized cross-correlation for defect detection in printed-circuit boards," *IEEE Trans. Semicond. Manuf.*, vol. 32, no. 2, pp. 199–211, May 2019.
- [43] J. Yan and M. Pollefeys, "A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2006, pp. 94–106.
- [44] R. Dragon, B. Rosenhahn, and J. Ostermann, "Multi-scale clustering of frame-to-frame correspondences for motion segmentation," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2012, pp. 445–458.
- [45] S. Rao, R. Tron, R. Vidal, and Y. Ma, "Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1832–1845, Oct. 2010.



Taotao Lai received the M.S. and Ph.D. degrees in computer science and technology from Xiamen University, China, in 2009 and 2016, respectively. He is currently a Professor with Minjiang University, China. He has published over 20 papers in the international journals and conferences, including *IEEE TRANSACTIONS ON CYBERNETICS (TCYB)*, *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (TITS)*, *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS (TIE)*, *Pattern Recognition (PR)*, and *Computer Vision and Image Understanding (CVIU)*. His research interests include computer vision and pattern recognition. He has been awarded the Best Ph.D. Thesis in Fujian, China.



Yizhang Liu received the B.S. degree in electronic and information engineering and the M.S. degree in computer science and technology from Fujian Agriculture and Forestry University, Fuzhou, China, in 2013 and 2017, respectively. He is currently pursuing the D.Eng. degree with the School of Software Engineering, Tongji University, Shanghai. His current research interests include computer vision and image matching.



Jie Chang received the M.S. degree in computer science and technology from Xiamen University, China, in 2009, and the Ph.D. degree in computer science from the University of Science and Technology of China, Anhui, China, in 2019. He is currently an Associate Professor with the Department of Medical Information, Wannan Medical College. His research interests include machine learning, medical image processing, deep neural networks, and graph neural networks.



Lifang Wei received the M.S. degree in biomedical engineering from Northwest Polytechnical University, Xi'an, in 2008, and the Ph.D. degree in communication and information system from Fuzhou University, Fuzhou, China, in 2013. She is currently an Associate Professor with the College of Computer and Information, Fujian Agriculture and Forestry University, Fuzhou. Her research interests include computer vision and image processing.



Zuo Young Li received the B.S. and M.S. degrees in computer science and technology from Fuzhou University, Fuzhou, China, in 2002 and 2006, respectively, and the Ph.D. degree from the School of Computer Science and Technology, Nanjing University of Science and Technology (NUST), Nanjing, China, in 2010. He is currently a Professor with Minjiang University, Fuzhou. He has published more than 70 articles in international/national journals. His research interests include image processing, pattern recognition, and deep learning.



Hamido Fujita (Life Senior Member, IEEE) received the master's and Ph.D. degrees in information engineering from Tohoku University, Sendai, Japan, in 1985 and 1988, respectively, the Ph.D. (Honoris Causa) degree from Obuda University, Budapest, Hungary, in 2013, and the Ph.D. degree from Timisoara Technical University, Timisoara, Romania, in 2018. He was an Adjunct Professor of computer science and artificial intelligence at Stockholm University, Stockholm, Sweden; the University of Technology Sydney, Ultimo, NSW, Australia; and others. He is currently a Distinguished Professor of artificial intelligence with Iwate Prefectural University, Takizawa, Japan. He is also a Research Professor with the University of Granada, Granada, Spain; the Ho Chi Minh City University of Technology (HUTECH), Vietnam; Harbin Engineering University, China; and others. He has supervised the Ph.D. students jointly with the University of Laval, Quebec City, QC, Canada; Oregon State University, Corvallis, OR, USA; and many others. He headed a number of projects, including intelligent HCI, a project related to mental cloning for healthcare system as an intelligent user interface between human users and computers, and SCOPE project on virtual doctor systems for medical applications. He was a recipient of title of the Honorary Professor from Obuda University in 2011 and the Honorary Scholar Award from the University of Technology Sydney in 2012. He is the Chairperson of the i-SOMET Incorporated Association. He was the Editor-in-Chief of the *Knowledge-Based Systems* from 2005 to 2020. He is currently an Emeritus Editor for the *Knowledge-Based Systems* and the Editor-in-Chief of the *Applied Intelligence* (Springer) and the *Journal of Healthcare Management* (Taylor & Francis). He has given many keynotes in many prestigious international conferences on intelligent system and subjective intelligence. He is also a Highly Cited Researcher in crossfield in 2019 and computer science in 2020, 2021, and 2022, respectively, by Clarivate Analytics.