

Optical Character Recognition (OCR) software is now available for handwritten and printed text of many languages. Matrix matching and feature analysis is the basic technologies used.

In this study an attempt is made to introduce OCR techniques to identify Sinhala characters. Templates for Matrix matching and algorithms for searching the scanned document to recognize characters are given.

When the document contains Sinhala characters with size shape, font typeface (bold or normal), the matrix method becomes inefficient. In such a case feature analysis of Sinhala Characters are done. Here thee scanned document is fed into a computer programme and all the characters in it are separated out.

Afterwards the features (such as the number of closed regions inside) are obtained by running feature extraction programmes. These features are then compared with those of the standard set of Sinhala characters in store. This way characters can be identified even though they very in size shape or style.

The algorithms for extracting several features of characters are also given.