

Contents

1. Introduction	1
1.1 Online Handwriting Recognition	1
1.1.1 Issues and Challenges	2
1.2 Context and Motivation	3
1.2.1 OHR of Assamese Characters	5
1.3 Objectives	5
1.4 Thesis Outlines	6
2. Literature Review	7
2.1 Handwriting Recognition	7
2.1.1 Online and Offline Handwriting Recognition	8
2.1.2 Writer Dependent and Writer Independent Handwriting Recognition	9
2.1.3 Constrained and Unconstrained Handwriting Recognition	9
2.1.4 Variability in Writing Styles	10
2.2 Data Acquisition and Datasets	10
2.2.1 Data Acquisition	10
2.2.1.1 Text Input Devices	11
2.2.1.2 Data Acquisition Methodology	13
2.2.2 Datasets	14
2.2.2.1 Standard Datasets in Western Scripts	14
2.2.2.2 Standard Datasets in Indian Scripts	15

2.3	Preprocessing	16
2.3.1	Preprocessing Steps	16
2.4	Features of Online Handwritten Characters	18
2.4.1	Geometrical Features	18
2.4.1.1	Resampled Horizontal and Vertical Coordinates	18
2.4.1.2	Pen-down and Pen-up Positions	18
2.4.1.3	Writing Direction	19
2.4.1.4	Curvature Feature	19
2.4.1.5	Eight-directional feature	20
2.4.2	Structural or Topological Features	21
2.4.2.1	Junction Point	21
2.4.2.2	Loop	22
2.4.2.3	Cusp	22
2.4.2.4	Dot	23
2.4.2.5	Head Line	23
2.4.2.6	Vertical Line	24
2.4.3	Statistical Features	24
2.4.3.1	Normalized horizontal and vertical coordinates	24
2.4.3.2	Zone	25
2.4.3.3	Mean of horizontal and mean of vertical coordinates	25
2.4.3.4	Standard Deviation	25
2.4.3.5	Number of Strokes	25

2.4.3.6	Stroke length	26
2.4.3.7	Aspect Ratio	26
2.4.3.8	Derivative feature	26
2.4.3.9	Distance	27
2.5	Qualitative Representation of Planar Outlines	27
2.5.1	What is Qualitative Reasoning?	27
2.5.2	Qualitative Orientation	28
2.5.3	Qualitative Outlines	30
2.5.3.1	Codon Theory of Richard and Hoffman	31
2.5.3.2	Leyton's Process Grammar for Shape	32
2.5.3.3	Qualitative Outline Theory of Meathrel and Galton	34
2.5.4	Approaches to Curvature Components	35
2.5.4.1	Quantitative Consideration of Curvature Components	35
2.5.4.2	Qualitative Curvature Components	36
2.5.4.2.1	Tripartite Line Tracks	37
2.5.4.2.2	Primitive Curve Tokens	39
2.6	Online Handwriting Classification Methods	42
2.6.1	Deformable Template Matching of Elastic Matching	43
2.6.2	Statistical Approach	44
2.6.2.1	Support Vector Machine (SVM)	45
2.6.3	Syntactic or Structural Approach	47
2.6.4	Neural Networks	50

3	A Dataset of Online Handwritten Assamese Characters	53
3.1	Materials and Methods	53
3.1.1.	Subjects and Experimental Setup	53
3.1.2.	Experimental Protocol	54
3.2	Data Acquisition and Verification	55
3.2.1	Data Acquisition	56
3.2.2	Data Verification	57
3.3	TU-OHAC Dataset	57
3.3.1	Dataset Description	58
3.3.1.1	Data Attributes	61
3.3.2	Usage of the dataset	62
3.4	Challenges Encountered	63
3.4.1	Choice of Size of the Alphabet	63
3.4.2	User Friendly Data Acquisition GUI	64
3.4.3	Human Factors	64
3.5	Conclusion	64
4	Recognition of Online Handwritten Assamese Characters: Initial Results	66
4.1	Preprocessing	66
4.1.1	Size normalization	67
4.1.2	Smoothing	68
4.1.3	Resampling	69
4.2	Feature Set	70

4.2.1	Feature Computation	70
4.2.1.1	Geometrical Features	71
4.2.1.2	Structural Feature	71
4.2.1.3	Statistical features	71
4.2.2	Initial Feature Vectors	71
4.2.2.1	Feature Vector FS:1	72
4.2.2.2	Feature Vector FS:2	72
4.2.2.3	Feature Vector FS:3	73
4.3	Classification	73
4.3.1	SVM based Classifier	74
4.3.2	Classification Results	75
4.3.2.1	Classification Results: Numerals	75
4.3.2.2	Classification Results: Basic Alphabetic Characters	79
4.3.2.3	Classification Results: <i>Juktakkhors</i>	86
4.3.2.4	Overall Recognition Rates	99
5	Qualitative Representation of Online Handwritten Assamese Characters	108
5.1	Motivation	108
5.2	Stroke and Qualitative Orientation	109
5.2.1	Qualitative Orientation Model	109
5.2.2	Direction of Stroke: Series of Orientation Symbols	110
5.2.3	Qualitative Encoding	111
5.3	Qualitative Analysis of Strokes	114

5.3.1	Curvature Components for the character set	116
5.3.2	Most frequently occurring stroke sequence for each character	116
5.4	Qualitative Curvature Components Signature	116
5.4.1	Stroke Sequence and Signature	118
5.5	Conclusion	118
6	Recognition of Online Handwritten Assamese Characters: Using Qualitative Feature Vectors	119
6.1	Feature Set	119
6.1.1	Feature Computation	120
6.1.1.1	Qualitative Encoding	120
6.1.1.2	Qualitative Curvature Component Signature	121
6.1.2	Qualitative Feature Vectors	122
6.1.2.1	Qualitative Feature Vector QFS:1	122
6.1.2.2	Qualitative Feature Vector QFS:2	122
6.2	Classification	123
6.2.1	SVM based Classifier	123
6.2.2	Classification Results	123
6.2.2.1	Classification Results: Numerals	124
6.2.2.2	Classification Results: Basic Alphabetic Characters	126
6.2.2.3	Classification Results: <i>Juktakkhors</i>	129
6.2.2.4	Overall Recognition Rates	132
6.3	Conclusion	134

7 Conclusion and Future Directions	136
7.1 Contributions	136
7.1.1 A Dataset of Online Handwritten Assamese Characters	136
7.1.2 Qualitative Representation of Online Handwritten Assamese Characters	137
7.1.3 Online Handwriting Recognition of Assamese Characters	138
7.2 Future Directions	139

Bibliography

Appendix A

Appendix B

Appendix C

Appendix D