DRAWING ORDER RECOVERY FOR HANDWRITING CHINESE CHARACTERS

Bocheng Zhao, Minghao Yang, Jianhua Tao

zhaobocheng2015@ia.ac.cn, mhyang@nlpr.ia.ac.cn, jianhua.tao@ia.ac.cn

ABSTRACT

Recover drawing orders from a Chinese handwriting image is a challenge issue. Most of English drawing order recovery(DOR) methods perform unsatisfactorily in Chinese. This paper proposes a novel image-to-sequence algorithm to deal with Chinese DOR problem. The proposed method utilizes two regression convolution neural network(CNN) models to generate two corresponding pen-tip movement heat-maps. To estimate pen-tip movement for most of the normal states in writing process, the algorithm analyzes the above two heat-maps with a specifically designed framework. Then the drawing order is restored through a simple iteration process based on the proposed framework. Experiments on public online handwriting database show that our method have got a remarkable result for Chinese DOR tasks. In addition, for English tasks, our method performs superiorly among state-of-the-art methods.

***Index Terms*—** Drawing order recovery, Chinese handwriting, Convolution neural network, image-to-sequence model

1. INTRODUCTION

Recovering the pen movement order from English or Arab digit symbol handwriting picture is a hot topic in the past few decades. DOR could be used to improve offline handwriting recognition, stroke segmentation and some other fields. Researchers had proposed some effective method for English DOR tasks [1, 2, 3, 4, 5, 6, 7, 8, 9, 10], Yu Qiao et al. [6, 7, 8], Y. Kato [9], and some other researchers [11, 12, 13] proposed certain widely recognized DOR algorithm in recent years, which had got excellent performances in some widely used public English handwriting database. These studies tend to configure some basic writing laws, such as stroke skeleton circle [11], end-point(EP) and branch-point(BP) regulars [6, 7, 8, 9]. In recent years, researchers tried to use more machine learning algorithms in DOR tasks, Lake B M [5] defined some stroke and sub-stroke libraries and utilized some probability models to choose best stroke orders, these orders are generated by pen tip random walking methods. A survey[14] of handwriting drawing restoration summarized most of existing methods. Some researchers [15, 16, 17] published their studies using deep neural network and other machine learning algorithms for DOR or other topics which contain DOR tasks. One of the major contributions of this paper is that we establish a novel way to solve DOR problem by a library-free image-to-sequence algorithm framework with CNN. This framework could also be used for other languages' and some special multi-stroke curves' DOR tasks in most cases. Another contribution is that we put forward a definition of writing energy to estimate pen tip moving for any instant writing states.

The rest parts of this paper are arranged as following: the two writing estimation networks are listed in Sec. 2. The algorithm framework for writing pattern with the corresponding writing energy network is presented in Sec. 3. Experiments settings and results are shown in Sec. 4. Finally we make a conclusion and expectations for our future work in Sec. 5.

2. WRITING ENERGY ESTIMATION NETWORK

Characters often have basic writing patterns. We name it static writing energy. Another one to model pen tip movement details for different instant writing states, such as meeting branch, end-point choosing or walking inline, we name it dynamic writing energy. The CNN models of these two writing energies will be presented in this section.

2.1. Static energy

Assume that a handwriting sample is normalized into a $l \times l$ image, We could configure the value of each elements in M where higher value of M_{ij} means that the corresponding dot in location (i, j) should be written earlier and lower value means written later, as shown in Fig. 1.Therefore, we designed a encoder base on CNN to encode Chinese character handwriting images to a heat-map image. We named it static energy network(SEN), the input of SEN is a character picture generated by online handwriting data while the output is the heat-map labeled by point sequences. Network structure and layers are (128,1,1)Conv-(64,3,2)Conv-(2,2)MP-(2048)FC-Relu-(1024)FC-Relu-Regression, as shown in Fig. 1. Theoretically speaking, SEN could finish DOR task individually. However, the experiment of SEN is not accurate enough to distinguish local point drawing sequences, in other words SEN could predict the basic drawing order trend but not details.Thus, it is imperative to develop another method

978-1-5386-4658-8/18/$31.00 ©2019 IEEE

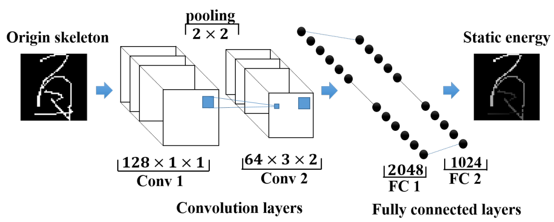

Fig. 1. Static energy network

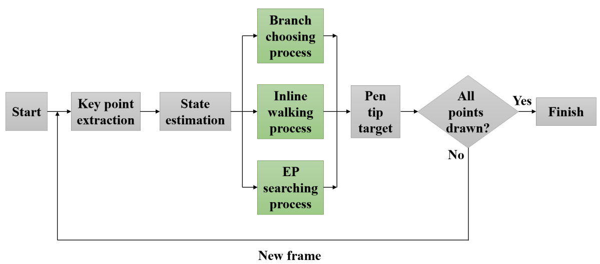

Fig. 2. Algorithm framework

to predict pen tip movement details. Then dynamic energy network(DEN) are proposed under this situation.

2.2. Dynamic energy

The Dynamic energy prediction network was firstly presented by us in paper [1]. In this paper, we changed some parameters in the network to fit the using situation. The input of DEN would be real-time handwriting images and the output should be a probability distribution on a l by l map for the next nib point. The network should have three major input information, nib instant location, nib instant moving direction and handwriting skeleton. We could consider the handwriting drawing process as a sequence of frames, each frame represents an instant drawing state. Then the inputs of DEN are three frames, the instant frame, the previous frame and whole handwriting skeleton. The output of DEN is a probabilistic heat-map that presents the probability of pen tip location for the next frame. Network layers are (128,2,2)Conv-(2,2)MP-(64,2,2)Conv-(2,2)MP-(32,2,2)Conv-(4096)FC-Relu-(2048)FC-Relu-(1024)FC-Relu-Regression.

3. ALGORITHM FRAMEWORK

As mentioned above, one of the key steps of character DOR is to decision the sequences of EPs and the skeleton branch choosing for BPs.Therefore, the three major aspects in our framework are End point searching, inline moving and Branch choosing.Thus the preparing work has been done.framework sketch map is shown in Fig. 2.

3.1. End point searching process

We have the following steps to take an overall consideration on EPs using SEN and DEN. Firstly, the program need to discriminate pen movement status. If the nib has not drawn any dot on skeleton or just finished a stroke, the program will start the EP choosing process. Secondly, the algorithm compute the energy of each EPs using (1) and (2)

$$EN(X,Y) = (heat_{SEN} + \gamma \cdot heat_{DEN}) \odot \phi(X,Y) \quad (1)$$

$$\phi(X,Y) = 1 - sigm\|(X - X_0) + (Y - Y_0)\| \quad (2)$$

Let l be the image size, where $(X,Y) \in [1,l]$ is the location of points in $heatmap_{SEN}$, $heatmap_{DEN}$ and origin image. (X_0, Y_0) is the location of initialize EPs.$\gamma \in [0,1]$ is a coefficient which controls the weight of DEN and SEN output. Hyper parameters in sigmoid function are configured by cross verification.

Thirdly, choosing the EP with the highest energy, drawing this point and its adjacent inline point. Remove these two new points in the origin skeleton and then we get a newly generated pseudo-skeleton.

3.2. Inline moving process

Generally speaking, inline pressing is believed a relatively simple work, so we utilize the pseudo-EP(pEP) to assist inline searching problems. Pseudo-skeleton is generated by removing drawn parts of origin skeleton. By taking pEPs and EPs into consideration together, we use (2) (3) and (4) to compute inline writing energies.

$$EN_{IL}(X,Y) = heat_{DEN} \odot \phi(X,Y) \quad (3)$$

$$EN_{IL}(X,Y) = EN_{IL}(X,Y) + \alpha, \forall (X,Y) \in pEP \quad (4)$$

Where $(X,Y) \in [1,l]$ is the location of points in $heatmap$ and origin image, α is a hyperparameter which could ensure the pen tip moving along origin orientation for normal inline walking status and "break down" in time if the output of the DEN shows it should change stroke, and α is set 0.35 after experiments.

3.3. Branch choosing process

Branch choosing is the key part of DOR studies, researchers could hardly evading this issue. Qiao [9] and many articles defined several BP kinds, such as "X","T","Y" BPs to solved this kind of problems in English and some alphabetic characters. Nevertheless, amount of BPs in Chinese is much more higher than English, and branch choosing problem is more complexity, so we set a new program framework to solve this issue with the following steps.

Firstly, monitor the drawing process, once the pen tip meet branch points, active branch choosing process. Secondly, drawing all adjacent branch points, this step will bring

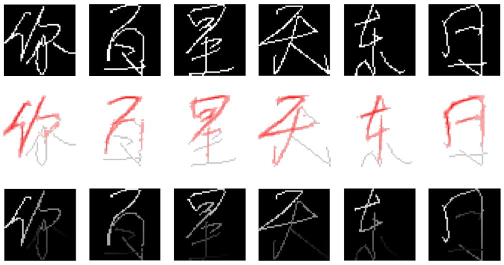

Fig. 3. SEN output, from top to bottom, the first row is SEN input, the second row shows SEN output and the last row is reference grayscale map

several new pseudo EPs(no less than two). Thirdly, record the new pEPs as a set S then use (2) (3) and (4) to choose a pEP in S for drawing. In addition, during whole drawing process, if pen tip have meet a point in S, repeat the third step until all points in S are drawn.

In summery, the proposed framework contains three major program processes, which are EP searching, inline walking and branch choosing. Each process utilizes SEN or DEN to finish the pen tip target prediction.

4. EXPERIMENT

We train the network and test the algorithm framework by using OLHWDB1.1 [18, 19] for Chinese task and UNIPEN [20, 21] for English and digit symbol task. Including 3755 Chinese characters written by 240 users. We take 10000 English letters and Arab digit symbols in UNIPEN , 8000 as training set and 2000 as testing set. To make a comparision, we compare our method with Zhao [1] for Chinese DOR task. Qiao[6] proposed a widely recognized method, although this work is published not recently, but the performance of it for the English database is excellent.Another comparison is work of Kha[3], which had got a same result in single-stroke English DOR with Qiao[6], but paper[3] studied multi-stroke and got an excellent result using tradition methods.

4.1. Network experiments

For training SEN, we use the above mentioned 15,000 input handwriting images and their corresponding heat-maps to train the network. For DEN, we set 30 frames for each sample, then with the 3 Channel inputs and totally utilized 1,350,000 images. So we separate DEN training into 10 subsets to ensure the program not out of memory. (Core i7 7700k, 16GB RAM ,GTX1080 8GB). As listed above, we use regression models in both networks and the loss function is mean-squared-error (MSE).

Fig.3 shows the prediction results of SEN, we could conclude that the restoration of SE is accuracy and excellent. To

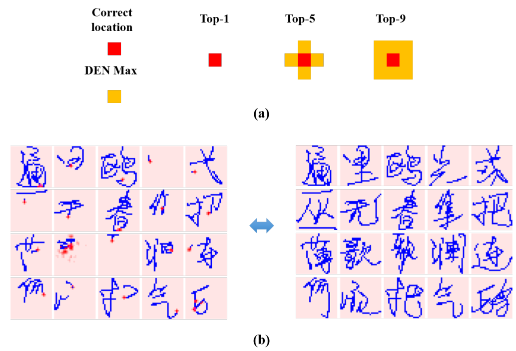

Fig. 4. (a): Three evaluations, top-1, top-5 and top-10, DEN max and correct location are coincide in top-1 so the color of DEN max is omitted. (b): Experiment result of DEN, left part shows 20 prediction results, where red dots are pen-tip probably moving target and blue dots are writing states, right parts are origin whole skeletons of these characters

Table 1. DEN experiment results.

Evaluation	Top-1	Top-5	Top-9
Accuracy	91.4%	96.8%	98.9%

evaluate DEN, we set three standard evaluations, top-1, top-5 and top-9 accuracy. Where top-1 is the lacation of maximum element in $heat_{DEN}$ is the exactly standard right location, top-5 means that the maximum element hits the 4 nearest dots on map for standard location and top-9 means hits the 9 nearset dots on map. As shown in Fig. 4a. We test DEN with the above conditions and the result is shown in Fig. 4b. and Table 1. In addition, we verified DEN by verifying data set for training data scale, network structure and other hyperparameter setting.

It can be read from Fig. 4b and Table 1 that the proposed DEN successfully finished the dynamic writing energy prediction work, for Top-9 the accuracy is 98.9 percent for outlier Chinese characters, which is a pretty satisfactory result. The performance of DEN has reached our respective. In addition, the performances for English letters are also very competitive, these results gives us confidence to deal Chinese DOR task by using the algorithm framework to fully tap the abilities of these two neural networks. For deeper analyzation, we observed the accuracy of our network in the three writing

Table 2. Accuracy of different writing status. IL is in line walking, BP is branch choosing and EP is end-point searching

State	IL	BP	EP	EP (SEN+DEN)
Accuracy	99.8%	92.1%	71.9%	85.3%

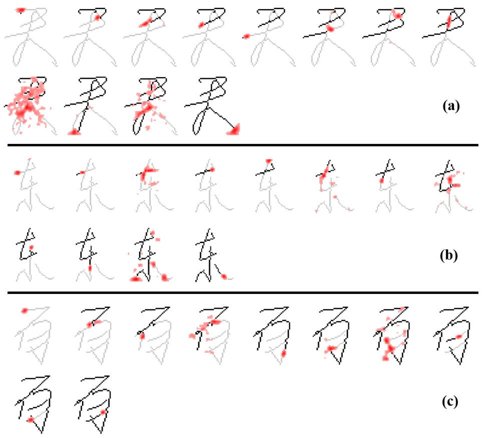

Fig. 5. The frames are listed by their template sequence and should be read from left to right in the figure row by row. Red pixels are DEN heatmap outputs, darker red means higher writing energy, black pixels are drawn stroke dots and gray pixels are stroke dots for drawing.

Table 3. DOR experiment accuracy. En(ss/ms)means single/multi-storke English character, digits in the table are accuraccy%

Data	$Cl.1$	$Cl.2$	$Cl.3$	$Cl.4$	$Cl.5$	$En(ss)$	$En(ms)$
SEN+DEN	83.1	68.2	50.5	35.6	23.3	98.1	95.5
DEN	70.1	55.5	40.8	18.8	9.0	96.8	94.0
Zhao[1]	65.6	49.1	33.1	12.6	5.8	96.0	94.0
Qiao[6]	–	–	–	–	–	96.0	–
Kha[3]	–	–	–	–	–	96.0	94.0

status separately, as shown in Table 2. With a further analysis, we found that SEN helps the EPs which should be written early in the writing sequence, but later EPs get little assistance from SEN. In addition, for Chinese characters with less than 3 strokes, both of the BP and EP accuracies are more than 98%, which indicates that for single-stroke and dual-stroke tasks our method will get a very excellent performance, however most of Chinese characters contain more than three strokes.

4.2. Drawing order recovery

By using SEN, DEN and the iteration framework introduced above, we test our method in DOR task. We compute the Dynamic Time Warping (DTW) between our result and standard sequence to evaluate the performance of DOR experiment. The testing database are classified into 5 complexity levels (marked $Cl.1$ to $Cl.5$). Chinese character maintains less than 5 strokes are configured $Cl.1$ and upgrades every 5 more strokes. It is worth introducing the DOR detail in Fig. 5. In this example we cut 12 frames for a Chinese handwriting, and this sample has severe stroke connection problem. The examples show 12 real-time status cover most of the problems existed in DOR. It can be concluded form Table 3. that our method have a improvement performance in simple Chinese character DOR task and a competitive result against the benchmark DOR methods in English task.In addition, Tab. 3 shows that with the help of SEN, the performance of DOR increased significantly, that is because SEN helps to find the correct start point of a character, and in our experiments, many mistakes happens in this area. Especially for simple characters with less than five end-points.

However, it is worth noticing that with the character complexity upgrades, the DOR correct rate deceases sharply, we believe that there are three following major reasons. The first reason is amount of BP in complex character is much higher. It is mentioned above that average branch choosing accuracy is 92.1%, $Cl.1$ character often contains one or two BPs in average, but $Cl.3$ or above character contains more than five BPs in average, which leads the correctly recover rate drops by BP mistakes. Another problem is that key points in complex characters are much more dense, because of the network output heat-maps are not precisely enough. It is an intractable problem to resolve several very close key points with same attribute(EP,BP,sEP). The third problem is that resolution ratio used in our algorithm is not high enough.We used to believe that 48×48 images is appropriate, but for complex characters, some strokes in this kind of image huddled together and pixels in these strokes are often removed by skeleton extraction process.

5. CONCLUSION

This paper introduced a novel framework for Chinese handwriting DOR. The proposed method utilizes two specific configured CNNs for the defined character writing energies, which are static energy and dynamic energy. Experiments on public online handwriting database demonstrate that SEN has the ability to predict the pen movement trend for handwriting samples and DEN could model instant writing states from real-time character images precisely. The most important contribution of this article is building an novel image-to-sequence framework for Chinese handwriting DOR tasks, which is a important and very challenge topic. Our algorithm could effectively recover drawing order from Chinese handwriting pictures. Besides, the proposed method performs superiorly among state-of-the-art methods on English letter and Arab symbol DOR tests.

6. ACKNOWLEDGE

This work is supported by the National Key Research Development Plan of China (No.2016YFB1001404) and the National Natural Science Foundation of China (NSFC) (No.61873269, No.61425017, NO.61332017).

7. REFERENCES

[1] B. Zhao, M. Yang, and J. Tao, "Pen tip motion prediction for handwriting drawing order recovery using deep neural network," in *Pattern Recognition (ICPR), 2018 International Conference on*. IEEE, 2018.

[2] M. Dinh, H.-J. Yang, G.-S. Lee, S.-H. Kim, and L.-N. Do, "Recovery of drawing order from multi-stroke english handwritten images based on graph models and ambiguous zone analysis," *Expert Systems with Applications*, vol. 64, pp. 352–364, 2016.

[3] V. A. Kha, H. H. Kha, and M. Blumenstein, "Extraction of dynamic trajectory on multi-stroke static handwriting images using loop analysis and skeletal graph model," *REV Journal on Electronics and Communications*, vol. 6, no. 1-2, 2016.

[4] M. Liwicki, Y. Akira, S. Uchida, M. Iwamura, S. Omachi, and K. Kise, "Reliable online stroke recovery from offline data with the data-embedding pen," in *Document Analysis and Recognition (ICDAR), 2011 International Conference on*. IEEE, 2011, pp. 1384–1388.

[5] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level concept learning through probabilistic program induction," *Science*, vol. 350, no. 6266, pp. 1332–1338, 2015.

[6] Y. Qiao, M. Nishiara, and M. Yasuhara, "A framework toward restoration of writing order from single-stroked handwriting image," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 11, pp. 1724–1737, 2006.

[7] Y. Qiao and S. Ong, "Arc-based evaluation and detection of ellipses," *Pattern recognition*, vol. 40, no. 7, pp. 1990–2003, 2007.

[8] Y. Qiao and M. Yasuhara, "Recovering drawing order from offline handwritten image using direction context and optimal euler path," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 2. IEEE, 2006, pp. II–II.

[9] Y. Kato and M. Yasuhara, "Recovery of drawing order from single-stroke handwriting images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 9, pp. 938–949, 2000.

[10] M. Diaz, M. A. Ferrer, G. S. Eskander, and R. Sabourin, "Generation of duplicated off-line signature images for verification systems," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 5, pp. 951–964, 2017.

[11] L. P. Cordella, C. De Stefano, A. Marcelli, and A. Santoro, "Writing order recovery from off-line handwriting by graph traversal," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 1896–1899.

[12] T. Steinherz, D. Doermann, E. Rivlin, and N. Intrator, "Offline loop investigation for handwriting analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 193–209, 2009.

[13] L. Rousseau, J. Camillerapp, and E. Anquetil, "What knowledge about handwritten letters can be used to recover their drawing order?" in *Tenth international workshop on frontiers in handwriting recognition*. Suvisoft, 2006.

[14] V. Nguyen and M. Blumenstein, "Techniques for static handwriting trajectory recovery: a survey," in *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*. ACM, 2010, pp. 463–470.

[15] A. Graves, "Generating sequences with recurrent neural networks," *arXiv preprint arXiv:1308.0850*, 2013.

[16] X.-Y. Zhang, F. Yin, Y.-M. Zhang, C.-L. Liu, and Y. Bengio, "Drawing and recognizing chinese characters with recurrent neural network," *IEEE transactions on pattern analysis and machine intelligence*, 2017.

[17] X.-Y. Zhang, G.-S. Xie, C.-L. Liu, and Y. Bengio, "End-to-end online writer identification with recurrent neural network," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 2, pp. 285–292, 2017.

[18] F. Yin, Q.-F. Wang, X.-Y. Zhang, and C.-L. Liu, "Icdar 2013 chinese handwriting recognition competition," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 1464–1470.

[19] T.-H. Su, T.-W. Zhang, D.-J. Guan, and H.-J. Huang, "Off-line recognition of realistic chinese handwriting using segmentation-free strategy," *Pattern Recognition*, vol. 42, no. 1, pp. 167–182, 2009.

[20] A. R. Ahmad, M. Khalia, C. Viard-Gaudin, and E. Poisson, "Online handwriting recognition using support vector machine," in *TENCON 2004. 2004 IEEE Region 10 Conference*. IEEE, 2004, pp. 311–314.

[21] E. H. Ratzlaff, "Methods, reports and survey for the comparison of diverse isolated character recognition results on the unipen database," in *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*. IEEE, 2003, pp. 623–628.