A Writer Identification System for Handwritten Malayalam Documents

Handwriting is an acquired tool used for communication of one's observations or feelings. Factors that influence a person's handwriting not only dependent on the individual's bio-mechanical constraints, handwriting education received, writing instrument, type of paper, background, but also factors like stress, motivation and the purpose of the handwriting. Despite the high variation in a person's handwriting, recent results from different writer identification studies have shown that it possesses sufficient individual traits to be used as an identification method.

Handwriting as a behavioral biometric has had the interest of researchers for a long time. But recently it has been enjoying new interest due to an increased need and effort to deal with problems ranging from white-collar crime to terrorist threats. The identification of the writer based on a piece of handwriting is a challenging task for pattern recognition. The main objective of this thesis is to develop a text independent writer identification system for Malayalam Handwriting. The study also extends to developing a framework for online character recognition of Grantha script and Malayalam characters.

The writer identification system proposed in this thesis comprises of different phases like image preprocessing, feature extraction, training and classification or identification. The feature extraction phase includes three schemes. One is at the grapheme level, next Character level and third at the document level. The performance of the overall system is measured using statistical measurements. In order to analyse the system performance, experiments are carried out with different classifiers like Naive Byes, KNN, SVM and Adaboost. The comparison of results are based on the identification rate of classifiers, stability of features and classifiers. consistency measurements, influence of single features, and cumulative features. From each of these schemes elegant /decisive features that distinguish a writer were obtained.

A system that can recognize online handwritten Malayalam characters utilizing the optimum decisive features obtained from above schemes was developed. Further comparisons were made with the system developed and systems using other features and classifiers. Results showed that the system developed with the decisive features performed better.

The analysis and recognition of historical documents have attracted interest recent years. This may be due to digitization drive for preservation of documents that embody the artistic, cultural and technical heritage of a country. With this in mind the thesis proposes a system to recognize handwritten Grantha script arid obtain its Malayalam equivalent. This has significance because the root script of Malayalam is the Grantha script. The system adopts the same framework as that of Online Malayalam Character Recognition.