2009 Fifth International Joint Conference on INC, IMS and IDC

Stroke Segmentation of Reconstruct Offline Handwriting Diagram Based On Continuous Hidden Markov Model

Liming WU, Yingmin ZHANG, Yaohua Deng, Xiuchun Tang
Faculty of Information Engineering, Guangdong University of Technology,
Guangzhou, 510006, China
jkyjs@gdut.edu.cn

Abstract- **In this paper, we experiment the capabilities of continuous density Hidden Markov Model (CHMM) to model the offline diagram sketch signals such as the electrical circuit diagram and the flowchart diagram . We attempt to imitate the online signals by extracting the offline diagram data as the time-varying coordinate sequence based on the Gradient Sharpening and Freeman code, considering that is generated by a two-level stochastic process. The underlying process governs the stroke production from a neuron-motor control point of view: go straight line, change direction line, produce a curve. A second stochastic process delivers the continuous density observed signal, which is the sequence of offline extracted points. A stroke segmentation technique based on CHMM architecture and geometric features is proposed. On a dataset of 180 hand-drawn sketches, the proposed method allows to classify correctly more than 90% of the points with respect to the connector and symbol classes.**

Keywords-Continuous Density Hidden Markov Model (CHMM); offline handwriting diagram; stroke segmentation; flowchart diagram; electric circuit diagram

I. INTRODUCTION

Sketch with pencil and paper is a convenient tool for designers and engineers to communicate and solve problem of multi language, culture and times. The sketch drawing process is also a new creative ideation. With the large application of scanner, many offline hand-drawn diagrams for engineering have been able to received and intelligibly interpreted as the digital format by computers. However, the signals obtained from the scanner are lattice image data, the stream of which is large, leading to the storage flexibility decreasing. In addition to this, especially regarding the temporal order and the dynamic information of the writing process, online handwriting signal contains more information on the writing process than the offline signal, hence the online writing recognition rate is superior. Several researchers reconstruct the temporal order of offline signals. This is often done by imitating the writing process direction, such as generating the left-right and top-bottom order performance as well as the minimal curvature at the crossings [1].

Hidden Markov Models (HMM) were first introduced in the late 1960s, and they are especially good at modeling sequential and temporal phenomenon. It has received many achievements in applications, such as speech recognition [2], human gesture [3] and facial expression [4] recognition, and handwritten text recognition [5]. However, fewer people attempt to spread HMM into sketching recognition research because it's not as regular as characters. Sezgin and Davis [6] attempt to make the sketch-recognition framework by grouping the symbols in object-level because one object would comprise multiple strokes or have shared strokes, learning users' drawing style. Artieres [7] introduced basic geometric shapes in Stroke-Level Representation (SLR) to deal with two-dimensional graphical shapes such as Latin and Asian characters. More generally, these 2-D graphical shapes can be used to encode patterns like gestures, symbols, small drawings. Herry and Wardhani [8] proposes to use chain-code line features to recognize isolated symbols. Corners here are detected simply by direction change, and it can only deal with line primitives, while our approach can recognize strokes mixing lines and arcs. Viard-Gaudin [5] reconstructs the temporal order of offline

978-0-7695-3769-6/09 $26.00 © 2009 IEEE
DOI 10.1109/NCM.2009.290

two-dimensional cursive word images into one-dimensional input signals, which is based on the simultaneous time ordering and recognition of the signal at the word level. But the word image has a large difference from diagram sketches.

These researches scribble the same symbol in different strokes numbers and orders with HMM, nevertheless they fasten on the online sketches. In this paper, we present HMM models that can segment the offline strokes in lines and arcs. The organization of the paper is as follows. Section II describes the reconstruct model from offline 2-D raw sketch image to 1-D input signals. Section III presents observation vector frame. The architecture and topology of the HMM models are described in Section Ⅳ. After discussing the experiments and results, we conclude the paper in section Ⅵ.

II. OFFLINE SKETCH IMAGE RECONSTRUCTION

Learnt from our former research [11], online handwriting signal contains more information on the writing process than the offline signal, especially regarding the temporal order and the dynamic information of the writing process. To be available for the Hidden Markov Model (HMM) segmentation, which is good at the time-varying sequence analysis, the offline 2-D sketch is first reconstructed into a true pen trajectory in this section. The transmit system is showed in Fig.1.

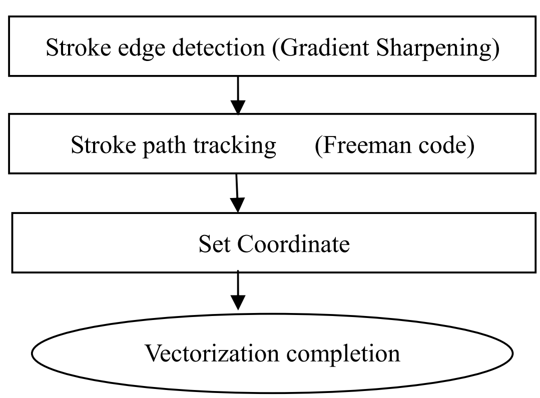


Figure 1 sketch image vectorization [9]

Gradient Sharpening, which is the first step, is to detect the edge based on the change rate of the pixel gray value. On the edge of the diagram track, the more evident the pixel gray value changes, the bigger the gradient value is. On the contrary, the lower change rate is, the less the gradient value is. In (1), f is the gray value function of the raw handwriting offline image.

$$G[f(j,k)] = |f(j,k) - f(j+1,k)| + |f(j,k) - f(j,k+1)| \quad (1)$$

Freeman chain code is introduced in the stroke path tracking step. It searches the next possible pixel based on the eight neighborhood pointed by the central pixel. The next tracking point is based on the former contour point, avoiding scanning all the image pixels and improving the stroke path tracking rate. More details of Freeman chain code are available in [10].

After the stroke path tracking, the stroke coordinate unit is still based on the grid pixel. With (2), it is transformed into the true pen trajectory.

$$\begin{cases} x_i = X_i / \mathrm{Re}\,solution \\ y_i = Y_i / \mathrm{Re}\,solution \end{cases} \quad (2)$$

With this data reconstruction, the offline diagram is suitable for the HMM model segmentation, which will be describe later.

III. OBSERVATION VECTOR STRUCTURE

From the reconstruct sequence of points, we derive a vector stream which will define the observation sequence of the HMM. In our model, two main features have been considered, that is local direction and local curvature. To achieve a specific symbol recognizer, additional geometric features will be included later.

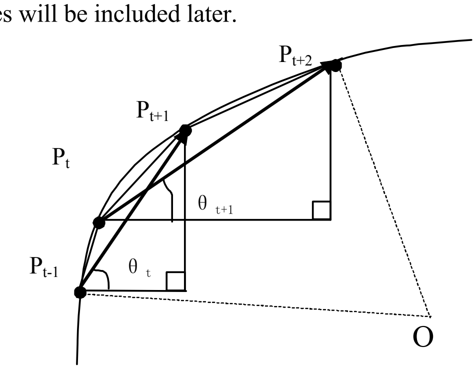


Figure 2. Direction and curvature

The local direction t is computed by the former components f1 and f2 of the observation vector. Take P_t for example, its local direction is defined by the cosine and sine value of the slant angle of line $P_{t-1}P_{t+1}$ that of line, equation (3) and (4).

$$f_1(t)=\cos\theta_t=\frac{x_{t+1}-x_{t-1}}{\sqrt{(x_{t+1}-x_{t-1})^2+(y_{t+1}-y_{t-1})^2}} \quad (3)$$

$$f_2(t)=\sin\theta_t=\frac{y_{t+1}-y_{t-1}}{\sqrt{(x_{t+1}-x_{t-1})^2+(y_{t+1}-y_{t-1})^2}} \quad (4)$$

The latter components f3 and f4 are linked to the local curvature, computed as the direction bias from line $P_{t-1}P_{t+1}$ to line P_tP_{t+2}. If these two segments are in alignment, the bias is very small, which can even reach zero. On the contrary, it increases up to the value of 2π. Equation (5) and (6) can derive from (3) and (4).

$$\begin{aligned} f_3(t) &= \cos(\Delta\theta)=\cos(\theta_{t+1}-\theta_t) \\ &= \cos\theta_{t+1}\cos\theta_t+\sin\theta_{t+1}\sin\theta_t \\ &= f_1(t+1).f_1(t)+f_2(t+1).f_2(t) \end{aligned} \quad (5)$$

$$\begin{aligned} f_4(t) &= \sin(\Delta\theta)=\sin(\theta_{t+1}-\theta_t) \\ &= \sin\theta_{t+1}\cos\theta_t-\cos\theta_{t+1}\sin\theta_t \\ &= f_2(t+1).f_1(t)-f_1(t+1).f_2(t) \end{aligned} \quad (6)$$

To adapt the HMM model, we define the Probability Density Function (PDF) of the four features as the observation emission probability, which will be described in section 4. We will suppose that these 4 components are independent so that the overall Probability Density Function (pdf) of the observation vector is the product of the respective four pdf components.

Ⅳ. CHMM TOPOLOGY

A HMM is a statistical model that describes a time-varying process, and consists of states with transitions between the states. A CHMM is that with continuous signals (or vector) observations. The continuous observation PDFS is model with a finite mixture, see (7). More details are in [12].

$$b_j(O)=\sum_{m=1}^{M}c_{jm}\Re[O,\mu_{jm},U_{jm}],1\le j\le N \quad (7)$$

Where, the PDF should be normalized as

$$\int_{-\infty}^{\infty}b_j(x)dx=1,1\le j\le N \quad (8)$$

There are N emitting states {q0, q1, . . . , qN-1} that have observation PDFs associated with them. Each emitting state qi is associated with a PDF pdfi(x(t)), where t ∈ {1, . . . , T} and i ∈{0, . . . , N-1}. Where x(t) denotes a D-dimensional feature vector at discrete-time instant t, and T is the length of the observed sequence.

A CHMM is expressed as:

λ = {A, { pdfi(x(t)), i = 1, . . . , N}}, (9)

where A is a matrix representation of the transition links and pdfi(x(t)) is the continuous observation PDF of qi evaluated at x(t), for i ∈ {1, . . . , N} and t ∈ {1, 2, . . . , T}, where x(t) is a D-dimensional vector.

Now, let us describe our segmental CHMM architecture. The role of CHMM model is to decode the offline imitated time-varying signals in the stroke level representation (SLR) process using the Viterbi algorithm (see Fig.3 and 4). In this level, the models implement time-varying functions by describing the time evolution of the tangent angle of the drawing.

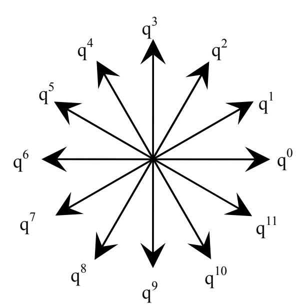


Figure 3. Straight line direction segments

Figure 4. Example of non-straight segments

Based on our team research, the pen-based basic arbitrary stroke dictionary [13], we use a fixed stroke alphabet of thirteen elementary strokes as primitives, which correspond to nine states for the CHMM (see Fig.3 and 4): twelve straight lines states and one single non-straight line state. The twelve straight line elements (q^0-q^{11}) express the twelve directions of a straight line (see Fig.3). The single non-straight line element models all other situations (see Fig.4), such as angle, round, arc, and curve. This alphabet is rich enough to represent a large collection of diagram symbols.

The CHMM model is proposed in Fig.5. In our research, the electronic circuit and flowchart diagrams are taken as the study object to study human's writing habits, leading to our CHMM model has special constrain to them based on the basic equal model [12]. We raise some transition probability based on the practical diagram situations, like the following the same direction and turning to right angle. Taking q0 of Fig.3 for example, it is a left-right horizontal direction line state. Compared to other states, it has a higher probability keeping in the same state or turning to right angle (see Fig.6). Other states have similar condition. With this special model we can encode user stroke order and habit in different kinds of diagram and increase the specificity of the modeling system.

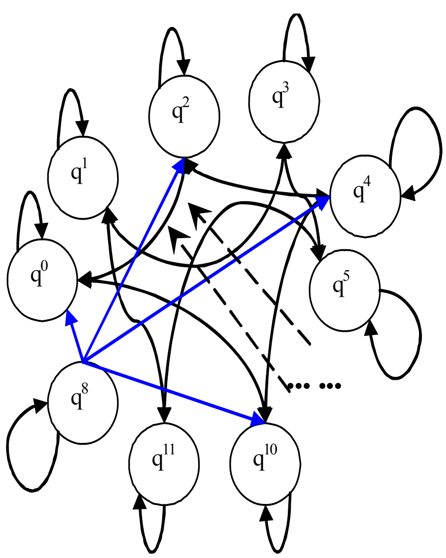


Figure 5. CHMM architecture

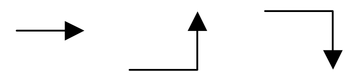


Figure 6. Turning directions for state q0

Since the PDF components are independent, the joint observation PDF of qi, evaluated at feature vector o(t), is given by:

$$pdf_i(t) = \prod_{j=1}^{4} pdf_{i,j}(o_j(t)) \tag{10}$$

Where, the feature pdf is assumed to have Gaussian distribution with a standard deviation corresponding to be on the frontiers between two consecutive directions. An observation sequence of points O is matched to λ using the Viterbi algorithm [12],and the result is the optimal hidden state sequence q = [q1, q2, . . . , qm]. The globally optimized likelihood of q, based on λ and O, is then given by:

$$\delta = a_{q_0 q_1} \prod_{t=1}^{T} a_{q_t q_{t+1}} pdf_{s_t}(\mathbf{o}(t)) \tag{11}$$

Where q0 = qs corresponds to the non-emitting initial state of λ, qT+1 = qf corresponds to the non-emitting terminating state of λ, and fst(o(t)) is described by (6). As alluded to above, is useful in our application to identify the label corresponding to straight lines (q^0 to q^{11}) and non-straight lines (q^{12}) in O.

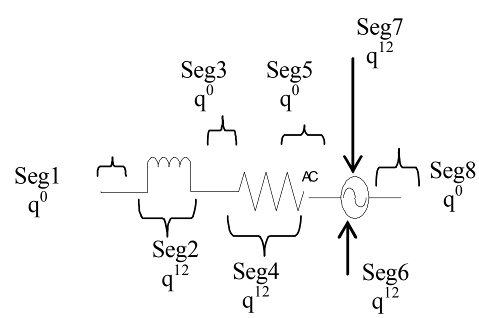


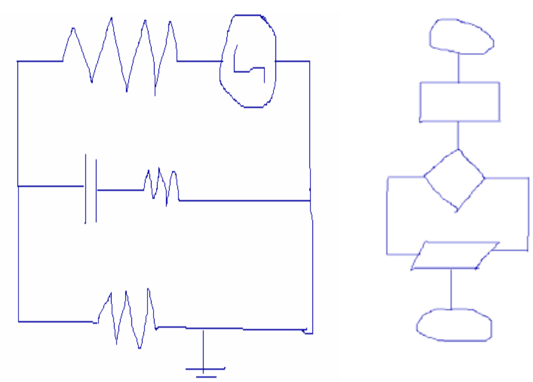
Figure 7. Example of stroke segmentation

In this figure, the Viterbi algorithm first finds a best state sequence according to the time-varying feature value, then groups the data with the same state value together and makes segmentation for different state segments at last.

V. SYSTEM EVALUATION

Our goal for the evaluation was to test the capability of our segmentation system on the imitated online sketch diagrams that were the corresponding offline diagrams. Electronic circuits and flowchart diagrams have been considered as two kinds of typical diagrams. Based on the segmentation carried out by the Viterbi algorithm, we aim to extract all the connectors, which are supposed to be only straight line segments, and consequently to increase the level of understanding of the sketch by providing a classification of the segments within connector or component classes.

The diagram images are a series of electric circuits or flowcharts. The sketches were first drawn very freely on ordinary document, and then scanned by a scanner. We ask 6 subjects to prepare the sketches for experiment. Each of them is asked to copy 30 diagrams, in which half is electric circuits and the rest is flowchart. Fig.8 are some examples of sketch images.



(a) an electric circuit diagram (b) a flowchart diagram

Figure 8. Sketch image examples

Experiment results are given in Fig.9 and 10. The two measurements are false-negative rate and false-positive rate. The definitions are given below.

$$false\_positive = \frac{no\#\_accepted\_false\_segment\_point}{no\#\_all\_accepted\_segment\_point} \times 100\% \tag{10}$$

$$false\_negative = \frac{no\#\_rejected\_true\_segment\_point}{no\#\_all\_true\_segment\_point} \times 100\% \tag{11}$$

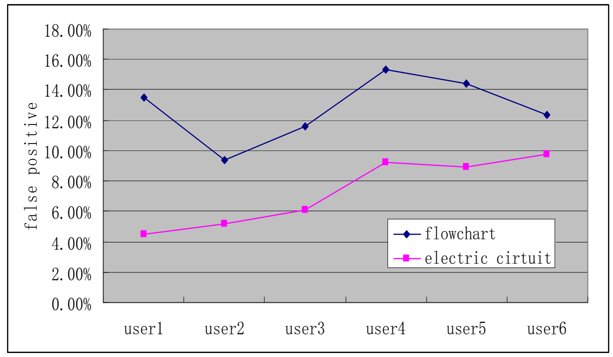


Figure 9. Recognition result of false positive

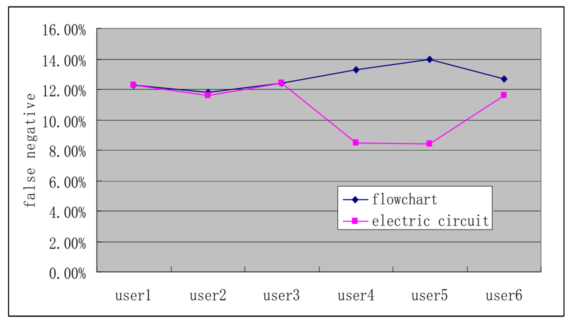


Figure 10. Recognition result of false negative

From the experiment it can be seen that for all the users, our approach has achieved a considerable result. The false negative rate in our approach is very low, less than 16%, while the false positive rate is also only around 10%. Especially, that of the electric circuit diagram, it's around 6% percent. Since CHMM based stroke segmentation is a task to analysis the time-varying sequence and try to find out the optimal matching models, our attempt shows that the reconstruct offline sketch diagram method is able to imitated online handwriting track. However, both the rates vary between users. Take the false positive rate of Fig.9 for example, user2 is lower than 9% percent, while user4 goes up to 16 % when drawing the same diagrams. This is due to the fact that different user has different drawing style. If the sketch order and slim, it is more easer to be recognized and reconstructed. If the sketch is scratchy, for example, the straight seems to be a curve or the angle seems to be a long arc, it will lead to the increase of the false positive recognition. Further, the size of the diagram has an indispensable effect on the segmentation. Different size diagram with the same detective and reconstructed distance results in the information quality between each two points is not equal.

Observed from the experimental result, we also found that the false rate of electric circuit diagram is less lower than the flowchart diagram. That is influenced by more subjective factors. One of the reasons is because the electric circuit components have more different features than the straight connector, while the flowchart diagram has less ones.

Ⅵ. CONCLUSION

Stroke segmentation for sketches is a burgeoning research in all over the world. It has many different features with the character, because it is an unbending and semantic blurry input that reflects a natural and direct intercommunication. Both in online and offline domain, it is a challenge task. In this paper, we present an offline sketch understanding based on continuous density Hidden Markov Model (CHMM), which imitates online handwriting diagram recognition by reconstructing the corresponding offline diagram. The model first reconstructs the diagram to be a continuous time-varying stroke stream, next does the segmentation by a fixed alphabet of nine elementary strokes of geometric features, then classifies the connectors and other components according to the best state sequence of the Viterbi algorithm. The preliminary experiment has achieved an average recognition rate of almost 90%. This demonstrates that the extension of the use of CHMM from online to offline handwritten sketch understanding is a feasible area of interest.

In the future, we plan to plan to solve problems when there is many overtraded oval happened in the diagram, and conduct a more convictive mode evaluation of the proposed approach.

REFERENCES

[1] P.M. Lallican, C. Viard-Gaudin, S. Knerr, “From off-line to on-line handwrting recognition”. In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 303-312

[2] Y. Luo, “ A human speech emotion recognition classifier based on hidden markov model and artificial neural network”. Control & Automation, vol.23, Dec. 2007,

[3] D.Willems, R. Niels, M. Gerven, L. Vuurpijl. Iconic and multi-stroke gesture recognition. Pattern Recognition (2009), doi: 10.1016/j.patcog.2009.01.030

[4] S. Zhou, X. Liang, C. Zhu, Q. Yang. “Facial Expression Recognition Based on Independent Component Analysis and Hidden Markov Model”. Journal of image and graphics. Vol.13, Dec. 2008.

[5] C. Viard-Gaudin, P.M., Lallican, S. Knerr. “Recognition-directed recovering of temporal information from handwriting images”. Pattern Recognition Letters 26 (2005) 2537-2548.

[6] T.M. Sezgin and R. Davis. “Sketch Interpretation Using Multiscale Models of Temporal Patterns”. Computer Graphics, 2007 (Vol. 27, No. 1) pp. 28-37.

[7] T. Artières, S. Marukatat and P. Gallinari. “On-line handwritten shape recognition using segmental Hidden Markov Models[J]”. In IEEE. Trans. on Pattern Analysis and Machine Intelligence. 29.2 (2007), pp. 205-217.

[8] D. Henry, A. Wardhani. “Diagram recognition using Hidden Markov Models”. Technical Report, Faculty of Information Technology, Queensland University of Technology, 2003.

[9] X. Pang，J. Yu, F. Liu. “ Research on vectorization processing based on ICT image” [J]，Machinery Design ＆ Manufacture，(2005)

[10] L. Tang, A. He, H. Chen. “Optimization vectorization method of image contour for quilting”, Electronic Measurement Technology, vol.31, Apr. 2008

[11] Y. Zhang, C. Viard-Gaudin, L. Wu. “An online hand-drawn electric circuit diagram recognition system using hidden markov models”. Proc.IEEE. International Symposium on Information Science and Engineering (ISISE2008), Shanghai, Dec. 2008, Vol. 2, pp.143-148

[12] L.R. Rabiner. “A tutorial on HMM and selected applications in speech recognition”. Proceedings of the IEEE 77,2 (1989), pp.257-286.

[13] G. Feng, C. Viard-Gaudin. An approach for stroke fragmentation based on geometry features and Hidden Markov Model. Technical Report, Ecole Polytechnique de l’Université de Nantes, 2007.