Classification of the Structure of Square Hmong Characters and Analysis of Its Statistical Properties

Li-Ping Mo(✉), Kai-Qing Zhou, Liang-Bin Cao, and Wei Jiang

College of Information Science and Engineering,
Jishou University, Jishou 416000, Hunan, China
zmx89@jsu.edu.cn

Abstract. Analysis of the character structure characteristics can lay an information foundation for the intelligent processing of square Hmong characters. Combined with the analysis of character structure characteristics, this paper presents a definition of the linearization of square Hmong characters, a definition of equivalence class division of the structure of square Hmong characters, and proposes a decision algorithm of structure equivalence class. According to the above algorithm, the structure of square Hmong characters is divided into eight equivalent classes. Analysis of the statistical properties, including the cumulative probability distribution, complexity, and information entropy of square Hmong characters appearing in practical documents, shows that, first, more than 90% of square Hmong characters appearing in practical documents are composed of two components, and more than 80% of these characters possess a left-right, top-bottom, or lower-left-enclosed structure, second, the number of mean components in a square Hmong character is slightly greater than 2, third, the information entropy of the structure of Hmong characters is within the interval (1.19, 2.16). Results reveal that square Hmong characters appearing frequently in practical documents follow the principle of simple structure orientation.

Keywords: Information entropy · Probability distribution
Square Hmong character · Statistical analysis

1 Introduction

Analysis of character structure characteristics is the basis of character attributes analysis, and plays an important role in the intelligent generation and character recognition. In the field of Chinese information processing, many methods for intelligent character generation, recognition and other related processing methods based on the analysis of character structure characteristics were proposed. Instituse put for-ward a refinement method for smoothing Chinese character images using the nearest neighbor pixel correlation function to filter out noise that affects the structure characteristics of Chinese characters [1]. This method uses the nearest neighbor pixel correlation function to filter out noise that affects the structure characteristics of Chinese characters. Moreover, according to the structure characteristics of Chinese characters, Shin et al. proposed a method for generating handwritten Chinese characters on the basis of stroke

M. Zhang et al. (Eds.): NLPCC 2018, LNAI 11109, pp. 154–165, 2018.
https://doi.org/10.1007/978-3-319-99501-4_13

correspondence [2]. Liu et al. presented a method for acquiring knowledge related to the basic elements required for intelligent generation of Chinese characters [3]. Tan et al. analyzed the structure characteristics of Chinese characters and presented a method of radical extraction using affine sparse matrix factorization for printed Chinese character recognition [4]. Dobres et al. discussed the influence of factors, such as font design difference, fine stroke difference, and color contrast polarity difference, on the scanning time of readable subtitles by analyzing the structure characteristics of Chinese characters [5].

In the past few decades, studies on the structure characteristics of Tibetan and Korean characters had also explored. Ai et al. carried out a statistical analysis of the glyph structure of Tibetan characters [6]. Cai et al. classified the structure of Tibetan characters and developed a statistical system model of glyph characters by analyzing the structure characteristics of Tibetan [7]. Kwon used the division of the Korean character structure to achieve rough classification in the process of syllable matching to preprocess alphabet division in the Korean character matching process [8]. Xu et al. utilized the rules of character structure to implement post-processing of handwritten Korean text [9]. Meanwhile, Cui et al. calculated the contribution of the information provided by letters at different positions in the spatial structure to the structure classification of Korean characters, further explored the distribution structure of Korean text information [10].

Square Hmong characters are ideograms with a fixed structure, and they are commonly used in the daily life of Hmong people in the Wuling mountain area of China. Information processing for square Hmong characters has not been extensively studied. Mo et al. conducted a number of studies on computer coding, keyboard input, glyph generation, and font creation of square Hmong characters [11–13]. However, the structure characteristics of square Hmong characters have not been investigated so far.

To address this limitation, this paper presents a classification method for square Hmong characters on the basis of structural distance. The method is implemented based on a linear representation of square Hmong characters and provides a convenient means to analyze the statistical characteristics of the structure of square Hmong characters in practical documents.

The rest of this paper is organized as follows. Section 2 introduces square Hmong characters and their linear representation. Section 3 introduces the proposed classification method and decision algorithm for the structure of square Hmong characters. Section 4 presents an analysis of the statistical characteristics of the structure of square Hmong characters in practical documents. Section 5 provides the conclusions.

2 Square Hmong Characters and Their Linear Representation

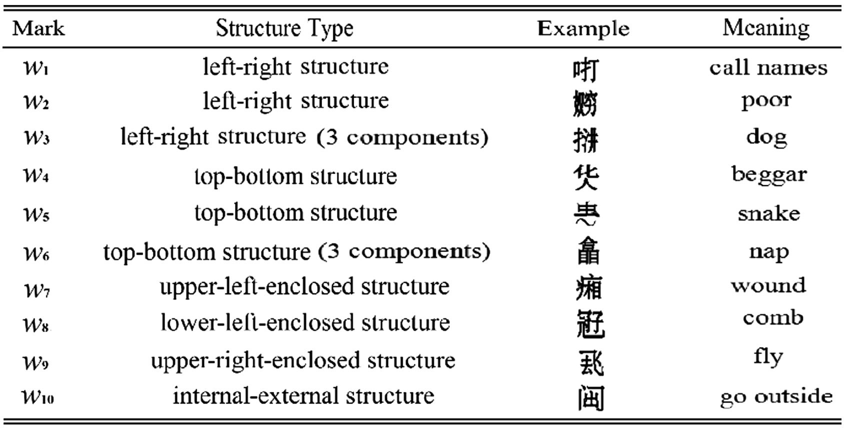
2.1 Word Information Principles of Square Hmong Characters

Square Hmong characters were created in the late Qing Dynasty for use by local Hmong people to record Hmong songs. According to [14, 15], the word information principles of square Hmong characters can be summarized from two aspects.

(1) Several simple Chinese characters, Chinese radicals, and symbols without pronunciation and meaning (such as ~ and X) are used to represent the component of phonetic-symbol, meaning-symbol, or shape-symbol.
(2) The "one word and one syllable" method is utilized to mark a morpheme or word.

The structure of square Hmong characters can be divided into four types, namely, left-right, top-bottom, part-enclosed, and internal-external structures. The part-enclosed structure can be further classified into upper-left-enclosed, lower-left-enclosed, and upper-right-enclosed types.

Several typical examples of square Hmong characters with different structures are listed in Fig. 1.



Mark	Structure Type	Example	Meaning
w_1	left-right structure	[illegible]	call names
w_2	left-right structure	[illegible]	poor
w_3	left-right structure (3 components)	[illegible]	dog
w_4	top-bottom structure	[illegible]	beggar
w_5	top-bottom structure	[illegible]	snake
w_6	top-bottom structure (3 components)	[illegible]	nap
w_7	upper-left-enclosed structure	[illegible]	wound
w_8	lower-left-enclosed structure	[illegible]	comb
w_9	upper-right-enclosed structure	[illegible]	fly
w_{10}	internal-external structure	[illegible]	go outside

Fig. 1. Examples of square Hmong characters in different structures.

According to the word information principle, if a square Hmong character consists of three or more components, two or three of these components can be combined into a simple Chinese character. This Chinese character is regarded as a component of the square Hmong character. A statistical analysis of the 1,129 square Hmong characters indicated that most Hmong characters are composed of two components, and only a few Hmong characters with left-right and top-bottom structures are composed of three components.

2.2 Linearization Representation of Square Hmong Characters

Definition 1 Linearization of square Hmong characters: The process of decomposing a square Hmong character into a uniquely identified component sequence according to its spelling order is called linearization.

To unify the description, we decompose all square Hmong characters into component sequences with a uniform length. When a component of a position in the sequence does not exist, the symbol "ε" is used to replace the missing component. According to the 16 types of components shown in Table 1, the uniform length of a component sequence can be 16 when a square Hmong character is linearized.

Table 1. Components of square Hmong characters

Number	Name	Mark
1	left component	C_{l}
2	right component	C_{r}
3	top component	C_{t}
4	bottom component	C_{b}
5	upper-left-external component	C_{olu}
6	lower-right-internal component	C_{ird}
7	lower-left-external component	C_{old}
8	upper-right-internal component	C_{iru}
9	upper-right-external component	C_{oru}
10	lower-left-internal component	C_{ild}
11	external component	C_{oa}
12	internal component	C_{ia}
13	right-left component	C_{rl}
14	right-right component	C_{rr}
15	bottom-left component	C_{bl}
16	bottom-right component	c_{br}

Considering that a two-component square Hmong character can be regarded as a three-component square Hmong character lacking a component, the right component (C_{r}) can be regarded as a right-left component (C_{rl}) when the right-right component (C_{rr}) is absent. The bottom component (C_{b}) can also be regarded as a bottom-left component (C_{bl}) when the right-bottom-right component (C_{br}) is absent. Therefore, only 14 types of components are required. That is, any square Hmong character can be decomposed into a 14-component combination sequence.

2.3 Linearization Function and Component Extraction Function

Definition 2 Linearization function of square Hmong characters: Assuming that Σ is a finite set of square Hmong characters, w is a square Hmong character ($w \in \Sigma$), and C_i (i = 1, 2, …, 14) is a finite set of components, the mapping (f) in Eq. (1) is called the linearization function of square Hmong characters.

$$f : \Sigma \rightarrow C_{\mathrm{l}} \times C_{\mathrm{rl}} \times C_{\mathrm{rr}} \times C_{\mathrm{t}} \times C_{\mathrm{bl}} \times C_{\mathrm{br}} \times C_{\mathrm{olu}} \times C_{\mathrm{ird}} \times C_{\mathrm{old}} \times C_{\mathrm{iru}} \times C_{\mathrm{oru}} \times C_{\mathrm{ild}} \times C_{\mathrm{oa}} \times C_{\mathrm{ia}} \quad (1)$$

A component combination sequence is practically a regular expression. Thus, f is a regular replacement from Σ to C_i that converts a square Hmong character into a uniquely identified component sequence. Assume that the components included in w are marked as s_{i} (i = 1, …, 14) and arranged in the order of "C_{l}-C_{rl}-C_{rr}-C_{t}-C_{bl}-C_{br}-C_{olu}-C_{ird}-C_{old}-C_{iru}-C_{oru}-C_{ild}-C_{oa}-C_{ia}." Then, $f(w)$ can be expressed by

$$f(w) = s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8 s_9 s_{10} s_{11} s_{12} s_{13} s_{14} \quad (2)$$

Considering that a square Hmong character contains only two or three components, valid components can only be found in two or three positions. Other locations are empty and marked as "ε". Figure 2 shows the linearization of the 10 characters in Fig. 1.

Square Hmong character	Linearization function	Components sequence													
		s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}
w_1	$f(w_1)$	布	兑	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε
w_2	$f(w_2)$	身	左	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε
w_3	$f(w_3)$	口	目	墨	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε
w_4	$f(w_4)$	ε	ε	ε	知	面	ε	ε	ε	ε	ε	ε	ε	ε	ε
w_5	$f(w_5)$	ε	ε	ε	化	人	ε	ε	ε	ε	ε	ε	ε	ε	ε
w_6	$f(w_6)$	ε	ε	ε	合	目	目	ε	ε	ε	ε	ε	ε	ε	ε
w_7	$f(w_7)$	ε	ε	ε	ε	ε	ε	厂	朗	ε	ε	ε	ε	ε	ε
w_8	$f(w_8)$	ε	ε	ε	ε	ε	ε	ε	ε	色	白	ε	ε	ε	ε
w_9	$f(w_9)$	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	飞	去	ε	ε
w_{10}	$f(w_{10})$	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	ε	门	出

Fig. 2. Examples of linear representation of square Hmong characters.

Definition 3 Component extraction function of square Hmong characters: The function of extracting the i-th component from linearization result $f(w)$ is called the extraction function of component s_i. It is denoted as $f_i(w)$ and expressed as

$$f_i(w) = s_i(i = 1, 2, \ldots, 14) \tag{3}$$

With the component extraction function, the linearization result $f(w)$ of w can also be expressed by

$$f(w) = f_1(w)f_2(w)f_3(w)f_4(w)f_5(w)f_6(w)f_7(w)f_8(w)f_9(w)f_{10}(w)f_{11}(w)f_{12}(w)f_{13}(w)f_{14}(w) \tag{4}$$

3 Equivalence Class of the Structure of Square Hmong Characters and Its Judgment

3.1 Definition of Structural Distance

Structural distance is used to measure structural differences between different characters. The structural distance of square Hmong characters is defined as follows.

Definition 4 Structural distance of square Hmong characters: Given square Hmong characters w_1 and w_2, the structural distance between w_1 and w_2 is expressed as $Distance(w_1, w_2)$. The calculation formula of $Distance(w_1, w_2)$ is shown in Eq. (5).

$$Distance(w_1, w_2) = \sum_{i=1}^{14} f_i(w_1)\theta f_i(w_2) \tag{5}$$

Where, operation θ is defined as

$$x\theta y = \begin{cases} 1, & \text{for } \forall(x,y) \in \{(x,y)|(x = \varepsilon \text{ and } y \neq \varepsilon) \text{ or } (x \neq \varepsilon \textit{ and } y = \varepsilon)\} \\ 0, & \text{for } \forall(x,y) \in \{(x,y)|(x \neq \varepsilon \text{ and } y \neq \varepsilon) \text{ or } (x = \varepsilon \textit{ and } y = \varepsilon)\} \end{cases} \tag{6}$$

Equations (5) and (6) show that the distance between two characters with the same structure is 0, and the distance between two characters with a different structure is a positive integer. For example, *Distance*(w_1, w_2) = 0, *Distance*(w_1, w_3) = 1, *Distance* (w_1, w_4) = 4, *Distance*(w_1, w_6) = 5, *Distance*(w_1, w_7) = 4, *Distance*(w_3, w_6) = 6, *Distance*(w_3, w_7) = 5, and *Distance*(w_4, w_5) = 0.

3.2 Structure Equivalence Class of Square Hmong Characters and Its Division

Definition 5 Equivalent structure relation: Given square Hmong characters w_1 and w_2, when *Distance*(w_1, w_2) = 0, w_1 and w_2 have an equivalent structure relation.

Definition 6 Structure equivalence class of square Hmong characters: As shown in Eq. (7), a given Hmong document (*D*) is divided into *m* disjoint subsets D_i (*i* = 1, 2, …, *m*) according to the equivalence structure relation.

$$\begin{cases} D = D_1 \cup D_2 \cup \ldots \cup D_m \\ D_i \cap D_j = \phi, \text{ for all } i \neq j \ (i,j = 1,2,\ldots,m) \end{cases} \tag{7}$$

If it exists,

$$Distance(w_1, w_2)\Bigg| \begin{matrix} = & 0, & \text{for } \forall w_1, w_2 \in D_i \\ > & 0, & \text{for } \forall w_1 \in D_i, \forall w_2 \in D_j \ (i \neq j;\ i,j = 1,2,\ldots,m) \end{matrix} \tag{8}$$

Then, D_i (*i* = 1, 2, …, *m*) is the *i*-th structure equivalence class. Square Hmong characters belonging to the same set D_i have the same structure. Square Hmong characters belonging to different D_i have different structures.

Eight equivalence classes of the structure of square Hmong characters can be calculated by using Eq. (8). The structure description of each equivalence class is shown in Table 2.

Table 2. Equivalence classes of the structure of square Hmong characters and structure description

Class number	Structure description	Linear representation	Structure type
Class 1	C_l-C_r	s_1s_2	left-right structure with two components
Class 2	C_l-C_{rl}-C_{rr}	$s_1s_2s_3$	left-right structure with three components
Class 3	C_u-C_d	s_4s_5	top-bottom structure with two components
Class 4	C_u-C_{dl}-C_{dr}	$s_4s_5s_6$	top-bottom structure with three components
Class 5	C_{olu}-C_{ird}	s_7s_8	part-enclosed structure with two components
Class 6	C_{old}-C_{iru}	s_9s_{10}	part-enclosed structure with two components
Class 7	C_{oru}-C_{ild}	$s_{11}s_{12}$	part-enclosed structure with two components
Class 8	C_{oa}-C_{ia}	$s_{13}s_{14}$	part-enclosed structure with two components

3.3 Judgment of Structure Equivalence Class of Square Hmong Characters

The position of each component in the spatial structure of a square Hmong character can be regarded as an attribute of the character. According to the linear representation of a square Hmong character, if $A(i)$ is used to represent the i-th attribute, each character has 14 attributes $A(1)$, $A(2)$, …, $A(14)$. This attribute value only indicates whether a certain type of component exists in a character, and its range is {0, 1}. "1" means that such a component exists, and "0" means that such component does not exist.

The process of determining the structure equivalence class of a square Hmong character is divided into two stages. In the first stage, the value of each attribute $A(i)$ (i = 1, 2, …, 14) is calculated. In the second stage, the structure equivalence class of this character is determined according to the combination of attribute values. The decision process is described in Algorithm 1.

Algorithm 1. Structure equivalence class decision algorithm
Input: Given square Hmong character *w*
Output: Number of structure equivalence class T_w
Steps:

Step 1: Linearize *w* to a uniquely identified component sequence *f*(*w*) using Equation (2).

Step 2: Extract the *i*-th component $f_i(w)$ (i=1,2,…,14) from *f*(*w*) according to Equation (3) and check whether $f_i(w)$ is "ε". If so, set $A(i)$=0; otherwise, set $A(i)$=1.

Step 3: Check the values of each attribute $A(i)$ (i=1,2,…,14) of *w* and determine T_w according to the combination of $A(i)$ values.

(1) If $A(i)$=1 (i=1,2) and $A(j)$=0 (j=3,5,…,14), then T_w=Class 1.

(2) If $A(i)$=1 (i=1,2, 3) and $A(j)$=0 (j=4,5,…,14), then T_w=Class 2.

(3) If $A(i)$=0 (i=1,2, 3), $A(j)$=1(j=4, 5), and $A(k)$=0 (k=6,7,…,14), then T_w=Class 3.

(4) If $A(i)$=0 (i=1,2,3), $A(j)$=1 (j=4,5,6), and $A(k)$=0 (k=8,9,…,14), then T_w=Class 4.

(5) If $A(i)$=0 (i=1,2,…,6), $A(j)$=1 (j=7,8), and $A(k)$=0 (k=9,…,14), then T_w=Class 5.

(6) If $A(i)$=0 (i=1,2,…,8), $A(j)$=1 (j=9,10), and $A(k)$=0 (k=11,…,14), then T_w=Class 6.

(7) If $A(i)$=0 (i=1,2,…,10), $A(j)$=1 (j=11,12), and $A(k)$=0 (k=13,14), then T_w=Class 7.

(8) If $A(i)$=1 (i=1,2, …,12) and $A(j)$=1 (j=13,14), then T_w=Class 8.

(9) In other cases, *w* is not a valid square Hmong character.

4 Analysis of the Statistical Characteristics of the Structure of Square Hmong Characters

4.1 Cumulative Probability Distribution of the Structure of Square Hmong Characters

The occurrence probability of the *i*-th equivalence class of the structure of square Hmong characters in a given document *D* can be calculated by

$$P_i = \frac{|D_i|}{|D|}\,(i = 1, 2, \ldots, m) \tag{9}$$

Where, $|D|$ and $|D_i|$ indicate the number of characters in documents D and D_i, respectively, and m is the number of equivalence classes.

Document D_1 consists of all square Hmong characters that have been collected, and documents D_2, D_3, and D_4 are composed of square Hmong characters appearing in [14], [15], and [16], respectively. The probability distribution of the equivalence class of the structure of square Hmong characters in each of the four documents (D_1, D_2, D_3, and D_4) is shown in Fig. 3.

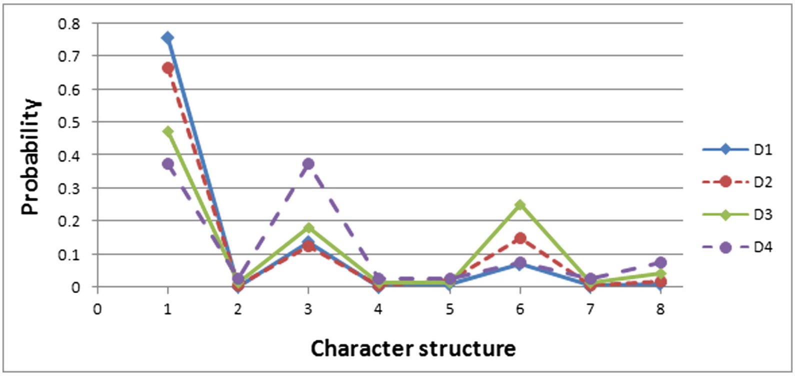


Fig. 3. Probability distribution of the equivalence class of the structure of square Hmong characters

The following points are determined from Fig. 3.

(1) Class 1 (C_l-C_r), Class 3 (C_u-C_d), and Class 6 (C_{old}-C_{iru}) have high probability characteristics. In documents D_1, D_2, D_3, and D_4, the cumulative probabilities of these three classes are as high as 0.9672, 0.9405, 0.9028, and 0.8250, respectively. The high probability of Class 1 (C_l-C_r) is particularly significant.

(2) Class 2 (C_l-C_{rl}-C_{rr}), Class 4 (C_u-C_{dl}-C_{dr}), Class 5 (C_{olu}-C_{ird}), Class 7 (C_{oru}-C_{ild}), and Class 8 (C_{oa}-C_{ia}) have a significantly low probability characteristic. In documents D_1, D_2, D_3, and D_4, the cumulative probabilities of these five classes are as low as 0.0328, 0.0595, 0.0972, and 0.1750, respectively.

(3) In documents D_1, D_2, D_3, and D_4, the cumulative probabilities of the occurrence of Class 2 (C_l-C_{rl}-C_{rr}) and Class 4 (C_u-C_{dl}-C_{dr}) are extremely low at only 0.0053, 0.0119, 0.0278, and 0.050, respectively.

Eight equivalence classes of the structure of square Hmong characters in the practical documents are not evenly distributed. More than 80% of the characters appearing in the practical documents belong to Class 1, Class 3, and Class 6, and more than 90% of them consist of two components.

4.2 Complexity of the Structure of Square Hmong Characters

The results on the cumulative probability distribution of the structure of square Hmong characters reflect the simplicity of the structure. This simplicity can be verified by calculating the complexity of the structure of square Hmong characters. Complexity can be expressed by the average number of components in each square Hmong character appearing in the practical documents. *Length* is calculated by

$$Length = \sum_{i=1}^{m} P_i l_i \qquad (10)$$

Where, P_i is the probability of occurrence of various structures derived from Eq. (9) and l_i is the number of components in the i-th equivalence class.

In documents D_1, D_2, D_3, and D_4, the *Length* values calculated with Eq. (10) are 2.0053, 2.0119, 2.0278, and 2.0500, respectively. This result fully confirms the simplicity of the structure of square Hmong characters.

4.3 Information Entropy of the Structure of Square Hmong Characters

Information entropy is used to describe the distribution characteristics of square Hmong characters with different structures in the same document. A large entropy value equates to a large number of structure equivalence classes included in the document and to a uniform distribution of square Hmong characters with different structures. On the contrary, a small entropy value equates to a few structure equivalence classes included in the document and to a concentrated distribution of square Hmong characters in a few classes.

Given document D, the information entropy of the structure of square Hmong characters is calculated as

$$Entropy(D) = -\sum_{i=1}^{m} P_i log_2 P_i \tag{11}$$

Where, P_i and m are similar to those in Eq. (9).

Figure 4 presents the information entropy of the structure of square Hmong characters in documents D_1, D_2, D_3, and D_4. As shown in Fig. 4, the structure of square Hmong characters in the four documents reveals a first-order approximation of the information. The entropy is small, falling in the interval (1.19, 2.16). The results also show that the probability of occurrence of various structures appearing in the practical documents is extremely uneven.

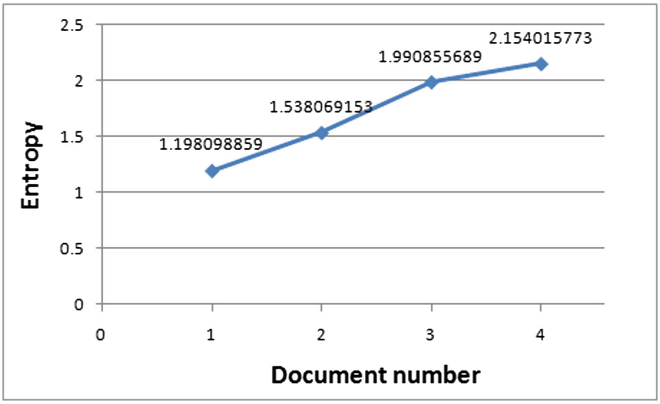


Fig. 4. Information entropy of the structure of square Hmong characters in the practical documents

5 Conclusion and Future Work

In this work, we study the linear representation of square Hmong characters and the classification method for the structure of square Hmong characters. We also analyze the statistical properties of the structure of square Hmong characters appearing in practical documents. The analysis results can be used as a basis for the design of a heuristic rule in the post-processing of an intelligent input and recognition system for square Hmong characters. The study provides a good foundation for the further investigation of the application of the structure characteristics of such characters.

Notably, this study focuses on character structure and not on specific characters. To comprehensively understand the application characteristics of square Hmong characters in different areas, a variety of practical documents with different themes, styles, and nature must be collected. An in-depth analysis of the differences in the use of specific characters should be conducted, and the effectiveness of specific characters should be evaluated.

Acknowledgement. This work is supported by the National Natural Science Foundation of China (Nos. 61462029 and 61741205).

References

1. Instituse, R.: A fast smoothing & thinning method based on character structure. J. Chin. Inf. Process. **4**(2), 49–55 (1990)
2. Shin, J., Suzuki, K., Hasegawa, A.: Handwritten Chinese character font generation based on stroke correspondence. Int. J. Comput. Process. Orient. Lang. **18**(3), 211–226 (2005)
3. Liu, M.Y., Duan, C.S., Pi, Y.G.: Basic elements knowledge acquisition study in the Chinese character intelligent formation system. J. Softw. Eng. Appl. **2**(5), 316–322 (2009)
4. Tan, J., Xie, X.H., Zheng, W.H., et al.: Radical extraction using affine sparse matrix factorization for printed Chinese characters recognition. Int. J. Pattern Recognit. Artif Intell. **26**(3), 211–226 (2012)
5. Dobres, J., Chahine, N., Reimer, B., et al.: The effects of Chinese typeface design, stroke weight, and contrast polarity on glance based legibility. Displays **41**, 42–49 (2016)
6. Ai, J.Y., Yu, H.Z., Li, Y.H.: Statistical analysis on Tibetan shaped structure. J. Comput. Appl. **29**(7), 2029–2031 (2009)
7. Cai, Z.J., CaiRang, Z.M.: Research on the distribution of Tibetan character forms. J. Chin. Inf. Process. **30**(4), 98–105 (2016)
8. Kwon, Y.B.: Hangul tree classifier for type clustering using horizontal and vertical strokes. In: Proceedings of the 16th International Conference on Pattern Recognition, pp. 228–231. IEEE, Quebec City (2002)
9. Xu, R.J., Liu, C.P.: Grapheme segmentation and recognition in machine printed Hangul characters. J. Chin. Inf. Process. **20**(2), 66–71 (2006)
10. Cui, R.Y., Kim, S.J.: Research on information structure of Korean characters. J. Chin. Inf. Process. **25**(5), 114–119 (2011)
11. Mo, L.P., Zhou, K.Q.: Formal description of dynamic construction method for square Hmong language characters. J. Comput. Appl. **34**(3), 861–864, 868 (2014)
12. Mo, L.P., Zhou, K.Q., Jiang, X.H.: Research on square Hmong language characters fonts based on OpenType technology. J. Chin. Inf. Process. **129**(2), 150–156 (2015)

13. Mo, L.P., Zhou, K.Q.: A dynamical glyph generation method of Xiangxi Folk Hmong characters and its implementation approach. Acta Scicentiarum Naturalum Universitis Pekinesis **52**(1), 141–147 (2016)
14. Zhao, L.M., Liu, Z.Q.: Xiangxi square Hmong characters. Minor. Lang. China **12**(1), 44–49 (1990)
15. Yang, Z.B., Luo, H.Y.: On the folk coinage of characters of the Miao People in Xiangxi area. J. Jishou Univ. (Soc. Sci. Edn.) **29**(6), 130–134 (2008)
16. Long, Z.H.: Re-study of the coinage method of square characters of Miao language in the Youshui river basin of Yu, Xiang and E. J. Chongqing Educ. Coll. **25**(5), 56–59 (2012)