TOPOLOGY PRESERVING GRAPH MATCHING FOR PARTIAL FACE RECOGNITION

Yueqi Duan[1,2,3], *Jiwen Lu*[1,2,3,*], *Jianjiang Feng*[1,2,3], *Jie Zhou*[1,2,3]

[1]Department of Automation, Tsinghua University, Beijing, China
[2]State Key Lab of Intelligent Technologies and Systems, Beijing, China
[3]Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing, China
duanyq14@mails.tsinghua.edu.cn; {lujiwen,jfeng,jzhou}@tsinghua.edu.cn

ABSTRACT

In this paper, we propose a topology preserving graph matching (TPGM) method for partial face recognition. Most existing face recognition methods extract features from holistic face images, yet faces in real-world unconstrained environments are usually occluded by objects or other faces, which cannot provide the whole face images for recognition. Latest keypoint-based partial face recognition methods only match on the detected keypoints to remove the occluded regions. However, they simply measure the node-wise similarity without higher order geometrical graph information, thereby depending heavily on descriptors which are susceptible to noises. To address this, our TPGM method estimates a non-rigid transformation encoding the second order geometric structure of the graph, so that more accurate and robust correspondence can be computed with the topological information. Experimental results on three widely used face datasets show that the proposed TPGM outperforms most existing state-of-the-art partial face recognition methods.

Index Terms— Partial Face Recognition, Graph Matching, Keypoint, Topology-Preserving

1. INTRODUCTION

Face recognition is a longstanding computer vision problem and a number of face recognition approaches have been proposed over the past three decades [1–6]. While most methods have achieved impressive performance under controlled conditions where frontal holistic face images are prealigned and normalized, there are still challenges for unconstrained face recognition in many real-world applications. Typical face recognition applications include smart surveillance systems and handheld devices, where faces may be occluded by objects or other faces under crowded scenes. In such scenarios, only partial faces can be obtained and face alignment may fail to work with facial landmarks occluded.

Most face recognition methods including the state-of-the-art CNN approaches [3–6] utilize the whole face images for recognition, which are not applicable to partial face recognition. On the one hand, they assume that each image have the same content of the aligned whole face and describe the holistic facial images for representation, yet the contents of images may be different even for the same person in partial face recognition, e.g. one image without the left eye and another without the mouth. On the other hand, the occluded objects are included in the representation, which harms the discriminative power. Therefore, it is required to design a partial face recognition method which directly recognizes partial faces and is robust to occlusions.

In order to better recognize partial faces, the occluded facial parts should be removed when computing the similarities between probe partial faces and gallery faces. Inspired by the motivations above, a few keypoint-based methods have been proposed by computing the similarities over the detected feature keypoints on the face images [7, 8], such as multi-keypoint descriptor with Gabor ternary pattern (MKD-GTP) [7] and robust point set matching (RPSM) [8], which achieve the state-of-the-art performance on partial face recognition. However, they only exploit node-wise similarity, which depend heavily on descriptors and are susceptible to illumination, deformation and other noises. Graph matching is an effective manner to address such unstableness by exploiting geometric structure of the graph, which enhances the robustness of feature matching in various of visual analysis tasks such as object categorization [9], feature tracking [10] and action recognition [11]. Inspired by the fact that graph matching delivers higher matching accuracy and stronger stableness, we propose a topology preserving graph matching (TPGM) method for partial face recognition by estimating a non-rigid transformation encoding the topological structure and measuring the correspondence between nodes and edges. Fig. 1 illustrates the pipeline of our proposed TPGM method.

* Corresponding author.

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB1001001, the National Natural Science Foundation of China under Grants 61672306, 61572271, 61527808, 61373074 and 61373090, the National 1000 Young Talents Plan Program, the National Basic Research Program of China under Grant 2014CB349304, the Ministry of Education of China under Grant 20120002110033, and the Tsinghua University Initiative Scientific Research Program.

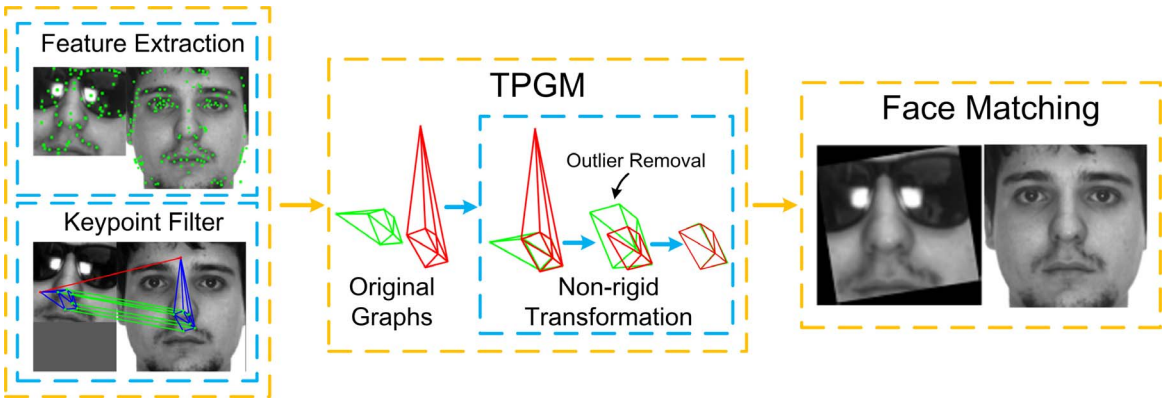

Fig. 1. The flowchart of our proposed TPGM approach for partial face recognition. For each pair of images, we first detect SIFT keypoint descriptors on both faces marked as green dots. Then, the keypoints are coarsely matched and selected based on Lowe's method, where green lines represent the correct matches while the red ones link the wrong matching keypoint pairs. We construct the graphs for both images by deploying Delaunay triangulation with blue lines. In TPGM procedure, a non-rigid transformation is estimated encoding the structural graph iteratively, with outliers removed during the iteration. Finally, the probe partial face is aligned and recognized with the transformation and the matching result.

Experimental results on LFW [12], AR [13] and PubFig [14] show the effectiveness of the proposed approach.

2. RELATED WORK

There have been extensive work on robust face recognition with occlusions in recent years [15, 16]. These methods require face alignment, which has been proved to be critical by Ekenel and Stiefelhagen [17]. However, these face alignment based methods fail to work well with unknown missing facial portions. A number of part-based representations have also been proposed, which can be mainly classified into subregion-based methods [1, 2] and component-based methods [18, 19]. However, as occluded face regions in real-world applications are exceedingly unstructured, both subregion-based methods and component-based methods may fail with incorrect occlusion detection or occluded facial components. More recently, a few keypoint-based approaches have been proposed [7, 8] which remove the occluded facial regions by computing on the detected feature keypoints. Liao *et al.* [7] have proposed a MKD-GTP method which presented a general formulation of the keypoint-based partial face recognition. Weng *et al.* [8] have applied point set matching method by considering the first-order compatibility between point sets, which achieved the state-of-the-art performance. However, these methods only exploit node-wise unary similarity without higher order geometrical graph information, which depend largely on descriptors and are susceptible to noises.

3. PROPOSED APPROACH

In this section, we first model the partial face recognition as a geometric graph matching problem. Then, we present our TPGM approach and the optimization details. Finally, we introduce how to use TPGM for partial face recognition.

3.1. Feature Extraction and Keypoint Filter

In order to present TPGM to match a pair of partial faces, we first extract features on both images, and then filter the candidate matching keypoints.

We follow [8] to extract the "SiftSurfSILBP" descriptor. More specifically, we first utilize Scale-Invariant Feature Transform (SIFT) to detect and describe the local keypoints [20], which is the most successful keypoint descriptor in the literature. We then concatenate the SIFT descriptor with the Speeded Up Robust Features (SURF) to enhance its robustness against illumination variations [21], and apply the Scale Invariant LBP (SILBP) features [22] to obtain more detailed facial textures with four different sets of {P,R} values including {8,1}, {8,2}, {16,2} and {16,3} to obtain four SILBP histograms. "SiftSurfSILBP" combines the discriminative power of SIFT, SURF and SILBP, presenting robustness to rotation, scale and illumination, which is critical for partial face recognition.

The number of keypoints detected by SIFT in a typical 128×128 face image can be hundreds, which suffer from heavy computational cost if we directly apply graph matching on the keypoints. Therefore, we apply Lowe's matching method to filter out obvious outliers at first, and then perform our TPGM method only on the candidate keypoint pairs.

3.2. Topology Preserving Graph Matching

In order to incorporate the pairwise information, we utilize Delaunay triangulation to construct a graph for each image, and present each graph as a 4-tuple $g = \{\mathbf{P}, \mathbf{Q}, \mathbf{T}, \mathbf{G}\}$. We

denote $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_n] \in R^{2\times n}$ as the set of nodes and $\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_m] \in R^{2\times m}$ as the set of edges, where edges are the coordinate differences between nodes, and n and m are the number of keypoints and edges, respectively. Also, we denote the textural features of each node as $\mathbf{T} = [\mathbf{t}_1, \ldots, \mathbf{t}_n] \in R^{d_t\times n}$, and the node-edge incidence matrix as $\mathbf{G} \in \{0,1\}^{n\times m}$ to encode the topology of the graph, where $g_{ic} = 1$ if the cth edge connects the ith node. In particular, $g^P = \{\mathbf{P}^P, \mathbf{Q}^P, \mathbf{T}^P, \mathbf{G}^P\}$ represents the 4-tuple from the probe set, and $g^G = \{\mathbf{P}^G, \mathbf{Q}^G, \mathbf{T}^G, \mathbf{G}^G\}$ is from the gallery.

Let $f^p(\cdot)$ be the non-rigid transformation function of nodes, and $f^q(\cdot)$ be the transformation function of edges, where $f^q(\mathbf{q}_c) = f^p(\mathbf{p}_i) - f^p(\mathbf{p}_j)$ for $\mathbf{q}_c = \mathbf{p}_i - \mathbf{p}_j$. Let $h^p(\cdot)$ be the matching map to connect each probe node to the corresponding gallery node, and $\mathbf{X} \in \{0,1\}^{n_p\times n_g}$ be the binary correspondence matrix between nodes, where $\mathbf{X}_{ij} = 1$ for the matched ith probe keypoint and jth gallery keypoint. We can represent the mapping as $h^p(\mathbf{p}_i^P) = \mathbf{P}^G\mathbf{X}_i^{\mathrm{T}}$. Similarly, we denote $h^q(\cdot)$ as the matching function of edges, and define $\mathbf{Y} \in \{0,1\}^{m_p\times m_g}$ as the binary correspondence matrix between edges. For $\mathbf{q}_i^P = \mathbf{p}_k^P - \mathbf{p}_l^P$ and $\mathbf{q}_j^G = \mathbf{p}_{k'}^G - \mathbf{p}_{l'}^G$, $\mathbf{Y}_{ij} = 1$ if $\mathbf{X}_{kk'} = 1$ and $\mathbf{X}_{ll'} = 1$:

$$\mathbf{Y} = \mathbf{1}\left[(\mathbf{G}^P)^{\mathrm{T}}\mathbf{X}\mathbf{G}^G > 1\right] , \tag{1}$$

where $\mathbf{1}(\text{true}) = 1$ and $\mathbf{1}(\text{false}) = 0$. The physical meaning of $(\mathbf{G}^P)^{\mathrm{T}}\mathbf{X}\mathbf{G}^G \in \{0,1,2\}^{m_p\times m_g}$ is the number of matching nodes connected with the corresponding edges, and edges are matching only if both the two pairs of connected nodes are matching.

With the denotations above, we formulate the objective function for the partial face matching as follows:

$$\begin{aligned}\min J &= K_t(\mathbf{t}^P, h^t(\mathbf{t}^P)) + \lambda_p K_p(f^p(\mathbf{p}^P), h^p(\mathbf{p}^P)) \\ &+ \lambda_q K_q(f^q(\mathbf{q}^P), h^q(\mathbf{q}^P)) ,\end{aligned} \tag{2}$$

where the first term is to minimize the differences of textural features between the matched nodes, the second term aims to minimize the distances between the transformed matched nodes and the physical meaning of the third term is distances between the transformed matched edges.

Textural Matching Cost: The first term in (2) measures the difference in texture feature. We sum up the textural difference of the corresponding feature descriptors to obtain the overall textural matching cost:

$$K_t = \sum_i^{n_p}\sum_j^{n_g}\mathbf{T}_{ij}\mathbf{X}_{ij} = tr(\mathbf{T}^{\mathrm{T}}\mathbf{X}) , \tag{3}$$

where $\mathbf{T}_{ij} = \sqrt{((\mathbf{t}_i^P - \mathbf{t}_j^G)^{\mathrm{T}}(\mathbf{t}_i^P - \mathbf{t}_j^G))}$.

Node Geometric Matching Cost: The second term in (2) measures the positional differences of transformed probe keypoints and corresponding gallery keypoints. In order to incorporate the non-rigid transformation, we apply the thin plate spline (TPS) model, which is widely used for representing flexible coordinate transformations. The bending energy of TPS is defined as follows:

$$\begin{aligned}K_p &= \sum_i^{n_g}\|f^p(\mathbf{p}_i^P) - h^p(\mathbf{p}_i^P)\|_2^2 \\ &+ \frac{\lambda}{\lambda_p}\iint\left[\left(\frac{\partial^2 f^p}{\partial x^2}\right)^2 + \left(\frac{\partial^2 f^p}{\partial x\partial y}\right)^2\right. \\ &+ \left.\left(\frac{\partial^2 f^p}{\partial y^2}\right)^2\right]\mathrm{d}x\mathrm{d}y ,\end{aligned} \tag{4}$$

where $f^p(\mathbf{p}_i^P)$ is computed as

$$f^p(\mathbf{p}_i^P) = \mathbf{A}_{2\times 2}\mathbf{p}_i^P + \mathbf{b}_{2\times 1} + \mathbf{W}\mathbf{\Phi}(i)_{n\times 1} ,$$

$$\mathbf{\Phi}(i)_{n_p\times 1} = \begin{bmatrix} \|\mathbf{p}_i^P - \mathbf{p}_1^P\|_2^2\log(\|\mathbf{p}_i^P - \mathbf{p}_1^P\|_2^2) \\ \vdots \\ \|\mathbf{p}_i^P - \mathbf{p}_{n_p}^P\|_2^2\log(\|\mathbf{p}_i^P - \mathbf{p}_{n_p}^P\|_2^2) \end{bmatrix} . \tag{5}$$

In (5), we denote $\mathbf{A}$ and $\mathbf{b}$ as the affine transformation matrix and the bias vector, $\mathbf{W}$ as the weight matrix and $\mathbf{\Phi}$ as the TPS kernel. We substitute L1 norm for the L2 norm for linearization and rewrite the node geometric matching cost as follows:

$$\begin{aligned}K_p &= ||\mathbf{A}\mathbf{P}_{2\times n_p}^P + \mathbf{b}\mathbf{1}_{n_p}^{\mathrm{T}} + \mathbf{W}\mathbf{\Phi}_{n_p\times n_p} - \mathbf{P}_{2\times n_g}^G\mathbf{X}_{n_g\times n_p}^{\mathrm{T}}||_1 \\ &+ \frac{\lambda}{\lambda_p}\|\mathbf{W}\mathbf{\Phi}\|_1 .\end{aligned} \tag{6}$$

Edge Geometric Matching Cost: The third term in (2) measures the geometric differences of transformed probe edges and corresponding gallery edges. We calculate the distance between edges by the squared L2 norm of vector difference. Let $\mathbf{q}_i^P = \mathbf{p}_{i1}^P - \mathbf{p}_{i2}^P$ and $\mathbf{q}_i^G = \mathbf{p}_{i1}^G - \mathbf{p}_{i2}^G$, and we can obtain:

$$K_q = \sum_i^{m_p}||(f^p(\mathbf{p}_{i1}^P) - f^p(\mathbf{p}_{i2}^P)) - (h^p(\mathbf{p}_{i1}^P) - h^p(\mathbf{p}_{i2}^P))||_2^2 . \tag{7}$$

Reformulating and linearizing the above equation into a matrix form result in:

$$\begin{aligned}K_q &= ||\mathbf{A}\mathbf{Q}_{2\times m_p}^P + \mathbf{W}_{2\times n_p}(\mathbf{\Phi}\mathbf{P}_i - \mathbf{\Phi}\mathbf{P}_j)_{n_p\times m_p} \\ &- \mathbf{Q}_{2\times m_g}^G\mathbf{Y}_{m_g\times m_p}^{\mathrm{T}}||_1 .\end{aligned} \tag{8}$$

3.3. Optimization Details

As the objective function is NP-hard with discrete constraints which cannot be efficiently solved, we perform a linearization and rewrite the objective function into a linear programming (LP) form:

$$\begin{aligned}\min J &= tr(\mathbf{T}^{\mathrm{T}}X) + \lambda_p\mathbf{1}_2^{\mathrm{T}}\mathbf{U}\mathbf{1}_{n_p} + \lambda\mathbf{1}_2^{\mathrm{T}}\mathbf{V}\mathbf{1}_{n_p} \\ &+ \lambda_q\mathbf{1}_2^{\mathrm{T}}\mathbf{M}\mathbf{1}_{m_p} - \lambda_x\mathbf{1}_{n_p}^{\mathrm{T}}\mathbf{X}\mathbf{1}_{m_p},\end{aligned} \tag{9}$$

subject to

$$
\begin{aligned}
&-\mathbf{U} \leq \mathbf{A}\mathbf{P}^P + \mathbf{b}\mathbf{1}_{n_p}^{\mathrm{T}} + \mathbf{W}\mathbf{\Phi} - \mathbf{P}^G\mathbf{X}^{\mathrm{T}} \leq \mathbf{U}\ , \\
&-\mathbf{V} \leq \mathbf{W}\mathbf{\Phi} \leq \mathbf{V}\ , \\
&-\mathbf{M} \leq \mathbf{A}\mathbf{Q}^P + \mathbf{W}(\mathbf{\Phi}\mathbf{P}_i - \mathbf{\Phi}\mathbf{P}_j) - \mathbf{Q}^G\mathbf{Y}^{\mathrm{T}} \leq \mathbf{M}\ , \\
&\mathbf{U} \geq 0,\ \mathbf{V} \geq 0,\ \mathbf{M} \geq 0\ , \\
&\sum_i \mathbf{X}_{ij} \leq 1,\ \sum_j \mathbf{X}_{ij} \leq 1,\ \mathbf{X}_{ij} \geq 0\ , \\
&\mathbf{A}_{1,1} = \mathbf{A}_{2,2},\ \mathbf{A}_{1,2} = -\mathbf{A}_{2,1}\ . \qquad (10)
\end{aligned}
$$

In (9), the binary constraint of $\mathbf{X}_{ij} \in \{0,1\}^{n_p \times n_g}$ is relaxed to a continuous domain $0 \leq \mathbf{X}_{ij} \leq 1$, and $\mathbf{U} \in R^{2\times n_p}$, $\mathbf{V} \in R^{2\times n_p}$ and $\mathbf{M} \in R^{2\times m_p}$ are auxiliary matrices representing upper bounds of K_{p1}, K_{p2} and K_q respectively. As there are the degenerate cases when the optimal solution of $\{\mathbf{A}, \mathbf{W}, \mathbf{X}, \mathbf{b}\}$ are all zeros, we add a penalty term of $\mathbf{X}$ in the formulation.

Region Shrinkage: We apply the successive trust region shrinkage [23, 24] to solve the LP model. For each keypoint $\mathbf{p}_i^P$ in the probe image, we set a trust region D_i in the gallery image, which is a circle with center point $f^p(\mathbf{p}^P)$ and a radius of r. Only keypoints inside the corresponding trust region are considered as the matching candidates, while others are excluded in the optimization process by setting $\mathbf{X}_{ij} = 0$. The trust region cover the entire gallery image at first, and then gradually shrinks to refine the matching candidates during the iteration, which is formulated as: $r^{(n+1)} = \alpha_1 \times r^{(n)}$, where $0 < \alpha_1 < 1$.

Parameter Shrinkage: As face images are globally rigid and locally non-rigid, we confine the transformation to be rigid in the global exploration period, and then gradually convert to non-affine transformation in the local exploitation period. As λ controls the energy of non-affine transformation, we initialize λ to a relatively large value, which is gradually reduced it during iterations. Specifically, we apply $\lambda^{(n+1)} = \alpha_2 \times \lambda^{(n+1)}$, where $0 < \alpha_2 < 1$.

Outlier Removal: We remove the outlier matching pairs during the iteration by calculating the summation of each row of $\mathbf{X}$. For an outlier probe $\mathbf{p}_j^P$, all elements in the j^{th} row of $\mathbf{X}$ are close to 0. Given a threshold $0 < \tau < 1$, if $\sum_{k=1}^{n_p} \mathbf{X}_{jk} < \tau$, we consider the $\mathbf{p}_j^P$ as an outlier and remove it from the candidate keypoints.

Algorithm 1 summarizes the detailed procedure of the proposed TPGM method. As the physical meaning of the first three lambdas are similar with [8], we directly followed the same parameters by setting $\lambda_p = 0.01$, $\lambda_x = min(C)$ and $\lambda = 5$ for simplicity and a fair comparison. λ_q is finally fixed to 0.05 for all the experiments.

3.4. Partial Face Recognition Using TPGM

We obtain the average matching cost $\bar{d} = J_{min}/(\sum_{i,j} \mathbf{X}_{ij})$ simply dividing the matching cost by the number of matching keypoints. We further define the distance between the probe and the gallery as follows:

$$d = \frac{\bar{d}}{\sum_{i,j} \mathbf{X}_{ij}} = \frac{J_{min}}{\left(\sum_{i,j} \mathbf{X}_{ij}\right)^2} = \frac{K_t + \lambda_p K_p + \lambda_q K_q}{\left(\sum_{i,j} \mathbf{X}_{ij}\right)^2}\ . \quad (11)$$

In (11), We separate the distance into two parts, the texture matching distance and the graph matching distance. It is proportional to the average matching cost which indicates the affinity of the matching difference, and it is inversely proportional to the number of matching pairs which indicates the area of similar parts in two images.

Algorithm 1: TPGM

Input: Probe image P, Gallery image G, iteration number T and parameters λ_p, λ_x, λ, λ_q, τ, α_1 and α_2
Output: $\mathbf{A}, \mathbf{b}, \mathbf{W}, \mathbf{X}, \mathbf{Y}$

```
Detect keypoints and extract SiftSurfSILBP features.
Using Lowe's method to filter the keypoints and construct the graph g^P and g^G.
Initialize r = r_init and set the constraint set Φ = ∅.
for t = 1, 2, ··· , T do
    Construct Φ.
    Obtain A, b, W, X, Y using (9).
    Clear the constraint set Φ to ∅.
    Find the outsiders of the trust region and construct the constraint set Φ.
    Shrinkage: r^(n+1) = α1 × r^(n) and λ2^(n+1) = α2 × λ2^(n+1).
    Remove p_i^P, X_i: , if Σ_j X_ij < τ.
    Remove p_j^G, X_:j , if Σ_i X_ij < τ.
end for
Binarize X.
return A, b, W, X, Y
```

4. EXPERIMENTS

We evaluated the performance of our proposed TPGM on LFW [12], AR [13] and PubFig [14], where LFW and PubFig are holistic face datasets and AR is a partial face dataset. We compared the proposed TPGM to several state-of-the-art methods on partial face recognition tasks, which illustrated the effectiveness and robustness of the proposed approach.[1]

For LFW and PubFig, we added random transformation (e.g. random occlusion, rotation) based on the original face images to evaluate the proposed method on partial face recognition tasks.[2] For AR, we tested on the original AR datasets with variant image sizes and disalignment. In the experiments, we compared against 5 baseline algorithms, where three of them are alignment-free matching methods and others are state-of-the-art partial face recognition approaches, including Locally Affine Invariant Robust Point

[1] As the key contribution of the proposed method is for partial face recognition, we only conduct experiments on partial face settings.

[2] We will provide the modified data along with the codes for reproduction and evaluation.

Table 1. Mean verification rate (VR) and corresponding standard deviations (%) comparison on LFW.

Method	VR $\pm S_E$
HDLBP	49.32 ± 1.09
CNN	71.27 ± 1.38
CPD-SiftSurfSILBP	61.62 ± 1.19
MKD-SRC-GTP	68.18 ± 1.77
MLERPM-SiftSurf	65.55 ± 1.53
MLERPM-SiftSurfLBP	67.22 ± 1.83
LAIRPM-SiftSurf	70.40 ± 1.02
LAIRPM-SiftSurfSILBP	70.73 ± 1.68
RPSM-SiftSurf	70.81 ± 1.46
RPSM-SiftSurfSILBP	71.65 ± 1.57
TPGM-SiftSurfSILBP	$\mathbf{73.48 \pm 1.12}$

set Matching (LAIRPM) [8], Robust Point Set Matching (RPSM) [8], Metric Learned Extended Robust Point Matching (MLERPM) [25], Multi-Keypoint Descriptors-Sparse Representation-based Classification-Gabor Ternary Pattern (MKD-SRC-GTP) [7] and Coherent point drift (CPD) [26].

4.1. Results on LFW

The Labeled Face in the Wild (LFW) database [12] consists of 13233 labeled faces of 5749 subjects. The images were obtained under unconstrained circumstances, causing large variations in scale, rotation, illumination and occlusion.

We evaluated the proposed method on the "View 2" of the LFW dataset, which includes 6000 pairs in total with half of them matched and the other mismatched. They are separated into 10 folds with 600 pairs for each fold. In order to obtain partial face images, we first applied the Viola-Jones face detector to detect and crop the facial regions of all images. Then, we randomly transformed those detected holistic facial images to obtain the arbitrary face patches, where both gallery and probe images were partial faces with some facial components cropped out, making it extremely difficult to match. We added high dimensional LBP (HDLBP) method [27] with face alignment by CFAN facial landmark detector and CNN [5] feature with VGG-16 network into comparison, which achieve the state-of-the-art performances on holistic face recognition tasks.

Table 1 shows the mean verification rates of different methods. Our TPGM outperforms other baseline methods which proves its effectiveness, while HDLBP presents the poorest performance on the partial face recognition task. HDLBP employs CFAN to detect 25 landmarks, yet it does not work on partial faces especially when some facial components are occluded. CNN exploits holistic facial images for representation through learning from large amount of outside data. However, the contents of each partial face image are different, which leads to a confusing description. The proposed TPGM directly matches on the image pairs without training procedure and achieves the highest accuracy on partial face recognition, which proves its effectiveness and robustness.

Table 2. Recognition accuracy (%) of different methods on the AR dataset.

Method	S1-G	S1-S	S2-G	S2-S
CPD	71.00	75.67	49.33	61.00
MLERPM	75.00	78.33	53.33	66.67
LAIRPM	87.33	88.33	56.33	81.33
MKD-SRC-GTP	82.33	83.33	57.67	76.33
RPSM	88.67	90.33	63.67	85.67
TPGM	**89.33**	**91.00**	**65.00**	**86.67**

Iteration Number and Computational Time: We tested the average iteration number and computational time of the proposed TPGM. Our hardware configuration comprises of a 2.8-GHz CPU and a 15G RAM. In each experiment, the time complexity varies with the number of initial matching keypoints, where the mean iteration number is 1.4 and the computational time is 0.94s on LFW. The proposed TPGM converges much faster than RPSM, because the geometric information accelerates the convergence.

4.2. Results on AR

The AR database [13] consists of 126 identities including 70 males and 56 females. It contains two sessions and there are 13 face images for each subject in a session, where three of them are taken under various illumination conditions, four with different expressions, three wearing sunglasses and the other three wearing scarves.

We evaluated the performance of our proposed approach on the face recognition task where the original images from the AR dataset were directly used as the probe set without crop or alignment. It is ubiquitous in real-life applications that probe images and gallery images are of different size and are not properly aligned. Table 2 shows the recognition accuracy of different methods. We see that our method obtained the best result, demonstrating the robustness against various scales. MKD-SRC-GTP utilizes Harris-Laplacian keypoint detector [7] which is sensitive to corners, so that a great number of improper keypoints were detected in hair regions, leading to unstable and discriminative descriptors.

4.3. Results on PubFig

PubFig dataset [14] contains 58797 images of 200 people from unconstrained conditions, which vary in poses, illuminations, expressions and background. The evaluation set of 140 people are used for the probe set and the remaining 60 people are regarded as the gallery set. Following the settings in [8], we select 5 images for each subject. The selected dataset well preserves the variation in poses, illuminations, expressions and background.

Table 3. Average recognition rates (%) of different methods with different ranks on PubFig.

Method	rank = 1	rank = 10	rank = 20
SiftSurfSILBP	25.00	49.29	57.86
CPD	28.36	51.93	62.29
MLERPM	27.86	52.86	64.29
MKD-SRC-GTP	38.57	62.14	72.14
LAIRPM	37.14	64.29	72.86
RPSM	42.86	65.00	74.29
TPGM	**43.57**	**66.43**	**75.71**

We randomly transformed the original images to obtain partial face patches. We first rotated each image randomly from -20° to 20°. Then, we cropped the rotated image to the 0.8 times of the original image size. Finally, the cropped image was randomly scaled ranging from 0.8 to 1.2. We conducted five-fold testing scheme by randomly splitting the cropped images into five subsets. Besides the baseline approaches, we also tested on the original SiftSurfSILBP where the Lowe's matching scheme was utilized for partial face recognition.

Table 3 shows that TPGM outperforms other state-of-the-art partial face recognition methods on PubFig. Compared with SiftSurfSILBP which simply applies Lowe's matching scheme, the proposed TPGM performs a graph matching procedure to encode the graphical information and obtains about 18% increases in recognition rates, which demonstrates the effectiveness of the proposed TPGM approach.

5. CONCLUSION

In this paper, we have proposed a topology preserving graph matching (TPGM) method for partial face recognition, where a non-rigid transformation is estimated encoding the geometric structure of the graph. With the geometric graph structure, TPGM obtains more accurate and robust correspondence, which leads to a stronger recognition ability. Moreover, the proposed TPGM directly matches on the partial image pairs, which does not require large amount of training data. The experimental results show the effectiveness of TPGM.

6. REFERENCES

[1] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, "Face description with local binary patterns: Application to face recognition," *TPAMI*, vol. 28, no. 12, pp. 2037–2041, 2006.

[2] Jiwen Lu, Venice Erin Liong, Xiuzhuang Zhou, and Jie Zhou, "Learning compact binary face descriptor for face recognition," *TPAMI*, vol. 37, no. 10, pp. 2041–2056, 2015.

[3] Yi Sun, Xiaogang Wang, and Xiaoou Tang, "Deep learning face representation from predicting 10,000 classes," in *CVPR*, 2014, pp. 1891–1898.

[4] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *CVPR*, 2014, pp. 1701–1708.

[5] Omkar M Parkhi, Andrea Vedaldi, and Andrew Zisserman, "Deep face recognition," in *BMVC*, 2015, vol. 1, pp. 1–12.

[6] Florian Schroff, Dmitry Kalenichenko, and James Philbin, "Facenet: A unified embedding for face recognition and clustering," in *CVPR*, 2015, pp. 815–823.

[7] Shengcai Liao, Anubhav K Jain, and Stan Z Li, "Partial face recognition: Alignment-free approach," *TPAMI*, vol. 35, no. 5, pp. 1193–1205, 2013.

[8] Renliang Weng, Jiwen Lu, and Yap-Peng Tan, "Robust point set matching for partial face recognition," *TIP*, vol. 25, no. 3, pp. 1163–1176, 2016.

[9] Olivier Duchenne, Armand Joulin, and Jean Ponce, "A graph-matching kernel for object categorization," in *ICCV*, 2011, pp. 1792–1799.

[10] Hao Jiang, Stella X Yu, and David R Martin, "Linear scale and rotation invariant matching," *TPAMI*, vol. 33, no. 7, pp. 1339–1355, 2011.

[11] William Brendel and Sinisa Todorovic, "Learning spatiotemporal graphs of human activities," in *ICCV*, 2011, pp. 778–785.

[12] Gary B Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Tech. Rep., Technical Report 07-49, University of Massachusetts, Amherst, 2007.

[13] Aleix M Martinez, "The AR face database," *CVC Technical Report*, vol. 24, 1998.

[14] Neeraj Kumar, Alexander C Berg, Peter N Belhumeur, and Shree K Nayar, "Attribute and simile classifiers for face verification," in *ICCV*, 2009, pp. 365–372.

[15] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma, "Robust face recognition via sparse representation," *TPAMI*, vol. 31, no. 2, pp. 210–227, 2009.

[16] Andrew Wagner, John Wright, Arvind Ganesh, Zihan Zhou, Hossein Mobahi, and Yi Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *TPAMI*, vol. 34, no. 2, pp. 372–386, 2012.

[17] Hazım Kemal Ekenel and Rainer Stiefelhagen, "Why is facial occlusion a challenging problem?," in *ICB*, 2009, pp. 299–308.

[18] Alex Pentland, Baback Moghaddam, and Thad Starner, "View-based and modular eigenspaces for face recognition," in *CVPR*, 1994, pp. 84–91.

[19] Bernd Heisele, Purdy Ho, Jane Wu, and Tomaso Poggio, "Face recognition: component-based versus global approaches," *CVIU*, vol. 91, no. 1, pp. 6–21, 2003.

[20] David G Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.

[21] Luo Juan and Oubong Gwun, "A comparison of sift, pca-sift and surf," *IJIP*, vol. 3, no. 4, pp. 143–152, 2009.

[22] Zhi Li, Guizhong Liu, Yang Yang, and Junyong You, "Scale-and rotation-invariant local binary pattern using scale-adaptive texton and subuniform-based circular shift," *TIP*, vol. 21, no. 4, pp. 2130–2140, 2012.

[23] Hongsheng Li, Xiaolei Huang, and Lei He, "Object matching using a locally affine invariant and linear programming techniques," *TPAMI*, vol. 35, no. 2, pp. 411–424, 2013.

[24] Hao Jiang, Mark S Drew, and Ze-Nian Li, "Matching by linear programming and successive convexification," *TPAMI*, vol. 29, no. 6, pp. 959–975, 2007.

[25] Renliang Weng, Jiwen Lu, Junlin Hu, Gao Yang, and Yap-Peng Tan, "Robust feature set matching for partial face recognition," in *ICCV*, 2013, pp. 601–608.

[26] Andriy Myronenko and Xubo Song, "Point set registration: Coherent point drift," *TPAMI*, vol. 32, no. 12, pp. 2262–2275, 2010.

[27] Dong Chen, Xudong Cao, Fang Wen, and Jian Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *CVPR*, 2013, pp. 3025–3032.