Evolution of Indic Scribal Handwriting

University of St Andrews

Vinodh Rajan
vrs3@st-andrews.ac.uk

Dr Mark-Jan Nederhof
mn31@st-andrews.ac.uk

sicsa*

Introduction

The myriad of modern Indian scripts that exist today were all derived from the same source script i.e. Brahmi. The initial shape of the Brahmi script was largely geometrical, but it has given rise to a wide variety of scripts over time due to inherent variations in human handwriting.

For any Indic script, we can derive an "almost" linear evolutionary line from Brahmi. Therefore, we have a unique opportunity to analyze script developments in terms of changes in handwriting behavior. We can investigate how the different handwriting features have evolved in terms of handwriting production and visual appearance.

Palaeographic Dataset

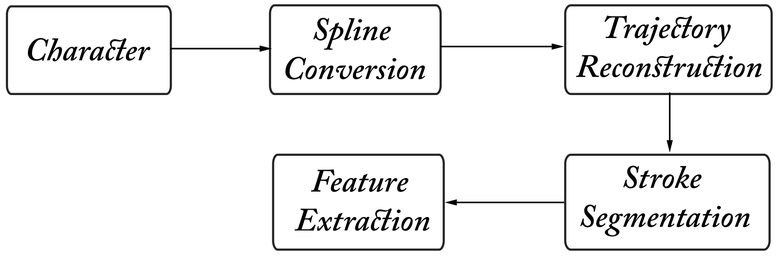
To obtain a comprehensive view of the script development process we have taken four major scripts belonging to the Brahmic family – Devanagari, Tamil, Kannada and Grantha. These scripts represent most of the important Brahmic scripts in India. We consider the scripts in six stages of evolution. A single stage of a script can be considered to represent ~300 years covering ~1800 years of development.

In total, we have 21 distinct scripts consisting of 730 distinct glyphs.

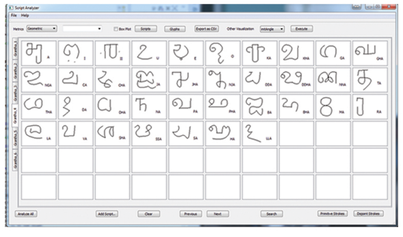
Evolution of Grantha-Tamil character VA

Data Extraction

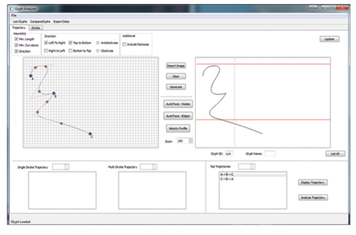
The scripts were digitized using a script analysis framework. Characters were first converted into splines, followed by reconstruction of their trajectories and then finally decomposed into their respective strokes. From this normalized stroke structure we extract two types of features, geometric features and production features, which were used for subsequent quantitative analyses. The geometric features consisted of 9 different features based on the static shape of characters and the production features consisted of 12 different features based on the written trajectories.



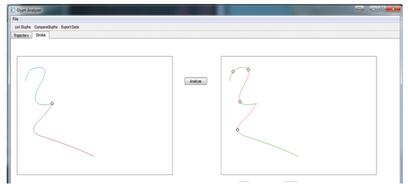
Script Analysis Framework



Script Repository

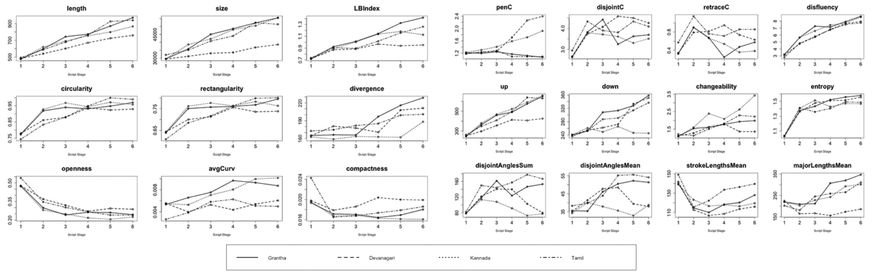


Digitized Character



Decomposed Character

Handwriting Trends

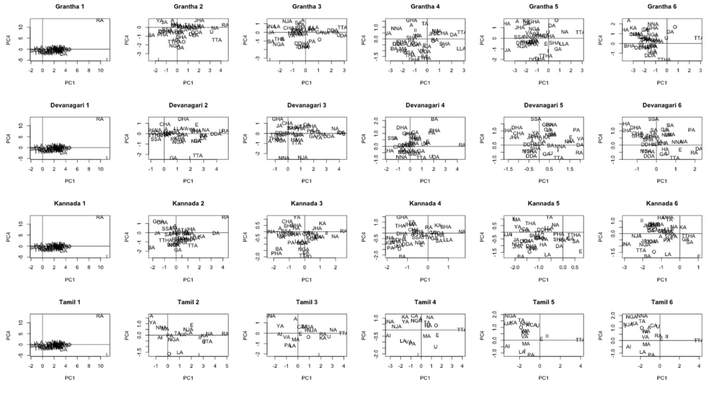


Trends in Geometric Features *Trends in Production Features*

We can see that many features show logarithmic growth. This shows that characters after an initial period of diversification began to stabilize slowly. Also one would expect that humans tend to reduce disfluency to increase writing speed but on a large scale it appears not to be the case. Writing appears to have gathered more disfluency, more disjoint strokes and an increase in entropy.

Spread of Variations

The original feature set consisting of 9+12 features is too large for individual character-wise analysis. Hence, we proceeded to perform Principal Component Analysis (PCA), which reduced the feature set and also resulted in descriptive aggregate features.



Plot of PC1 vs PC4 for scripts in all 6 stages of development.