
Binarization of Degraded Brahmi Script Estampage Images

Mr. Aniket S. Nagane

Asst. Professor, Dept. of Comp. Sci., MIT ACSC ALANDI(D)
8855097424, aniket.nagane@gmail.com

Dr. Shankar M. Mali

Assoc. Professor, Dept. of Comp. Sci., MIT-WPU, PUNE
9011092541, shankarmali007@gmail.com

Dr. C. H. Patil

Assoc. Professor, Dept. of Comp. Sci., MIT-WPU, PUNE
9881204044, chpatil.mca@gmail.com

Abstract:

Separation of text as foreground from degraded document images is a challenging task and existing methods may not work for all images. This gives scope to develop new binarization techniques. This paper presents binarization method for degraded Brahmi script Estampage images. The proposed technique is based on combination of multiple existing binarization methods.

Keywords: *inarization; Brahmi; Estampage; Thresholding*

I. Introduction:

The oldest script used in Indian subcontinent and central Asia is "Brahmi". It was commonly used between 400 B.C. and 500 A.D. Considering such a long time period, medium used for writing and geographical region where Brahmi was typically practiced, lot of variations have been found in the scripts of

Brahmi [1]. There are many scripts that are derived from Brahmi. The rich heritage of Indian subcontinent and central Asian history has also been found engraved in "Brahmi" script on rock-cut edicts. There are also other forms like *Bhurjpatra*, *Tamrapat*, etc. on which this script is written. A large number of documents are available in these forms, many of which are unread [2,3]. As it is not a script that is commonly used today, expertise is required to understand and interpret it. All the data in Bramhi script is preserved in the form of estampages (an impression of an inscription made on inked paper). [4]

Archaeological Survey of India, Mysore maintains and preserves digitized copies of the estampages so as to make it available to the historians for their purpose. These images are degraded and hence difficult to read. In Optical Character Recognition, image pre-processing using binarization is one of the most important steps. Binarization is useful to distinguish between the foreground text from the background. [5]