

Chapter 5

Qualitative Representation of Online Handwritten Assamese Characters

This chapter introduces a Qualitative Representation of Online Handwritten Assamese Characters. The representation is based on qualitative representation of planar outline. The method introduced in this chapter is based on two Qualitative schemes, namely Qualitative Orientation and Qualitative Curvature Components. Qualitative Orientation captures the change of direction of strokes leading to a Qualitative Encoding of strokes; and Qualitative Curvature Components describe curvature components of strokes leading to a Curvature Component Signature for each character.

5.1 Motivation

The characters in Assamese are of various shapes. The shapes of some characters are simple and others have complex shapes in Assamese alphabet. Generally Assamese numerals have the simplest shapes and basic alphabetic characters are of relatively simple shapes. Majority of Assamese conjunct consonants or Juktakkhors have complex shapes. It is clear from the recognition rates of online handwritten Assamese basic alphabetic characters and Juktakkhors presented in Chapter 4 that the recognition rates of online handwritten Assamese basic alphabetic characters and Juktakkhors based on quantitative description are not very encouraging. Moreover, in case of two handwritten instances of the same character with small difference in writing pattern quantitative description may not be able to recognize them to be in the same class due to the differences in writing

pattern of the character. However, shapes can be described in both quantitative and qualitative ways. The qualitative description of the shape can be obtained as a string of symbols which represent the string of shape primitives, describing the complete shape. Symbolic schemes for describing shapes are considered preferable to existing quantitative techniques due to the advantage of qualitative representation in dealing with abstract shape outlines. We explore qualitative description for online handwritten Assamese characters with an aim to improve the recognition rates of basic alphabetic characters and Juktakkhors.

5.2 Stroke and Qualitative Orientation

Online handwritten characters can be seen as a series of (x, y) positions on a planar surface. Online handwritten characters are captured as a function of time. A stroke of an online handwritten character is a sequence of points captured between consecutive pen-down and pen-up events. An online handwritten character consists of one or more strokes, which are captured as a function of time. Accordingly, strokes are ordered with respect to time and together represent the complete character. The constituent strokes of an online handwritten character are of different shapes. Qualitative representation of the change of direction of the stroke for a series of (x, y) positions is possible. Qualitative Change of direction of a time series can be modeled through a Qualitative Orientation Calculus. A model of Qualitative Orientation Calculus QDA_8 is discussed in Chapter 2. Here, we shall discuss an adaption of the same for Qualitative Encoding of strokes.

5.2.1 Qualitative Orientation Model

Figure 5.1 illustrates the qualitative orientation model. The 360 degree angle is divided into eight equally spaced angles of 45 degree each. The angles (positive values) are measured in anticlockwise direction from a reference line. A unique

symbol is assigned to the corresponding angle range as per the scheme¹ shown in Table 5.1.

Table 5.1: Qualitative Orientation angle range and symbols

Direction	Symbol	Angle Range	Direction	Symbol	Angle Range
1	P	[0, 0]	7	G	[180, 180]
2	Q]0, 45]	8	V]180, 225]
3	R]45, 90[9	W]225, 270[
4	S	[90, 90]	10	X	[270, 270]
5	T]90, 135]	11	Y]270, 315]
6	U]135, 180[12	Z]315, 360[

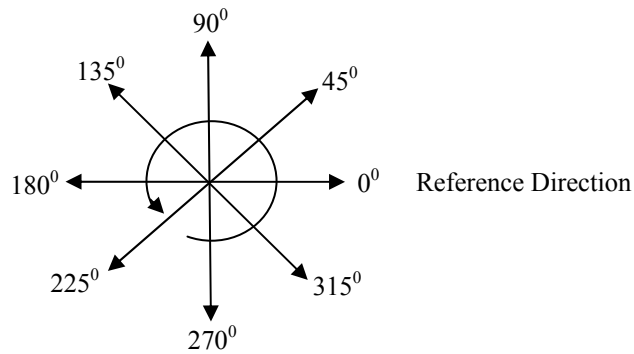


Figure 5.1: Qualitative orientation model

5.2.2 Direction of Strokes: Series of Orientation Symbols

This section introduces the method for the assignment of a series of qualitative symbols to a stroke based on the Qualitative Orientation model discussed above. For each constituent stroke of a character, the set of points are parsed with three consecutive points at a time. The orientation angle is computed at the middle

¹ [x,y] includes both the bounding values x and y,]x,y] includes y but not x and]x,y[doesn't include the bounding values x and y.

point and it is symbolized as per the scheme shown in Table 5.1. A series of qualitative symbols is obtained for each stroke.

5.2.3 Qualitative Encoding

The qualitative representation of stroke through a series of orientation symbols is termed as qualitative encoding. Qualitative encoding of a complete character is composed of a chain or chains of such qualitative symbols, for a single stroke character or multiple stroke character respectively. An example of qualitative encoding of stroke is shown in the Figure 5.2. The stroke is sampled at the points $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}$ and P_{11} . For three consecutive points P_{i-1}, P_i and P_{i+1} in the stroke, the orientation angle is computed at the point P_i and the corresponding symbol is assigned to it based on the angle range shown in the Table 5.1. In this example, the chain of qualitative orientation symbols for the stroke can be approximately written as TTTTUVWWW.

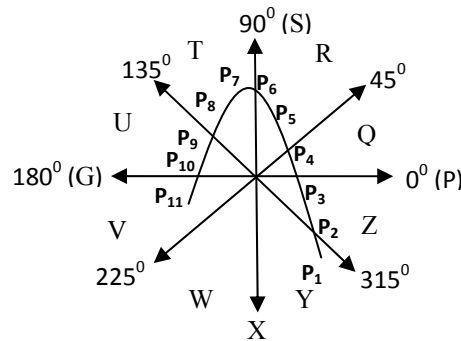
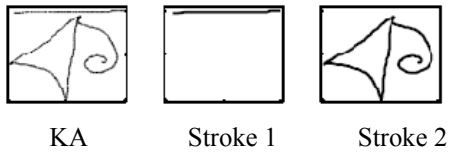


Figure 5.2: An example stroke: TTTTUVWWW

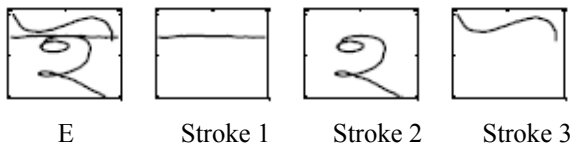
The qualitative encoding of strokes for few characters is shown in Figure 5.3 through Figure 5.7. Figure 5.3 shows the character KA which consists of two constituent strokes. The combination of encoded stroke 1 and encoded stroke 2 (stroke 1 followed by stroke 2) represent the qualitative encoding of the complete character KA.



Encoded Stroke 1:
 VUUUVUVUVUVUVUVUUU
 Encoded Stroke 2:
 VZYYYYZZZZZZZUUVVVVVVVVWWWWWTTTTTTTTTTTTTTTTTTTTTYVVVVVV
 WWWWWYYYZZZQQRTUUUVW

Figure 5.3: The Character KA and its Encoded Strokes

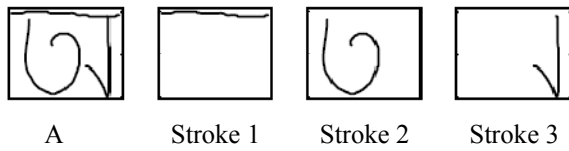
Figure 5.4 shows the character E which consists of three constituent strokes. The combination of encoded stroke 1, encoded stroke 2 and encoded stroke 3 (stroke 1 is followed by stroke 2 and stroke 2 is followed by stroke 3) represent the qualitative encoding of the complete character E.



Encoded Stroke 1:
 WUVUVUVUVUUUVUUUVUV
 Encoded Stroke 2:
 WQQRUUUVVVVVWYYYYYYYYZZZZZZZZZQTUUUVVVVVVVVVUUUUUU
 Encoded Stroke 3:
 VRTRRQQZZZZZZZQQQQQR

Figure 5.4: The Character E and its Encoded Strokes

Figure 5.5 shows the character A which consists of three constituent strokes. The concatenation of encoded stroke 1, encoded stroke 2 and encoded stroke 3 (stroke 1 is followed by stroke 2 and stroke 2 is followed by stroke 3) represent the qualitative encoding of the complete character A.



Encoded Stroke 1:

VUUUVUVVVUVVUVVVVVVVUVUUVU

Encoded Stroke 2:

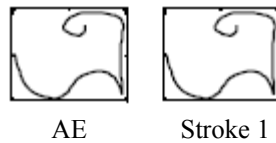
VTUUUVVWWWWWWYYYYYYYYZZZQQQRRRRRTTTTTT

Encoded Stroke 3:

VVWWWWWWWTTRTTTRRRTRT

Figure 5.5: The Character A and its Encoded Strokes

Figure 5.6 shows the character AE which consists of only one stroke. The encoded stroke 1 represents the qualitative encoding of the complete character AE.

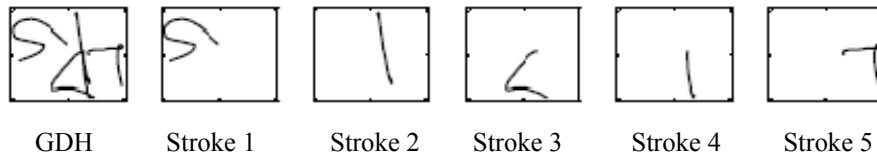


Encoded Stroke 1:

VYYZZQQQRTTTTTUUUUUVVVVVVVWWWWYYYYYYYWYWYWYWWWWWYZ
RRRRRRQQQZZZZZZYYYYYYZZZZZQQQQQRRRRRRRRRRR

Figure 5.6: The Character AE and its Encoded Stroke

Figure 5.7 shows the character GDH which consists of five constituent strokes. The combination of encoded stroke 1, encoded stroke 2, encoded stroke 3, encoded stroke 4 and encoded stroke 5 (in the order from stroke 1 to stroke 5) represent the qualitative encoding of the complete character GDH. The character in Figure 5.7 is a conjunct consonant or Juktakkhor. A Juktakkhor has typically more number of constituent strokes as compared to basic alphabetic characters which are presented in Figure 5.3 through Figure 5.6.



Encoded Stroke 1:
WUUUUTTRQQQQQRRRTTTUUUVVVWWWW

Encoded Stroke 2:
UYWWWWWWWWWWRT

Encoded Stroke 3:
WZYYYYYYYWVUUUVVVVVVW

Encoded Stroke 4:
WYWYWWWWW

Encoded Stroke 5:
VTUUUUURQYYWYW

Figure 5.7: The Character GDH and its Encoded Strokes

5.3 Qualitative Analysis of Strokes

An analysis on strokes of online handwritten Assamese characters was carried out. All the 8235 (=183 Characters×45 Writers) samples of online handwritten Assamese characters available in the dataset were considered. Since an online handwritten character is a sequence of strokes, the number and order of strokes in the sequence making up the character varies for different writers. A single character is normally written by different writers in different ways. Some characters have single stroke for some writers and have multiple strokes for others in the corresponding stroke sequence. Examples of stroke sequences for three characters; one for each character, are shown in the Table 5.2. Examples of multiple stroke sequences for a single character considering four stroke sequences are shown in the Table 5.3. We have analyzed each online handwritten Assamese character using its plots of the stroke sequences. Exhaustive inspection of the plots of stroke sequences of each character is performed and from this analysis we have identified the general pattern in which a particular Assamese character is

written by most of the writers. This general pattern is in the context of number of constituent strokes and the order in which the constituent strokes are written. Lists of all stroke sequences for two sample characters are included in Appendix C.

Table 5.2: Characters with corresponding stroke sequences

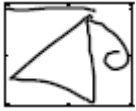



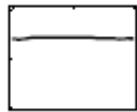

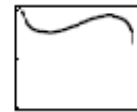
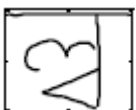

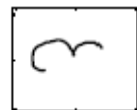

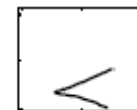
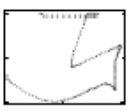
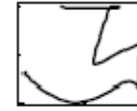
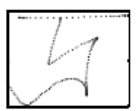
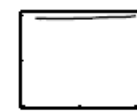

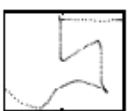




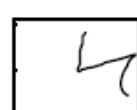

Character	Stroke Sequence				
					
					
					

Table 5.3: Examples of multiple stroke sequences of the same character

Character	Stroke Sequence			
				
				
				
				

5.3.1 Curvature Components for the character set

From the stroke analysis of all 8235 online handwritten Assamese characters we have identified a set of qualitative curvature components using which any stroke in the dataset can be represented. The set of curvature components for the character set is listed in Table 5.4.














5.3.2 Most frequently occurring stroke sequence for each character

It was observed from the stroke sequences explained in section 5.3.1, that a similarity of writing patterns of the same character exists among at least 50% of writers. This similarity is in terms of number and order of strokes in the stroke sequence making up the character. This similar or identical sequence of strokes is the most frequently occurring stroke sequence of the character. The most frequently occurring stroke sequence of a character is chosen and this sequence is considered as the representation of the character. Figures of all the 183 individual characters with their representative stroke sequences are included in Appendix D.

5.4 Qualitative Curvature Components Signature

A character consists of one or more number of strokes. A stroke can be broken into one or more components based on qualitative curvature. Using distinct symbols for each of the qualitative curvature component, a sequence can be obtained. The sequence of symbols can be termed as a qualitative stroke and will represent the online handwritten characters qualitatively. This scheme is explained in the following sub-section.

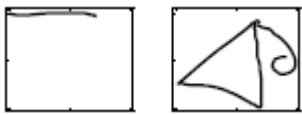

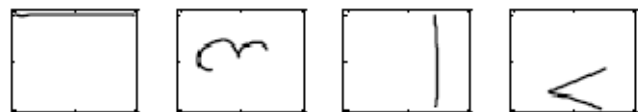
Table 5.4: Curvature Components with corresponding symbols

Curvature Component	Descriptive Name	Symbol
	Horizontal	H
	Vertical	I
	Spherical	O
	Convex	C ⁺
	Concave	C ⁻
	Concave Up	U ⁺
	Convex Down	U ⁻
	Downward Pointing Angle	V ⁺
	Upward Pointing Angle	V ⁻
	Left Pointing Angle	A ⁺
	Right Pointing Angle	A ⁻
	Curl Vertical	S ⁺
	Curl Horizontal	S ⁻

5.4.1 Stroke Sequence and Signature

The strokes in the representative sequence are composed of curvature components shown in the Table 5.4. An online handwritten character can be represented as a sequence of curvature components of its representative stroke sequence. Table 5.5 shows samples of representative stroke sequence and sequence of symbols of curvature components. The sequence of curvature components corresponding to a character can be termed as a qualitative curvature component signature of the character.

Table 5.5: Stroke sequence and signature

Stroke Sequence	Signature
	H A ⁺ I C ⁺ O
	H O U ⁻ A ⁺ S ⁻
	H U ⁻ U ⁻ I A ⁺

5.5 Conclusion

A Qualitative Representation of Planar Outlines is introduced for Online Handwritten Assamese Characters. Using the qualitative representation of online handwritten Assamese characters we put forward a Qualitative Feature vector for Online Handwriting Recognition of Assamese characters. The description of qualitative feature vectors for online handwritten Assamese characters and character recognition experiments based on qualitative feature vectors are presented in Chapter 6.