Abstract

Content extraction from badly degraded ancient document is very challenging task due to the different causes of degradation. Ancient documents are of great importance to us and accumulate a significant amount of human heritage over times. These ancient documents are in the degraded form containing vital and valuable information but the contents of the document not recognized easily. There are many causes of degradations such as environmental factors improper handling and the poor quality of the materials. In recent period, the rising interest in the historical document image analysis many researchers are attracting towards extraction of contents from historical document and preserve their contents for future generations. Numerous methods exist but they are not suitable for all types of degraded documents. The
proposed method is simple, robust and based on the phase binarisation model. The proposed method is divided into Preprocessing, Post processing and extraction. The Preprocessing helps to separate foreground and background. The post processing enhances the document image and Extraction helps to extract the content from the document image.

References

- Hossein Ziaei Nafchi, Reza Farrahi Moghaddam Member, IEEE and Mohamed Cheriet, Senior Member, IEEE, &quot;Phase-based binarization of ancient document images: Model and applications&quot; 10. 1109/TIP. 2014. 2322451, IEEE Transactions on Image Processing.
- Prashali Chaudhary, B. S. Saini, &quot;An Effective And Robust Technique For The Binarization Of Degraded Document Images&quot;; International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321-7308 Jun-2014.
- Brij Mohan Singh, Mridula, &quot;Efficient binarization technique for severely degraded document Images &apos;&apos; @ CSI Publications 2014.

Index Terms

Computer Science Image Processing
Keywords
Degradation  Phase Binarization  Local And Global Thresholding