A Unified Network-based Approach for Online Recognition of Multi-Lingual Cursive Handwritings

Jay J. Lee and Jin H. Kim

Department of Computer Science & Center for AI Research, KAIST
373-1 Kusong-dong Yusong-gu, Taejon 305-701, Korea
joony@ai.kaist.ac.kr, jkim@cs.kaist.ac.kr

Abstract

Although several studies have focused on recognition of individual language, no attempt has been seriously made for online recognition of handwritten script in multiple languages. In this paper, a network-based approach is proposed for recognizing sequences of words in multiple languages. Viewing handwritten script as an alternating sequence of words and interword ligatures, a hierarchical hidden Markov model(HMM) is constructed by interconnecting HMMs of ligature and word of multiple languages. In turn, those HMMs are constructed from HMM of lower level components. Given such a construction, recognition corresponds to finding optimal path in the network which can be searched efficiently by Viterbi algorithm. Although combining component languages, recognition accuracy of each language drops negligibly little, and recognition result of intermixed usage is acceptable.

1 Introduction

By development of pen computing technology, handwriting could be one of the major modes of human and computer interaction. Although several products which use pen for the sole medium of user input are on the market, their handwriting recognition capabilities are somewhat limited [1]. Therefore, recognizing unconstrained handwritings still remains unsolved problem which is being attacked by several innovative approaches such as neural networks, hidden Markov models, and fuzzy theories.

As long as English is used as international language, intermixed use of English words in their native non-Roman text is inevitable in most of the Asian countries. For instance, Hangul text is often used intermixedly with Chinese characters and English words. Figure 1 shows a Hangul text in which English words are embedded. Since no attempt has been seriously made to develop a recognizer to handle texts written in more than one languages, somewhat awkward interface would be introduced - such as clicking a button before writing or using specially reserved area for a particular language - to indicate the language currently in use.

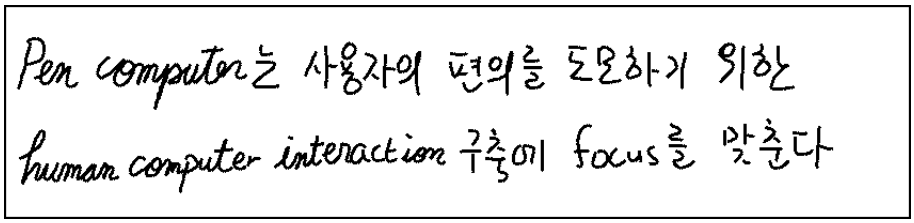


Figure 1: **Typical Hangul Text with English Words**

In this paper, we present a unified network-based approach for recognition of multi-lingual handwriting. This approach is based on the hierarchical network of hidden Markov models(HMM). For each language, basic components are modelled by HMM, and composed to network structure according to the language composition rules. The unified network for multiple language is constructed by linking each language's HMM network. Recognition is formulated as searching an optimal path in this network. The details for searching steps, implementation issues, and experimental results are shown in the following sections.

2 Unified HMM Network for Recognition of Multi-Lingual Handwriting

2.1 Three Approaches for Multi-Lingual Recognizer

Several approaches are possible to develop a system which can recognize handwritten text in intermixed languages. The first approach is to develop individual recognizers for each language, and apply them parallel for all handwritings. Recognition trials are then evaluated, and one of them is selected as a recognition result. This approach may easily utilize individually developed recognizers, but it may not be easy to compare scores produced from recognizers with different scales. The second considerable approach is to develop a preclassifier which classifies unknown handwritings into a language class by examining a certain set of features. Then, an appropriate recognizer does its job. This second approach seems efficient if the decision of language category is reliable. But it requires possibly more difficult task to develop a robust preclassifier. Finally, we propose a unified recognizer for the entire character set of target languages. It uses same modeling and decoding method together for each languages. Therefore, it is possible to construct one recognizer for the job. This approach can avoid the problems of the first and second approaches. However, conquering the different features of different languages, and unifying them into one recognizer are not easy tasks.

2.2 HMM Network-based Approach

We can consider handwritten word as an alternating sequence of characters and intermediate ligatures by encoding both pen-down and pen-up movements into observation symbols. Therefore, words in a language are modeled by a network of interconnected character and ligature models. For instance, Hangul syllables, which are composed by two or

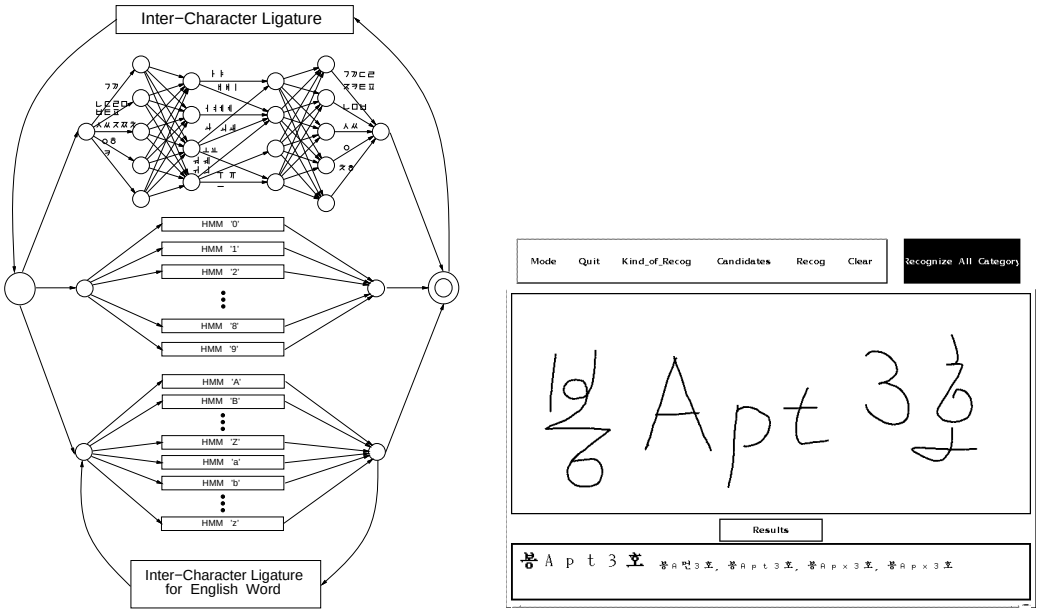
three character words, are modeled by a 5-layer finite state network called BongNet [3], and English words of indefinite length are modeled by circularly interconnected models of character and ligature [4]. Because of the spatiotemporal variability on the handwriting, hidden Markov modeling technique [2] is adopted to construct the character and ligature models.

In such a word network, a path spanning from the initial node to the final node, possibly with a number of cycles on the way, represents a word. Consequently, the recognition problem is a matter of finding the most likely path from the initial node to the final node, consuming all symbols of the given observation sequence. The Viterbi algorithm is developed for this purpose. We have modified the original Viterbi to keep several highest tracks of transitions so that we can check with language models and other source of knowledge for postprocessing. From the maximal probability path, optimal character and ligature segmentation, and associated character labels can be obtained simultaneously.

2.3 Unified Network for Multi-Lingual Handwriting

Such a HMM network mentioned above is constructed by grouping and interconnecting basic component's HMM as character composition rules of corresponding language. In this regard, we may call such construction as a hierarchical HMM, where, the high level HMM models sequence of basic characters, while the lower level HMM models their components. This structure can be easily extended to multi-lingual case. HMM network for each language may be regarded as a middle component HMM, and these are combined into a single network structure. In addition, ligatures among words of each language, which are mostly pen-up movement, are modeled as a circular path from the global terminal node to the global initial node. Because each HMM network produces a same scale probability associated with a path, and these all paths are joined into the global node of the network, there is no scaling problem from language combining. As a result, a *unified network* is constructed for recognition of intermixed usage of multiple language. In this paper, unified network is constructed by BongNet which models Hangul syllables, the interconnected HMMs for English words, and the simple network for digit as shown in Figure 2-(a). As done in individual networks for single language recognition, all available paths in the unified network from the global initial node to the global terminal node are competing for the label of the observation sequence. Therefore, if we cut the path into a network of specific language, various kinds of multi-lingual recognizer can be easily made. Any language combinations can be handled within the framework providing proper model training.

Various kinds of knowledge can be easily integrated with such a network based approach. Knowledge from language models is integrated by checking bigram or trigram at the entrance of each character HMM. In addition, the regularity of Hangul construction rule is used for the efficiency of search. Checking validity of intermediate results by transition automata between languages is also added in the circular path for preventing improper usage. Finally, structural constraints, either within character or between characters, are confirmed when exiting from each character HMM. Furthermore, subtle differences among some digits and Roman alphabets, such as the difference in printed style 'h' and 'n', are emphasized by a pairwise discrimination at the end of the character HMMs.



(a) (b)

Figure 2: **(a) Unified Network for Intermixed Use of Hangul Syllable, English word, and Digit (b) Recognition Example**

3 Implementation and Experiments

Data used in the training and experiments were collected with no constraint on writing style (but keeping writing order) by using a digitizer connected to a SPARC workstation. For Hangul training, we used about 52,000 Hangul syllables written by 59 Korean writers without any constraints. For English, about 50,500 handwriting alphabets and ligatures were collected from 23 Korean writers (in case of uppercase alphabet, we modelled only printed style). Also 4,100 digits from 12 writers were obtained for training of digit HMM. For experiment, we constructed test data set by different writers who did not participate the training data set.

Data were preprocessed for removing noise and reducing variablities at first [5]. Specially for slant correction, sophisticated method was needed because there exist a lot of of diagonal strokes in Hangul. Delayed strokes on English word were carefully extracted and registered for later usage in search process. Pen-down movement was encoded as a chain of 16-directional codes and pen-up movement, invisible stroke, was encoded to another 16-directional codes.

The recognition results of word data are shown in Table 1. 'Multiple Recog' row shows recognition results by the proposed multi-lingual recognizer, and 'Singular Recog' shows results of the particular language recognizer without interfering other language recognizers. For testing handwritings of intermixed language usage, we used multi-lingual texts of Hangul, English and digits. The result shows the degradation of recognition ratio by adding other languages is about 1 percent, which provide us confidence to claim that recognition accuracy for the intermixed handwritings is acceptable.

Table 1: Word recognition result

	number of writers	number of test data	Avg. word Length	Multiple Recog (correctness, %)	Singular Recog (correctness, %)
Hangul	7	1012	2.8	80.73	81.82
English	9	1710	4.5	82.11	82.75
Intermixed	8	785	4.1	83.82	

4 Discussion

We have proposed a unified network approach for handwritten multiple language recognition. By hierarchical HMM construction, multiple languages are combined into one recognizer under the probability framework. Recognition accuracies of each language categories are sustained for the intermixed usage. But more powerful search methods and machines are in demand to make this approach practical. We believe this approach can be easily applied to any combination of writing systems including Arabic, Thai, and Japanese.

Language combining causes the increment of basic characters, and this surely brings the rise of confusing pairs of similar shapes. Therefore more powerful heuristics such as pairwise discrimination are needed.

References

[1] Michael Nadeau, *"PDAs Bounce Back,"* BYTE, Vol.20, No.6, pp. 147–156. Jun. 1995.

[2] L. R. Rabiner, *"A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,"* Proc. of the IEEE, Vol.77, No.2, pp.257–286, 1989.

[3] B. K. Sin, J. H. Kim, *"A Statistical Approach with HMMs for On-line Cursive Hangul (Korean Script) Recognition,"* Proceedings of the Second International Conference on Document Analysis and Recognition, Zukuba, Japan, pp147–150, Oct. 1993.

[4] J. Y. Ha, S. C. Oh, J. H. Kim, *"Recognition of Unconstrained Handwritten English Word with Character and Ligature Modeling,"* Pattern Recognition and Artificial Intelligence, Vol.9, 1995

[5] Charles C. Tappert, Ching Y. Suen, Toru Wakahara, *"The State of the Art in On-Line Handwriting Recognition,"* IEEE Tr on PAMI, pp 787–808, Aug. 1990.