Motion Consistency Guided Robust Geometric Model Fitting With Severe Outliers

Hanlin Guo , Yang Lu , *Member, IEEE*, Haosheng Chen , Hailing Luo ,
Guobao Xiao , *Member, IEEE*, Haifang Zhang , and Hanzi Wang , *Senior Member, IEEE*

Abstract—**Geometric model fitting has been widely applied in the electronic industry. However, it remains as a challenging task when handling the data corrupted by a large number of false matches (i.e., severe outliers) between two-view images. In this article, we propose a novel motion consistency guided fitting method (MCF) to robustly and efficiently estimate the parameters of model instances in data involving severe outliers. Specifically, from input data, we first generate a series of neighborhood sets, in each of which gross outliers that are inconsistent in motions can be effectively filtered, according to motion consistency among true matches (i.e., inliers). Then, we propose an effective sampling algorithm to sample minimal subsets from the generated neighborhood sets. In this way, the model hypotheses computed from the sampled minimal subsets can cover all model instances with a high probability. Furthermore, by taking advantages of the generated hypotheses and neighborhood sets, we propose a novel model selection algorithm to estimate the number and the parameters of model instances. For fitting evaluation, we also build a new dataset, in which the images are collected from a fundus camera. Experiments on a variety of electronic industrial applications show that the proposed MCF achieves higher fitting accuracy at a much lower computational cost than several state-of-the-art fitting methods.**

Index Terms—**Geometric model fitting, guided sampling, model selection, motion consistency, outliers.**

Hanlin Guo, Yang Lu, Haosheng Chen, Hailing Luo, and Hanzi Wang are with the Fujian Key Laboratory of Sensing and Computing for Smart City, School of Informatics, Xiamen University, Xiamen 361005, China (e-mail: hanlinguo@stu.xmu.edu.cn; luyang@xmu.edu.cn; haoshengchen@stu.xmu.edu.cn; luohailing@stu.xmu.edu.cn; hanzi.wang@xmu.edu.cn).

Guobao Xiao is with the College of Computer and Control Engineering, Minjiang University, Fuzhou 350108, China (e-mail: gbx@mju.edu.cn).

Haifang Zhang is with the Department of Ophthalmology, Hebei General Hospital, Shijiazhuang 050051, China (e-mail: zhangzh@hebcm.edu.cn).
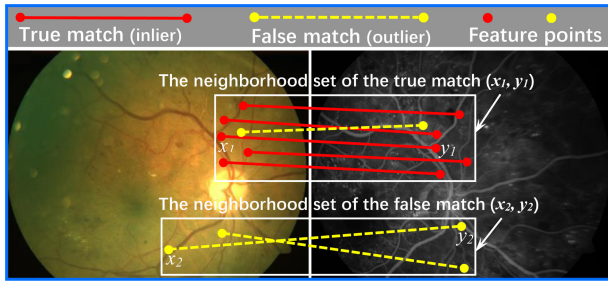
I. INTRODUCTION

ROBUST fitting of a geometric model among the data corrupted by noise and outliers plays an extremely important role in computer vision and robotics, due to the limitations of data acquisition systems and preprocessing techniques in many industry applications. Without robustness against outliers, the estimated model parameters may be biased severely, leading to the failure of many applications (such as robotics navigation [1], [2], remote sensing [3], [4], and motion segmentation [5], [6]) in the field of the electronic industry.

To illustrate the problem of robust geometric model fitting, we can see the example in Fig. 1, where only a part of the data (i.e., tentative feature matches between the two-view images) are plotted for clarity. In this example, the goal of robust model fitting is to estimate the parameters of a model instance (also called a "structure," which is an affine matrix) in the data, and identify the inliers (with respect to the estimated affine matrix) and outliers (caused by various uncertainty factors like cluttered scenes and false matches). As a result, the estimated parameters of the affine matrix can be used to align the two retinal images.

The problem of robust model fitting is generally solved in a two-step manner, as done in the well-known model fitting method (i.e., RANdom SAmpling Consensus, RANSAC [7]). First, model hypotheses are generated by sampling minimal subsets from input data. Here, a minimal subset represents the minimal number of data required to generate a model hypothesis. Second, based on the generated hypotheses, the parameters of model instances are estimated by using a model selection criterion. However, in many industrial applications, input data are likely to be corrupted by a large number of outliers, and they may also contain multiple structures, where the data of one structure act as pseudo-outliers to the other structures. Note that, gross outliers are the outliers that do not belong to any structures. Therefore, it is critical to develop a highly robust method for model fitting to estimate the parameters of multiple model instances from input data, even in the presence of a large number of outliers.

Based on the two key steps of robust model fitting, some fitting methods have been proposed to improve fitting performance from these two aspects: model hypothesis generation and model selection. On one hand, many guided sampling algorithms (such as [8] and [9]) have been proposed to increase the probability of hitting an all-inlier minimal subset (whose elements are all inliers of a model instance) to generate a good hypothesis.

Fig. 1. Example of the neighborhood sets with motion consistency. For the retinal image pair, the left one is a color fundus image, and the right one is a fundus fluorescein angiography image. Here, a motion of each match corresponds to a vector of each match. As can be seen, the true matches are more likely to have consistent motion patterns, and they show the similar motion magnitudes and directions with their neighboring matches. Moreover, the inliers tend to reside in a neighborhood set with the larger number of matches. Hence, the neighborhood set of the inlier (x_1, y_1) not only has a higher inlier ratio, but also contains more matches than that of the outlier (x_2, y_2).

As a result, they can reduce the number of sampled minimal subsets while covering all structures in data within a reasonable time. However, these sampling algorithms may not be able to efficiently generate a sufficient number of good hypotheses when input data contain severe outliers.

On the other hand, some model selection algorithms (such as [10], [11]) mainly focus on clustering based strategies to estimate the parameters of model instances especially in handling multistructural data, i.e., they formulate the model selection problem as a clustering problem. In this case, the parameters of different model instances are estimated based on clustering results, by which the problem of fitting multistructural data becomes tractable in practice. However, these model selection algorithms are less effective when dealing with data containing a number of outliers and/or structures. This is because that some existing clustering algorithms used for model selection are either sensitive to outliers or hard to adaptively estimate the number of clusters.

To overcome the above problems, in this article, we propose a novel geometric model fitting method based on motion consistency among input data, aiming at robustly and efficiently handling multistructural data with severe outliers. The proposed model fitting method mainly consists of a new guided sampling algorithm based on motion information and a novel model selection algorithm based on robust clustering analysis. More specifically, we observe that the inliers of a model instance usually tend to have consistent motion patterns, and they show similar motions with their neighboring data. In contrast, gross outliers tend to be randomly scattered across images. Based on such observations, we generate a neighborhood set, whose elements are consistent in motions, for each of input data by leveraging an effective motion consistency constraint. An example of the neighborhood sets with motion consistency is shown in Fig. 1. After that, we propose a novel sampling algorithm, which utilizes the data in the neighborhood sets with motion consistency to derive the effective sampled minimal subsets, to efficiently generate good model hypotheses for model selection.

Furthermore, to estimate the parameters of all model instances from input data, we propose a novel model selection algorithm, which analyzes the similarity values between input data and weighting scores of input data for identifying different clusters. The main contributions of this article are summarized as follows:

1) We introduce an effective constraint of motion consistency to explore the neighboring relationship among feature matches for model fitting. We show that the motion consistency information is a powerful clue that can be used to effectively alleviate the influence of outliers and improve the robustness of the proposed fitting method.
2) We present an effective guided sampling algorithm, which exploits the motion consistency information to guide the sampling of effective minimal subsets and to improve the computational efficiency for model fitting.
3) We propose a novel model selection algorithm based on robust clustering analysis to effectively estimate the parameters of all model instances from input data involving a large number of outliers, without requiring a user-specified number of structures in advance.

In addition, we also build a new dataset suitable for electronic industrial applications. Experimental results in various industrial application scenarios show the significant advantages of the proposed method over several other competing methods.

The rest of this article is organized as follows: In Section II, we give an overview of the related work. In Section III, we describe the proposed fitting method in detail. In Section IV, we present the experimental results on both the electronic industrial application scenarios and the images from publicly available datasets. In Section V, concludes this article.

II. RELATED WORK

Many robust geometric model fitting methods have been proposed during the past few decades [7], [12], [13] and comprehensive surveys can be found in [14] and [15]. According to the key steps (i.e., model hypothesis generation and model selection) of a model fitting method, we roughly divide the related work into the following three categories.

Guided Sampling Based Methods in Model Fitting: The methods in this category aim to effectively generate model hypotheses by using different sampling algorithms (e.g., [7]–[9], [16], [17]) to improve fitting performance. Many guided sampling algorithms have been proposed to improve the quality of the generated model hypotheses based on various types of prior information, such as matching scores [12], spatial coherence information [9], and residual sorting information [13]. These guided sampling algorithms can efficiently generate good model hypotheses for simple data. However, for data with severe outliers, the prior information adopted in these sampling algorithms become less useful for sampling. In contrast, we adopt the motion consistency information in order to capture the reliable neighborhood sets of data to guide the sampling process, and thus we can obtain better sampling results for fitting. Unlike some methods that formulate motion consistency by using complex optimization strategies (such as optical flow estimation [8]),

we introduce a simple yet effective formulation of the motion consistency for model fitting.

Clustering Based Methods in Model Fitting: The methods in this category aim to develop different model selection algorithms (e.g., [10], [11], [18]–[22]) to improve fitting performance. They formulate the model selection problem as a clustering problem, and the related methods can be further divided into two kinds, i.e., model hypothesis clustering based methods and data clustering based methods. The methods of the first kind (e.g., AKSWH [20], MSHF [21]) cluster model hypotheses into different classes, each of which indicates a model instance in data. The methods of the second kind (e.g., T-Linkage [10], HOMF [11]) cluster data into different classes, and then estimate the parameters of different model instances from different classes. In addition, the collective density clustering method [23], which is specifically designed for motion segmentation, attempts to detect model instances from the point trajectories under heavy noise by finding density peaks of clusters as in [24]. Indeed, these clustering based methods have achieved promising progress on model selection, but with the increases of structure number and outlier ratio in input data, estimating the parameters of multiple model instances becomes problematic. In this article, we propose a novel model selection algorithm, which can effectively exploit the similarity values and weighting scores among input data for clustering data, to robustly handle multistructural data with a large number of outliers.

Alternating Optimization Based Methods in Model Fitting. The methods in this category alternatingly and continuously implement the two steps of model hypothesis generation and model selection to improve fitting performance. The main advantage of the methods (such as [5], [25]) in this category is that they can use the optimal results obtained by one step to optimize the other step, and thus improve the final fitting results. However, severe outliers in data may affect the effectiveness in both steps in practice, leading to suboptimal solutions. In contrast, the proposed method reduces the influence of outliers by using the neighborhood sets with motion consistency, thus improving the robustness in handling outliers.

III. PROPOSED METHOD

In this section, we first propose an effective guided sampling algorithm to generate model hypotheses. Then, we propose a novel model selection algorithm to estimate the parameters of all model instances in data. Finally, we summarize the complete proposed model fitting method.

A. Proposed Guided Sampling Algorithm

For two-view images of the same object, the neighboring feature matches usually have similar motions even in the presence of large viewpoint changes, since their corresponding feature points in one image often lie in an object or structure [19], [26]. Therefore, the magnitudes and directions of the motions of the neighboring true matches (inliers) from a model instance tend to vary within a limited range, while those of false matches (outliers) tend to vary randomly across images. Based on this

observation, we introduce an effective motion consistency constraint, which is used to find the neighboring data with motion consistency for each of input data, in order to accelerate the process of sampling all-inlier minimal subsets.

1) Neighborhood Sets With Motion Consistency: Suppose that we have obtained the input data $\mathcal{S} = \{\boldsymbol{s}_i\}_{i=1}^N$ of N feature matches, where $\boldsymbol{s}_i = (\boldsymbol{x}_i, \boldsymbol{y}_i)$ is a feature match (i.e., a datum), and $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are the pixel coordinates of two corresponding feature points in two-view images, respectively. Here, the motion vector of a feature match is denoted as $\boldsymbol{v}_i = \boldsymbol{y}_i - \boldsymbol{x}_i$. We generate neighborhood sets from $\mathcal{S}$ based on a motion consistency constraint in the following two steps.

In the first step, we aim to generate an initial neighborhood set for each datum from $\mathcal{S} = \{\boldsymbol{s}_i\}_{i=1}^N$ based on their motion magnitudes. For an image pair, let a datum $\boldsymbol{s}_i = (\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathcal{S}$ be a reference datum. We first, respectively, search for two K-nearest neighbors $\widetilde{\mathcal{N}}_{\boldsymbol{x}_i}$ and $\widetilde{\mathcal{N}}_{\boldsymbol{y}_i}$ for the two feature points $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ of the reference datum $\boldsymbol{s}_i$ under the Euclidean distance measure. If there is a datum $\boldsymbol{s}_j$, whose two corresponding feature points $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$, respectively, fall into $\widetilde{\mathcal{N}}_{\boldsymbol{x}_i}$ and $\widetilde{\mathcal{N}}_{\boldsymbol{y}_i}$, the lengths of the motion vectors of the data $\boldsymbol{s}_j$ and $\boldsymbol{s}_i$ are close to each other. Consequently, the data $\boldsymbol{s}_j$ and $\boldsymbol{s}_i$ can be regarded as the neighboring data with similar motion magnitudes. We then define the initial neighborhood set $\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}$ for $\boldsymbol{s}_i$ from $\mathcal{S}$ as

$$\widetilde{\mathcal{N}}_{\boldsymbol{s}_i} = \{\boldsymbol{s}_j = (\boldsymbol{x}_j, \boldsymbol{y}_j) | \boldsymbol{x}_j \in \widetilde{\mathcal{N}}_{\boldsymbol{x}_i}, \boldsymbol{y}_j \in \widetilde{\mathcal{N}}_{\boldsymbol{y}_i}, \boldsymbol{s}_j \in \mathcal{S}\} \quad (1)$$

where $|\widetilde{\mathcal{N}}_{\boldsymbol{x}_i}| = |\widetilde{\mathcal{N}}_{\boldsymbol{y}_i}| = K$ and $0 \leq |\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}| \leq K$.

In the second step, we aim to obtain the data with similar motion directions from the generated $\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}$ in order to generate the final neighborhood set for the reference datum $\boldsymbol{s}_i$. For this purpose, we use the cosine similarity measure to compute the similarity value of the motion directions between the reference datum $\boldsymbol{s}_i$ and one of its neighboring datum $\boldsymbol{s}_j$ from $\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}$ as follows: $c(\boldsymbol{v}_i, \boldsymbol{v}_j) = \frac{<\boldsymbol{v}_i, \boldsymbol{v}_j>}{\|\boldsymbol{v}_i\| \times \|\boldsymbol{v}_j\|}$, where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively, denote the inner product and the induced norm, and $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, respectively, denote the motion vectors of the data $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$. If the computed similarity value is larger than a threshold λ, the data $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ from $\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}$ are regarded as the neighboring data with similar motion directions. Based on the similar magnitudes and directions of the motions of the data from $\widetilde{\mathcal{N}}_{\boldsymbol{s}_i}$, for the datum $\boldsymbol{s}_i$, its final neighborhood set $\mathcal{N}_{\boldsymbol{s}_i}$, whose elements are consistent in motions, is defined as

$$\mathcal{N}_{\boldsymbol{s}_i} = \{\boldsymbol{s}_j | c(\boldsymbol{v}_i, \boldsymbol{v}_j) > \lambda, \boldsymbol{s}_j \in \widetilde{\mathcal{N}}_{\boldsymbol{s}_i}\}. \quad (2)$$

According to (1) and (2), we can generate a series of neighborhood sets with motion consistency (i.e., $\{\mathcal{N}_{\boldsymbol{s}_i}\}_{i=1}^N$) for all the data from $\mathcal{S}$. With the motion consistency constraint, the gross outliers in each neighborhood set are filtered as many as possible, while the inliers are reserved effectively. Thus, the data in the same neighborhood set have a high probability of belonging to the inliers of the same model instance. Therefore, we aim at sampling the minimal subsets based on the neighborhood sets with motion consistency for effectively and efficiently generating model hypotheses.

Algorithm 1: The Proposed Guided Sampling Algorithm.

1:    **Input:** Data $\mathcal{S} = \{s_i\}_{i=1}^N$, the minimal subset size p, parameters K and λ.

2:    **Output:** The generated model hypothesis set $\Theta = \{\theta_\ell\}_{\ell=1}^L$.

3:    Obtain N neighborhood sets $\{\mathcal{N}_{s_i}\}_{i=1}^N$ among $\mathcal{S}$ based on the motion consistency constraint by using (1) and (2).

4:    Generate a set of L seed data $\mathcal{I}_{seed} = \{\tilde{s}_\ell\}_{\ell=1}^L \in \mathcal{S}$ (described in Section III-A2).

5:    Initialization: $\Theta = \varnothing$.

6:    **for** $\ell = 1$ to L **do**

7:      Select a seed datum $\tilde{s}_\ell$ from $\mathcal{I}_{seed}$ as the first datum u_1 of a sampled minimal subset $\mathcal{P}_\ell$.

8:      Sample the datum that has the largest distance value from $\tilde{s}_\ell$ in the neighborhood set $\mathcal{N}_{\tilde{s}_\ell}$ as the second datum u_2 of $\mathcal{P}_\ell$.

9:      Randomly sample the other data $[u_3, \ldots, u_p]$ of $\mathcal{P}_\ell$ from the remaining data in the neighborhood set $\mathcal{N}_{\tilde{s}_\ell}$.

10:      Compute a model hypothesis θ_ℓ using the sampled minimal subset $\mathcal{P}_\ell = [u_1, u_2, \ldots, u_p]$.

11:      $\Theta = \Theta \cup \{\theta_\ell\}$.

12:    **end for**

---

2) Model Hypothesis Generation: In this section, we propose a novel guided sampling algorithm to sample minimal subsets based on the neighborhood sets with motion consistency. For sampling each of the minimal subsets, the proposed sampling algorithm consists of the two steps: It chooses the first datum (i.e., the seed datum) of a minimal subset by effectively selecting a potential inlier from input data. Then, it chooses the remaining data of the corresponding minimal subset in the neighborhood set of the chosen seed datum.

Specifically, for the two-view based model fitting problem, an inlier usually tends to be associated with more neighboring data in a corresponding neighborhood set than an outlier (see Fig. 1 for an example). Thus, we introduce a threshold ε to select the data from input data $\mathcal{S}$, whose numbers of neighboring data are larger than ε, as a potential inlier set $\mathcal{I}_{seed} \in \mathcal{S}$. We set the threshold ε as the minimal subset size p (e.g., $p = 8$ data for fundamental matrix estimation), in order to obtain the sufficient number of the data in a neighborhood set to yield a minimal subset. To improve the effectiveness of a sampled minimal subset, the seed datum of a minimal subset is chosen from $\mathcal{I}_{seed}$ during the sampling process.

Then, to select the remaining data of the minimal subset, we focus on the selection in the neighborhood set of the corresponding seed datum. Thus, in a corresponding neighborhood set, we select the datum that has the largest Euclidean distance value from the seed datum as the second datum of the minimal subset. This is because that a minimal subset with large spans will be more effective for computing a model hypothesis [27]. At last, the random sampling technique as done in RANSAC is used to select the other data of the minimal subset from the remaining

data in the corresponding neighborhood set, due to its simplicity and efficiency of implementation.

In the proposed sampling algorithm, we only sample the minimal subsets from the neighborhood sets with motion consistency, thus precluding a majority of gross outliers from being chosen. Meanwhile, the number of the sampled minimal subsets can be adaptively determined based on the number of the chosen potential inliers corresponding to the seed data, whose number is much smaller than that of the input data. Thus, we can effectively obtain good model hypotheses with a few sampled minimal subsets, by which the computational efficiency is significantly improved. We present the novel guided sampling algorithm as summarized in Algorithm 1.

B. Proposed Model Selection Algorithm

For model selection, some fitting methods introduce different clustering techniques to cluster the data of the same model instance, e.g., T-Linkage [10] uses an agglomerative clustering technique and HOMF [11] uses a spectral clustering technique. However, these methods are usually sensitive to severe outliers and hard to adaptively estimate the number of model instances. To relieve the above limitations, we propose a novel model selection algorithm by using robust cluster analysis based on the similarity between data and the weighting scores of the data. In general, a cluster center corresponding to a true model instance usually has a higher local weighting score than its neighbors, and a relatively smaller similarity value between a datum (that has a higher weighting score than the cluster center) and the cluster center itself. Thus, we first calculate the similarity values and local weighting scores of input data based on the generated hypotheses and neighborhood sets. Then, we search for the cluster centers and cluster the data by analyzing the weighting scores and the similarity for model selection.

To calculate the similarity values between two data, we first compute a residual value $r(s_i, \theta_\ell)$ between a datum s_i from $\mathcal{S}$ and a hypothesis θ_ℓ from the L generated model hypotheses $\{\theta_\ell\}_{\ell=1}^L$. The computed residual value is then mapped to a preference value using the preference function in [28] as

$$f_\ell^{(i)} = \begin{cases} \exp(-r^2(s_i, \theta_\ell)/\delta^2), & \text{if } r(s_i, \theta_\ell) < \tau_\ell \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where τ_ℓ is the inlier scale estimated by IKOSE [20], and δ is a normalization constant as in [28]. Consequently, the preference vector $f^{(i)}$ between a datum s_i and the L model hypotheses is written as $f^{(i)} = [f_1^{(i)}, f_2^{(i)}, \ldots, f_L^{(i)}]$. An $N \times L$ preference matrix $\mathbf{F}$ is then defined as $\mathbf{F} = [f^{(1)}, f^{(2)}, \ldots, f^{(N)}]^\mathsf{T}$. Based on the tensor theory [29], we then multiply the matrix $\mathbf{F}$ with its transpose $\mathbf{F}^\mathsf{T}$ to construct a similarity matrix $\mathbf{H} = \mathbf{F} * \mathbf{F}^\mathsf{T}$, whose element $h(s_i, s_j)$ denotes the similarity value between two data (i.e., two feature matches) s_i and s_j from $\mathcal{S}$.

To reduce the influence of gross outliers, the local weighting score of a datum $s_i \in \mathcal{S}$ is measured based on the similarity values between s_i and its neighboring data in a neighborhood set as: $\rho(s_i) = \frac{1}{T} \sum_{s_j \in \mathcal{N}_{s_i}} h(s_i, s_j)$, where $\mathcal{N}_{s_i}$ refers to a neighborhood set (described in Section III-A1), and T is a constant for the normalization of weighting scores. Since the

---

Algorithm 2: The Proposed Motion Consistency Guided Fitting (MCF) Method.

1: **Input:** Data $\mathcal{S} = \{s_i\}_{i=1}^N$, the minimal subset size p, parameters K, λ, τ and η.
2: **Output:** The number and the parameters of model instances.
3: Generate a model hypothesis set Θ by Algorithm 1.
4: Calculate a similarity matrix $\mathbf{H} = [h(s_i, s_j)]$ of the data $\mathcal{S}$ based on the generated hypothesis set Θ.
5: Estimate the weighting score $\rho(s)$ for each datum s of $\mathcal{S}$.
6: Sort the data in $\mathcal{S}$ by the descending order of ρ to obtain $\mathcal{S}^* = \{s^*\}_{i=1}^N$ and define s_1^* as the initial cluster center.
7: **for** $i = 1$ to N **do**
8:     Obtain the nearest neighbor $\hat{s}_j^*$ of s_i^* by using (4).
9:     **if** $h(s_i^*, \hat{s}_j^*) > \tau$ **then**
10:         Assign s_i^* to the cluster that $\hat{s}_j^*$ belongs to.
11:     **else**
12:         Assign s_i^* to be a new cluster center.
13:     **end if**
14: **end for**
15: Refine the clusters to obtain the final clustering results and distinguish inlier clusters from outlier ones (Section III-B).
16: Estimate the number and the parameters of model instances from input data based on the inlier clusters.

---

number of the data in a neighborhood set is not more than K, the value of T is set to the value of K used in Section III-A1.

In order to effectively find the cluster centers, we first sort the input data $\mathcal{S} = [s_1, s_2, \ldots, s_N]$ by the descending order of weighting scores. The sorted data are denoted as $\mathcal{S}^* = [s_1^*, s_2^*, \ldots, s_N^*]$. Then, we define the nearest neighbor datum $\hat{s}_j^*$ of a datum $s_i^* \in \mathcal{S}^*$, by finding the maximum similarity value between s_i^* and any datum $s_j^* \in \mathcal{S}^*$ that has a higher weighting score than s_i^*, which is written as follows:

$$\hat{s}_j^* = \begin{cases} \underset{s_j^* \in \mathcal{S}^*}{\mathrm{argmax}} \ h(s_i^*, s_j^*), \ \exists \ s_j^*, \ s.t. \ \rho(s_j^*) > \rho(s_i^*) \\ \underset{s_j^* \in \mathcal{S}^*}{\mathrm{argmin}} \ h(s_i^*, s_j^*), \ \text{otherwise.} \end{cases} \quad (4)$$

After that, each datum from $\mathcal{S}^*$ is assigned to a corresponding cluster center based on its nearest neighbor datum and their similarity values to obtain the clustering results. Specifically, the initial cluster center is identified as the datum $s_1^* \in \mathcal{S}^*$ with the maximum density value. In addition to the datum that is the initial cluster center, the other data from $\mathcal{S}^*$ are processed effectively by the descending order of weighting scores. That is, for a datum $s_i^* \in \mathcal{S}^*$, it is assigned to the cluster that its nearest neighbor $\hat{s}_j^*$ belongs to, when their similarity value $h(s_i^*, \hat{s}_j^*) > \tau$. Otherwise, the datum s_i^* is used as a new cluster center during the clustering process. As a result, the number of clusters can be determined automatically.

After the above clustering process, the cluster of the inliers belonging to a model instance tends to have a larger size with more data than the ones mainly consisting of outliers. Therefore, we regard the data in the clusters, whose sizes are less than η, as outliers. In contrast, the other clusters are regarded as the inlier clusters, each of which corresponds to a true model instance in data. Moreover, since a model instance may correspond to multiple clusters, we follow the common process as in [10] and [20] and refine the clusters by merging redundant clusters based on the residual distribution of the data in the clusters. In summary, the proposed model selection algorithm is able to adaptively estimate the number of model instances and in parallel estimate the parameters of each model instance by the clustering and merging processes.

We note that both [21] and our method employ the mode-seeking-based clustering strategy to handle the robust model fitting problem. However, there are significant differences between them: 1) the authors in [21] seek modes from the generated model hypotheses in the parameter space to represent the estimated model instances. However, such a way may lead to suboptimal fitting results, especially when the generated model hypotheses do not contain all model instances. In contrast, we seek modes directly from input data to obtain data clustering results, from which we effectively estimate the parameters of all model instances. Thus, our method is able to alleviate the influence of the generated insignificant model hypotheses on the performance of model fitting, and it is more effective for the estimation of multiple model instances. 2) In [21], the authors derive inliers/outliers dichotomies corresponding to each of the estimated model instances according to the user-specified inlier noise scales. In contrast, we directly segment inliers and outliers according to the data clustering results, by which the sensitivity to inlier noise scales for the segmentation of data can be significantly alleviated. Therefore, our method can obtain more accurate fitting and segmentation results in various industrial applications for data with severe outliers.

C. Complete Method

With all the components developed in the previous sections, we summarize the complete proposed motion consistency guided fitting (MCF) method in Algorithm 2. The proposed MCF starts from generating a series of neighborhood sets from input data by leveraging an effective motion consistency constraint. We also propose a novel guided sampling algorithm, which efficiently samples a small number of effective minimal subsets from the obtained neighborhood sets, to generate high-quality model hypotheses for model selection. Finally, by analyzing the similarity values and weight scores of data, we propose a novel model selection algorithm to effectively find inlier clusters, from which we can estimate the parameters of all model instances in data.

The computational complexity of our MCF is mainly governed by searching the K-nearest neighbors of each feature point from input data $\mathcal{S}$ during the sampling process (i.e., Step 1 of Algorithm 1) and computing the similarity matrix of the input data during the model selection process (i.e., Step 2 of

Algorithm 2). In contrast, the other steps of MCF take much less time than the above two steps. For these two steps, the time complexity of searching the K-nearest neighbors for each feature point in $\mathcal{S}$ and computing the similarity matrix of the input data are approximately $O((K + N) \log N)$ and $O(LN^2)$, respectively. Here, N is the number of feature matches, and L is the number of the generated model hypotheses. Thus, the total complexity of MCF is about $O((K + N) \log N) + O(LN^2)$.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, to evaluate the performance of the proposed fitting method (MCF), we first design two experiments for retinal disease diagnosis and industrial defect inspection. Then, we evaluate the proposed MCF on two publicly available datasets. In addition, we also conduct experiments on various image pairs for analyzing the components of the proposed MCF. In the above experiments, the proposed MCF is evaluated with several other state-of-the-art fitting methods (including T-Linkage [10], RCMSA [9], RansaCov [18], MSHF [21], HOMF [11], and HRMP [22]). In order to evaluate fitting results, following the common practice in [18] and [22], the fitting error is computed by measuring the ratio between the number of mislabeled data and the total number of data.

A. Experiments on Fundus Retinal Images

Image registration is an important task in medical imaging systems, where patients are examined and diagnosed by images. In retina analysis, microscopic differences in fundus retinal images are significant in making better medical decisions. The purpose of retinal image registration is to assist ophthalmologists to obtain more comprehensive details of the retinal structure for disease diagnosis. Therefore, it is of great significance to obtain accurate registration results from retinal images. Although many intensity-based methods and feature-based methods have been proposed recently, the registration of retinal images is still a challenging task due to the complex nature of retinal images (e.g., unhealthy regions) [30]. In this article, we mainly focus on feature-based methods, which first extract feature-based matches between image pairs and then apply a geometric transformation model for registration.

To evaluate the performance of the proposed MCF on retinal images, we develop an experimental system for retinal disease diagnosis, as shown in Fig. 2. The experimental system mainly includes a Zeiss FF 450plus Fundus Camera for collecting fundus retinal images, and a workstation with an Intel i7-7700 CPU @3.6GHz and 16GB RAM for running algorithms. An example of the image registration result obtained by MCF on a pair of multimodality retinal images is also shown in Fig. 3. The pair of multimodality images are captured with the color fundus photograph and the fluorescein angiography by using the Zeiss fundus camera device [see Fig. 3(a)]. It is worth noting that color fundus images mainly reveal the whole appearance of the retinal surface [see Fig. 3(c)]. In contrast, fluorescein angiography images mainly provide the information of blood vessels [see Fig. 3(d)]. In the pair of multimodality retinal images, we first extract features from each image to generate a set
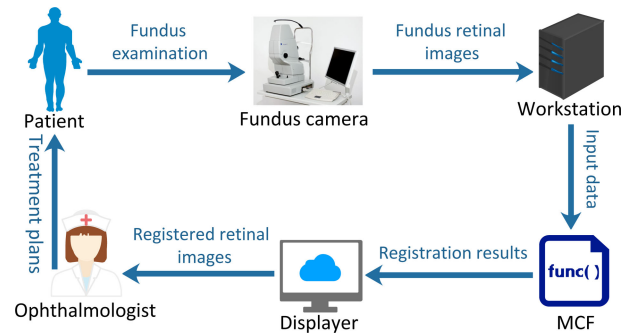

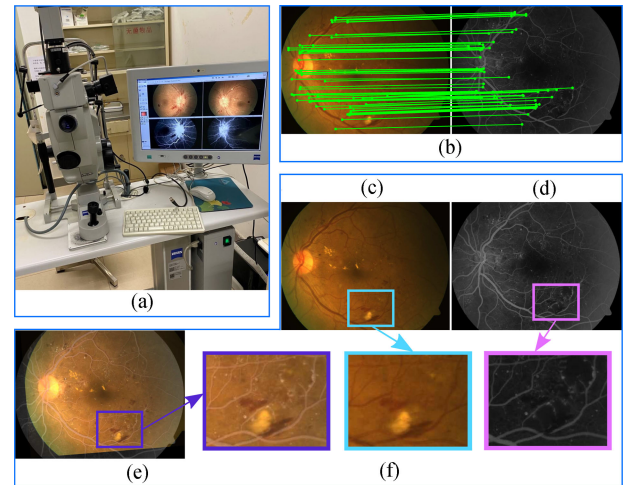Fig. 2.   Experimental system for fundus retinal disease diagnosis.

Fig. 3.   Example of image registration on a pair of multi-modality retinal images collected from a Zeiss FF 450plus Fundus Camera.

of feature matches. Then, we apply the proposed MCF to identify the correct matches for model estimation. The fitting result is shown in Fig. 3(b). After that, we estimate the parameters of an affine matrix from the identified correct matches (inliers) in order to obtain the corresponding geometric transformation between the image pair. Finally, by the robust estimation of the affine matrix, we can obtain the accurate registration result, as shown in Fig. 3(e). With accurate registration, the information gained from the pair of multimodality retinal images can be used in a complementary manner to provide additional insight for retinal disease analysis [as shown in Fig. 3(f)].

To obtain quantitative results, we compare the proposed MCF with the aforementioned six state-of-the-art fitting methods on retinal images. Since there are few publicly available datasets for model fitting on multimodality retinal images, we built a new multi-modality retinal image (MMRI) dataset for evaluation. The MMRI dataset consists of 15 pairs of retinal images that involve a number of challenges, such as large textureless regions and significant illumination variations. As a result, the obtained feature matches between the image pairs contain a large number of outliers. As retinal images are usually captured under different conditions, the image pairs in the MMRI dataset are divided into three categories, i.e., Category $\mathcal{A}$ containing five pairs of images

TABLE I
QUANTITATIVE RESULTS ON THE RETINAL IMAGE PAIRS
FROM THE MMRI DATASET

Data	Outliers		T-Linkage	RCMSA	RansaCov	MSHF	HOMF	HRMP	MCF
Retina_A1	43.17%	Mean	16.55	4.17	6.47	7.05	7.16	5.61	**1.44**
		Time	17.27	2.26	1.14	11.97	2.10	11.62	**0.20**
Retina_A2	78.23%	Mean	21.59	10.70	8.82	5.27	9.71	4.52	**1.37**
		Time	55.23	4.84	7.20	25.82	4.64	31.99	**0.70**
Retina_A3	47.46%	Mean	13.14	5.59	5.17	6.36	16.42	7.29	**2.92**
		Time	33.88	3.58	2.67	17.59	3.16	18.08	**0.50**
Retina_A4	54.21%	Mean	24.92	8.88	5.14	4.95	5.39	4.30	**2.94**
		Time	28.71	3.26	2.50	16.64	2.89	16.60	**0.36**
Retina_A5	52.68%	Mean	20.24	12.50	7.62	9.76	5.81	**5.24**	5.71
		Time	55.35	4.76	6.25	24.54	4.57	25.17	**0.73**
Retina_B1	72.30%	Mean	15.63	10.05	10.14	8.33	6.94	7.37	**5.79**
		Time	25.76	3.28	2.52	15.61	2.82	16.56	**0.35**
Retina_B2	55.61%	Mean	28.78	6.24	7.32	5.56	5.50	9.46	**2.49**
		Time	23.54	3.12	2.37	15.68	2.77	16.16	**0.36**
Retina_B3	85.53%	Mean	46.67	17.19	13.79	16.68	6.62	7.91	**4.51**
		Time	30.68	3.66	2.96	17.88	3.50	17.71	**0.36**
Retina_B4	85.00%	Mean	19.72	24.33	18.17	21.17	14.44	14.58	**12.33**
		Time	28.57	3.73	2.99	18.42	3.51	18.71	**0.52**
Retina_B5	44.01%	Mean	27.55	18.22	12.31	11.24	**8.10**	9.67	9.42
		Time	102.83	6.96	12.15	36.62	5.74	35.20	**1.21**
Retina_C1	83.46%	Mean	25.20	25.98	9.76	7.24	15.98	6.46	**4.41**
		Time	14.16	2.19	0.99	10.60	1.99	10.78	**0.20**
Retina_C2	87.32%	Mean	6.88	21.09	8.19	13.04	7.39	4.64	**4.13**
		Time	37.11	3.97	4.20	19.67	3.98	20.93	**0.48**
Retina_C3	59.16%	Mean	27.12	11.51	9.07	8.23	16.87	7.97	**4.21**
		Time	46.40	4.59	5.33	22.12	4.32	23.30	**0.56**
Retina_C4	63.96%	Mean	16.41	11.78	10.15	6.09	8.37	7.11	**5.38**
		Time	22.63	3.16	2.20	15.15	2.75	18.45	**0.28**
Retina_C5	72.55%	Mean	15.69	12.55	11.24	9.54	11.59	7.45	**5.69**
		Time	17.90	2.45	1.33	12.00	2.66	12.69	**0.23**
Total		Avg.	21.74	13.39	9.56	9.37	9.75	7.31	**4.85**
		Std.	9.21	6.65	3.41	4.57	4.20	**2.62**	2.90

The best results are boldfaced.

with partial overlaps, Category $\mathcal{B}$ containing five pairs of images with different modalities, and Category $\mathcal{C}$ containing five pairs of images with different modalities and partial overlaps.

We repeat each experiment 50 times on the image pairs from the MMRI dataset for the seven competing methods, and show the fitting results in Table I. As can be seen, MCF achieves the lowest mean fitting errors in 13 out of 15 image pairs, and it uses the lowest CPU time (in seconds) for each test data among all the seven competing methods. MCF also achieves the lowest average fitting error and the second lowest standard variance of fitting errors for the overall results. This is because that MCF takes advantage of the motion information among feature matches to significantly reduce the influence of severe outliers for robust model fitting, and thus it achieves better fitting accuracy than the other six competing methods.

B. Experiments on Printed Circuit Board (PCB) Images

PCBs are commonly used in electronic devices in the industrial field. Defect inspection is a critical part in the production process of PCBs. Automatic optic inspection systems, in which image registration plays an important role, are widely applied in the defect inspection of PCBs due to their efficiency and reliability [31]. However, electronic PCBs often include complex scenes with repeating elements (see Fig. 4). As a result, the obtained sets of feature matches between image pairs often contain a large number of feature matches and could be corrupted heavily by outliers. Therefore, it requires a fitting method to be highly robust to outliers and efficient to estimate accurate model parameters for image registration in defect inspection of PCBs.

TABLE II
QUANTITATIVE RESULTS ON THE ADELAIDERMF DATASET FOR
FUNDAMENTAL MATRIX ESTIMATION AND ON THE SNU DATASET FOR AFFINE
MATRIX ESTIMATION. THE BEST RESULTS ARE BOLDFACED

Datasets		T-Linkage	RCMSA	RansaCov	MSHF	HOMF	HRMP	MCF
Adelaide-RMF	Avg.	13.43	8.05	14.72	10.71	6.59	7.37	**4.01**
	Std.	11.78	4.65	8.88	8.30	5.53	6.90	**4.64**
	Time	39.76	1.22	15.54	17.10	2.30	16.14	**0.56**
SNU	Avg.	16.40	11.60	13.87	15.31	13.02	12.76	**4.57**
	Std.	7.84	3.21	3.86	4.13	4.29	7.45	**2.70**
	Time	95.65	3.38	10.42	33.96	5.79	35.25	**1.51**

To obtain qualitative results, we apply the aforementioned seven fitting methods on two representative image pairs of PCBs (called *PCBs _ A* and *PCBs _ B*) for image registration. Some qualitative results are shown in Fig. 4. Note that the number of feature matches (and outlier ratios) on the image pairs of *PCBs _ A* and *PCBs _ B* are 2514 (83.25%) and 2157 (14.60%), respectively. For the *PCBs _ A* image pair, the proposed MCF obtains the best fitting results, and it can correctly identify the inliers in the test data. HRMP obtains the second best fitting results, due to the fact that some false matches are incorrectly identified as inliers (an example is shown in Fig. 4). From the results, the proposed MCF can accurately estimate the parameters of an affine matrix and it precisely aligns the *PCBs _ A* image pair with satisfactory registration (which are highlighted with the purple rectangles in Fig. 4), compared to HRMP. The registration results obtained by HRMP still includes distortions and "ghosting" (which are highlighted with the red ellipses in Fig. 4). For the other five methods, the registration results obtained by HOMF are slightly better than those of T-Linkage, RCMSA, RansaCov, and MSHF, but less accurate than those of MCF. For the *PCBs _ B* image pair, the best and the second best fitting results are, respectively, obtained by MCF and HRMP, and there are no false matches (outliers) that are incorrectly identified as inliers. However, the proposed MCF can detect more inliers than HRMP, as shown in Fig. 4. Thus, the registration process in the proposed MCF is more effective than that of HRMP. The registration results obtained by the seven methods on the *PCBs _ B* image pair are not shown, due to the space limit. In short, the proposed MCF can robustly estimate the parameters of the geometric model for image registration of *PCBs* and it is highly robust to outliers.

C. Experiments on Two Publicly Available Datasets

We also compare the proposed MCF with the six state-of-the-art methods mentioned above for the task of fundamental matrix estimation on the AdelaideRMF dataset [13] and for the task of affine matrix estimation on the SNU dataset [32].

We report the results on the AdelaideRMF and SNU datasets in Table II, and we show some fitting results obtained by MCF in Fig. 5. As we can see, the proposed MCF achieves the lowest average fitting errors and the lowest standard deviations of fitting errors on the two datasets among all the competing methods. Moreover, MCF has significantly improved the computational efficiency compared to the other methods. Specifically, the average CPU time used by MCF (i.e., 0.56/1.51*s*) is much less than that used by the second fastest RCMSA (i.e., 1.22/3.38*s*)
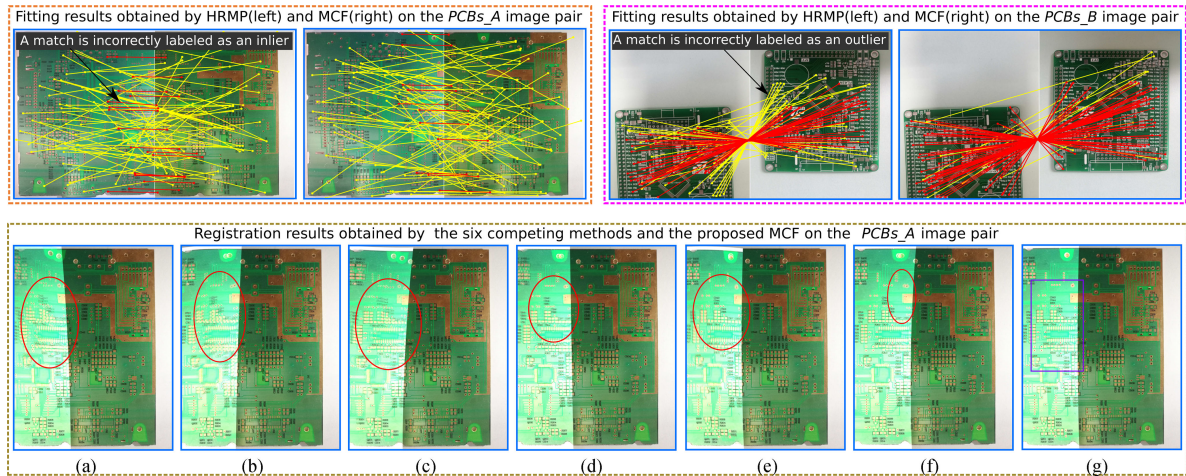
Fig. 4. Results obtained by the seven fitting methods on two image pairs (i.e., *PCBs _ A* and *PCBs _ B*). In the first row, we show the best and second best fitting results, respectively, obtained by MCF and HRMP on the image pairs. For clarity, only 80 matches including inliers (red lines) and outliers (yellow lines) are randomly selected for each case. In the second row, we only show the registration results obtained by the seven methods on the *PCBs _ A* image pair, due to the space limit. The red ellipses highlight distortions or errors. The purple rectangles highlight satisfactory registration.
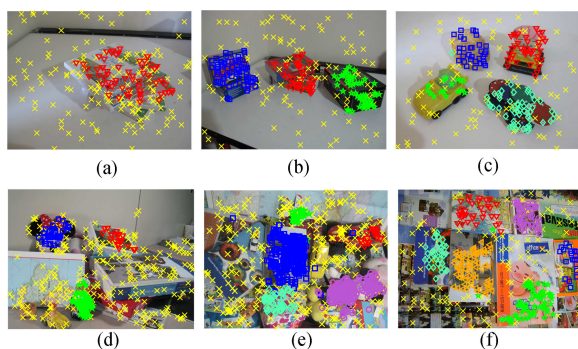

Fig. 5. Some fitting results obtained by the proposed MCF. (a)–(c) and (d)–(f), respectively, show the results on the image pairs from the AdelaideRMF dataset for fundamental matrix estimation and on the image pairs from the SNU dataset for affine matrix estimation (only one of two views is shown). The inliers of different model instances are marked in different colors. The outliers are marked in the yellow color.
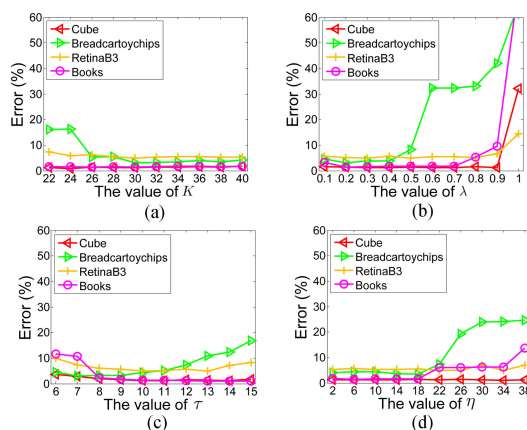

Fig. 6. Mean fitting errors obtained by the proposed MCF with different parameter values on the four representative image pairs.
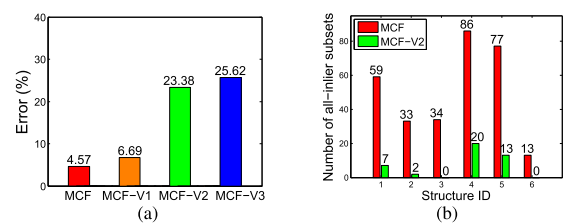

Fig. 7. Comparison results for MCF and its three variants on the SNU dataset. (a) The overall average fitting errors obtained by the four methods. (b) The number of all-inlier subsets for each structure obtained by MCF and MCF-V2 on the *Books* image pair, respectively.

on the two datasets, respectively. The reason is that MCF only generates a small number of model hypotheses with a high ratio of good model hypotheses to cover all true model instances in data, and it also benefits from the effectiveness of the proposed model selection algorithm. T-Linkage, RansaCov, MSHF, and HRMP are significantly slower than the proposed MCF on both datasets. This is because that it is difficult for these methods to sample all-inlier minimal subsets during the sampling processes when the data contain severe outliers, and these methods need to sample a large number of minimal subsets in order to hit a true model instance. As a result, these methods spend much more CPU time than the proposed MCF during the sampling processes for model fitting. Compared to T-Linkage (which uses an agglomerative clustering technique for model selection) and HOMF (which uses a spectral clustering technique for model selection), MCF uses the proposed model selection algorithm based on robust cluster analysis, and it achieves great improvements in fitting accuracy. The comparison results further show the superiority of the proposed MCF over the other six methods in both fitting accuracy and computational speed. It is worth pointing out that only T-Linkage, RCMSA, MSHF, HRMP and the proposed MCF can automatically estimate the number of
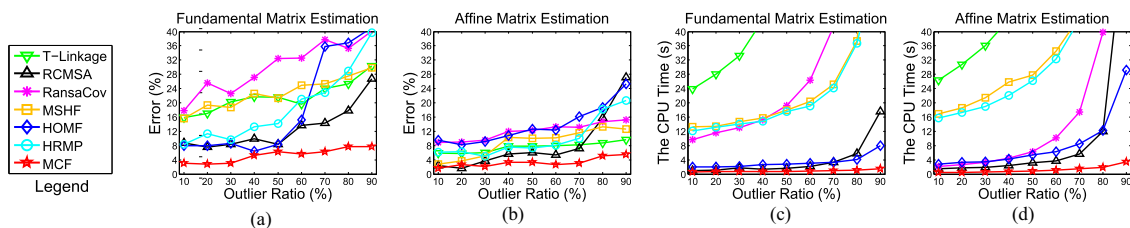
Fig. 8. Mean fitting errors and the mean CPU time obtained by the competing methods on two image pairs with different outlier ratios. (a) and (c) show the quantitative comparison on the *Breadcartoychips* image pair. (b) and (d) show the quantitative comparison on the *Minnies* image pair.

model instances. In contrast, the remaining two methods (i.e., RansaCov and HOMF) require a user-specified model instance number, which is not known in advance in many practical industrial applications.

D. Analysis for the Proposed MCF

1) Parameter Analysis and Settings: There are four parameters used in the proposed MCF, i.e., K is for the generation of the neighborhood sets with motion consistency, λ is for the selection of the neighboring data with similar motion directions, τ is the cut-off threshold for clustering, and η is the minimal cluster size. We test different values of the parameters for fundamental matrix estimation on the two image pairs (i.e., *Cube* and *Breadcartoychips*) and for affine matrix estimation on the other two image pairs (i.e., *RetinaB3* and *Books*). We show the experimental results obtained by MCF in Fig. 6.

As we can see, when the value of K is set to be too small, MCF obtains large fitting errors on some image pairs, mainly due to the lack of sufficient data in a neighborhood set for sampling. However, when the value of K is set to be too large, a generated neighborhood set will include sufficient data but it also contains many outliers for sampling. Overall, MCF can obtain low fitting errors on all image pairs, when the value of K ranges from 30 to 40.

For the parameter λ, we can see that MCF achieves low fitting errors on the four image pairs when the value of λ is within 0.2–0.4. In contrast, when its value is larger than 0.4, MCF may filter out too many feature matches, including inliers during the process of generating neighborhood sets. This situation will lead to yielding the insufficient effective sampled minimal subsets for model fitting.

For the cut-off threshold τ, we can see that if we set the value of τ too large or too small, MCF obtains relatively high fitting errors for most cases. The reason is that when τ is too large, many inliers may be wrongly removed in the clustering process, while when τ is too small, clusters may undesirably involve many outliers, leading to bad fitting results.

Finally, for the minimum clustering size η, the fitting errors remain relatively stable on all four image pairs when its value is increased from 2 to 18, while the fitting errors become relatively high when the value of η is larger than 18. The results show that the proposed MCF can accurately distinguish the inlier clusters from the outlier ones that consist of few data by using a proper threshold. Therefore, based on the experimental results, we set

the default values of these parameters as $K = 30$, $\lambda = 0.3$, $\tau = 10$, and $\eta = 10$ for the proposed method in all the experiments.

2) Ablation Study: Three variants are presented to analyze the influence of different components of MCF. The three variants of MCF are 1) MCF-V1: MCF without using the local weighting scores in the proposed model selection algorithm. It means that MCF-V1 measures the weighting score of each datum based on the whole data set. 2) MCF-V2: MCF without using the proposed sampling algorithm. MCF-V2 samples minimal subsets from the whole dataset by random sampling. 3) MCF-V3: MCF without using both of the abovementioned components. MCF-V3 implements the steps of model selection and model hypothesis generation without using the motion consistency constraint.

We evaluate the performance of MCF and its three variants on the SNU dataset and report the quantitative results as shown in Fig. 7(a). From the experimental results, we can see that MCF-V1 that does not use the local weighting scores has an increase in the average fitting error by 2.12%, while MCF-V2 that does not use the proposed sampling algorithm has an increase in the average fitting error by 18.81%. Therefore, the proposed sampling algorithm contributes more to improve the fitting performance. To show the effectiveness of the proposed guided sampling algorithm, we provide an example about the number of the sampled all-inlier minimal subsets obtained by MCF and MCF-V2 on the *Books* image pair, as shown in Fig. 7(b). In addition, MCF-V3 that does not use the motion consistency constraint has an increase in the average fitting error by 21.05%. These comparison results show that the adopted motion consistency constraint among data is the most crucial information used in MCF and it plays an important role in improving the performance of model fitting.

3) Influence of Different Gross Outlier Ratios: To show the robustness of the proposed fitting method, we evaluate the performance of the seven competing methods on multistructural data with different ratios of gross outliers. We change the number of gross outliers to each test dataset and fix the number of inliers for experiments. We report the quantitative results on two image pairs with different gross outlier ratios, as shown in Fig. 8. From Fig. 8(a) and (b), we can see that MCF does not show large fluctuations in the mean fitting errors and it still achieves low mean fitting errors at high outlier ratios compared with the other fitting methods. This is because that the proposed model selection algorithm can effectively handle multistructural data with severe outliers. From Fig. 8(c) and (d), we can see that the CPU time used by MCF shows no significant changes when

the outlier ratios are increased, and MCF is much faster than the other competing methods, which can be mainly attributed to the effectiveness and efficiency of the proposed guided sampling algorithm. In summary, these experimental results show the robustness of MCF with regard to different gross outlier ratios.

V. CONCLUSION

In this article, we proposed a robust and efficient MCF method for geometric model fitting. In contrast to previous methods, the proposed MCF can significantly reduce the influence of outliers by leveraging the motion consistency among feature matches for model fitting. Specifically, from input data, we first generated a series of neighborhood sets, whose corresponding elements were consistent in motions. Based on the obtained neighborhood sets, we then proposed a novel sampling algorithm, which can increase the probability of sampling all-inlier minimal subsets to efficiently generate good model hypotheses. Moreover, by robust clustering analysis, we proposed a novel model selection algorithm, which can effectively estimate the number and the parameters of model instances in data. Benefiting from both the proposed sampling algorithm and the proposed model selection algorithm, MCF was able to robustly deal with multistructural data even in the presence of severe outliers. Extensive experiments in various industrial application scenarios showed the superiority of the proposed MCF over several other state-of-the-art model fitting methods.

REFERENCES

[1] M. L. Balter, A. I. Chen, T. J. Maguire, and M. L. Yarmush, "Adaptive kinematic control of a robotic venipuncture device based on stereo vision, ultrasound, and force guidance," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 1626–1635, Feb. 2017.

[2] E. P. van Horssen, J. A. van Hooijdonk, D. Antunes, and W. Heemels, "Event-and deadline-driven control of a self-localizing robot with vision-induced delays," *IEEE Trans. Ind. Electron.*, vol. 67, no. 2, pp. 1212–1221, Feb. 2020.

[3] X. Liu, Y. Ai, B. Tian, and D. Cao, "Robust and fast registration of infrared and visible images for electro-optical pod," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1335–1344, Feb. 2018.

[4] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.

[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] Y. Wang, Y. Liu, X. Li, C. Wang, M. Wang, and Z. Song, "Gorflm: Globally optimal robust fitting for linear model," *Signal Process.: Image Commun.*, vol. 84, 2020, Art. no. 115834.

[7] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[8] K. Ni, H. Jin, and F. Dellaert, "Groupsac: Efficient consensus in the presence of groupings," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 2193–2200.

[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] 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.*, 2014, pp. 3954–3961.

[11] S. Lin, G. Xiao, Y. Yan, D. Suter, and H. Wang, "Hypergraph optimization for multi-structural geometric model fitting," in *Proc. Assoc. Adv. Artif. Intell.*, 2019, pp. 8730–8737.

[12] O. Chum and J. Matas, "Matching with prosac-progressive sample consensus," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 220–226.

[13] H. S. Wong, T.-J. Chin, J. Yu, and D. Suter, "Dynamic and hierarchical multi-structure geometric model fitting," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 1044–1051.

[14] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "Usac: A universal framework for random sample consensus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 2022–2038, Aug. 2012.

[15] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, no. 1, pp. 23–79, 2021.

[16] Y. Jiao, Y. Wang, X. Ding, B. Fu, S. Huang, and R. Xiong, "2-entity ransac for robust visual localization: Framework, methods and verifications," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4519–4528, May 2021.

[17] 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.

[18] L. Magri and A. Fusiello, "Multiple model fitting as a set coverage problem," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3318–3326.

[19] X. Jiang, J. Ma, J. Jiang, and X. Guo, "Robust feature matching using spatial clustering with heavy outliers," *IEEE Trans. Image Process.*, vol. 29, pp. 736–746, Aug. 2020.

[20] H. Wang, T.-J. Chin, and D. Suter, "Simultaneously fitting and segmenting multiple-structure data with outliers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1177–1192, Jun. 2012.

[21] 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. 2019.

[22] S. Lin, X. Wang, G. Xiao, Y. Yan, and H. Wang, "Hierarchical representation via message propagation for robust model fitting," *IEEE Trans. Ind. Electron.*, doi: 10.1109/TIE.2020.3018074.

[23] Y. Wu, Y. Ye, C. Zhao, and Z. Shi, "Collective density clustering for coherent motion detection," *IEEE Trans. Multimedia*, vol. 20, no. 6, pp. 1418–1431, Jun. 2018.

[24] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[25] K. H. Lee and S. Wook Lee, "Deterministic fitting of multiple structures using iterative maxfs with inlier scale estimation," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 41–48.

[26] J. Bian *et al.*, "GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence," *Int. J. Comput. Vis.*, vol. 128, no. 6, pp. 1580–1593, 2020.

[27] 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, 2014.

[28] M. Tepper and G. Sapiro, "Nonnegative matrix underapproximation for robust multiple model fitting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2059–2067.

[29] V. M. Govindu, "A tensor decomposition for geometric grouping and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 1150–1157.

[30] J. Ma, J. Jiang, C. Liu, and Y. Li, "Feature guided gaussian mixture model with semi-supervised em and local geometric constraint for retinal image registration," *Inf. Sci.*, vol. 417, pp. 128–142, 2017.

[31] M. Annaby, Y. Fouda, and M. 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.

[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.*, 2008, pp. 144–157.

Hanlin Guo is currently working toward the Ph.D. degree in computer science and technology with the School of Informatics, Xiamen University, Xiamen, China.

His research interests mainly focus on geometric model fitting, image registration, video analysis and motion detection. His research interests also focus on computer vision and pattern recognition.

Yang Lu received the M.S. degree in software engineering from Macau University, Macau, China, in 2014, and the Ph.D. degree in computer science from Hong Kong Baptist University, Hong Kong, in 2019.

He is currently an Assistant Professor with the Department of Computer Science, School of Informatics, Xiamen University, China. His research interests include imbalanced data learning, clustering, ensemble learning, and online learning.

Haifang Zhang received the M.S. degree in ophthalmology from Hebei Medical University, Shijiazhuang, China, in 2004.

She is currently an Associate Chief Physician with the Department of Ophthalmology, Hebei General Hospital, China. Her research interests include retinal fundus image analysis, and retinal disease diagnosis.

Haosheng Chen received the M.S. degree in software engineering from Zhengzhou University, Zhengzhou, China, in 2014. He is currently working toward the Ph.D. degree in computer science and technology with the School of Informatics, Xiamen University, Xiamen, China.

His research interests include computer vision and intelligent systems.

Hanzi Wang (Senior Member, IEEE) received the Ph.D. degree in computer vision from Monash University, Melbourne, VIC, Australia.

He has been awarded the Douglas Lampard Electrical Engineering Research Prize and Medal for the best Ph.D. thesis at Monash University. He is currently a Distinguished Professor of "Minjiang Scholars" in Fujian province and a Founding Director of the Center for Pattern Analysis and Machine Intelligence (CPAMI) with Xiamen University, China. His research interests are concentrated on computer vision and pattern recognition including visual tracking, robust statistics, object detection, video segmentation, model fitting, optical flow calculation, 3-D structure from motion, image segmentation and related fields.

Hailing Luo received the M.S. degree in computer science and technology from Xiamen University, Xiamen, China, in 2020.

Her research interests mainly focus on geometric model fitting and image registration. Her research interests also focus on computer vision and machine learning.

Guobao Xiao received the B.S. degree in information and computing science from Fujian Normal University, Fuzhou, China, in 2013 and the Ph.D. degree in computer science and technology from Xiamen University, Xiamen, China, in 2016.

He is currently a Professor with Minjiang University, Fuzhou, China. He has authored and coauthored more than 30 papers in the international journals and conferences including IEEE TPAMI/TIP/TITS/TGRS/TIE, IJCV, PR, ICCV, ECCV, AAAI, etc. His research interests include computer vision and pattern recognition. He has been awarded the best Ph.D. thesis award in China Society of Image and Graphics (a total of ten winners in China).

Prof. Xiao also served on the Program Committee (PC) of CVPR, ICCV, ECCV, AAAI, WACV, etc.