Nonrigid Point Matching of Chinese Characters for Robot Writing

Bocheng Zhao
1 National Laboratory of Pattern Recognition, Institute of Automation
2 School of Computer and Control Engineering
1 Chinese Academy of Sciences
2 University of Chinese Academy of Sciences
Beijing, China
zhaobocheng2015@ia.ac.cn

Minghao Yang
National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences,
Beijing, China
mhyang@nlpr.ia.ac.cn

Hang Pan
Guangxi Key Laboratory of Trusted Software
Guilin University of Electronic Technology
Guilin, China
hang6016@foxmail.com

Qingjie Zhu
Guangxi Key Laboratory of Trusted Software
Guilin University of Electronic Technology
Guilin, China
zhuqj21@foxmail.com

Jianhua Tao
1 National Laboratory of Pattern Recognition, Institute of Automation
2 CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation
3 School of Computer and Control Engineering
1,2 Chinese Academy of Sciences
3 University of Chinese Academy of Sciences
Beijing, China
jianhua.tao@ia.ac.cn

***Abstract*—Point matching is a key step in robot writing-learning process. There are three major challenges that weakened most of the existing point matching algorithms to match points for Chinese characters, including nonlinearly deformation, connected strokes and geometrically dispersive. We propose a novel algorithm based on constrained global energy function (CGE) in the matching process to cope with the abovementioned challenges in this paper. We utilized a global spatial distribution energy function (EF) to evaluate relationship among point sets. Then we could solve problems with point registration by minimizing the energy function. To evaluate the matching results of energy function, we defined a vector that describe local spatial information to optimize the algorithm presentation. In addition, to avoid divergence, we designed an operator based on sigmoid function, using literation numbers as inputs to constrain the vector. We have conducted experiments on an extensive human handwriting database, and our algorithm performed competitively against the state-of-the-art point matching algorithm in terms of accuracy and stability.**

Keywords—point match, Chinese character, energy function, local spatial information, constrain operator

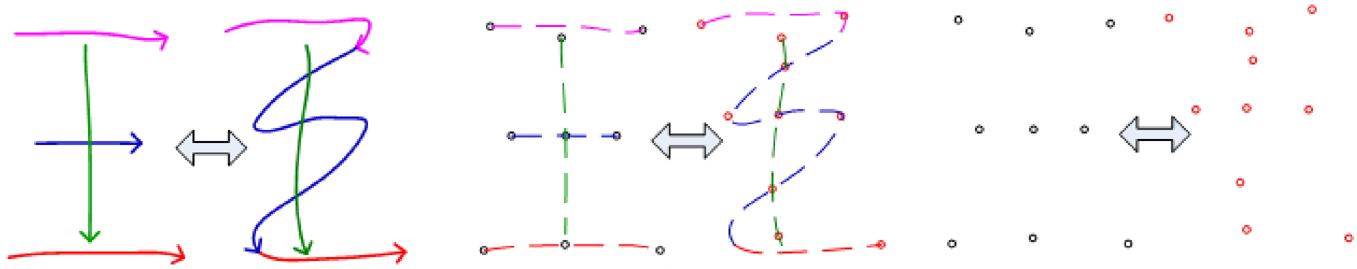
I. Introduction

Robot writing is considered as a useful tool for children to improve their writing ability through interactions with robots. To ameliorate children's handwriting, the robot firstly evaluate children's writing. We would evaluate Chinese character handwriting by comparing its stroke deformation, organization and order with a standard font or calligraphy master's handwriting. Accordingly, the robot should have the ability of stroke extraction and registration of different Chinese handwritings. However, because of the timeliness requested in interactive robot learning, it is not possible to use all of the points in Chinese character trajectory to complete the match task. We need to extract interest points in one Chinese character examples for matching. Furthermore, these interest points should represent important informations of the Chinese character trajectory, including stroke shape, direction, location and connection relationship. There are numerous methods [17-19] to extract feature points of Chinese character strokes, which are end-points, turning-points and cross-points. We can also rebuild a new holonomic Chinese character trajectory by these feature points. Thus we convert the stroke extraction and matching task to point matching task.

Surveys of point matching and registration can be found in [2-5]. At present, most of the state-of-the-art algorithms for point matching focus on global position distribution and local similarity to solve deformation and noise points problems, such as iterative closest point (ICP) methods [8], [9]; the robust L_2E estimator [7]; coherent point drift (CPD) method [6] and others [10-13]. Especially, asymmetric point matching (APM) algorithm [1] presented by Wei *etc.*, had an outstanding performance for matching limited nonlinear deform and quantity-uncertain point sets. These point matching algorithms are widely and successfully used in 3D reconstruction, image segmentation and target recognition tasks.

However, comparing with common point registration tasks, Chinese character feature point matching task has three major problems and challenge:

- Nonlinear spatial transformation. The challenge cause the spatial distribution of the two matching point sets to



(a) (b) (c)

Fig. 1.

(a): all the strokes in two images are severely nonlinear deformed and curved.
(b): purple, blue and red strokes are connected, there are two additional points in the right character example that leads to amount error.
(c): local dispersive. The two point sets (black and red) are dispersive and useful local similarity are very difficult to find out for point matching.

disturb severely. Furthermore, this kind of nonlinear deformation reoccurs randomly. Fig.1a shows that all the four strokes in the different character handwriting are severely curved, deformed and transformed.

- Connected strokes. Some people used to connect different strokes for convenience while writing, causing great amount errors in point sets. It can be seen from Fig.1b that because of the purple, blue and red strokes are connected, there are two additional points in the right character example that lead serious amount error.

- Lack of local similarity information. Unlike the point sets used in some other matching tasks, Feature points of Chinese characters are vary geometrically dispersive, so it is difficult to use local spatial distribution similarity to match points. We can see from Fig. 1c that the two point sets (black and red) are dispersive and useful local similarity are very difficult to find out for point matching.

Due to these difficulties. The performances of methods described above are somewhat unsatisfactory in terms of Chinese characters point set matching tasks. Thus we put forward a new method used in point matching for robot writing learning task. We proposed an energy function to solve the three major issues, which is based on global spatial distribution of point set. And then we put forward a vector feature to enhance the performance of energy function. Our method got a desired result in the Chinese character trajectory feature point matching. The proposed algorithm overcame the influence of severe distortions and connected strokes, suppressed some serious errors by tradition methods. This paper presents the detail of Energy Function in Sec. II, the local information describe method in Sec. III, parameter setting and experimental comparison in Sec. IV and a summary in Sec. V.

II. Energy Function

A. Trajectory Distortion and Outlier Points

As discussed above, two handwriting Chinese character trajectories may have great distortion from each other, distribution of point sets are randomly deformed and amount of points is altered. Accordingly, it is difficult to match point sets through methods using points' local structure or shape similarity. However, no matter how severely does a handwriting Chinese character deform or how many strokes are connected, globally relative positional relation among strokes remains the same. The two characters in Fig. 1 look dissimilar, actually, the blue stroke always upper to the red one and lower to the pink one. Based this characteristic, we designed a mathematical model based on global relation of point sets to cope with the challenges listed in Sec. I.

B. Drag Vector and Energy Function

i) Drag vector $\vec{v}_{ij}$

Assume that two point sets $X = (x_1, \dots, x_n) \in \mathbb{R}^{D\times n}$ and $Y = (y_1, \dots, y_m) \in \mathbb{R}^{D\times m}$ (*D*=2) are feature points extracted from handwriting trajectories Ψ_1 and Ψ_2 and each point is a *D* by 1 vector. May wish to set up $n < m$. A drag vector is a 2 dimensional vector start in x_i and pointing to y_j.

$$\vec{v}_{ij} = x_i - y_j \tag{1}$$

ii) Matching policy Π

Global matching policyΠcould be seen as a series of sub-policies: $\pi_{1\xi_1}, \pi_{2\xi_2}, \dots, \pi_{n\xi_n}$. A sub matching policy π_{ij} means matching point x_i to point y_j indicates dragged point set X to a new temporary location X' along Drag Vector $\vec{v}_{ij}$. As shown in Fig. 2.

Where Drag Vector and new location are:

$$x_i' = x_i + \vec{v}_{ij} \tag{2}$$

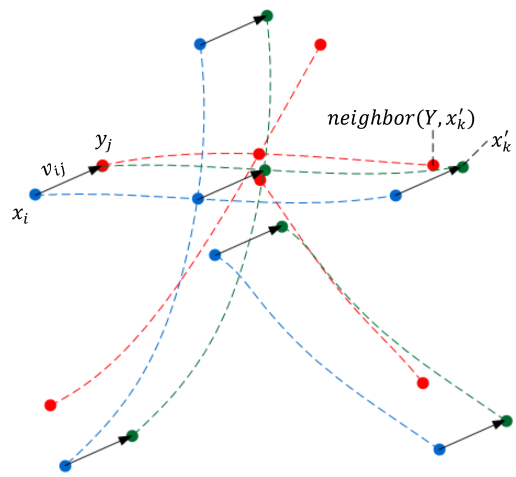
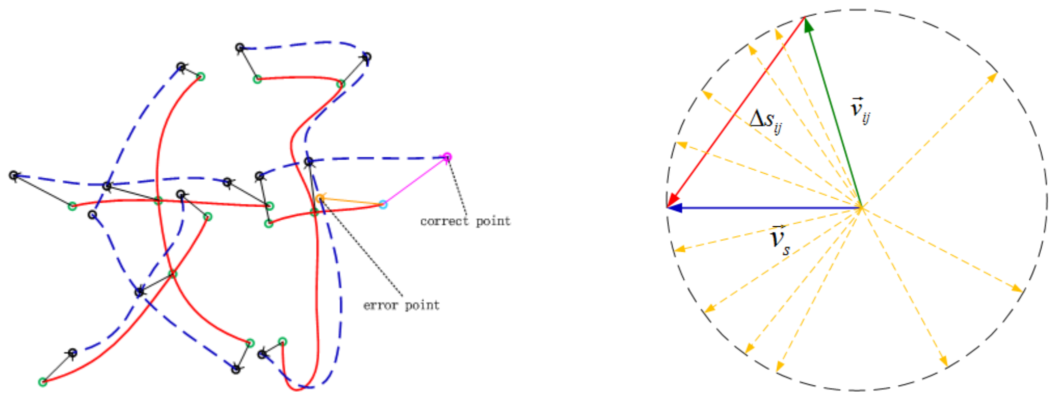


Fig. 2.

All the points in $X^{D\times n}$ (blue points) are dragged to new location X'(green ones) by vector $\vec{v}_{ij}$ and set $Y^{D\times m}$ are the red points



(a) (b)

Fig. 3. Vector Feature

(a): natural limitation of EF, EF tend to match points with the vector (yellow) whose direction is similar to average vectors' (black), but the correct vector is the pink one

(b): $\vec{v}_s$ is the average direction of all the drag vectors of energy function (orange and green), $\vec{v}_{ij}$ is the drag vector of policy π_{ij} and Δs_{ij} is the extracted feature vector

iii) Energy function

Energy function could be seen as mathematical representations of matching policies. The global energy function is a $n \times m$ matrix $Energy \in \mathbb{R}^{n\times m}$ of $X^{D\times n}$ to $Y^{D\times m}$ points matching task. Elements of matrix $Energy$ could be calculated by formula $Energy_{ij} = \varepsilon(\pi_{ij})$ and sub energy En, where the sub energy En for policy π_{ij} of point x_k is:

$$En(\pi_{ij}, x'_k) = \|x'_k - neighbor(x'_k, Y)\|, k \in [1, n] \qquad (3)$$

Which means that: after dragged by $\vec{v}_{ij}$, point x_k came to a new location x'_k, $En(\pi_{ij}, x'_k)$ is the energy of x'_k. Where $neighbor(x'_k, Y)$ is a point in Y which is the nearest point from x'_k, as shown in Fig. 2. By summarizing these sub energy function, we get an element of the global energy function (EF):

$$\varepsilon(\pi_{ij}) = \sum_{k=1}^{n} En(\pi_{ij}, x_k) \qquad (4)$$

Thus, through repeat the process above, we can calculate global energy matrix $Energy \in \mathbb{R}^{n\times m}$. There are $n \times m$ policies for $X^{D\times n}$ to $Y^{D\times m}$ task in summary, and in $Energy \in \mathbb{R}^{n\times m}$ we got $n \times m$ presentations of these policies. We use $Energy\_Row \in \mathbb{R}^{1\times m}$ to present a row of matrix $Energy$. Then the index of output matching result ξ_j for x_j is:

$$\xi_j = \mathrm{argmin}(Energy\_Row(j)) \qquad (5)$$

After repeating n times, we end up with a tentative global matching policy $\Pi : \pi_{1\xi_1}, \pi_{2\xi_2}, \ldots, \pi_{n\xi_n}$

III. LOCAL SPATIAL INFORMATION DESCRIBE

It can be seen from Fig. 3a that most of the feature points (black point) of blue character is on the left side of green ones, which are re character's feature points. When using EF, dragging the blue point to the error target (yellow point) will get the best energy. Due to the natural limitation of the energy function, we find that some of the points would be dragged to reverse positions if the Drag Vectors between these points and their correct matching points are opposite to the average vector of the whole point set but the correct point is the pink one. It is not hard to recognize that when the point set move along the yellow vector, the energy function will decrease. This means that when dealing with and individual point matching relation, the EF algorithm tends to force the exploration direction as consistent as possible with average drag vectors, so we would need a method to avoid such situation.

A. Spatial Information Describe Vector

We put forward a novel method which utilizes vector feature of matching policies to describe the natural mistakes of EF and correct tentative results gradually by iteration.

i) Vector feature:

For a tentative matching policy π_{ij} and corresponding drag vector, we defined a vector feature Δs_{ij} to judge the similarity between π_{ij} and other policies in $\Pi : \pi_{1\xi_1}, \pi_{2\xi_2}, \ldots, \pi_{n\xi_n}$ and their corresponding drag vectors $\vec{v}_{1\xi_1}, \vec{v}_{2\xi_2}, \ldots, \vec{v}_{n\xi_n}$.

$$\vec{v}_s = \frac{1}{n}\sum_{i=1}^{n} \frac{\vec{v}_{i\xi_i}}{|\vec{v}_{i\xi_i}|} \qquad (6)$$

$$\Delta s_{ij} = \frac{\vec{v}_{ij}}{|\vec{v}_{ij}|} - \vec{v}_s \qquad (7)$$

Where $\vec{v}_s$ is the mean vector which is same as the blue vector shown in Fig. 3b.

ii) Result optimization process

We used this vector feature Δs_{ij} as feedback to optimize algorithm presentation, including the following three steps:

Step1, find out the multi-matched and unmatched points both in X and Y, put them into a new tentative point set $\dot{X}$ and $\dot{Y}$ for literation.

Step2, calculate the average vector $\vec{v}_s$, and Δs_{ij} for each of the rest policies π_{ij}, if:

$$|\Delta s_{ij}| > |v_0| \qquad (8)$$

put points x_i, y_j into $\dot{X}, \dot{Y}$ and v_0 here is the judgment vector and a unit vector by default.

Step3, use $\dot{X}, \dot{Y}$ as the input to energy function algorithm and repeat steps above.

B. Constrain Operator

It is not hard to understand that using fixed v_0 will sometimes lead algorithm not convergent. To ensure algorithmic convergence, we build up an operator σ, which is based on the literation times:

$$\sigma = \lambda(1 - \text{sigmoid}(t, p, c)) \tag{9}$$

Where t is current literation time, λ, p and c are hyper parameters and could be get by cross validation. Then we get the fixed vector Δv and replace v_0 in (8) with Δv.

$$\Delta v = \sigma v_0 \tag{10}$$

C. Algorithm Framework

After EF and VC, we get a relatively reliable matching result. The total algorithm of CGE could be presented as a recurrent framework in summary and is shown in Algorithm1.

Algoritm1. The CGE Algorithm.

Input: two point sets X and Y, parameter λ, p and c
Output: result policy $\Pi : \pi_{1\xi_1}, \pi_{2\xi_2}, \ldots, \pi_{n\xi_n}$.
Initialize $\varepsilon(\pi_{ij}) = 0$
Repeat
  Energy -step
    Drag all points in X along vector v_{ij} to X'
    Compute matrix $Energy_{ij} = \varepsilon(\pi_{ij})$
    Get tentative result $\Pi : \pi_{1\xi_1}, \pi_{2\xi_2}, \ldots, \pi_{n\xi_n}$
  Constrain-step
    Compute Δv
    Remove policy $\pi_{i\xi_i}$ when $|\Delta s_{i\xi_i}| > |\Delta v|$
    Put the points of removed policies into $\dot{X}, \dot{Y}$
    Update X and Y by $\dot{X}, \dot{Y}$
    Consolidate tentative results
Until all points in X are matched

IV. Experiment

A. Data

In this section, we show the matching result and analyses of the proposed CGE and APM method for point matching tasks. The data used in this paper is from HWDB 1.1 [14-16], which is an extensively adopted and recognized handwriting database. We use algorithms in [17-19] to extract the stroke skeletons and feature points of Chinese character. We choose 500 representative Chinese characters and each character has randomly chosen 3 handwriting examples. We defined a series of "complexity level" base on point quantity, where Chinese character has less than 10 feature points is classified "level 1"; and the level upgrades every 5 more points. If the point quantity is greater than 25, the corresponding level is 5. And 5 is the max level. Correct matching results are manually marked, and the point set used CGE are normalized to a 300× 300 pixel image.

B. Coefficient Setting

We reviewed the effects of coefficient λ, p and c on algorithm performance by variable-controlling method, we defined a coefficient vector $[\lambda, p, c]$ to store value. As with sigmoid function, λ, p and c represent linear and nonlinear aspects of vector feature method respectively. Through cross validation, we confirmed that to get the best result, the value of λ, p and c need to vary with number of points to be matched, deformation degree and structure complexity. Deformation degree here could be represented by Energy matrix at the first round of literation in Algorithm 1. λ could be recognized as the tolerance of drag vector difference; p is the convergence rate of tolerance and c is the duration of high tolerance value. When the number of points to be matched gets lower, the value of λ needs to be higher; when the deformation degree gets larger, the value of p needs to be higher; when the structure complexity gets lower, the value of p needs to be higher, and vice versa. It is concluded that to get the best result, in simple level tasks (level 1-2), value of coefficient vector is suggested to be [1.2, 1.1, 2.7]; in normal tasks (level 3-4) coefficient vector should be [0.9, 1.9, 1.9]; in complicated tasks (level 5), coefficient vector should be [0.6, 2.7, 0.7]. More experiment data and detail is listed in Fig. 4.

C. Comparison

We used APM [1] algorithm as comparison, which is a state-of-the-art and outstanding algorithm for point matching, the source code of which is available online (http://www4.comp.polyu.edu.hk/~cslzhang/APM.htm). To

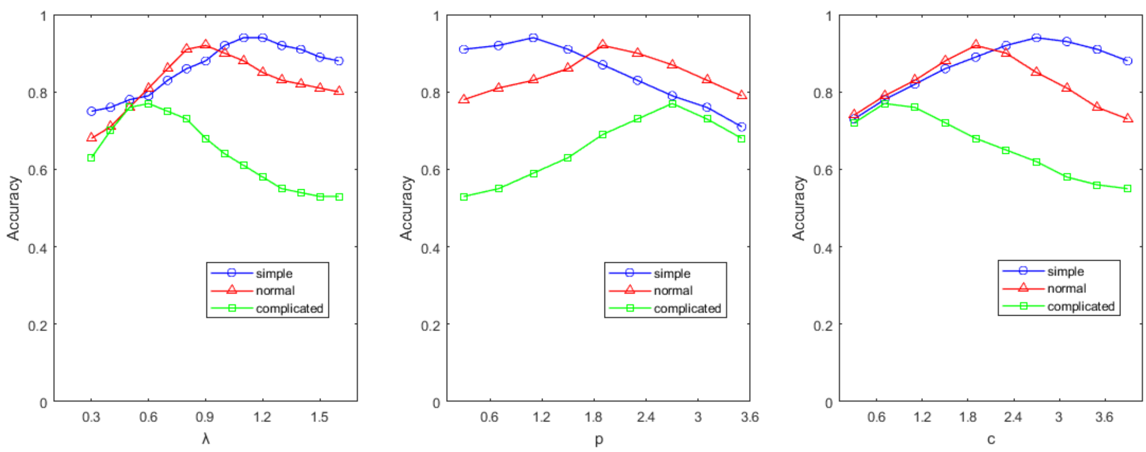


Fig. 4. Experiments on the choice of λ, p and c

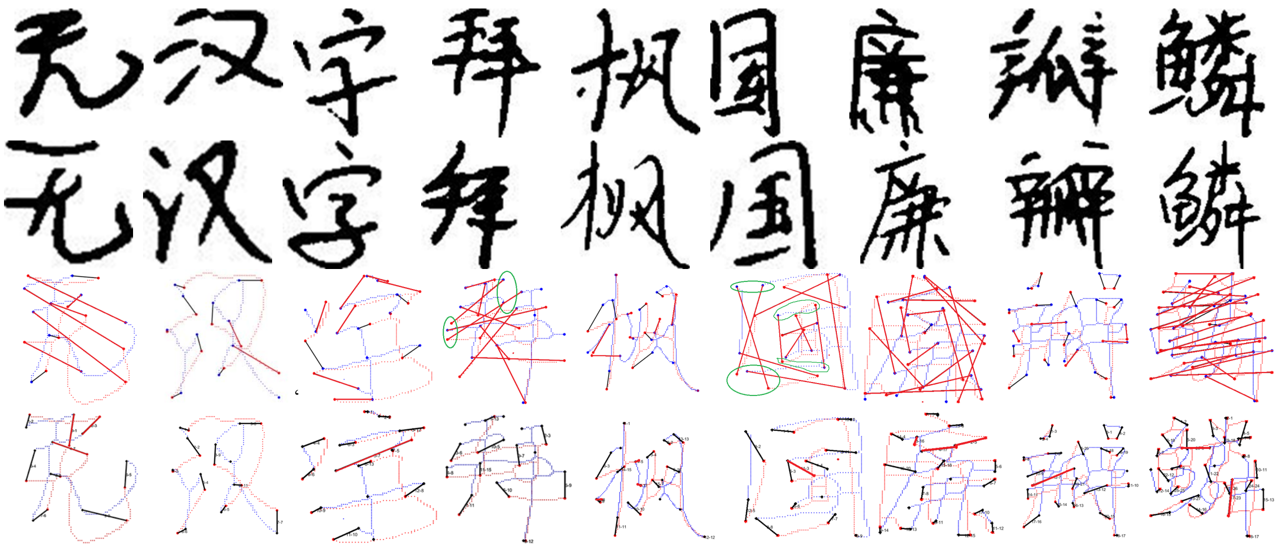


Fig. 5. Comparison experiment results on accuracy and stability

From top to bottom: origin characters, target characters, results of APM, results of CGE. From left to right: complexity level 1 to level 5 characters. The black line indicate correct results and the red line indicate wrong results.

provide a quantitative and comprehensive evaluation of our algorithm, we evaluate experiment results of two aspects: 1) Matching accuracy on point sets of different complexities. 2) Algorithm stability. We tested these two aspects of data difference and algorithm parameter to all the 1500 character samples. The results are reported in Fig. 5 and Fig. 6a.

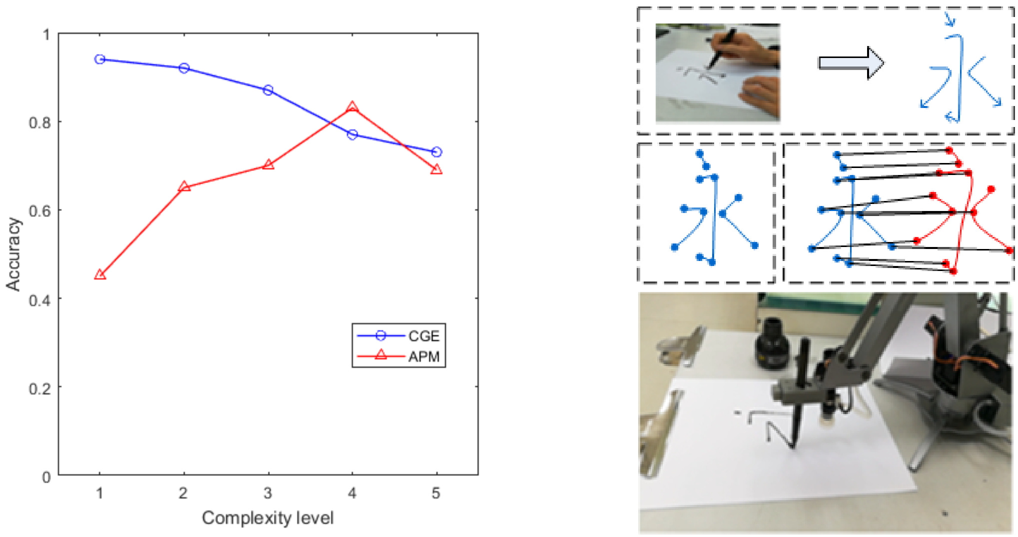
i) Point matching accuracy

It can be seen from Fig. 5 that results of the proposed CGE are well and stable. Even if some outlier and error points occurred along with the character to be matched becomes more complexity, CGE could welly suppress interference and get a satisfactory performance.

By comparison, experiments in Fig. 5 indicate that APM gets an unsatisfied result when the input point sets are too simple that local geography relations are hardly to find out. But then, with the increase of input point amount, more local minutia emerge, the results of APM are improved by these deformation details.

ii) Algorithm stability

We found in Fig. 5 that experiment method APM is very sensitive at point sets' local deformation that performance of APM is quietly unstable and variance is large. If some of the points accidentally share similar local structures, as marked out by the green circles in the last row of Fig. 5, they will be wrongly matched together by APM. That is because in this case, local deformation similarity are meaningless and in a sense could be believed as noises. This kind of noise would sometimes be recognized as local relations by APM and then output error results. We have concluded that APM performs better in high complexity tasks because point sets in these tasks



(a) (b)

Fig. 6.

(a): Experimental results statistics

(b): Interactive robot writing learning, the robot is writing a standard Chinese character with the stroke order of the user by step 1,2,3 and 4.

have more local information. However, more detail means more noise, with the points increase further, stability of APM goes worse. The third row of Fig. 5 shows different kind of local deformation similarity noises.

In contrast, because of the global energy function and Semi-Geometry method used in CGE, the error of local similarity is ignored and make the algorithm not sensitive at character natural structure at all.

iii) Empirical result

Statistics of the results indicate that in simple tasks CGE method have got a favorable performance against APM. Especially in level 1 task, CGE got a 94% accuracy. CGE works stably during all the matching tasks. In contrast, APM performs unsatisfactory in simple tasks and need correct local similarity informations to improve its performance. When the point sets accidentally have local similarity noises, it would take noises as useful information mistakenly. Such problem lead APM performs instability in Chinese character point matching task. The consequence of experiment results are shown in Fig. 6a.

In addition, we built an interactive robot writing system to apply our algorithm. Our system learns human handwriting by the follow steps: 1) Record the user's writing stroke sequence by a camera. 2) Extract feature points of the user's handwriting and a standard character example. 3) Match the feature points of two character examples by our algorithm. 4) Write the standard character with the stroke consequence of the user. As shown in Fig. 6b. Our robot writing system could learn the stroke order, shape and relationship remarkably due to the outstanding robust of the proposed CGE algorithm.

V. CONCLUSION

We proposed a novel point matching algorithm in this paper. The proposed algorithm creates a global deformation based energy function to solve the two major problem in feature point of Chinese character handwriting matching task. By using vector feature, we took three hyper-parameters as linear, nonlinear and convergence rate index respectively to fix the natural errors in energy function. The resulting algorithm scales well with data size HWDB. Experimental results on all of the choose handwriting data showed that the proposed CGE algorithm performs favorably in terms of matching accuracy against the state-of-the-art method APM in Chinese character feature points matching task.

VI. ACKNOWLEGEMENT

This work is supported by the National Key Research & Development Plan of China (No. 2016YFB1001404), and the Strategic Priority Research Program of the CAS (Grant XDB02080006), the National Natural Science Foundation of China (NSFC) (NO.61332017, NO.61273288, No.61233009, No.61375027, No.61425017), Beijing Municipal Science and Technology Commission(No. Z161100000216124), Guangxi Colleges and Universities Key Laboratory of Intelligent Processing of Computer Images and Graphics (No GIIP201602), Guangxi Key Laboratory of Trusted Software (No kx201601).

REFERENCES

[1] Lian W, Zhang L, Yang M H. An efficient globally optimal algorithm for asymmetric point matching[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(7): 1281-1293.

[2] Brown L G. A survey of image registration techniques[J]. ACM computing surveys (CSUR), 1992, 24(4): 325-376.

[3] Zitova B, Flusser J. Image registration methods: a survey[J]. Image and vision computing, 2003, 21(11): 977-1000.

[4] Tam G K L, Cheng Z Q, Lai Y K, et al. Registration of 3D point clouds and meshes: a survey from rigid to nonrigid[J]. IEEE transactions on visualization and computer graphics, 2013, 19(7): 1199-1217.

[5] Parekh H S, Thakore D G, Jaliya U K. A survey on object detection and tracking methods[J]. International Journal of Innovative Research in Computer and Communication Engineering, 2014, 2(2): 2970-2979.

[6] Myronenko A, Song X. Point set registration: Coherent point drift[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(12): 2262-2275.

[7] Ma J, Zhao J, Tian J, et al. Robust estimation of nonrigid transformation for point set registration[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 2147-2154.

[8] Besl P J, McKay N D. A method for registration of 3-D shapes[J]. IEEE Transactions on pattern analysis and machine intelligence, 1992, 14(2): 239-256.

[9] Zhang Z. Iterative point matching for registration of free-form curves and surfaces[J]. International journal of computer vision, 1994, 13(2): 119-152.

[10] Nguyen Q, Gautier A, Hein M. A flexible tensor block coordinate ascent scheme for hypergraph matching[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 5270-5278.

[11] Deng Y, Rangarajan A, Eisenschenk S, et al. A Riemannian framework for matching point clouds represented by the Schrodinger distance transform[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 3756-3761.

[12] Ma J, Zhao J, Yuille A L. Non-rigid point set registration by preserving global and local structures[J]. IEEE Transactions on image Processing, 2016, 25(1): 53-64.

[13] Ma J, Zhao J, Ma Y, et al. Non-rigid visible and infrared face registration via regularized Gaussian fields criterion[J]. Pattern Recognition, 2015, 48(3): 772-784.Liu C L, Yin F, Wang D H, et al. Chinese handwriting recognition contest 2010[C]//Pattern Recognition (CCPR), 2010 Chinese Conference on. IEEE, 2010: 1-5.

[14] Yin F, Wang Q F, Zhang X Y, et al. ICDAR 2013 Chinese handwriting recognition competition[C]//Document Analysis and Recognition (ICDAR), 2013 12th International Conference on. IEEE, 2013: 1464-1470.

[15] Su T H, Zhang T W, Guan D J, et al. Off-line recognition of realistic Chinese handwriting using segmentation-free strategy[J]. Pattern Recognition, 2009, 42(1): 167-182.

[16] Bunke H, Bengio S, Vinciarelli A. Offline recognition of unconstrained handwritten texts using HMMs and statistical language models[J]. IEEE transactions on Pattern analysis and Machine intelligence, 2004, 26(6): 709-720.

[17] Pham T A, Delalandre M, Barrat S, et al. Accurate junction detection and characterization in line-drawing images[J]. Pattern Recognition, 2014, 47(1): 282-295.

[18] Liu K, Huang Y S, Suen C Y. Identification of fork points on the skeletons of handwritten Chinese characters[J]. IEEE transactions on pattern analysis and machine intelligence, 1999, 21(10): 1095-1100.

[19] Liu C L, Kim I J, Kim J H. Model-based stroke extraction and matching for handwritten Chinese character recognition[J]. Pattern Recognition, 2001, 34(12): 2339-2352.