Abstract

Document Image Analysis (DIA) is the subfield of digital image processing that aims at converting document images to symbolic form for modification, storages, retrieval, reuse and transmission. It helps the transition from bookshelves and filing cabinets to the paperless and perhaps even wireless world. Preprocessing is the first stage in document image analysis. In Document Image Analysis, Preprocessing activity involves Representation, Noise reduction, Binarization, Skew estimation/detection, Zoning, Character segmentation. This paper focuses on the major challenges that are to be faced in preprocessing of document images for document image analysis.

References


**Index Terms**

Computer Science  
Emerging Trends in Technology

**Keywords**

Document Image Analysis  
Information Retrieval  
Binarization  
Skew Detection  
Character Segmentation