Robust Anatomical Correspondence Detection by Hierarchical Sparse Graph Matching

Yanrong Guo,
School of Computer and Information, Hefei University of Technology, Hefei 230009, China and also with the Department of Radiology and Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, NC 27599 USA

Guorong Wu,
Department of Radiology and Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, NC 27599 USA

Jianguo Jiang, and
School of Computer and Information, Hefei University of Technology, Hefei 230009, China

Dinggang Shen [Senior Member, IEEE]
Department of Radiology and Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, NC 27599 USA

Yanrong Guo: gyr0716@gmail.com; Guorong Wu: guorong_wu@med.unc.edu; Jianguo Jiang: jgjiang@hfut.edu.cn; Dinggang Shen: dinggang_shen@med.unc.edu

Abstract

Robust anatomical correspondence detection is a key step in many medical image applications such as image registration and motion correction. In the computer vision field, graph matching techniques have emerged as a powerful approach for correspondence detection. By considering potential correspondences as graph nodes, graph edges can be used to measure the pairwise agreement between possible correspondences. In this paper, we present a novel, hierarchical graph matching method with sparsity constraint to further augment the power of conventional graph matching methods in establishing anatomical correspondences, especially for the cases of large inter-subject variations in medical applications. Specifically, we first propose to measure the pairwise agreement between potential correspondences along a sequence of intensity profiles which reduces the ambiguity in correspondence matching. We next introduce the concept of sparsity on the fuzziness of correspondences to suppress the distraction from misleading matches, which is very important for achieving the accurate, one-to-one correspondences. Finally, we integrate our graph matching method into a hierarchical correspondence matching framework, where we use multiple models to deal with the large inter-subject anatomical variations and gradually refine the correspondence matching results between the tentatively deformed model images and the underlying subject image. Evaluations on both synthetic data and public hand X-ray images indicate that the proposed hierarchical sparse graph matching method yields the best correspondence matching performance in terms of both accuracy and robustness when compared with several conventional graph matching methods.

Index Terms

Graph matching; hierarchical correspondence detection; line patch; multi-models; sparsity

I. Introduction

Robust anatomical correspondence detection is very important in many medical image applications, such as deformable image registration [1] and organ motion correction [2]. Although many successful local image descriptors have been proposed in the last decades, correspondence detection still remains a major challenge in subjects with large anatomical differences.

Recently, graph matching has emerged as a robust correspondence detection approach by modeling not only the point-to-point correspondence [3]–[5], but also the pair-to-pair matching consistency in a graph [6]–[8]. Specifically, each possible correspondence is considered as a node in the graph, and the pairwise agreement between any two possible correspondences is described as a link in the graph. An affinity matrix is specially used to encode the matching graph, with the diagonal elements representing the graph nodes (for all possible correspondences) and the off-diagonal elements denoting the graph links (for all inter-pair agreements). Then, the problem of correspondence matching becomes an optimization problem, i.e., finding the cohort of one-to-one correspondences that produces the maximal pairwise agreements in the affinity matrix. Spectral matching [6] and quadratic programming [9] are the typical solutions to this optimization problem as reported in the literature. Graph matching based methods have also been successfully applied to the medical images [10]–[12]. For example, Im *et al.* [10] proposed a quantitative method for automatically comparing and analyzing sulcal patterns between individuals by using a graph matching approach. Recently, graph matching technique has also been applied to lung/liver CT images for tree pruning and clique augmentation [11]. Applications to fluorescence microscopy can also be found in [12], where graph matching was used to segment and track cells for quantitative analysis of cell cycle behavior.

In general, the success of graph matching relies on two aspects. The first is the robust measurement of matching degree and inter-pair agreement, which is unfortunately still an open problem in the correspondence detection. Although many state-of-the-art image features [13], [14] have been proposed to establish correspondence between one point in the model image and another point in the subject image, only the simple geometric relationship is generally used to measure the matching coherence between two possible correspondences in the current graph matching methods. This is also the case in some learning-based methods [15], [16]. The second important aspect in graph matching is the optimization of one-to-one correspondence from the possible multiple-correspondence affinity matrices. A greedy solution is often used in the spectral matching methods by sequentially determining the one-to-one correspondence according to the order of matching confidence obtained from the eigenvector of affinity matrix with the largest eigenvalue. Although the relaxation from one-to-many to one-to-one constraint has been incorporated in [7], the solution is usually suboptimal due to the lack of discriminative power in measuring each possible correspondence.

To further improve the matching performance of the graph matching approach for medical images, we present a novel, hierarchical graph matching method with sparsity constraint to augment the power of conventional graph matching methods in establishing anatomical correspondences, especially in the case of large inter-subject variations in medical images. Our contributions are threefold. First, we propose a robust appearance measurement to characterize the pairwise agreement on each graph link. Specifically, for any two possible matches (with two starting points in the model image and two ending points in the subject image), a sequence of local intensity profiles (referred to as *line patches*) along the line connecting the two starting points in the model image, or the two ending points in the

subject image, is constructed. The appearance discrepancy between these two line patches is computed to measure the pairwise agreement between these two possible matches. By adding this novel measurement, our method becomes more robust to ambiguous matches than conventional graph matching methods, which generally use only simple geometric compatibility. Second, inspired by the discriminative power of sparse representation in machine learning and pattern recognition [17], [18], we apply a sparsity constraint on the possibilities of multiple correspondences, i.e., to seek only a small number of qualified correspondence candidates for each feature point. This reduces the risk of obtaining ambiguous matches, especially for determining one-to-one correspondences in the end of correspondence matching procedure. Based on these two novel improvements, we construct a new objective function and derive an efficient solution via quadratic programming to jointly estimate correspondence for all feature points. Finally, we integrate the proposed sparse graph matching method into a hierarchical multi-model based matching framework to deal with cases with large inter-subject variations. Importantly, multiple models encompass the anatomical variability, thus alleviating the difficulty of using a single model image to find the corresponding structure yet with large shape difference in the subject image. In order to fuse the matching results from different models into a final matching result, the mean locations of multiple candidates are iteratively computed by excluding more and more mismatching candidates. Since the optimization of anatomical correspondence between model and subject images is usually very challenging, the entire matching procedure is performed iteratively between the subject image and tentatively deformed model images (based on the previously-established correspondences from our method).

Evaluations on both synthetic data and publicly available hand X-ray images indicate that the proposed hierarchical sparse graph matching method (HSGM) yields the best matching performance in terms of both accuracy and robustness, when compared with three conventional graph matching methods: spectral matching with affine constraint (SMAC) [7], probabilistic hyper-graph matching (h-GM) [8], and the integer projected fixed point method initialized by spectral matching (SM + IPFP) [19]. SM + IPFP was reported to have one of the best matching performance among conventional graph matching methods [19].

We will present our hierarchical sparse graph matching method in Section II. In Section III, we will describe the clinical application of our proposed method such as in the hand X-ray images. In Section IV, our proposed method will be evaluated on both simulate data and hand X-ray images, and further compared with three other graph matching methods (SMAC, h-GM, and SM + IPFP). A discussion on different methods and their results is provided in Section V. Finally, we conclude in Section VI.

II. Method

Correspondence detection is usually defined as finding the correspondences between the feature points of the model image and subject images, where the feature points can be identified either by human expert manually or by key point detection methods automatically [20], [21]. Considering feature point sets $T = \{t_i | i = 1, …, N\}$ in the model image and $S = \{s_i | i = 1, …, N\}$ in the subject image, our goal is to find an assignment matrix $X = [X_{i,i}]_{N \times N}$ ($X_{i,i}$ {0,1}) between these two point sets. Here, each assignment $X_{i,i}$ indicates whether a feature point t_i in the model image is matched with a feature point s_i in the subject image, with "1" denoting matched and "0" denoting unmatched. Since it is very difficult to optimize X when each element in X is either "1" or "0," we have to convexify X by relaxing each $X_{i,i}$ to a continuous value between 0 and 1, $X_{i,i}$ 1. Fig. 1 schematically illustrates the main idea behind graph matching based correspondence detection method, using hand X-ray images as example. Given the model feature point set T [Fig. 1(a)] and the subject feature point set S [Fig. 1(b)], all possible correspondences between T and S are indicated by

the white lines in Fig. 1(c). An $NN \times NN$ affinity matrix M [Fig. 1(d)] can be constructed to describe the confidence of all established correspondences, as well as the pairwise agreement between any two possible matches. Specifically, each diagonal element [shown with boxes in Fig. 1(d)] in the affinity matrix M represents the point-wise similarity between two feature points t_i T and s_i S. Each off-diagonal element [shown with pink triangle in Fig. 1(d)] measures the pairwise agreement between two possible matches [with (i,i) indicated by the red box and (j,j) indicated by the dark blue box in Fig. 1(d)], where we propose to use the appearance-based line patch, along with simple geometric relationship [7], to robustly characterize their coherence. The continuous relaxed assignment matrix $X = [X_{i,i}]_{N \times N}$ ($X_{i,i}$ [0,1]) can be optimized by finding the cluster of correspondences among the diagonal elements of M while maximizing off-diagonal pairwise agreements. To alleviate the potential ambiguity in determining one-to-one correspondences [Fig. 1(e)] directly from the one-to-many assignment matrix X, the sparsity constraint is further applied to X to suppress distraction from ambiguous matches during the correspondence detection procedure.

A. Limitation of Conventional Graph Matching Method

In conventional graph matching methods, such as the SMAC method, the coherence between possible matches (i,i) and (j,j) is measured only by the geometric distance $|d(i,j) - d(i,j)|/\min(d(i,j), d(i,j))$ and the angle between two matches/correspondences (i,i) and (j,j) and. Here, $d(\cdot, \cdot)$ denotes the Euclidean distance between points. However, as we will show an example in Fig. 2, geometric information alone is not sufficient to prevent the ambiguous correspondences.

In the SMAC method, the energy function is defined to maximize the following quadratic score function of x:

$$J(x) = \frac{x^T M x}{x^T x} \ s.t. \ Ax = 1^T \ \text{and} \ x \in [0,1]^{NN'} \quad (1)$$

where assignment vector x is a NN column vector after concatenating each row of X. Each element $x_m (m = 1, \ldots, NN)$ in the vector x is associated with a particular correspondence (i,i) in the assignment matrix X, $x_m = X_{i,i}$. Since optimization of $J(x)$ is NP-hard, to derive an approximate solution to this problem, each element in x is relaxed to be a continuous value between 0 and 1. Additionally, the objective function $J(x)$ is subject to the affine constraint $Ax = 1^T$ (as in [3]) to enforce one-to-one correspondence. A is a $(N + N) \times NN$ selection matrix applied to vector x (vectorization of X^T) to ensure the sum of each column or each row of X is equal to 1, $\sum_{i=1}^{N} X_{i,i'} = 1$ or $\sum_{i'=1}^{N'} X_{i,i'} = 1$. A spectral relaxation technique can be used to maximize the energy function in (1).

Fig. 3(a) shows the optimized assignment matrix X by the SMAC method. It can be observed that the distribution of assignment in most rows (or columns) of X is not sharp [see an example of $X_{i,i}$ values along the pink line shown in the top of Fig. 3(c)], indicating that it is still very difficult to determine the one-to-one correspondence for each feature point based on the one-to-many correspondences (each with similar likelihood). A good solution is to keep only the large assignments while suppressing distractions from ambiguous matches. To achieve this, we propose to: 1) use appearance-based line patches to exclude incorrect matches when constructing the affinity matrix M and 2) further apply sparsity on the assignment matrix X during the optimization procedure for suppressing the influence from ambiguous matches. These strategies will be described independently below.

B. Construction of a Robust Affinity Matrix With Similarity of Line Patches

It is clear that matching performance is largely dependent on the affinity matrix M, especially for the off-diagonal elements which characterize the pairwise agreement between possible correspondences (i,i) and (j,j). However, conventional graph matching methods consider only the geometric coherence between (i,j) and (i,j). Although local image descriptors can be used to measure the appearance similarities between feature point t_i and s_i, as well as between t_j and s_j, they can still fail to discriminate unreasonable matches as shown in Fig. 2. In this example, there are two model feature points (t_1 and t_2) and three subject feature points ($S_1 \sim S_3$). The subject feature points S_1 and S_2 (blue circles) are the correct matches of model feature points t_1 and t_2 (white circles), while s_3 (blue triangle) is an incorrect match to t_2. However, neither geometric coherence nor local descriptor based measurement is able to distinguish the incorrect correspondence $(2,3)$ from the correct one $(2,2)$ in the affinity matrix M, which affects the optimization of assignment matrix X in (1).

To solve this problem, we define the *line patch* which uses a sequence of intensity profiles along the line connecting the two feature points in the model or subject image, for measuring inter-pair agreement, along with the geometric distance as used in the conventional graph matching method. In Fig. 2, the image intensity profiles along the lines $\overline{s_1 s_2}$, $\overline{t_{1'} t_{2'}}$, and $\overline{t_{1'} t_{3'}}$ are displayed as blue, green, and white stripes, respectively. The collection of intensity profiles along the underlying stripes is referred to as the *line patch* in our method and is a measure of the pairwise agreement of possible matches. Specifically, we use normalized cross correlation to measure the similarity between two line patches. As shown in the Fig. 2(c), the pairwise agreement measured by the line patches is able to distinguish between the correct and incorrect matches in this situation, even though the inter-pair geometric distances are small. Here, we note that the radius of intensity profiles is set to five pixels and we uniformly sample 60 local intensity profiles along each line patch. Therefore, the total number of intensity values included in the line patch of our method is 11×60.

C. Sparse Constraint on Assignment Vector

Although the one-to-many correspondence strategy ensures detection of all possible matches for each feature point, it also introduces many ambiguous matches, which can affect the final one-to-one correspondences as shown in Fig. 3(a). Inspired by the discriminative power of sparse representation, we apply the l_1-norm on the assignment vector x to require the number of nonzero elements in x to be as small as possible. Since the affine constraint $Ax = \mathbf{1}^T$ in (1) specifies each feature point to have at least one correspondence, the l_1-norm regularization term on the entire vector x eventually leads to the sparsity on the possible matches for each feature point.

The advantage of using l_1-norm regularization $\|x\|_1$ is demonstrated in Fig. 3(b). Compared with the assignment matrix obtained by SMAC without l_1-norm constraint, the distribution of assignments along each row and each column of matrix X is much sharper by our method. It is easier to finally apply the Hungarian algorithm [22] or more advanced IPFP method [19] to binarize X and obtain the one-to-one correspondences. In order to evaluate only the effectiveness of including l_1-norm regularization in correspondence detection, both SMAC and our sparse graph matching methods are performed on the same affinity matrix.

D. New Energy Function for Sparse Graph Matching

By incorporating the two novel improvements as proposed above, our new energy function for graph matching becomes

$$F(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}-\gamma\cdot\|\boldsymbol{x}\|_1,\ s.t.\ \boldsymbol{A}\boldsymbol{x}=\boldsymbol{1}^T\text{and }\boldsymbol{x}\in[0,1]^{NN'}\quad(2)$$

where   is the scalar value used to control the strength of sparsity constraint.

The first term in (2) is the graph matching score, which measures the overall agreement among all possible correspondences. By maximizing this score, we find a set of correspondences that best preserve pairwise agreement. The first term is similar to (1), except that: 1) the affinity matrix $\boldsymbol{M}$ in our method is constructed by using our proposed line patch to enhance the measurement of the pairwise agreement (i.e., off-diagonal elements in $\boldsymbol{M}$); and 2) we use the simplified term $\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}$, instead of $\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}/\boldsymbol{x}^T\boldsymbol{x}$ in (1) following the classical graph matching methods [6], [9] [19]. The second term $\|\boldsymbol{x}\|_1$ is the sparsity constraint term, which is used to suppress the ambiguous matches. This l_1-norm term regularizes the assignment vector $\boldsymbol{x}$, which keeps the number of nonzero elements in $\boldsymbol{x}$ small. In this way, the solution will exclude the ambiguous matches, by keeping only correspondences with relatively high confidence.

It is worth noting that the line patch is used to better describe the pairwise agreement among possible correspondences, which can also facilitate other conventional graph matching methods in terms of building a better affinity matrix $\boldsymbol{M}$. After obtaining the affinity matrix $\boldsymbol{M}$, we will propose our l_1-norm based optimization approach for (2) in Section III.

E. Optimization for the Sparse Graph Matching

In order to solve (2), we incorporate the affine constraint $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{1}^T$ into the energy function by using the l_1-norm regulation as below

$$F(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}-\gamma\cdot\|\boldsymbol{x}\|_1-\lambda\cdot\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{1}^T\|_2^2$$
$$s.t.\ \boldsymbol{x}\in[0,1]^{NN'}\quad(3)$$

where   and   are the two weights used to adjust the proportions of l_1-norm and l_2-norm, respectively. Practically, the values of   and   are set by training and then fixed for all the testing experiments (i.e.,   = 20 and   = 10 used in this paper). The above energy optimization problem is nonlinear and usually nonconvex, which makes it difficult to solve in its original form without good initialization. However, because the value of each element in assignment vector $\boldsymbol{x}$ is non-negative, the energy function $F(\boldsymbol{x})$ in (3) turns into a quadratic

assignment problem by simplifying $\|\boldsymbol{x}\|_1$ to $\|\boldsymbol{x}\|_1=\sum_{m=1}^{NN'}x_m=\boldsymbol{1}^T\boldsymbol{x}$ and $\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{1}^T\|_2^2$ as $(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{1}^T)(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{1}^T)$. Then, the maximization of $F(\boldsymbol{x})$ becomes the following constrained, indefinite, quadratic programming problem

$$\arg\max_{\boldsymbol{x}}F(\boldsymbol{x})=\arg\min_{\boldsymbol{x}}(-F(\boldsymbol{x}))$$
$$=\arg\min_{\boldsymbol{x}}\left(-\boldsymbol{x}^T\boldsymbol{M}\boldsymbol{x}+\gamma\cdot\|\boldsymbol{x}\|_1+\lambda\cdot\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{1}^T\|_2^2\right)$$
$$=\arg\min_{\boldsymbol{x}}\left(\boldsymbol{x}^T(\lambda\cdot\boldsymbol{A}^T\boldsymbol{A}-\boldsymbol{M})\boldsymbol{x}+(\gamma\cdot\boldsymbol{1}^T-2\lambda\cdot\boldsymbol{A})\boldsymbol{x}+\lambda\cdot\boldsymbol{1}\cdot\boldsymbol{1}^T\right)$$

which can be efficiently solved by the trust region reffective algorithm [23]. This optimization algorithm can be guaranteed to find a stationary, local solution for our nonconvex function. Finally, we can adopt the simple Hungarian method [22] to determine the one-to-one correspondence from the assignment vector $\boldsymbol{x}$ which maximizes the objective function $F(\boldsymbol{x})$.

Equipped with line patch and sparsity constraint, we can improve the matching performance of the conventional graph matching method. In order to reduce the challenge of large inter-subject variations, which is common in many medical applications, we go one step further and integrate our improved sparse graph matching method into a hierarchical correspondence matching framework as described below.

F. Hierarchical Correspondence Matching Framework for Sparse Graph Matching

To further reduce the risk of mismatches, we integrate our sparse graph matching method into a hierarchical correspondence matching framework. Specifically, we only use a small number of distinctive feature points with strict selection criteria at the beginning of correspondence detection. After applying our sparse graph matching method, we use thin-plate splines (TPS) to interpolate the dense deformation field according to the tentatively detected correspondence set. Since some feature points with large deformations might fail to find correct correspondence, we resort to using an interpolated deformation field to gradually warp the model image to the subject image domain and then apply correspondence detection again by sparse graph matching. In our hierarchical framework, we repeat this procedure, adding more and more feature points by relaxing the selection criteria until the matching results do not change. Since we detect the correspondence between the warped model image and the subject image, the matching result becomes more robust than directly finding correspondences between the original model and subject images. Hereafter, we call our hierarchical sparse graph matching method as HSGM.

III. Clinical Application of Hierarchical Sparse Graph Matching

In this section, we describe one clinical application of our hierarchical sparse graph matching method in hand X-ray images. Our method can also be applied to many other applications in medical imaging area, as well as in computer vision area.

Accurate correspondence detection in hand X-ray images plays an important role in a computer-assisted bone age assessment system. For instance, it is tedious and time consuming for radiologists to assess skeletal maturity through visual inspection. To assist radiologists in performing a more accurate assessment, various bone age assessment methods have been developed [24], [25] that compare the morphological patterns of the subject image at specific landmarks with a model image [26]. These anatomical landmarks are often located at bone joints and extremes. Fig. 4(a) displays 30 important landmarks in hand X-ray, with 14 points on phalanges, 10 points on metacarpals, one point on carpals, two points on ulna, and three points on radius.

Thus, the goal in this application is to efficiently identify these 30 landmarks in a new hand X-ray image by using our HSGM method. In order to deal with the large inter-subject anatomical variations as shown in Fig. 4(b), we further extend our hierarchical sparse graph matching method into a multi-model framework, which has been demonstrated to be more robust than single model framework in this area. Given the hand X-ray image, we apply our hierarchical sparse graph matching method to each model image, and identify the correspondence pairs for each landmark by the following steps: 1) calculate the median location of all detected spatial correspondences in both model images; 2) select K nearest correspondences; 3) set $K = K - 1$ and iteratively repeating step 1) and 2) until K decreases to 2; 4) use the mean location of K selected correspondences as the identified location of underlying landmark in subject image.

IV. Experiments

To evaluate the performance of our HSGM method, we first comprehensively examine each component of our HSGM method using the hand X-ray images. Then, we demonstrate the overall performance of our HSGM method on both synthetic and real data, comparing with three state-of-the-art graph matching methods: SMAC, *h*-GM, and SM + IPFP. Note that all parameters in HSGM are fixed in all experiments. Also, for other three methods under comparison, we obtained their implementations from authors' websites.[1]

A. Evaluation of the Components of Our HSGM Method on Hand X-ray Images

A publicly available USC hand dataset[2], including 43 left-hand radiographs, was used for performance evaluation. The resolution for each image is 0.1 mm × 0.1 mm [27]. Thirty land-marks were manually placed on each model image, as shown in Fig. 4. Since each hand image has manually placed landmarks, matching performance is evaluated by calculating the discrepancy between the ground-truth and the estimated correspondences by the graph matching methods. We first examine the contribution of each novel component proposed in our method in this section, and then report the overall performance of our HSGM method by comparison with SMAC, h-GM, and SM + IPFP in the Section IV-C.

Before correspondence matching, affine registration is first performed to align each subject image to the selected model image. For the model image, thirty manually placed landmarks are used as its feature points. For each subject image, we use the automatic landmark detection method in [28] to select around 450 candidate feature points. In all three comparison methods, the normalized cross correlation between local intensity patches (patch size 35 × 35 is used to measure the point-wise similarities for all possible correspondences (i.e., diagonal elements in affinity matrix M).

1) Contribution of Sparsity Constraint—To evaluate the contribution of this novel sparsity component in correspondence detection, we degrade our HSGM method by keeping only the sparsity constraint, *without using line patch, multiple models, and hierarchical matching*, as will be evaluated next. Two hand images are randomly selected as model images. We use SMAC, *h*-GM, SM + IPFP, and our degraded HSGM method to identify correspondences between each model image with the other 42 hand images. Table I shows the mean and standard deviation of matching errors between the ground-truth and the estimated correspondences by SMAC, *h*-GM, SM + IPFP, and our degraded HSGM method. In this experiment, our degraded method yields a lower matching discrepancy than the methods with which it is being compared.

To further evaluate the matching robustness according to the difficulty level of each subject image, we classify the subject images into three groups (with small, medium, and large landmark distances) based on the average distance of 30 landmarks between model and subject images after affine alignment. Fig. 5(a) shows the mean distance errors in each group by SMAC (blue), *h*-GM (green), SM + IPFP (brown), and our degraded HSGM method (red). Our degraded HSGM method performs best in the group with large inter-subject variations. Next, we roughly divide each subject image into two partitions, fingers and palm. Since the image contrast is poorer in the palm than in the finger, landmarks in the palm are more challenging to identify than landmarks in the fingers. We separately calculate the average matching errors in these two regions, namely the regions of lower and higher

contrast, to evaluate performance. The results by all four methods are shown in Fig. 5(b). Again, our degraded HSGM method achieves the best performance in the palm region.

2) Contribution of Line Patch—Since it is straightforward to incorporate our line patch innovation into other conventional graph matching methods to aid in the construction of a more robust affinity matrix, we show the matching performance of each graph matching method, before and after being equipped with the line patch method. All results are summarized in Table II. As we can see, the inclusion of the line patch significantly improves the matching performance for SMAC, *h*-GM, and SM + IPFP.

After returning our degraded HSGM (evaluated in Table I) to its described state with line patch, its performance is also further improved. The statistics of matching error is improved from 1.20 ± 1.50 mm (Table I) to 0.98 ± 1.08 mm for the model image #1 and from 1.36 ± 1.89 mm (Table I) to 1.07 ± 1.28 mm for the model image #2, respectively, as shown in the last two columns of Table II.

3) Contribution of Multiple Models—In this experiment, we evaluate the contribution of multiple model images by calculating the matching error as a function of the number of model images used. We use a maximum of eight model images. Fig. 6 shows the curve of mean matching error versus the number of model images used in our HSGM method. The mean matching error is computed from the remaining 35 subject images. The average matching error decreases with an increase in the number of model images. The decrease in matching error is especially rapid with the addition of the first few additional model images.

4) Contribution of Hierarchical Correspondence Matching Framework—
Hierarchical matching is able to improve the matching performance by gradually deforming the model image to the subject image and iteratively detecting correspondences between deformed model image and subject image. Fig. 7 demonstrates the contribution of our hierarchical matching framework by showing the decrease in matching error over the course of the matching procedure. Eight hand images serve as model images in this experiment. The mean matching error of all 35 subjects at the beginning, middle, and end of correspondence matching procedure are shown in blue, green, and red bars in Fig. 7, respectively. The matching result at the beginning stage can be considered as the degraded HSGM method *without hierarchical matching*. For each subject, the matching error decreases at each time point, indicating the advantage of using hierarchical strategy in correspondence matching.

B. Performance Evaluation on Synthetic Data

Following the evaluation performed in [7], we generate a large number of simulated affinity matrices in two ways: 1) Adding increasing amounts of noise to affinity matrices. The noise follows the uniform distribution $U(0, \tau)$, where $\tau = 1, \ldots, 10$. For each τ, we simulate 30 instance of affinity matrix. 2) Adding a different percentage of outliers to the subject feature points data. The outlier-to-data ratio ranges from 0 to 1. Similarly, 30 instances of affinity matrix are simulated with same number of outliers. Examples of an original affinity matrix, a degraded affinity matrix with added noise, and an enlarged affinity matrix with added outliers are shown in Fig. 8(a)–(c).

Four graph matching methods, SMAC, *h*-GM, SM + IPFP, and our HSGM method, are applied to each simulated affinity matrix to estimate the assignment matrix X. Since the affinity matrices have been already constructed, our HSGM does not take advantage of the line patch component in this experiment, which is called degraded HSGM as in Fig. 9. In this figure, the top and bottom rows show the results of noise and outlier cases, respectively,

with the ground truth assignments shown in the first column. The second through fifth columns show the estimated assignment matrices X obtained by SMAC, h-GM, SM + IPFP, and HSGM method, respectively. The distribution of correspondence assignment estimated by our HSGM method is much sharper than SMAC and h-GM methods, along both each row and each column of X. Although SM + IPFP can also obtain sparse result, the estimated results are less accurate than our HSGM method, compared with the ground-truth assignment in the first column.

Given the ground-truth correspondences, we can also calculate the number of mismatches after binarizing the estimated assignment matrix X by the four graph matching methods. Fig. 10(a) shows the curve of average mismatch ratios versus noise level by SMAC (blue), h-GM (green), SM + IPFP (brown), and HSGM method (red). With the same legend, Fig. 10(b) displays the average mismatch ratio versus the number of outliers. Again, our HSGM method consistently yields the lowest mismatch ratio in both cases (even though line patch is not used in this experiment).

C. Performance Evaluation on Hand X-ray Images

Here, we demonstrate the overall matching performance of our HSGM method on real hand X-ray images. Line patch is applied to improve inter-pair agreement during the construction of the affinity matrix; a sparsity constraint is enforced during estimation of the assignment matrix, in order to improve the matching quality; and our sparse graph matching method is integrated into a hierarchical correspondence matching framework. Additionally, eight hand images are used as multiple models to deal with the large model-subject anatomical variations found in real clinical applications. We apply SMAC, h-GM, SM + IPFP, and HSGM method to the rest 35 subject images for performance comparison.

The overall matching results by SMAC, h-GM, SM + IPFP, and our HSGM method are shown in Table III. HSGM method significantly outperforms other three state-of-the-art graph matching methods. Typical correspondence matching results by the four methods are also shown in Fig. 11, where correct matches are indicated with solid cyan lines and incorrect matches are indicated with dashed pink lines. Our HSGM method is able to correctly identify all 30 correspondences, while SMAC method fails at two landmarks (#3 and #30), h-GM fails at one landmark (#2), and SM + IPFP fails at two landmark (#20 and #30). For better illustration, we also zoom in to the selected regions containing the landmarks #2, #3, #20, and #30 (shown by white rectangles in the original images).

We also quantitatively show the performance of SMAC, h-GM, SM + IPFP, and our HSGM method over different levels of model-subject anatomical variation and in different regions of hand images in Fig. 12(a) and (b), respectively. HSGM method achieves the lowest mean matching error in all conditions.

V. Discussion

A. Computational Cost and Complexity Analysis

In the graph matching method, the major computation costs are related to the construction of affinity matrix M from a set of possible correspondences, and the optimization of correspondence assignments. Since the optimization of (3) is a quadratic programming problem, its computational complexity highly depends on the size of affinity matrix M. In general, the upper bound of the computational complexity is $O((NN)^2)$ for a $NN \times NN$ affinity matrix M. However, it is worth noting that the actual computation time is much less than $O((NN)^2)$, since M is a very sparse matrix (i.e., nearly 99% elements of M are zeros in the experiment for the hand X-ray images). For example, in the hand X-ray images (with 30 and ~450 feature points in the model and subject images, respectively), the computation

time for constructing the affinity matrix M is 580.48 s. The average computation time for optimization part by SMAC, h-GM, SM + IFPF, and our HSGM method are 3.58, 1.69, 23.12, and 89.59 s, respectively. All the algorithms were implemented in Matlab 7.12/Linux without using parallelism technique, and run on a workstation equipped with 8 dual-core 2.8 GHz CPU and 32GB memory. The construction of affinity matrix M is the most time-consuming part in all the examined graph matching methods. Fig. 13 shows the average time consumption by SMAC, h-GM, SM + IFPF, and our HSGM method. Although our method is the most computationally costly, the overall computation time is still comparable with other conventional graph matching methods. Our code is currently implemented by MATLAB, but we believe the computation time can be reduced significantly with optimization.

Fig. 14 demonstrates the change in average computation time when increasing the number of subject feature points from 100 to 800, where the number of model feature points is fixed at 50. The computation burden increases fastest with the SM + IPFP method. Considering its significant improvement in correspondence matching, our HSGM method provides the best results in terms of both efficiency and accuracy.

B. Techniques for Binarizing Correspondence Assignments

Given the estimated correspondence assignment x, it is important to appropriately discretize them, which is an important postprocessing step in all graph matching methods. In the literature, Hungarian method [22] is commonly used, due to its efficiency. Recently, IPFP (Integer Projected Fixed Point) method was also proposed in [19] with potentially better performance. Since the evaluation of this postprocessing step is meaningful, we would like to test the performances of both conventional Hungarian and more advanced IPFP method when they are integrated with our HSGM method. The matching accuracy by HSGM + Hungarian and HSGM + IPFP on the X-ray hand images is shown in Table IV. It can be observed that HSGM + Hungarian has the comparable results with HSGM + IPFP, which indicates that the gain of matching accuracy comes mainly from our HSGM method.

C. Future Work

Our future works include reducing the computation time and extending our HSGM method to other medical image applications. To speed up our method, we will re-implement and optimize our method in C/C++. Our proposed graph matching method has wide applications in the field of medical imaging, such as deformable registration of MR brain images and motion estimation in lung 4D-CT data. Although it is straightforward to apply our method to these applications, we still need to handle large number of feature points in these applications. For example, there could be as many as ~100,000 feature points in the MR brain images. To address this challenge, our proposed solution is to first partition the entire brain image into a number of overlapping blocks. A small number of feature points will be selected in each block and correspondence will be established by our HSGM method, and the correspondences in the entire brain will be the integration of correspondence detection in all overlapping blocks. By decomposing a large-scale problem into a number of small-scale problems, which can be solved one by one efficiently, our HSGM method can be made to handle a large number of feature points.

VI. Conclusion

In this paper, we have proposed a hierarchical sparse graph matching method to improve the accuracy in correspondence detection. Our contributions are threefold: 1) a new concept of line patch is proposed to robustly characterize the pairwise agreement between two possible matches/correspondences; 2) a sparsity constraint is further introduced to suppress the

influence from ambiguous matches during correspondence assignment; and 3) a hierarchical, multi-model framework is built based on our proposed graph matching method. Better results have been achieved on both synthetic data and hand X-ray images when compared to the state-of-the-art graph matching methods: SMAC, h-GM, and SM + IPFP. In the future, we will extend our method to other medical applications, such as deformable brain registration and motion estimation for lung 4D-CT data.

Acknowledgments

References

1. Chui HL, Rangarajan A. A new point matching algorithm for non-rigid registration. Comput Vis Image Understand. 2003; 89(2–3):114–141.

2. Castillo E, et al. Four-dimensional deformable image registration using trajectory modeling. Phys Med Biol. 2010; 55(1):305–327. [PubMed: 20009196]

3. Maciel J, Costeira JP. A global solution to sparse correspondence problems. IEEE Trans Pattern Anal Mach Intell. Feb; 2003 25(2):187–199.

4. Jiang H, Drew MS, Li ZN. Matching by linear programming and successive convexification. IEEE Trans Pattern Anal Mach In-tell. Jun; 2007 29(6):959–975.

5. Shen D, Wong W, Ip H. Affine-invariant image retrieval by correspondence matching of shapes. Image Vis Comput. May; 1999 17(7):489–499.

6. Leordeanu M, Hebert M. A spectral technique for correspondence problems using pairwise constraints. Int Conf Comput Vis. 2005:1482–1489.

7. Cour T, Srinivasan P, Shi J. Balanced graph matching. Adv Neural Inf Process Syst. 2006:313–320.

8. Zass R, Shashua A. Probabilistic graph and hypergraph matching. IEEE Conf Comput Vis Pattern Recognit. 2008:1–8.

9. Torresani L, Kolmogorov V, Rother C. Feature correspondence via graph matching: Models and global optimization. Proc 10th Eur Conf Comput Vis: Part I. 2008:596–609.

10. Im K, et al. Quantitative comparison and analysis of sulcal patterns using sulcal graph matching: A twin study. NeuroImage. 2011; 57(3):1077–1086. [PubMed: 21596139]

11. Metzen JH, et al. Matching of anatomical tree structures for registration of medical images. Image Vision Comput. 2009; 27(7):923–933.

12. Li F, et al. Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. IEEE Trans Med Imag. Jan; 2010 29(1):96–105.

13. Lowe DG. Distinctive image features from scale-invariant key-points. Int J Comput Vis. 2004; 60(2):91–110.

14. Calonder M, et al. BRIEF: Computing a local binary descriptor very fast. IEEE Trans Pattern Anal Mach Intell. Jul; 2012 34(7):1281–1298.

15. Leordeanu M, Sukthankar R, Hebert M. Unsupervised learning for graph matching. Int J Comput Vis. 2012; 96(1):28–45.

16. Caetano TS, et al. Learning graph matching. IEEE Trans Pattern Anal Mach Intell. 2009; 31(6):1048–1058. [PubMed: 19372609]

17. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B. Jun; 1996 58(1):267–288.

18. Wright J, et al. Robust face recognition via sparse representation. IEEE Trans Pattern Anal Mach Intell. Feb; 2009 31(2):210–227. [PubMed: 19110489]

19. Leordeanu M, Hebert M, Sukthankar R. An integer projected fixed point method for graph matching and MAP inference. Proc Neural Inf Process Syst. 2009:1114–1122.

20. Zheng Y, et al. Landmark matching based automatic retinal image registration with linear programming and self-similarities. Int Conf Inf Process Med Imag. 2011:674–685.

21. Xue Z, Shen D, Davatzikos C. Determining correspondence in 3-D MR brain images using attribute vectors as morphological signatures of voxels. IEEE Trans Med Imag. Oct; 2004 23(10): 1276–1291.

22. Burkard, R.; Dell'Amico, M.; Martello, S. Assignment Problems. Philadelphia, PA: SIAM; 2012.

23. Nocedal, J.; Wright, SJ. Numerical Optimization. 2. New York: Springer Verlag; 2006.

24. Martín-Fernández MÁ, et al. Automatic articulated registration of hand radiographs. Image Vis Comput. 2009; 27(8):1207–1222.

25. Gertych A, et al. Bone age assessment of children using a digital hand atlas. Comput Med Imag Graph. 2007; 31(4–5):322–331.

26. Pietka E, et al. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/ metaphyseal ROI extraction. IEEE Trans Med Imag. Aug; 2001 20(8):715–729.

27. Cao F, et al. An image database for digital hand atlas. Proc SPIE Med Imag: PACS Integrated Med Inf Syst: Design Evaluat. 2003:461–470.

28. Zhang P, Cootes T. Automatic construction of models for initialising groupwise registration. IEEE Trans Med Imag. Feb; 2012 31(2):341–358.
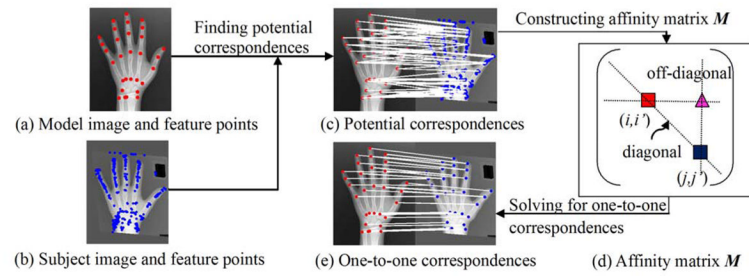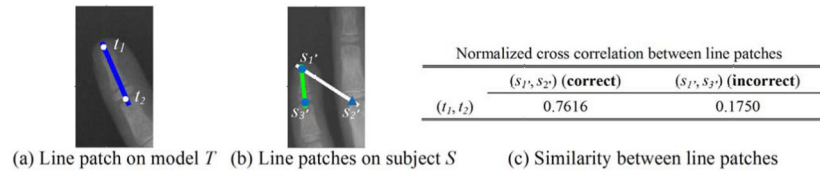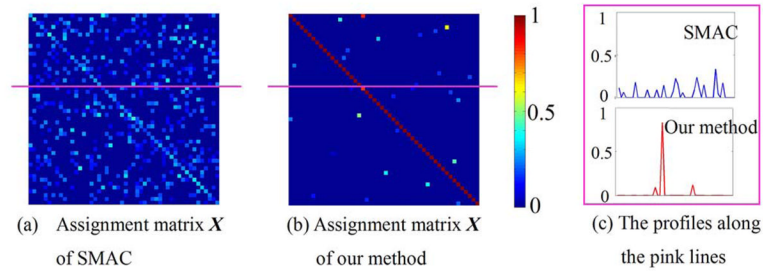
Fig. 1.
The scheme of the proposed anatomical correspondence detection by graph matching. After detecting thefeature points in the modelimage [red points in (a)] and the subject image [blue points in (b)], the one-to-one correspondence (e) is optimized from the affinity matrix *M* (d), which is constructed based on the established multiple correspondences (c).

	Normalized cross correlation between line patches	
	$(s_{1'}, s_{2'})$ **(correct)**	$(s_{1'}, s_{3'})$ **(incorrect)**
(t_1, t_2)	0.7616	0.1750

(a) Line patch on model T (b) Line patches on subject S (c) Similarity between line patches

Fig. 2.
Demonstration of using line patches in distinguishing incorrect matches. Three possible correspondences are shown, (t_1, s_1') (correct), (t_2, s_2') (correct), (t_2, s_3') (incorrect). The pairwise agreement between correct matches (t_1, s_1') and (t_2, s_2'), is measured by the similarity of blue and green line patches, while another pairwise agreement between (t_1, s_1') and (t_2, s_3') is measured by blue and white line patches for the case of incorrect match. Since each line patch uses the intensity profile between two feature points, it is able to suppress the effects of incorrect matches in the affinity matrix, as quantitatively measured by the normalized cross correlation provided in the right panel of this figure.

(a) Assignment matrix X of SMAC. (b) Assignment matrix X of our method. (c) The profiles along the pink lines.

Fig. 3.
The assignment matrix X optimized from the same affinity matrix by SMAC method (without sparsity constraint) and our graph matching method (with sparsity constraint). (a) Assignment matrix X of SMAC. (b) Assignment matrix X of our method. (c) The profiles along the pink lines.
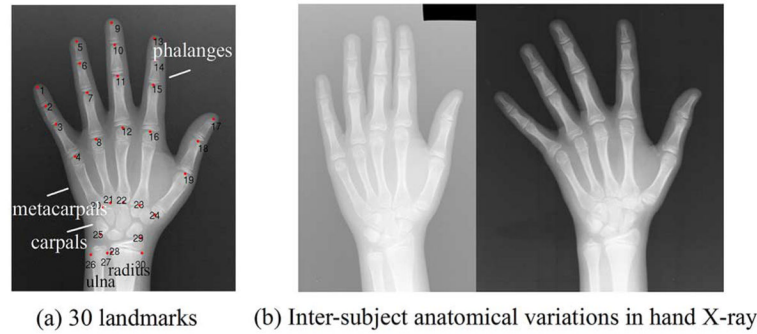
(a) 30 landmarks   (b) Inter-subject anatomical variations in hand X-ray

Fig. 4.
Hand X-ray images. (a) 30 important landmarks in a hand X-ray image, including 14 points on phalanges, 10 points on metacarpals, one point on carpals, two points on ulna, and three points on radius. (b) Two typical X-ray images with high inter-subject anatomical variations.

(a) Different levels of model-subject variations   (b) Different regions

Fig. 5.
Mean matching errors of SMAC, *h*-GM, SM + IPFP, and our degraded HSGM method (a) under different levels of model-subject variations, and (b) at different regions of hand images.

Fig. 6.
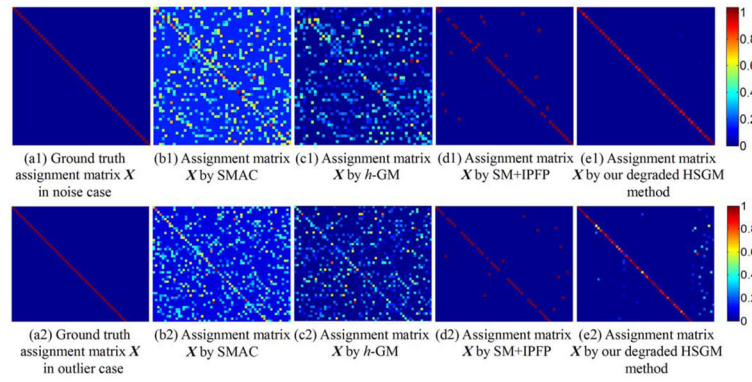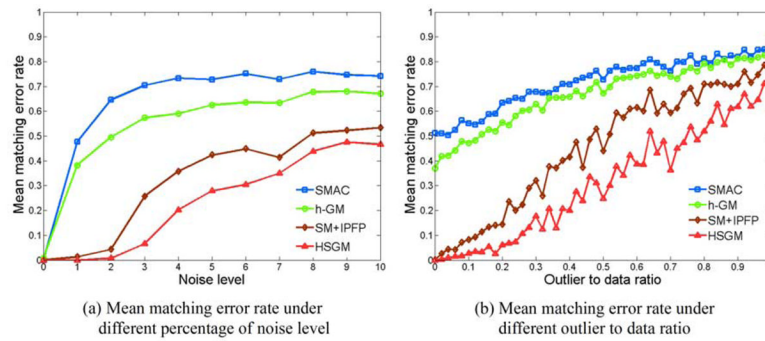The curve of mean matching error versus the number of model images used in our HSGM method.

Fig. 7.
The evolution of mean matching error during the hierarchical correspondence matching procedure in our HSGM method.
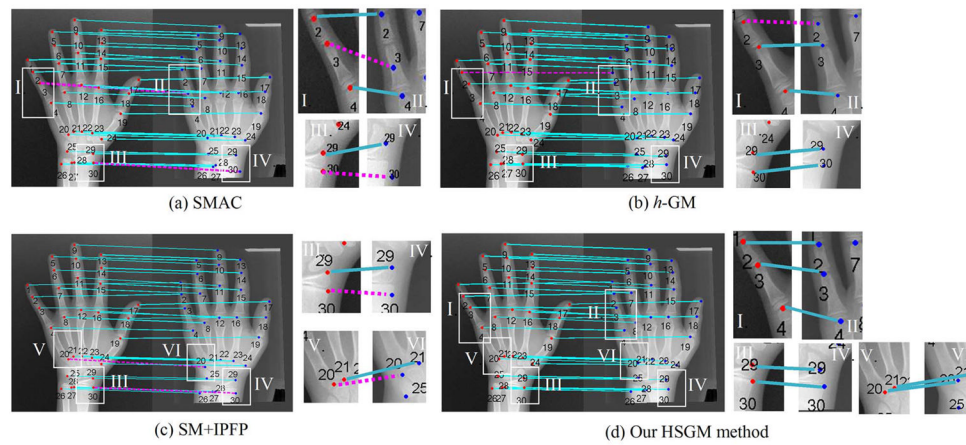
(a) Original affinity matrix *M*   (b) Simulated affinity matrix *M* with noise (c) Simulated affinity matrix *M* with outliers

Fig. 8.
Two typical instances of simulated affinity matrix by adding noise (b) and outliers (c) to the original affinity matrix *M* (a).

Fig. 9.
The estimated assignment matrices *X* by SMAC (second column), *h*-GM (third column), SM + IPFP (fourth column), and our HSGM (fifth column). Compared with the ground truth (first column), our HSGM method achieves the closest estimation results with the sharpest assignment distribution among all graph matching methods, in both noise (top) and outlier (bottom) cases.

(a) Mean matching error rate under different percentage of noise level

(b) Mean matching error rate under different outlier to data ratio

Fig. 10.
Mean matching error rate with respect to (a) noise level and (b) the ratio between outlier and data numbers, by SMAC (blue), *h*-GM (green), SM + IPFP (brown), and our HSGM method (red).

Fig. 11.
Matching results by (a) SMAC, (b) *h*-GM, (c) SM + IPFP, and (d) our HSGM method, with solid cyan lines showing correct matches and the dashed pink lines showing incorrect matches. In the right of (a–d), regions I, III, V represent the enlarged views at landmarks 2, 3, 20, and 30 of the model image, and regions II, IV, VI represent the corresponding enlarged views of the subject image.
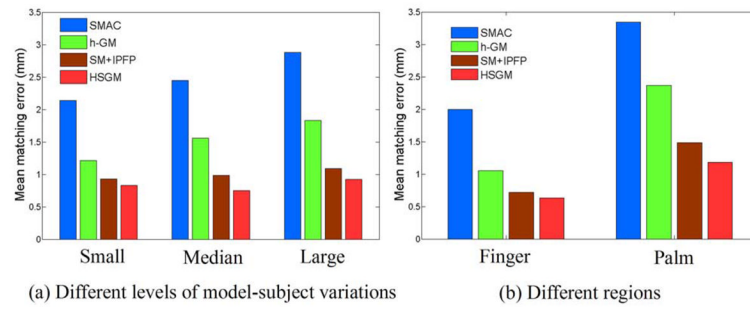
Fig. 12.
Mean matching errors of SMAC, *h*-GM, SM + IPFP, and our HSGM method (a) under different levels of model-subject anatomical variations, and (b) at different regions of hand images.
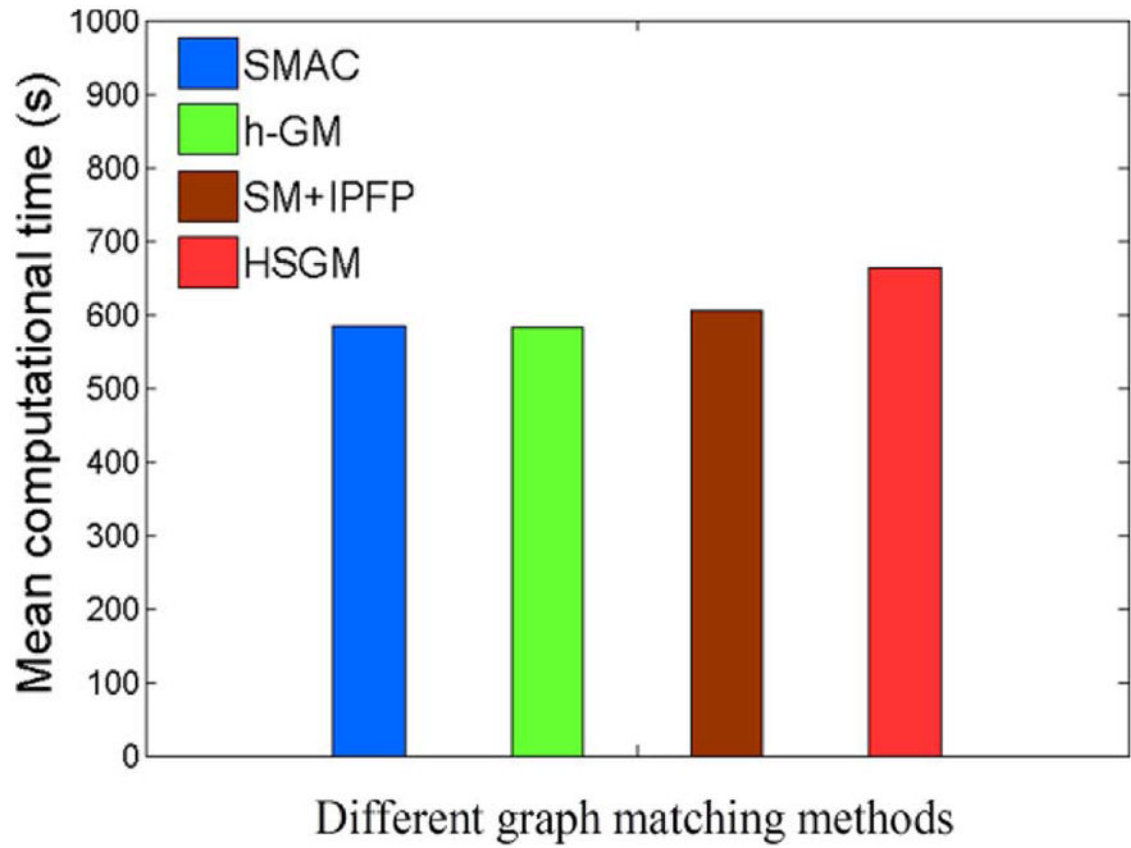
Fig. 13.
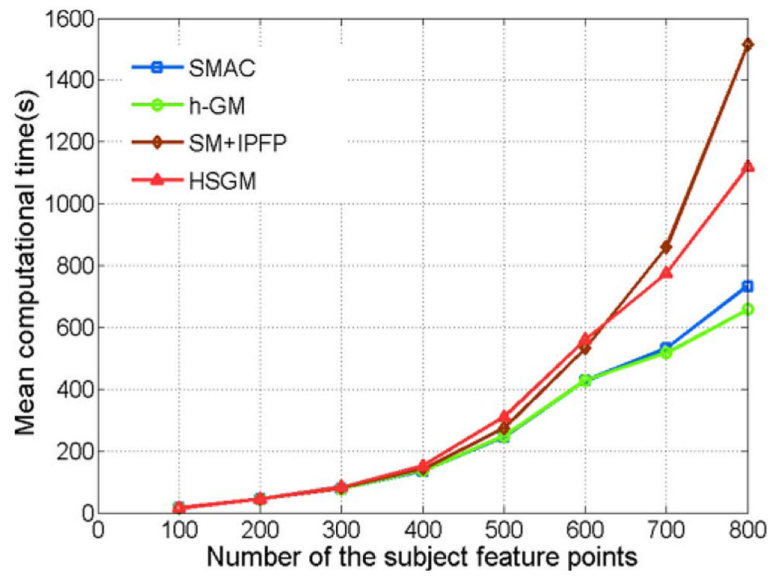The average computation time by SMAC (blue), *h*-GM (green), SM + IPFP (brown), and our HSGM method (red).

Fig. 14.
The change of average computation time with respect to different number of subject feature points by SMAC (blue), *h*-GM (green), SM + IPFP (brown), and our HSGM method (red).

TABLE I

Mean and Standard Deviation of Matching Discrepancy Between the Manual Ground-Truth and the Estimated Correspondences by SMAC, *h*-GM, SM + IPFP, and our Degraded HSGM Method (Using Only the Sparsity Constraint). (Unit: mm)

	SMAC	**h-GM**	**SM+IPFP**	**Degraded HSGM**
Model image #1	1.78±2.54	1.38±2.43	1.41±2.13	**1.20±1.50**
Model image #2	2.12±4.57	1.62±2.71	1.46±2.35	**1.36±1.89**

TABLE II

Mean and Standard Deviation of Matching Errors Between the Ground-Truth and the Estimated Correspondences by SMAC, h-GM, SM + IPFP, and our HSGM, With and Without Line Patch. (Unit: mm)

	Model image #1		Model image #2	
	No line patch	**+line patch**	**No line patch**	**+line patch**
SMAC	1.78±2.54	1.33±1.79	2.12±4.57	1.78±3.96
h-GM	1.38±2.43	1.08±2.18	1.62±2.71	1.32±2.57
SM+IPFP	1.41±2.13	1.11±1.66	1.46±2.35	1.13±1.42
HSGM	1.20±1.50	0.98±1.08	1.36±1.89	1.07±1.28

TABLE III

Mean and Standard Deviation of Matching Errors Between the Manual Ground-Truth and the Estimated Correspondences by SMAC, h-GM, SM + IPFP, and our HSGM Method. (Unit: mm)

	SMAC	h-GM	SM+IPFP	HSGM
8 models	2.43±2.60	1.54±3.07	1.12±1.91	**0.81±0.89**

TABLE IV

Mean and Standard Deviation of Matching Errors Between the Ground-Truth and the Estimated Correspondences by HSGM + Hungarian and HSGM + IPFP, Respectively. (Unit: mm)

	HSGM+Hungarian	HSGM+IPFP
Model image #1	0.98 ± 1.08	0.98 ± 1.24
Model image #2	1.07 ± 1.28	1.03 ± 1.23