Abstract

Online handwriting recognition is a major research topic because of thrust towards natural human-computer interaction and emerging technologies of handheld computing devices. These devices are based on digitizing technology. Online Handwriting Recognition refers to the methods and techniques dealing with the automatic processing of a character as it is written using a digitizer. Methods and approaches for online handwriting recognition have been proposed and tested for Chinese, Japanese and Western Scripts. Research for online handwriting recognition has also emerged for few Indian scripts which include Tamil, Telugu, Bengali, Devanagari and Gurmukhi. Online Handwriting Recognition for Assamese - a major language of North Eastern India is in its nascent stages. This thesis is an experimental study and analysis of Online Handwriting Recognition for Assamese characters.

For character recognition, the availability of standard datasets is of paramount importance. Therefore, acquisition and distribution of datasets have increasingly gained importance. A contribution of this thesis is development of such a dataset for Online Handwritten Assamese characters. TU-OHAC Dataset - a dataset of online handwritten Assamese characters has been developed. TU-OHAC Dataset contains a total of 8,235 characters collected from 45 subjects. TU-OHAC Dataset contains Assamese numerals, basic alphabetic characters, and conjunct consonants (*Juktakkhors*).

For recognition of online handwritten Assamese characters the effectiveness of standard features reported in the literature (for other scripts) is explored. Three different combinations of features based on geometrical, structural and statistical features is used for classification. Overall recognition of 73.21% is observed for the best performing feature set. *Juktakkhors* and certain complex shaped basic

alphabetic characters exhibited very low recognition rates. In order to distinguish *Juktakkhors* and basic alphabetic characters involving complex shaped strokes, the thesis explores qualitative representation.

Qualitative schemes for describing shapes are considered preferable to quantitative techniques as qualitative representation more succinctly captures abstract shape outlines. The thesis develops a qualitative representation for online handwritten Assamese characters. Using a *Qualitative Orientation Model* qualitative direction of strokes for a character is proposed. Further, a qualitative analysis of 8,235 characters is done based on *curvature components*; and a qualitative curvature component signature for online handwritten Assamese character is evolved. Two different combinations of features based on above qualitative representation are used for classification. Overall recognition of 93.93% is observed for the best performing qualitative feature set.

Keywords- *Online Handwriting, Assamese, Character Recognition, Dataset Collection, Data Verification, Support Vector Machine, Qualitative Features*