



Compounding Local Invariant Features and Global Deformable Geometry for Medical Image Registration

Jianhua Zhang^{1*}, Lei Chen¹, Xiaoyan Wang¹, Zhongzhao Teng², Adam J. Brown³, Jonathan H. Gillard², Qiu Guan¹, Shengyong Chen^{1*}

1 College of Computer Science, Zhejiang University of Technology, Hangzhou, Zhejiang, China, **2** Department of Radiology, University of Cambridge, Cambridge, United Kingdom, **3** Division of Cardiovascular Medicine, University of Cambridge, Cambridge, United Kingdom

Abstract

Using deformable models to register medical images can result in problems of initialization of deformable models and robustness and accuracy of matching of inter-subject anatomical variability. To tackle these problems, a novel model is proposed in this paper by compounding local invariant features and global deformable geometry. This model has four steps. First, a set of highly-repeatable and highly-robust local invariant features, called Key Features Model (KFM), are extracted by an effective matching strategy. Second, local features can be matched more accurately through the KFM for the purpose of initializing a global deformable model. Third, the positional relationship between the KFM and the global deformable model can be used to precisely pinpoint all landmarks after initialization. And fourth, the final pose of the global deformable model is determined by an iterative process with a lower time cost. Through the practical experiments, the paper finds three important conclusions. First, it proves that the KFM can detect the matching feature points well. Second, the precision of landmark locations adjusted by the modeled relationship between KFM and global deformable model is greatly improved. Third, regarding the fitting accuracy and efficiency, by observation from the practical experiments, it is found that the proposed method can improve 6~8% of the fitting accuracy and reduce around 50% of the computational time compared with state-of-the-art methods.

Citation: Zhang J, Chen L, Wang X, Teng Z, Brown AJ, et al. (2014) Compounding Local Invariant Features and Global Deformable Geometry for Medical Image Registration. PLoS ONE 9(8): e105815. doi:10.1371/journal.pone.0105815

Editor: Jie Tian, Institute of Automation, Chinese Academy of Sciences, China

Received: May 7, 2014; **Accepted:** July 20, 2014; **Published:** August 28, 2014

Copyright: © 2014 Zhang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. Data may be found using the link <http://pan.baidu.com/s/1gd9ra51>.

Funding: This study was supported by Zhejiang Provincial Natural Science Foundation of China grants R1110679, and Natural Science Foundation of China grants 61305021 (<http://www.nsf.gov.cn>), and 61325019. The funder of Zhejiang Provincial Natural Science Foundation of China had a role in the data collection and analysis. The funder of Natural Science Foundation of China had a role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: zjh@ieee.org (JHZ); sy@ieee.org (SYC)

Introduction

The majority of medical image processing tasks, including fMRI analysis [1], 3D reconstruction [2] [3], and medical image segmentation [4] [5] [6], rely heavily on accurate image co-registration. To date, many methods exist that can help perform image co-registration (e.g. [7] [8] [9] [10] [11]). A comprehensive review summarizing these methodologies has been performed in [12], which divides them into two main categories; geometric image features-based and voxel similarity measures-based. Among these registration methods, active contour models (ACM) [13], active shape models (ASM) [14] and active appearance models (AAM) [15] are particularly important methods. They are widely employed and have greatly improved medical image analysis, due to their excellent non-rigid deformable performance [16] [17] [18] [19].

However, there are still several critical issues to be resolved. Two recent surveys, [20] and [21], summarized the deformable model and medical image segmentation respectively, and identified problems, such as model initialization, and the robustness and accuracy for inter-subject anatomical variability, both of which need to be improved.

The results of deformable model rely heavily on its initial position. If the initial position of the model is close to the region of interest, fitting results will be more accurate and the evolution time will be shorter. On the other hand, if the initial position is far away, it can lead to failure, due to the limited local searching algorithm of deformable models. Therefore, it may be worth investigating how to guarantee that the initial model is within the searching range. However, most attempts at improving deformable models have ignored this problem and left users to manually initialize it. Few research projects involve automatic initialization, and are limited to special applications ([22] [23] [24]) or unique deformable models ([25] [26]). Garrido and Qerez proposed a reformulated Hough transform to initialize the deformable model for cell segmentation [22], but there were some cells which could not be correctly found in the image, and the algorithm did not consider the performance of initialization when the target image was deformed. In [23], the authors implemented a fast 3D generalized Hough transform based on prior information of heart position, which is designed for a cardiac model. Wang et al. [24] proposed to use irises to initialize ASM for use in facial recognition. The initialization presented in [25] is limited to ACM and can not be extended to ASM or AAM. In [26], the

initializing method is constrained to an improved ASM, resulting in an exhaustively search of a whole image to obtain good initialization. In [27], a method of automatically initializing a 3D deformable model is presented by compounding a pictorial structures model [28] and AAM. However, features used in this method are less distinctive than that used in our model. The method proposed in [29] only used Hough transform to determine the initializing model by a bounding box, where only rough positions of each landmarks can be located.

Inter-subject anatomical variability may abolish one-to-one image correspondence between subjects. Thus global deformable geometrical model will not fit objects well, particularly when the same anatomical structures are absent in different subjects. Only a few methods have investigated this problem. [30] improved the AAM through outlier detection techniques. Toews et al. also employed local invariant features to train a statistical part-based model [31]. This model learns the probability of a feature presence (or absence) from a large image set and matches local invariant features to a new image allowing fitting. The advantage of this model is that it does not rely on exact anatomical similarity between subjects to permit co-registration, as it will typically locate some corresponding local structures. However, the methodology used ensures that the images produced may not have clear anatomical meanings.

In an attempt to address the problems mentioned above, we propose a novel compounding model by combining a deformable geometric model and local invariant features. Local invariant features can extract local structures robustly and consequently match them more accurately to images from different subjects. In this paper, we exploit the scale invariant feature transform (SIFT) [32] to construct our compounding model. It is worth noting that the proposed model can be easily extended to other local invariant features. We first build the deformable geometric model, which is the point distribution model (PDM) in this study, by manually labeling landmarks and aligning the shape represented by landmarks in all images of the training set. There are two stages to choose landmarks. At the first stage, we choose those key points that correspond to anatomic structures and some special structures, such as T-junction. At the second stage, we uniformly choose points between two adjacent landmarks selected in the previous stage. Then, we extract the SIFT features from the training images and align them to the PDM. Two PDMs used in our experiments both have 38 landmarks, in which 8 are key points and 30 are other landmarks uniformly selected. We have developed a strategy to preserve the highly-repeatable and highly-robust SIFT features. Based on these selected SIFT features, an efficient and automatic learning strategy is developed to build a Key Features Model (KFM). The KFM is then connected to the PDM by modeling their geometrical relationship. Based on this relationship, we develop an automatic strategy, precisely initializing the position of PDM. After the initialization of PDM, more precise localization of each landmark is also achieved. Finally, the ultimate positions of landmarks are obtained by iterative search processes.

Unlike other initialization methods mentioned previously, the proposed method can initialize global deformable geometrical models in more extensive ranges, due to its accuracy in matching local parts. Unlike Toews's model, the proposed model connects the local parts with explicit anatomical structures represented by PDM. Therefore the proposed method can find local parts among variable inter-subject anatomical tissues, as well as the meaningful anatomical structures.

Extensively experiments have been carried out to evaluate the performance of the proposed method. We have collected data

from a large number of cardiac MRI images, consisting of 500 slices from 16 subjects, 10 male and 6 female, aged from 45–64 years. Firstly, we demonstrate the matching accuracy of the proposed KFM with respect to its repeatability and robustness. Then the accuracy of initialization and fitting is evaluated. Finally, our fitting results show the 6.2~8.3% improvements by comparison with an improved ASM [16].

The main conclusions of this paper is three-fold: 1) A KFM is built to match repeatable local parts among subjects with anatomical variance, even when no exact one-to-one correspondence exists. 2) An automatically initializing strategy is developed for the initialization of PDM. 3) An efficient strategy is developed to precisely localize landmarks of PDM.

Methods

Ethics statement

This study was approved by the Research Ethics Committee of Zhejiang University of Technology and Hangzhou Shaoyifu Hospital, China. All subjects used in this paper have been informed that their data may be published and provided written informed consent.

Model building

The new model consists of a PDM, a KFM and their geometric relationship. The PDM is built in an ASM first and the gray-levels of pixels around the landmarks in the PDM are also modeled. The KFM is built by extracting SIFT features and choosing high-repeatable and high-robust ones. This section explains how to build PDMs and KFMs, and model their geometric relationship.

Point distribution model. A PDM consists of n landmarks that are the important points to represent the particular parts or boundary of objects in the training set. The collection of these landmarks is used to define a shape vector as $\bar{X} = [x_0, y_0, x_1, y_1, \dots, x_n, y_n]$. These shapes in all training images must be aligned with respect to a set of axes. The required alignment is achieved by scaling, rotating, and translating the training shapes so that they agree as closely as possible. From all aligned shapes, a mean shape is computed by $\bar{X} = \frac{1}{N} \sum_{i=1}^N X'_i$, where X'_i indicates an aligned shape. Shape dimensions and variations are reduced by using the Principal Components Analysis (PCA). Thus a new shape is obtained by:

$$X_{new} = \bar{X} + \Phi b \quad (1)$$

where b is a $t(t \leq 2n)$ dimensional vector containing the PDM's parameters, and computed by:

$$b = \Phi^T (X - \bar{X}). \quad (2)$$

A new shape X_{new} is obtained by varying b within a constrained range $\pm \alpha \sqrt{\lambda_i}$, where α is usually set to a value between two and three. And the new shape should be transformed into a new and better location on the target image.

To find the accurate position of the new shape, the gray-level appearance should also be modeled. For each landmark, gray-level values are drawn from a line. This line passes through the landmark and is perpendicular to the boundary, which is formed by the landmark and its neighboring pixels. For j -th landmark in the i -th image, n_p extracted gray-level values centered at j -th

landmark form a gray-level profile, denoted by $g_{ij} = [g_{ij_0}, g_{ij_1}, \dots, g_{ij_{p-1}}]$. Thus a mean profile is defined as:

$$\bar{y}_{ij} = \frac{1}{N} \sum_{i=1}^N dg_{ij}, \quad (3)$$

where $dg_{ij} = [g_{ij_1} - g_{ij_0}, \dots, g_{ij_{p-1}} - g_{ij_{p-2}}]$. The final new shape is obtained by finding such a shape that the profiles of landmarks in new image have minimal difference with respect to their mean profiles and its variations do not exceed the range of b .

Key features model. Given total W images in the training set, W groups of SIFT features need to be extracted. However, there are too many SIFT features, even if there are only a few objects in one image. These features should be selected carefully and only those highly-robust and highly-repeatable key points can be used to establish the KFM. There are three steps to choose the suitable SIFT features.

The first step is to align all W groups of SIFT features to the mean shape obtained when building the PDM. When building the PDM, we align shapes on different training images as closely as possible. This procedure provides a set of transforming parameters. By applying these parameters to corresponding SIFT features, they can also be aligned to the mean shape of the PDM. If SIFT features in different images are associated with the same position, they will be closely located after being aligned.

The second step is to match the SIFT features to obtain highly-robust matching pairs. Usually, the appearance description of SIFT features is used to match SIFT features. SIFT features are local invariant features and their descriptions only represent their local image information. It may result in mismatched pairs, if there are similar local structures in two image but at different positions. After aligning SIFT features, an extra geometric constraint can be used to improve the matching correctness. Assuming two images in the training set are denoted as I_i and I_j and their SIFT features are denoted as f_i and f_j , respectively. For one SIFT feature in the image $I_j, f_{i,n}$, its matched SIFT feature in the image I_j can then be determined if they meet the following conditions:

$$\begin{aligned} \min(\|A_{f_{i,n}} - A_{f_j}\|_2) &< Th^d \\ \min(\|A_{f_{i,n}} - A_{f_j}\|_2) &< 0.8 \times \min^{second}(\|A_{f_{i,n}} - A_{f_j}\|_2) \\ \|O_{f_{i,n}} - O_{f_{j,m}}\|_2 &< Th^o \end{aligned} \quad (4)$$

where $A_{f_{i,n}}$ and $O_{f_{i,n}}$ denote the appearance description and aligned geometric position of $f_{i,n}$. The previous two conditions, suggested in [32], are usually used for SIFT feature matching. Here an extra condition based on aligned SIFT features is added. Assuming $O_{f_{j,m}}$ and $O_{f_{i,n}}$ are the aligned geometric position of $f_{j,m}$ and $f_{i,n}$, if they meet the previous two conditions, they are considered as the potential matching pair. If the geometric distance between $O_{f_{i,n}}$ and $O_{f_{j,m}}$ is less than a threshold, they are regarded as a matching pair. Some examples is shown in Fig. 1, from which we can observe that the third condition will greatly improve the matching accuracy.

By using these conditions, SIFT features of each image are matched to all other images in the training set and a set of matched pairs can be obtained. We denote $M_{i,n,j,m}$ as the n th SIFT feature in the i th image being matched to the m th SIFT feature in the j th image.

The third step is to choose highly-repeatable SIFT features by a matching propagation procedure. Given a SIFT feature, $f_{i,n}$, we

will choose all SIFT features corresponding to the same position in different images to form a collection G_k . Firstly we initialize G_k by adding $f_{i,n}$ into it. And we choose all SIFT features in other images that match with $f_{i,n}$ to be members of G_k . For each SIFT feature in G_k , we then choose all SIFT features in the other image that matches with it and then add them into G_k ; repeating this procedure, until there are no other SIFT features in all images that are matched with any features in G_k . Thus a collection of SIFT features in all images corresponding to the same position as $f_{i,n}$ can be obtained. Performing this procedure is necessary and important. One SIFT feature in one image may not find matched features from other training images. When applying matching propagation, we can obtain as many SIFT features as possible, corresponding to the same position in all images of the training set.

Supposing we finally obtain K collections and each collection has t_k SIFT features after the above matching propagation procedure. Since there is only one SIFT feature associated with one position in one image, t_k can reflect the number of images, in which robust SIFT features can be found at a certain similar position. The larger value t_k is, the higher repeatability the collection G_k is. To guarantee high-repeatability, the ratio between t_k and the total number of training images must be larger than a constant value R . R can be adjusted to balance the number of key points in KFM and the repeatability of the picked SIFT features. If R is set to a larger value, the number of key points will be decreased and the repeatability will be increased, and vice versa. In this study, according to the experiments in the sub-section of Evaluation of Different Picking Ratios, we suggest an adequate ratio (e.g. 0.6). An example is shown in Fig. 2, where the features with higher occurrence ratios are more likely located at meaningful positions, such as red and green feature points.

According to the three steps detailed above, the highly-robust and highly-repeatable SIFT features are obtained effectively. We can then use these SIFT features to build a KFM. A KFM consists of K sets of SIFT features. In each set, SIFT features locate at the close positions. The average appearance description of each SIFT features set, G_k , is computed as:

$$A_{G_k} = \frac{\sum_{c=1}^{t_k} A_{G_k,c}}{t_k} \quad (5)$$

where $A_{G_k,c}$ means the appearance description of the c th SIFT features in G_k .

If a SIFT feature, f_{n_i} , in a test image meets the following conditions simultaneously, we consider it match to a collection G_k in KFM:

$$\begin{aligned} \bar{d}_{n,G_k} &= \|A_{f_{n_i}} - A_{G_k}\|_2 = \arg \min_j (\|A_{f_{n_i}} - A_{G_k}\|_2) \\ d_{n,G_k^c} &= \|A_{f_{n_i}} - A_{G_k^c}\|_2 = \arg \min_{j,c} (\|A_{f_{n_i}} - A_{G_k^c}\|_2) \end{aligned} \quad (6)$$

where $A_{G_k^c}$ is the appearance description of the c -th SIFT feature in G_k . Thus the first equation in (6) means f_{n_i} has the most similar appearance description with A_{G_k} with respect to other SIFT features in the test image. The second equation in (6) means the minimal distance between $A_{f_{n_i}}$ and appearance descriptions of SIFT features in G_k is also the minimal distance between the appearance descriptions of any SIFT features in the test images and appearance descriptions of SIFT features in G_k .

Statistical geometric relationship between KFM and PDM. The relationship between the KFM and the PDM is built through relative displacements between aligned positions of

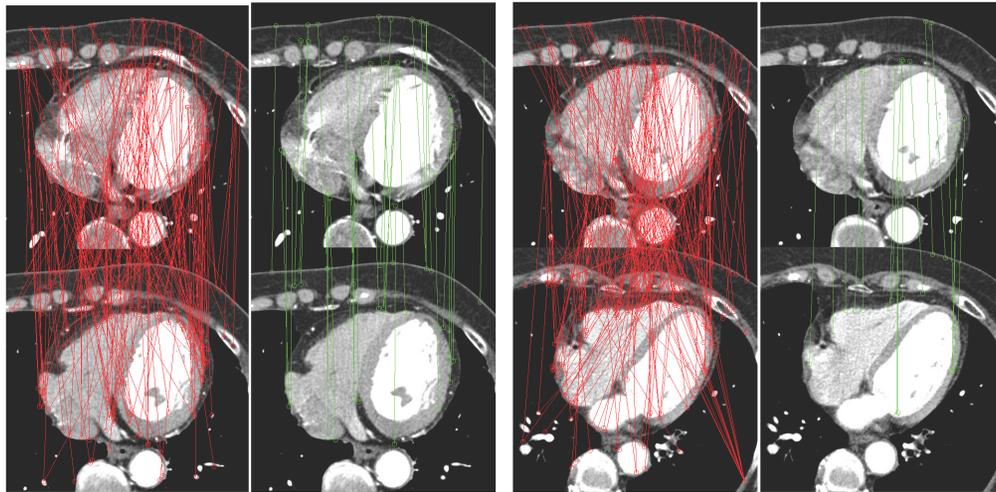


Figure 1. Two examples of matching results of the SIFT features. Picture marked by red points and lines illustrates the results of using only previous two conditions. Pictures marked by green points and lines illustrates the results of using three conditions.
doi:10.1371/journal.pone.0105815.g001

SIFT features in the KFM and aligned landmarks in the PDM. We denote the geometric position of the c th SIFT feature in the k th SIFT feature set of KFM as $O_{k,c} = \{x_{O_{k,c}}, y_{O_{k,c}}\}$ and the position of the l th landmark in PDM as $L_l = \{x_l, y_l\}$. The displacement between $O_{k,c}$ and L_l is then denoted as $E_{L_l, O_{k,c}} = \{e_{L_l, O_{k,c}}^x, e_{L_l, O_{k,c}}^y\}$. From the statistical perspective, given one landmark L_l , we have:

$$\begin{aligned} p(G_{k,c}|L_l) &= p(\{A_{k,c}, O_{k,c}\}|L_l) = p(A_{k,c}|L_l)p(O_{k,c}|L_l) \\ &= \begin{cases} 0 & \text{if } p(A_{k,c}|L_l) = 0 \\ p(O_{k,c}|L_l) & \text{if } p(A_{k,c}|L_l) = 1 \end{cases} \end{aligned} \quad (7)$$

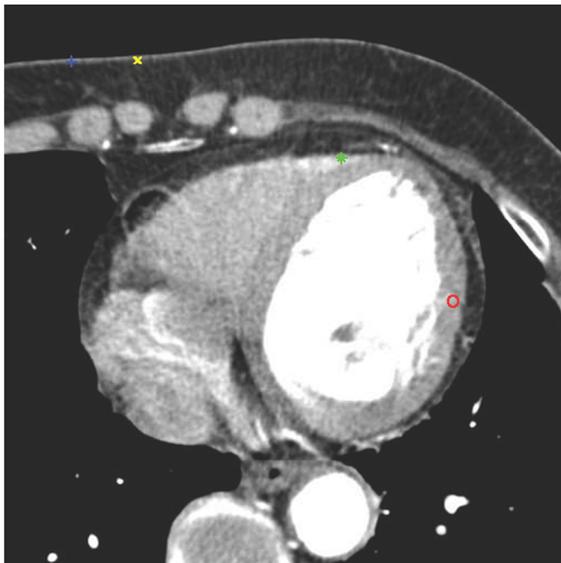


Figure 2. Features in one slice with different occurrence ratios. The ratios of red, green, blue and yellow features are 0.85, 0.73, 0.52 and 0.31, respectively.
doi:10.1371/journal.pone.0105815.g002

Given a training image, the SIFT feature $G_{k,c}$ may be present or not, therefore the probability $p(A_{k,c}|L_l)$ is equal to 1 or 0. The position of this SIFT feature can be denoted by its displacement relative to the given landmark L_l , and therefore we have:

$$p(O_{k,c}|L_l) = p(E_{G_k, L_l}|L_l) = \mathcal{N}(\mu_{G_k, L_l}, \Sigma_{G_k, L_l}) \quad (8)$$

where the displacement of one SIFT feature in one set G_k is assumed to be subjected to bivariate Gaussian distribution, and E_{G_k, L_l} denote the displacement between G_k and landmark L_l . Given a test image, in which the positions of landmarks are unknown, if the u -th SIFT feature f_{nu} in this image is found to match to one SIFT feature in G_k , the matched SIFT features can be used to determine positions of landmarks.

$$\begin{aligned} p(L_l|f_{nu}) &= p(L_l|A_{f_{nu}}, O_{f_{nu}}) = p(L_l|O_{f_{nu}}) \\ &= p(L_l|E_{G_k, L_l}) \propto p(E_{G_k, L_l}|L_l) \end{aligned} \quad (9)$$

Because in (8) positions of SIFT features in G_k can be determined by its displacement relative to the l th landmark, the position of this landmark in a test image can accordingly be determined by its displacement relative to the matched SIFT feature.

For each set of SIFT features in KFM, there are the relations denoted by (8) and (9), the position of the landmark in a test image is therefore determined by all SIFT features each of which is matched to one of the SIFT feature sets. It is worth noting that there may not be a matched SIFT feature for every G_k . Finally, the position of the landmark is calculated by maximizing the joint distribution determined by all observed matched SIFT features:

$$L_l = \arg \max_{L_l \in \mathbb{R}^2} (\prod_{k=1}^z p(L_l|E_{G_k, L_l})) \quad (10)$$

where we assume z SIFT features, each of which matched to one G_k , are found. By using the logarithm of (10) and solving its extrema, it is possible to obtain the position of landmarks.

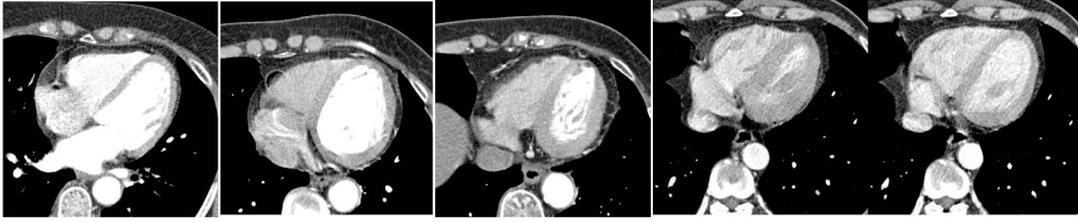


Figure 3. Some examples of cardiac MRI images used in our experiments.
doi:10.1371/journal.pone.0105815.g003

Extend to other local invariant features and deformable models. In the proposed model, we only assess the appearance description and geometric position of local invariant features. However, it does not depend on special local invariant features. Therefore the proposed model can be easily integrated with other local invariant features, such as the maximally stable extremal regions (MSER) [33] and the speeded up robust features (SURF) [34]. In the matching propagation steps mentioned in the subsection of Key Features Model the appearance description of the SIFT features can be replaced by another correspondence description and their matching algorithm can also be followed with an extra geometric constraint.

The proposed model can also be extended to the other deformable models with little difficulty. What the proposed model needs is a reference framework, just like the manually labeled landmarks in ASM. By using the reference framework, those extracted local invariant features can be normalized and the extra geometric constraint can be applied to improve the matching correctness. The AAM is similar to ASM and naturally has a reference framework since it also contains a manual labeling step. If using AAM instead of ASM in the proposed model, we can follow all the steps detailed previously. For ACM, an extra learning procedure is required. In this learning procedure, several reference points need to be labeled to normalize the local invariant

features and the relationship between reference points and local invariant features will also be learned. The initial position of ACM can then be determined by the relationship between those matched local invariant features and interesting objects.

Fitting process

Initialization. For initializing the model in a test image, the SIFT features are extracted first and matched to the KFM. The first two matching conditions in (4) are used, which are similar to the strategies suggested in [32]. Due to the highly robust SIFT feature sets in the KFM, there are only a few mismatching points. We employed the method proposed in [35] to eliminate the mismatching SIFT feature pairs. After obtaining the accurately matched SIFT feature pairs, we can compute transformation parameters, \bar{T} , \bar{S} and $\bar{\Theta}$ that transform the average positions of SIFT features in the KFM towards the positions of SIFT features in the test image. Thus, by applying these parameters to the mean shape of PDM, we can initialize the PDM model to an appropriate position:

$$X' = F(\bar{S}, \bar{\Theta})X + \bar{T} \quad (11)$$

where X is the mean shape in the PDM and X' is the initialized model.

Accurately localized landmarks. After obtaining an initialized model X' , each landmark in X' is still located at its rough position, as the transformation in the previous subsection is a type of rigid transformation that maintains the original contour of mean shape X . The position of each landmark can be modified according to the relationship between the KFM and the PDM built in the sub-section of Iterative Process to obtain the precise position.

Assuming there are z matched SIFT features in a new image, the position of each landmark can be determined by (11). According to the relationship between the KFM and the PDM, we can further improve landmark positions by using (10). Then this landmark will move towards this new position from the old position in M' . The PDM, however, is a global constrained model and the modification of the position of one landmark will influence other landmarks' positions, which is very important to keep the whole model from being abnormal. The positions are modified according to (12), which is similar to the elastic model. That is to say, if other landmarks are closer to one landmark, the forcing movement of these landmarks is larger than that of landmarks further away. Assuming the new position of the l th landmark is $\{x''_{O_l}, y''_{O_l}\}$ and its old position in M' is $\{x'_{O_l}, y'_{O_l}\}$, the modified position of the l th landmark caused by the movement of the l th landmark can be computed by:

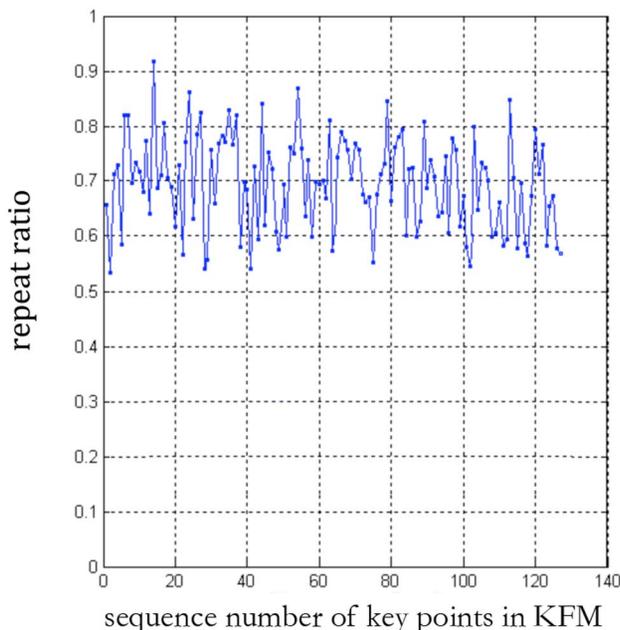


Figure 4. Repetition Ratio of all Key Points in KFMs.
doi:10.1371/journal.pone.0105815.g004

Table 1. Number of matching SIFT features.

Group index	Maximum number	Minimum number	Average number
1	22	10	18.2
2	23	8	16.8
3	25	13	20.6
4	16	6	12.5
5	16	8	11.9

doi:10.1371/journal.pone.0105815.t001

$$\begin{aligned}
 x_{O_{12}}^n &= \frac{\min(\text{dist}(\mathbf{O}'_M, \mathbf{O}'_{l1}))}{\text{dist}(\mathbf{O}'_{l2}, \mathbf{O}'_{l1})} (x_{O_{l1}}^n - x'_{O_{l2}}) + x'_{O_{l2}} \\
 y_{O_{12}}^n &= \frac{\min(\text{dist}(\mathbf{O}'_M, \mathbf{O}'_{l1}))}{\text{dist}(\mathbf{O}'_{l2}, \mathbf{O}'_{l1})} (y_{O_{l1}}^n - y'_{O_{l2}}) + y'_{O_{l2}}
 \end{aligned}
 \tag{12}$$

where O'_{l1} means the position of the $l1$ th landmark in the initialized PDM. O'_{l1} is the precise location of this point. \mathbf{O}'_M is the points set in PDM and $\min(\text{dist}(\mathbf{O}'_M, \mathbf{O}'_{l1}))$ means the minimal distance between O'_{l2} and other points in PDM. The distance between O'_{l2} and O'_{l1} is denoted as $\text{dist}(O'_{l2}, O'_{l1})$. $\{x'_{O_{l2}}, y'_{O_{l2}}\}$ is the coordinate of point O'_{l2} . Hence, $\frac{\min(\text{dist}(\mathbf{O}'_M, \mathbf{O}'_{l1}))}{\text{dist}(O'_{l2}, O'_{l1})}$ means the ratio of distance between the $l2$ th landmark and $l1$ th landmark related to the minimal distance between other landmarks and $l1$ th landmarks. Therefore, when O'_{l1} is modified to a precise pose, it will influence other landmarks' positions. Thus, more accurate location for the whole PDM can be

obtained after pinpointing all landmarks.

Iteratively fitting process. After the precise position of each landmark in the PDM is computed, the same iterative process to ASM can be used to obtain the ultimate position. After obtaining a new shape X_{i-1} in the $i-1$ th iterative step, the i th new shape is computed by the following steps:

1. For one landmark, e.g. the j th landmark, in X_{i-1} , we extract $\beta \times n_p$ points' gray-level values to compute gray profile $y_{i-1,j}$, where β is set to 2-5;
2. Then we compute the difference between the mean profile \bar{y}_j , obtained by (3) in the training procedure, and any continuous $n_p - 1$ items in $y_{i-1,j}$ to find one point centered at which the difference between continuous $n_p - 1$ items and \bar{y}_j is minimum;
3. Repeating the pervious two steps and obtaining a new set of points, denoted by X'_i ;
4. Next we transform X_{i-1} to X_i according to (1) to make sure that X_i is closest to X'_i .

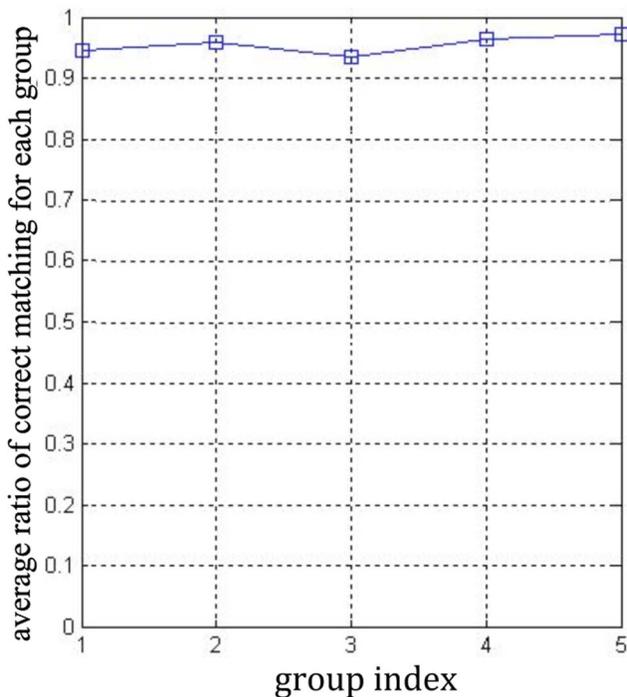


Figure 5. Average ratio of correct matching of all five groups of KFM.

doi:10.1371/journal.pone.0105815.g005

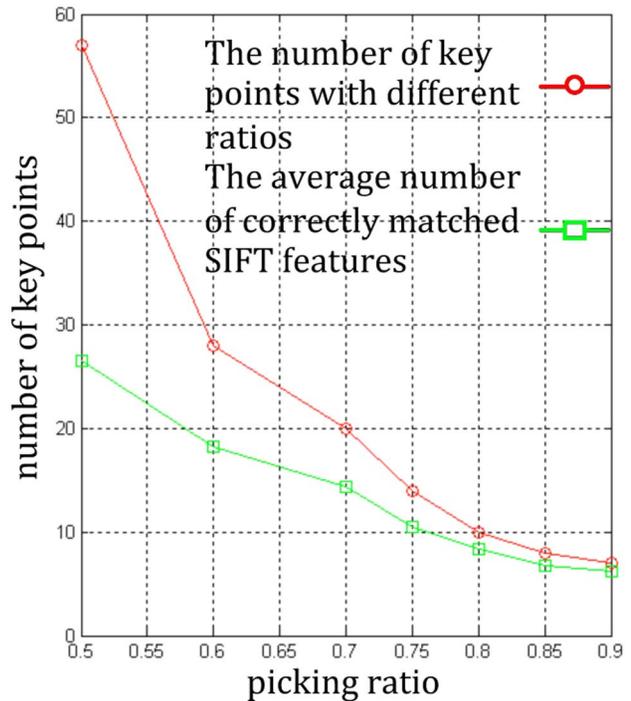


Figure 6. Number of key points in KFM and SIFT features under different picking ratios.

doi:10.1371/journal.pone.0105815.g006

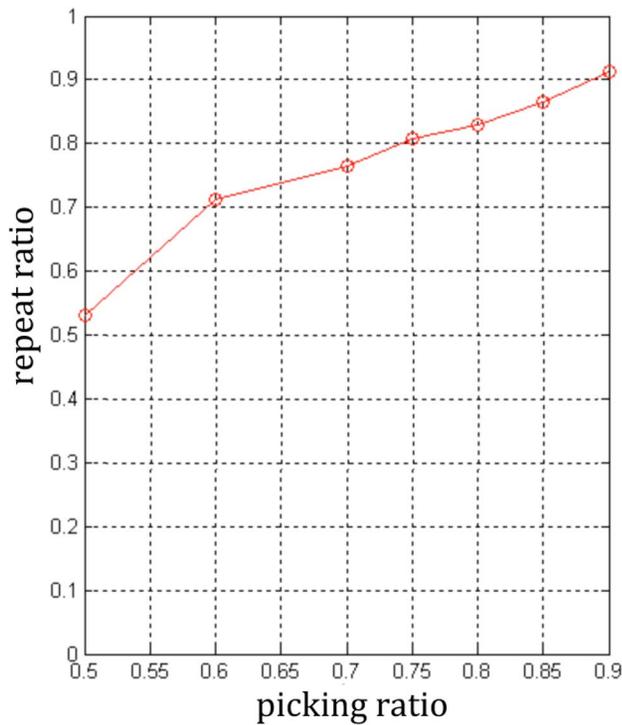


Figure 7. Repetition ratio under different picking ratios.
doi:10.1371/journal.pone.0105815.g007

These iterative steps usually are time-consuming and need a hundred iterations to converge if the initial position of PDM is not ideal. By applying the proposed method, however, the number of

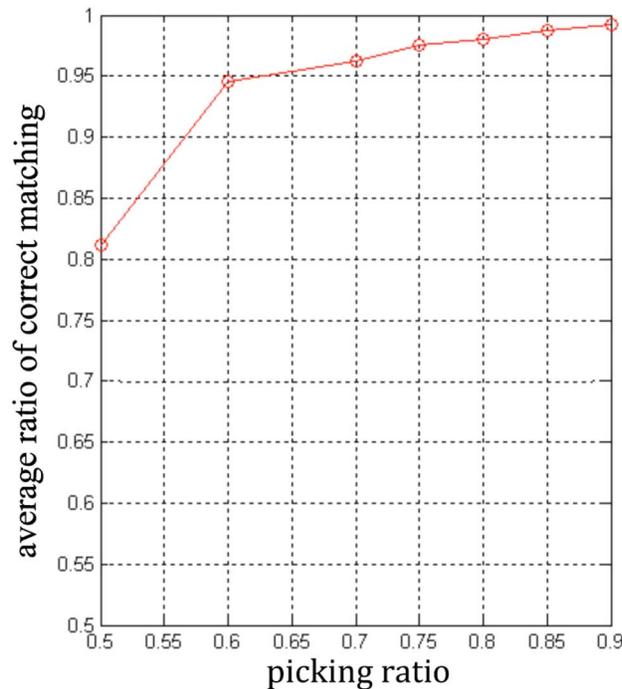


Figure 8. Average ratio of correct matching under different picking ratios.
doi:10.1371/journal.pone.0105815.g008

Table 2. Accuracy of initial position.

Group index	Ratio of correct initialization	Average error
1	0.982	2.98
2	0.976	3.06
3	0.968	3.15
4	0.973	3.04
5	0.965	3.16

doi:10.1371/journal.pone.0105815.t002

iterations will significantly decrease because each landmark was located at its precise position. Thus the time cost will be sharply reduced.

Experiments

We evaluate the proposed method in a set of cardiac MRI images. Experiments have been carried out from three different aspects. First, we tested the repetition and robustness of the KFM, which is the base for the proposed model, determining the accuracy of initialization and fitting results. Second, the accuracy of initializing PDMs was evaluated. The accurate initialization will decrease iterative times and reduce the computational cost. Finally, we evaluated the speed and accuracy of co-registering cardiac images. In this study, the SIFT implementation used in our experiments is publicly available in <http://www.cs.ubc.ca/~lowe/keypoints/>.

Datasets

A large number of cardiac MRI images were collected for the experiments. The data consisted of 200 slices from 16 subjects, 10 male and 6 female, aged from 45–64 years. Five groups of experiments were carried out based on five randomly selected image sets. In each group, we randomly select 10 subjects and approximately 120 slices. Around 60 slices from 5 subjects were used for model training, and the remainder for testing. Examples of the experimental images are shown in Fig. 3.

Matching repeatability and robustness of KFM

As the proposed model is based on the highly-repeatable and highly-robust key points in the KFM, the repetition and robustness of the KFM are investigated first. In five groups of experiments, the numbers of key points in each KFM are 28, 26, 31, 20, and 22 respectively. To evaluate the repeatability of the KFM, the SIFT features were extracted from all images tested, and these features matched to the corresponding KFM. The number of times that correctly matching key points in KFM with test images were accumulated, allowing calculation of the repeated ratio for these key points. The minimum repetition ratio was 0.53 and the maximum repetition ratio was 0.96. The repetition ratios of 127 key points in all five KFMs are shown in Fig. 4, proving that enough repetitions of the KFM could be obtained.

Additionally, we calculated the maximum, minimum and average numbers of matching SIFT features in a new image. Table 1 illustrates these numbers for the five experimental groups. The number of matching SIFT features for each test image ranged from 10 to 22, with an average of 18.2 in the first experimental group. This means that at least 10 exactly matched SIFT features could be found in a new image within the first experimental group. According to these matched SIFT features, we can accurately

Table 3. Evaluation of pinpointing landmarks.

Group index	Max number	Min number	Average number	Average error
1	34	18	28.5	2.56
2	30	18	27.3	2.34
3	33	17	26.2	2.83
4	30	20	25.1	2.67
5	35	21	29.8	2.72

doi:10.1371/journal.pone.0105815.t003

initialize the PDM in accordance with the proposed initializing strategy.

For further evaluating the matching correctness of KFM, we compute the average ratio of correct matching, which is defined as the ratio of correct matching points to all matching points for all test images. In all five KFMs used in the experiments, the maximum and minimum average ratios of correct matching are 1 and 0.91, respectively. We illustrate the average ratios in Fig. 5, which demonstrate the high-robustness of KFM.

From Table 1, we observe that we can obtain enough matching key points when using the KFM to match SIFT features extracted from a medical image. SIFT features can achieve good performance of matching the same object at different images in a natural scenario. However, because there are lots of duplicated local structures and flexible deformation of hearts in different patients' MRI images, most SIFT features are not as robust as obtained in natural images. But through the proposed KFM, the most robust SIFT features can be identified and matched accurately. As shown in Fig. 5, the accuracy of matching results using KFM is at the same level as that of matching objects in nature scenarios with SIFT features. Therefore, our KFM can represent a powerful tool to find accurately matched points in different patients' cardiac images. From Fig. 4 and Fig. 5, it is obvious that each key point in the KFM has a high matching accuracy with considerable repeatability.

Toews's model also matches features extracted from the new image [31]. However, it has to match all SIFT features extracted from all training images, which takes a much longer time. Furthermore, since all SIFT features extracted from training images are used, more matching errors may be generated. By comparison with Towe's model, the proposed KFM only consists of selected highly-repeatable and highly-robustly local invariant features, and has potential to reduce the computational power required while improving matching accuracy.

Evaluation of different picking ratios

As mentioned in the sub-section of Key Features Model, the picking ratio is an important parameter that balances the repetition and amount of key points in the KFM. It must be tested carefully and only the most proper ratio can be adopted. We evaluate several different picking ratios, such as 0.5, 0.6, 0.7, 0.75, 0.8, 0.85 and 0.9. The number of key points and the repeated ratio under different picking ratios were computed respectively. All experiments in this section are based on the image dataset from the first experimental group.

Fig. 6 illustrates the number of key points under different ratios in KPMS and Fig. 7 shows the repeated ratio of key points under different picking ratios. The average number of matching SIFT features in the test image under different picking ratios is shown in Fig. 8.

In Fig. 6, when the ratio rises from 0.5 to 0.6, the number of key points sharply decreases, but the decrease of the average number of matching SIFT features is less. When the ratio keep increasing, the key points' amount and the average number are gradually approaching. This figure implies that the KFMs with low ratio have low-repeatable key points. Moreover, the low ratio implies the low repeatability and low accuracy of matching results. Fig. 7 and Fig. 8 illustrate these results.

From Fig. 7 and Fig. 8, when the ratio is far below 0.6, the repeatability and correct matching rate is too low to ensure that there are enough correct matching SIFT features in the new image. If the number of correctly matching SIFT features is not enough, the initialization will fail. On the other hand, when the ratio is much larger than 0.6, the accuracy of matching can be high enough, but the number of matching SIFT features is too low. Only when the ratio is set to about 0.6, both the repeatability and the matching accuracy achieves a good level.

Accuracy of automatic initialization

As the first step of the fitting process, the initialization is very important. The accuracy of initialization directly affects the fitting results. In the standard ASM, there is no strategy to accurately and

Table 4. Performance comparison between an improved ASM and our model.

Group index	Average error of [6]	Average error of our model	Improvement
1	2.16	1.98	8.3%
2	2.19	2.03	7.3%
3	2.23	2.09	6.2%
4	2.07	1.93	6.7%
5	2.11	1.95	7.6%

doi:10.1371/journal.pone.0105815.t004

Table 5. Iterative times comparison between an improved ASM and our model.

Group index	Iterative times [16]	Iterative times (our model)
1	120.2	63.1
2	104.7	60.2
3	114.0	58.7
4	98.6	50.8
5	107.5	52.9

doi:10.1371/journal.pone.0105815.t005

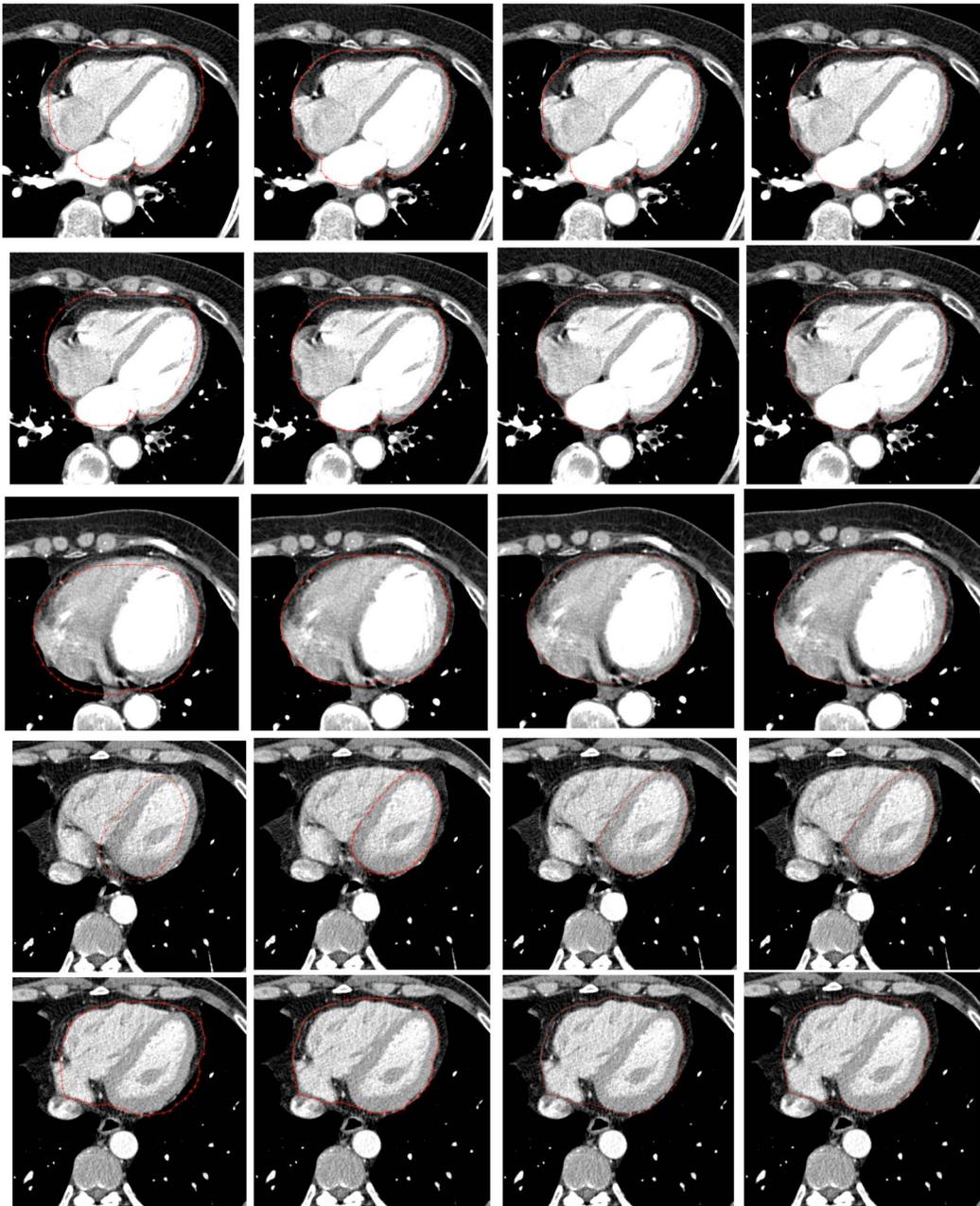


Figure 9. Some fitting results of our model. (Left) Pre-initialized model;(Middle left) Initialized model; (Middle right) Pinpointed model, note that there are some points that have been accurately located; (Right) Fitting result.
doi:10.1371/journal.pone.0105815.g009

automatically initialize the PDM. In our experiments, we achieve enough matched key points with a high enough accuracy to initialize the PDM by using the KFM. For evaluating the accuracy of initialization, we employ the average error which is defined in [24] to calculate the average distance between the manually labeled shapes and the initialization results.

$$E = \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{n} \sum_{l=1}^n \text{dist}(O_{li}, O'_{li}) \right) \quad (13)$$

where E is the average error, N is the number of test images in one group of experiments, and n is the number of landmarks. O_{li} is the l th landmark point in the manually labeled shape of the i th test image manually labeled, O'_{li} is the l th landmark point in the resulting shape of initialization for the i th test image.

The ratio of correct initialization is defined as the ration between correct initialization of the PDM and the number of all test images. The correct initialization of the PDM is determined manually. Actually, only these initialized PDMs are treated as correct initialization, if the number of landmarks, whose distance to the corresponding manually labeled landmarks is less than five pixels, is more than half of landmarks in the PDM. Table 2 shows the average accuracy of initialization.

Precision of pinpointing landmarks

During the process of pinpointing landmarks, the evaluation of precision is dependent on the distance between the position of manually labeling landmarks and pinpointing results by the proposed method. If the distance is less than three pixels, we believe that a landmark is precisely pinpointed. In all five experimental groups, the average number of precisely locating landmarks and the average precision are shown in Table 3. We also employ (13) to compute the average precision. In all models, there were 38 landmarks. From Table 3, the maximal and minimal numbers of precise pinpointing are 35 and 17, respectively. The average numbers are all more than 25. To the whole model with 38 landmarks, these precisely located landmarks are enough to adjust the whole model to a more accurate position by (12). By comparing with the initialized PDM, we found that the distance between the new position of the PDM and the manually labeled landmarks is smaller. As shown in Tabel 3, the reduced average errors with respect to the average errors obtained in the previous experiments are ranged from 0.3 to 0.7, which demonstrates that the process of pinpointing landmarks is effective.

Accuracy of fitting results

We now evaluate the accuracy of the fitting results of our model by comparison to an improved ASM proposed in [16]. Table 4

shows their fitting results, in which the accuracy is computed by (13).

Firstly, as Table 4 shows, the average fitting errors resulting from our model are less than the improved ASM in [16]. The improvement of automatic initialization ranges from 6.2% to 8.3%. This proves that our proposed model can achieve more accurate results than that of the improved ASM. Furthermore, due to accurate initialization and precise pinpointing, it will also decrease the running time of the fitting process. The average iterative times can decrease about 50% in the experiments. This means the computational cost of the proposed model only needs half the time compared with the improved ASM. We list the average iterative times in Table 5 for each group of experiment. Of course, it can also reduce the multi-resolution levels without significant loss in accuracy for the result. Some typical results are illustrated in Fig. 9.

Conclusions

This paper proposed a method of a deformable model utilizing local invariant features for accurate and rapid registration of cardiac images. The highly-repeatable and highly-robust local invariant features are chosen from a training set to develop a KFM. The the relationship between the KFM and the PDM is obtained statistically. During the fitting process, the PDM is transformed to its initial position according to the transformed parameters, by aligning the accurately matched SIFT features between the KFM and new images. The precise position of the PDM is further achieved by modifying the landmark coordinates in the PDM according to the relationship between the KFM and the PDM. Finally, the ultimate location of the PDM is obtained by an iterative process with less subsequent reduction in computational load.

Experiments demonstrate that this model outperforms traditional SIFT, in the sense of feature matching accuracy and robustness in cardiac MR images. Furthermore, it also surpasses other ASM regarding fitting accuracy and speed. Our experiments demonstrate that the proposed model improves the fitting accuracy by approximately 6–8% with only half of the computational cost of a state-of-the-art ASM.

Author Contributions

Conceived and designed the experiments: JZ JHG ZZT QG SYC. Performed the experiments: JZ LC XW QG. Analyzed the data: JZ AJB ZZT. Contributed reagents/materials/analysis tools: QG SYC. Contributed to the writing of the manuscript: JZ AJB JHG ZZT.

References

- Pajula J, Kauppi JP, Tohka J (2012) Inter-subject correlation in fmri: Method validation against stimulusmodel based analysis. *PLoS One* 7: e41196.
- Xia R, Li X, He B (2010) A comparison study of three different image reconstruction algorithms for mat-mi. *IEEE Transactions on Biomedical Engineering* 57: 708–713.
- Almukhtar A, Ju X, Khambay B, McDonald J, Ayoub A (2014) Comparison of the accuracy of voxel based registration and surface based registration for 3d assessment of surgical change following orthognathic surgery. *PLoS One* 9: e93402.
- Hishida H, Suzuki H, Michikawa T, Ohtake Y, Oota S (2012) Ct image segmentation using fem with optimized boundary condition. *PLoS One* 7: e31116.
- Wang L, Shi F, Yap PT, Gilmore JH, Lin W, et al. (2012) 4d multi-modality tissue segmentation of serial infant images. *PLoS One* 7: e44596.
- Wang L, Shi F, Li G, Shen D (2013) 4d segmentation of brain mr images with constrained cortical thickness variation. *PLoS One* 8: e64207.
- Zacharakis EI, Hoge CS, Biros G, Davatzikos CA (2008) Comparative study of biomechanical simulators in deformable registration of brain tumor images. *IEEE Transactions on Biomedical Engineering* 55: 1233–1236.
- Fang B, Tang YY (2006) Elastic registration for retinal images based on reconstructed vascular trees. *IEEE Transactions on Biomedical Engineering* 53: 1183–1187.
- Cerveri P, Mitiche A, Ayed IB (2010) Effective level set image segmentation with a kernel induced data term. *IEEE Transactions on Image Processing* 19: 220–232.
- Salah MB, Manzotti A, Vanzulli A, Baroni G (2014) Local shape similarity and mean-shift curvature for deformable surface mapping of anatomical structures. *IEEE Transactions on Biomedical Engineering* 61: 16–24.
- Su P, Yang J, Lu K, Yu N, Wong ST, et al. (2013) A fast ct and ct-fluoroscopy registration algorithm with respiratory motion compensation for image-guided lung intervention. *IEEE Transactions on Biomedical Engineering* 60: 2034–2041.

12. Makela T, Clarysse P, Sipilac O, Pauna N, Pham QC, et al. (2002) A review of cardiac image registration methods. *IEEE Transactions on Medical Imaging* 21: 1011–1021.
13. Kass M, Witkin A, Terzopoulos D (1987) Snakes—active contour models. *International Journal of Computer Vision* 1: 321–331.
14. Cootes TF, Taylor CJ, Cooper D, Graham J (1995) Active shape models—their training and application. *Computer Vision and Image Understand* 61: 38–59.
15. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 23: 681–685.
16. Chen SY, Zhang J, Guan Q, Liu S (2011) Detection and amendment of shape distortions based on moment invariants for active shape models. *IET Image Processing* 5: 273–285.
17. Saha PK, Liang G, Elkins JM, Coimbra A, Duong LT, et al. (2011) A new osteophyte segmentation algorithm using the partial shape model and its applications to rabbit femur anterior cruciate ligament transection via micro-ct imaging. *IEEE Transactions on Biomedical Engineering* 58: 2212–2227.
18. Davatzikos C, Tao XD, Shen DG (2003) Hierarchical active shape models, using the wavelet transform. *IEEE Transaction on Medical Imaging* 22: 414–423.
19. Papandreou G, Maragos P (2007) Multigrid geometric active contour models. *IEEE Transaction on Image Processing* 16: 229–240.
20. Heimann T, Meinzer HP (2009) Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis* 13: 543–563.
21. Petitjean C, Dacher JN (2011) A review of segmentation methods in short axis cardiac mr images. *Medical Image Analysis* 15: 169–184.
22. Garrido A, Qerez NP (2000) Applying deformable templates for cell image segmentation. *Pattern Recognition* 33: 821–832.
23. Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, et al. (2008) Automatic model-based segmentation of the heart in ct images. *IEEE Transaction on Medical Imaging* 27: 1189–1702.
24. Wang W, Shan S, Gao W, Cao B, Yin B (2002) An improved active shape model for face alignment. In: *IEEE International Conference on Multimodal Interfaces*. pp.523–528.
25. Li B, Acton ST (2008) Automatic active model initialization via poisson inverse gradient. *IEEE Transaction on Image Processing* 17: 1406–1420.
26. Liu J, Udupa JK (2009) Oriented active shape models. *IEEE Transaction on Medical Imaging* 28: 571–584.
27. Babalola K, Cootes TF (2010) Using parts and geometry models to initialise active appearance models for automated segmentation of 3d medical images. In: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. pp.1069–1072.
28. Felzenszwalb PF, Huttenlocher DP (2005) Pictorial structures for object recognition. *International Journal of Computer Vision* 61: 5579.
29. Zambal S, Schollhuber A, Buhler K, Hladuvka J (2008) Fast and robust localization of the heart in cardiac mri series. In: *International Conference on Computer Vision Theory and Applications*. pp.341–346.
30. Beichel R, Bischof H, Leberl F, Sonka M (2005) Robust active appearance models and their application to medical image analysis. *IEEE Transaction on Medical Imaging* 24: 11511169.
31. Toews M, Collins DL, Arbel T (2007) A statistical parts-based model of anatomical variability. *IEEE Transaction on Medical Imaging* 26: 497–508.
32. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60: 91–110.
33. Matas J, Chum O, Urban M, Pajdla T (2002) Robust wide baseline stereo from maximally stable extremal regions. In: *British Machine Vision Conference*. pp.384–393.
34. Bay H, Tuytelaars T, van Gool L (2008) Surf: Speeded up robust features. *Computer Vision and Image Understand* 110: 346–359.
35. Zhang J, Chen SY, Liu S, Guan Q (2008) Normalized weighted shape context and its application in featurebased matching. *Optical Engineering* 47: 097201.