Online Handwritten Lao Character Recognition by MRF

Latsamy SAYSOEURINHONG(a), Nonmember, Bilan ZHU†, Member, and Masaki NAKAGAWA†, Fellow

SUMMARY  This paper describes on-line recognition of handwritten Lao characters by adopting Markov random field (MRF). The character set to recognize includes consonants, vowels and tone marks, 52 characters in total. It extracts feature points along the pen-tip trace from pen-down to pen-up, and then sets each feature point from an input pattern as a site and each state from a character class as a label. It recognizes an input pattern by using a linear-chain MRF model to assign labels to the sites of the input pattern. It employs the coordinates of feature points as unary features and the transitions of the coordinates between the neighboring feature points as binary features. An evaluation on the Lao character pattern database demonstrates the robustness of our proposed method with recognition rate of 92.41% and respectable recognition time of less than a second per character.

key words: online handwritten character recognition, Lao character recognition, Markov random field

1. Introduction

In the past decade, the research on character recognition in both offline and online systems has revived attention due to the development of handheld devices and proliferation of the devices into all over the world. Not just western languages but Japanese, Korean and Chinese, Arabic, Indic and their derivatives are gathering interest in research and development. Many methods have been published to improve their derivatives are gathering interest in research and development. Many methods have been published to improve recognition rate, to remove writing constraints, to expand the languages to recognize, to speed up recognition, to make the recognizer compact and so on.

In online recognition, hidden Markov models (HMMs) have been dominant in many languages. HMMs were first described in the series of statistical papers [1] and they were applied to speech recognition [2], [3] in the middle of 1970s. Then, they were applied to online handwriting recognition of western languages. Recently, they were applied to Indian scripts [4]–[6]. However, HMMs have less degree of freedom than MRFs to express relations among multiple feature points explicitly, with the result of limited recognition accuracy [7], [23].

The template-based elastic matching algorithm working on handwritten character recognition is a recognition method that works well with online handwriting recognition for user dependent systems. Both dynamic time warping and elastic matching algorithms that work on Indian scripts [8]–[10] and on alphanumeric characters [11] have also been reported. Furthermore, many researches working on Indian scripts [12]–[15] had also been reported to achieve an acceptable result. However, the method does not consider the distributions of training patterns, with the result of limited recognition accuracy.

Despite many years of research in the fields of handwritten recognition, it has not reached the majority of local languages, such as Lao. As one of many countries under development in South East Asia, Laos has been opened to the new era and has adopted new technologies in daily life and the needs for research in this field has become one of many urgent issues to bring Laos into the IT world.

Bounnady [16] has introduced a template-based research for online handwriting Lao character. At first they segmented an input character pattern according to the changes in directions of curves (clockwise or counter clockwise). From each segment, the following features are gathered: cumulative change in direction of curve, cumulative curve length, cumulative length of left to right, that of right to left, that of top to bottom and that of bottom to top. The experimental result shows the promising recognition rate of 94.62%. However, in this paper they included only 27 single consonants (see Fig. 1 (a)) in the experiment and did not mention about the time consumption during testing process which is an important factor for the online handwritten character recognition.

We have introduced a template-based approach for online handwritten Lao character recognition [17] for consonants, vowels and tone marks, 52 characters in total. The system matches feature points from the input pattern to each prototype of each character class based on the dynamic programming algorithm. It does not consider the distributions for each feature point and only uses the unary features to calculate the distances between matched pairs of feature points of the input pattern and each prototype, and then sum those distances to evaluate the similarity between the input pattern

Fig. 1  Consonants (a) Single consonants and (b) Compound consonants.
and each prototype.

The MRF is described by an undirected graph in which a set of random variables have Markov property and, MRFs can effectively integrate the information among neighboring feature points such as binary features and trinary features [18] and they have been successfully applied to off-line handwritten character recognition [19] and on-line stroke classification [20]. Wolf, C. [21] has successfully employed MRF for document images binarization. Moreover, the MRF-based approach [22] has been shown to perform well in the context of enhancement of degraded historical typewritten document image. Current on-line handwritten character recognition tends to use HMM-based models (note that HMMs can be viewed as a specific case of MRFs). Zhu et al. [23] proposed MRF based framework for on-line handwritten Japanese character recognition with a higher recognition rate achieved.

In this research, we apply the MRF models for Lao character recognition and compare the performance with the previous method. The model effectively integrates unary and binary features and extracts feature points along the pen-tip trace from pen-down to pen-up and matched those feature points with states for character classes probabilistically based on this model. Experimental results on the collected handwritten Lao character database demonstrate the superiority of our method.

The rest of this paper is organized as follows. Section 2 gives an introduction to Lao scripts, and Sect. 3 presents the MRF model for Lao character recognition. Section 4 describes the collection of handwritten Lao character patterns used in the experimental process. Section 5 shows the result of experiments and Sect. 6 draws our conclusion.

2. Lao Scripts

Lao language is the official language of Lao People Democratic Republic (Laos). The script was derived from Sanskrit scripts; which is the classical standard language of ancient India. However, the modern Lao script has been modified both the character shape and the number of scripts but still have the same writing system as the original.

In the Lao language, the word contains syllables composed of vowel, consonant, and tone mark (A tone mark is used to characterize the change of syllable sound to short, medium, low or high tone.) The consonant consists of 27 single and 6 mixed consonants as shown in Fig. 1. Table 1 show the vowels along with tone marks and the place where they appear in sentences when writing the Lao language.

2.1 Consonants

The consonants including 27 single characters and 6 compound characters (the combination of two single consonant; sometimes the form is changed after the combination, as shown in Fig. 1) are divided into three tone classes; high, middle and low, which help to determine the tone of a syllable [24].

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Type</th>
<th>Appear in</th>
</tr>
</thead>
<tbody>
<tr>
<td>ใ ใл ใъ</td>
<td>Vowel</td>
<td>after consonant</td>
</tr>
<tr>
<td>๐ ๒ ๓</td>
<td>Vowel</td>
<td>above consonant</td>
</tr>
<tr>
<td>๔ ๕ ๖ ๗</td>
<td>Vowel</td>
<td>below consonant</td>
</tr>
<tr>
<td>๘ ๙ ๑๐</td>
<td>Vowel</td>
<td>before consonant</td>
</tr>
<tr>
<td>๑๑ ๑๒ ๑๓</td>
<td>Tone mark</td>
<td>above consonant or vowel</td>
</tr>
</tbody>
</table>

Fig. 2 Samples of the words whose meanings change when different tone marks appear. (a) Without the tone mark means “horn”, (b) With the tone mark means “knees”, (c) With the different tone mark means “rice”.

The sounds represented by some consonants change when they are used at the end of a syllable. The consonants can all be used at the beginning of a syllable but only some can be used at the end of the syllable.

2.2 Vowels and Tone Marks

Lao has a set of vowels with 18 classes as shown in Table 1 that are distinguished between vowel lengths (short or long) and vowels position (front to back). Vowel signs are always written around consonants.

However, combination of vowels themselves produces a new vowel and the combination of a vowel and a consonant creates another vowel sound. When writing Lao, a vowel can be placed either in front, or above, or right or below a consonant in the word (see the details in Table 1). We do not include the combination forms of vowels in our database because we can obtain them by concatenating those single vowels.

A tone mark is a symbol which is used to characterize the change of syllable sound to short, medium, low or high tone. There are four tone marks in Lao as shown in Table 1. A tone mark can only be placed above a consonant or vowel. With the present of the tone mark in a word, the meaning of the word is changed according to the tone mark that has been used. Figure 2 illustrates how the meaning of one word is changed to the appearance of the tone mark and Fig. 2 (c) has another meaning when using a different tone mark. Figure 3 shows a sample of a Lao sentence in printed form (a) and in its handwritten form (b).
3. MRF for Handwritten Lao Character

3.1 Stroke Normalization and Feature Point Extraction

The distinctive characteristic of Lao scripts is that each character is mostly consisting of only one stroke (as we can see from Fig. 1). However, the number of input points for each character collected varies according to the writer. We normalize an input pattern linearly by converting the pen trace pattern to a standard size, preserving the horizontal-vertical ratio.

After the normalization, we extract feature points using the method by Ramner [25]. Ishigaki et al. employed a similar method [26], [27]. First, the starting point and the end of every stroke are picked up as the feature points. Then, the most distance point from the straight line between adjacent feature points is selected as a feature point if the distance to the straight line is greater than a threshold value. This is applied recursively until no more feature point is selected. The feature points extracting process is shown in Fig. 4.

3.2 Markov Random Field Models

We apply the MRF to model Lao Character Recognition. The modeling method is similar to the one described in Ref. [23] for online handwritten Japanese character recognition. For the purpose of completeness we include a brief description here. For any further details the reader is referred to Ref. [23].

We set feature points from an input pattern as sites \( S = \{s_1, s_2, s_3, \ldots, s_I\} \) and states of a character class \( C \) as labels \( L = \{l_1, l_2, l_3, \ldots, l_J\} \). The system recognizes the input pattern by assigning labels to the sites to make the matching between the input pattern and each character class \( C \) such as \( F = \{s_1 = l_1, s_2 = l_1, s_3 = l_2, \ldots, s_9 = l_8, s_{10} = l_8, \ldots, s_{14} = l_{11}\} \) as shown in Fig. 5. \( F \) is called configuration and denotes a mapping from \( S \) to \( L \).

The feature vectors of the feature points from the input pattern constitute the observation set \( O \). The problem of recognition is to minimize the following energy function and to obtain the best matching:

\[
E(O, F|C) = E(O|F, C) + E(F|C)
\]

where \( E(O, F|C) \), \( E(F|C) \) and \( E(O|F, C) \) are called the energy function. For simplicity, we consider only single-site cliques \( c_1 = \{s_i\} \) and pair-site cliques \( c_2 = \{s_i, s_j\} \) in our system.

We define cliques as follow:

- Single-site: \( c_1 = \{s_1, s_2, s_3, \ldots, s_{10}, \ldots\} \)
- Pair-site: \( c_2 = \{s_1, s_2\}, \{s_2, s_3\}, \{s_3, s_4\}, \ldots, \{s_9, s_{10}\}, \ldots \)

The neighborhood system is according to the successive adjacent feature points in writing order. We define a linear-chain MRF for each character class as shown in Fig. 6, where each label has a state and each state has three transitions.

Therefore, the energy functions can derive as follow:

\[
E(O, F|C) = E(O|F, C) + E(F|C)
\]

\[
= \sum_{j=1}^{I} \left[ -\log P(O_{s_j}|F, C) - \log P(O_{s_{j-1}}|F_{s_j}, l_{s_{j-1}}, C) \right]
\]

where \( I \) is the number of feature points, \( l_{s_j} \) is the label of a class \( C \) assigned to \( s_j \), \( O_{s_j} \) is the unary feature vector extracted from a site \( s_j \), \( O_{s_{j-1}} \) is the binary feature vector extracted from the combination of \( s_j \) and \( s_{j-1} \).

The smaller is the value of the energy function in Eq. (2), the larger is the similarity between the input pattern and a character class \( C \).

Each character class has a linear-chain MRF. The system uses the Viterbi search to match feature points of an input pattern with states for the MRF of each character class and then find the matching path with the smallest value of the energy function in Eq. (7) for each character class.

The unary feature vector, \( O_{s_j} \), is consisting of \( X \) and \( Y \)-coordinates of site \( s_j \) and the binary feature vector \( O_{s_{j-1}} \) has two element; which are: \( X \) coordinate of \( s_j - X \) coordi-
nate of $s_{i-1}$ and Y coordinate of $s_i$. The Gaussian function is applied to estimate $P(O_{s_i} | l_{s_i}, C)$ and $P(O_{s_{x_{i+1}}} | l_{s_{x_{i+1}}}, l_{s_{x_{i+1}}}, C)$. $P(l_{s_i} | l_{s_{x_{i+1}}}, C)$ is estimated as follows:

$$P(l_{s_i} | l_{s_{x_{i+1}}}, C) = \frac{\text{Number of transition from } l_{s_{x_{i+1}}} \text{ to } l_{s_i}}{\text{Number of sites assigned } l_{s_{x_{i+1}}}}$$

$$P(l_{s_i} | l_{s_{x_{i+1}}}, C) = \frac{\text{Number of } s_i \text{ assigned } l_{s_i}}{\text{Number of } s_1 \text{ assigned } l_{s_1}}$$ (3)

Furthermore, to train the MRF for each character class, we first initiate feature points of an arbitrary character pattern among training patterns of the character class as states of the MRF, and secondly set each unary feature vector of each feature point as the mean of Gaussian function for each single-state, and then set each binary feature vector between two adjacent feature points as the mean of Gaussian function for each pair-state, and finally, initiate the variances of those Gaussian functions and the state transition probabilities as 1.

After that, we apply the Viterbi algorithm to match the features points of training patterns of the character class to the states of the MRF, and then estimate the parameters of the MRF (the means and the variances of Gaussian functions and the state transition probabilities) by using these matched feature points. Then we update the parameters and re-apply the Viterbi algorithm to match the features points of training patterns to the states of the MRF. We repeated the matching and the estimation of parameters three times.

4. Handwritten Lao Character Database

To support this research we have collected handwritten Lao character patterns from 50 native and 1 non-native speakers who learned Lao language, in which each person wrote each character 10 times. The database used for this research[28] has 26,520 samples, and 52 character classes consisting of 30 consonants (27 single consonants and 3 compound consonants), 18 vowels and 4 tone marks. We did not include three other (compound) consonants in the collection because those characters can be formed from other two consonants.

Figure 7 shows the interface for collecting the Lao handwritten character patterns where each character is written in the square box under its printed version shown above.

5. Experiments

In the experiment, the sample patterns from the handwritten Lao character database[28] were divided into 5 groups

![The interface for collecting Lao character patterns.](image)

Fig. 7 The interface for collecting Lao character patterns.

where each group includes 10 persons’ handwriting except the last group including one more person’s handwriting. We follow the cross validation method to measure the recognition rate and select one group among the 5 groups as the testing data and merge all the remaining groups as the training data. We use training data to estimate the parameters of each MRF model and use testing data to test the system performance. For discussing the recognition rates on testing patterns we take the average of the 5 testing data. The experiments were implemented on an Intel(R) Core(TM) 2 Duo CPU 2.40 GHz with 0.99 GB memory.

We compare the performance of three recognition models: MRFs and two template-based models presented in [17] (LTM and DPM). LTM and DPM extract the same feature points from on-line patterns as the MRFs and then match those feature points from the input pattern with those of each prototype of each character class. LTM and DPM does not consider the distributions for each feature points and only uses the following unary features $x$, $y$, $dir$: $x$: X coordinates of $s_i$, $y$: Y coordinates of $s_i$, $dir$: $\tan^{-1}(dy/dx)$

LTM uses linear-time elastic matching method to match feature points from the input pattern with those of each prototype of each character class while DPM uses the dynamic programming matching method to match those feature points. Table 2 shows the results from a cross-validation experiment on various models.

As we can see from the experimental result, the MRF Model brought better recognition performance although it consumed slightly more processing time and larger memory space compared with the template-based models (LTM and DPM). The MRF model did not take as much processing time and did not use as larger memory space as Japanese character recognition. This is because the Lao language has a very small set of character classes. We consider that the processing time and memory space usage for the MRF are acceptable. The experimental result has shown the recognition rate of 92.41% with highest rate of 99.63% when considered the correct character is in the top 10 candidates (see Figs. 8 and 9).

When we considered the results with only consonants, the recognition rate is 93.06% as the average result (with highest of 99.40% and lowest of 82.80%). The recognition...
rate for vowels was 91.1% as the average result (with highest of 98.60% and the lowest of 65.35%). The recognition rate for tone marks was 93.4% as the average result. The individual recognition results for each character are shown in Tables 3 and 4.

In Tables 3 and 4, we have noticed that the characters with lower recognition rates have unique shapes.

Tables 5 and 6 show the details of misrecognitions for the characters with a little lower recognition rates.

Figures 8 and 9 show the top rank candidates recognition results for each character that consider the correct characters are in the top rank candidates, where B01, B05 and B10 mean the correct results are in the 1st rank candidates, in the top 5 rank candidates and in the top 10 rank candidates, respectively. From the results, we can see that the recognition rates of B05 and B10 are higher than those of B01, and we can dramatically improve the recognition rate.
We observed two major sources causing the recognition errors from Tables 5 and 6.

(1) The problem of similarities between character classes: as we have mentioned before, in Lao scripts many character classes have similar shapes that can cause a major misrecognition in the experiment. To solve this problem, we should take into account other important feature factors such as size and position of each character. Exploiting linguistic context can dramatically reduce such misrecognitions. If we look more closely into the unique features of Lao scripts, we can see that there are a number of characters which are similar to other ones. It is difficult to distinguish them because with handwriting people do not pay much attention of the small details, such as head or tail of that character. For example, for the last two characters in Table 5, even the human eyes are sometimes not able to distinguish one from another or take time to do so.

Furthermore, when writing in the Lao language, the sizes of consonants, vowels and tone marks are different; that is, consonants usually have a bigger size when compared to tone marks or some vowels. However, in this research we do not taken into account the size of character.

(2) The problem of misrecognition: we need to increase the number of training patterns to improve the character recognition accuracy. Furthermore, as we have mentioned in (1), by considering other factors in the recognition process we can also improve the system performance.

Table 5: The details of misrecognized consonants.

<table>
<thead>
<tr>
<th>Char</th>
<th>Recognition Rate (%)</th>
<th>Details of misrecognized character: %character</th>
</tr>
</thead>
<tbody>
<tr>
<td>ຘ</td>
<td>90.8</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຜ</td>
<td>86.8</td>
<td>Incorrect: 0.55( ), 1.75( ), 1.33( ), 0.55(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>90.6</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ພ</td>
<td>86.8</td>
<td>Incorrect: 0.55( ), 1.75( ), 1.33( ), 0.55(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>92.8</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ພ</td>
<td>90.6</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>89.2</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>92.2</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>94.6</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>94.6</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>96.4</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>92.6</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>82.8</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>85.2</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>77.4</td>
<td>Correct: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
<tr>
<td>ຝ</td>
<td>22.6</td>
<td>Incorrect: 1.89(w), 1.33(r), 0.67( ), 0.04(b)</td>
</tr>
</tbody>
</table>

Table 6: Misrecognized vowels and tone marks.

<table>
<thead>
<tr>
<th>Char</th>
<th>Recognition Rate (%)</th>
<th>Details of misrecognized character: %character</th>
</tr>
</thead>
<tbody>
<tr>
<td>ພ</td>
<td>86</td>
<td>Correct: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>92.6</td>
<td>Incorrect: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>92.8</td>
<td>Correct: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>93.4</td>
<td>Incorrect: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>89.2</td>
<td>Correct: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>92</td>
<td>Incorrect: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>90.6</td>
<td>Correct: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>78.16</td>
<td>Incorrect: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>65.35</td>
<td>Correct: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
<tr>
<td>ພ</td>
<td>34.65</td>
<td>Incorrect: 1.56( ), 1.21( ), 0.96( ), 0.87( )</td>
</tr>
</tbody>
</table>

6. Conclusions

In this research we presented an approach for on-line handwritten Lao character recognition in which the system performed the task by comparing the input character pattern with the character classes to determine the similarities by adopting the MRF models. The experimental result has shown the robustness of the system with a promising 92.41% recognition rate.

In the future, recognition of on-line handwritten words and sentences need to be considered.

Acknowledgments

We would like to thank the students and colleagues from the National University of Laos and friends, who helped us to compile the database of handwritten Lao character patterns.

References

2007.


