

A Document Imaging Technique for Implementing Electronic Loan Approval Process

J. Manikandan, C.S. Celin, V.M. Gayathri

Department Of Computer Science and Engineering,
Saveetha School of Engineering, Saveetha University, Chennai
e-mail: manikandancse8@gmail.com, celin.cs7@gmail.com, vmg188@gmail.com

Abstract

The image processing is one of the leading technologies of computer applications. Image processing is a type of signal processing, the input for image processor is an image or video frame and the output will be an image or subset of image [1]. Computer graphics and computer vision process uses an image processing techniques. Image processing systems are used in various environments like medical fields, computer-aided design (CAD), research fields, crime investigation fields and military fields. In this paper, we proposed a document image processing technique, for establishing electronic loan approval process (E-LAP) [2]. Loan approval process has been tedious process, the E-LAP system attempts to reduce the complexity of loan approval process. Customers have to login to fill the loan application form online with all details and submit the form. The loan department then processes the submitted form and then sends an acknowledgement mail via the E-LAP to the requested customer with the details about list of documents required for the loan approval process [3]. The approaching customer can upload the scanned copies of all required documents. All this interaction between customer and bank take place using an E-LAP system.

Keywords: image processing, document imaging, E-LAP system

1. Introduction

These documents include many common types: business letters, forms, engineering drawing & maps, text books, technical manual, music notations & other symbolic data. Though paper was the limited medium in history, many documents now begin on the computers & often dwell entirely in electronic form. In spite of this it is ambiguous whether the computer has decreased/increased the amount of paper document formed, as these are printed out for reading, distribution, markup predictions of paperless offices made so frequently during the early 1980 has given way to the recognition that the objective is not eradication of paper but the ability to deal with the flow of electronic & paper document in effective & incorporated way. Document processing in any association whether having its operations manual or automated, forms an important activity in its functioning's. Within document processing, the key activity prior to all other actions is the identification of documents and hence their classification [1, 7] [10]. Several superior solutions are present for document processing and analysis, this paper tries to give general idea for document processing and the various steps/methods used for that.

Loan processing is one of the tedious process in the banking and financial industries. A loan approval system (E-LAP) is developed specially to support banks and other financial industries loan claim processing needs. It employ workflow technology to control and monitor the various steps in loan processing and uses digital imaging technology to reduce the delays and overhead associated with paper documents. This paper explains about enabling company's loan department to process the loan applications in a structured manner and helps in improving the response time and overall throughput time of loan processing process. The proposed system automates the loan approval process by enabling online requests, modification and it's tracking [4].

The proposed web-based application will be accessed by the applicants and loan officers over the internet. The loan applicants will be able to apply for a loan and track responses and action items. Whereas the loan team at the finance company will be able to view, respond and process loan applications. E-LAP will provide a secure user name and secret code based secured login mechanism to access its services. The proposed E-LAP system will deliver: Easy online loan application access for loan seekers. Better control over turnaround

time for loan application processing. Many financial industries have a test to manage competition. Aspects like customer service, relationship, short turnaround time have become critical factors, Data security, real time availability are some of the critical to success factors.

2. Document Image Processing

The document image processing techniques has become very familiar for information retrieval process in past years. The objective of Document Image analysis is to recognize the text & graphics components in image of documents & to extract intended information from them. Two categories of document image analysis can be defined. First one is Text Processing it Deals with the textual components of a document image. The task of text processing is to determine the skew in the document and identifying relevant attributes such as font, size in documents. The second category is graphical processing it deals with non-textual elements such as symbols, images, logos.

The paper documents are now can be converted into electronic documents for easy storage and better processing of information [6]. The document imaging processing is used to attain a document with a high value, precision and fast recovery. Document image processing technique involves extracting the exact content in a computerized method. This document image processing is an important mission for many organizations for many applications. Document classification is a vital mission in document Processing. Document classification permits the automatic allocation or archiving of documents. For instance, after classification of Business letters according to sender and message type the letters are sent to the concern departments for processing. Document classification improves indexing efficiency in Digital Library construction. The document image processing includes three following stages: pre-processing, feature mining and sorting. In pre-processing stage includes image attainment, Linearization, detection, layout processing. In feature mining stage, the important features in documents are preserved and additional features are included. In categorization stage performs office computerization, digital libraries and other document image processing applications [7]. The steps which are used to process the documents are shown in Figure 1. The objective of pre-processing step is to improve the quality of the document images and find the required content from the data. The feature mining step is used to identify the unique features of the document. In categorization step the common features are grouped for easy identification and retrieval process.

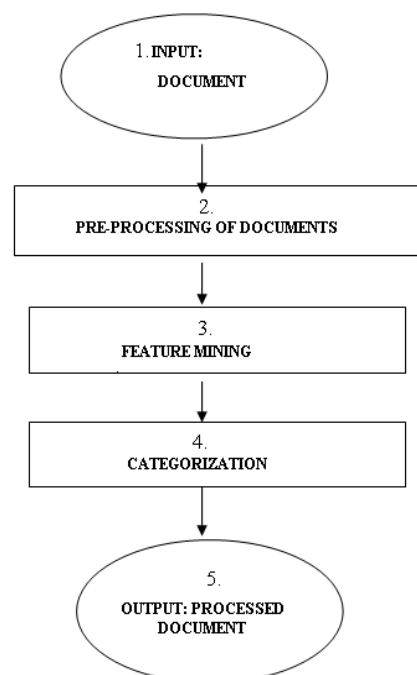


Figure 1. Document Image Processing Steps

The pre-processing step performs removal of unwanted noise in an image, reducing gray scale or color image to binary image, segmentation of different components in image. The given input image is then converted to gray scale or color image or binary formats. Then linearization is performed to differentiate the foreground and background information. In last step de-skewing process is performed. The deviation from the bottom line of page from horizontal direction is called skew. This process involves arranging the paper layout properly, so that it will be ideal for scanning process [16]. The process of de-skewing is shown in Figure 2(a) and Figure 2(b). In feature mining step, only the significant information are extracted from the given input document image. There are two types of features namely, local features and global features. The features extracted from particular block or subsection of image is known as local image. The feature extracted from the entire image is known as global image. Geometric, textual, content and structural based are some of the types of feature mining. One of the main benefits of feature mining is that it highlights only the significant information in an image. Finally in categorization technique, the similar information in images are grouped, so it reduces the number of searches required to find the documents and percentage of error rate in the documents, easy reorganization of documents. The categorization process main role in document image processing and it enhance the document image analysis and indexing features of the documents.

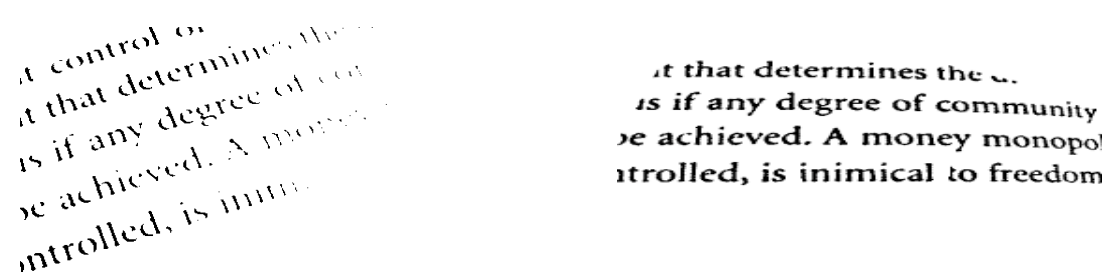


Figure 2(a). Document Imaging Process Flow Figure 2(b). Document Imaging Process Flow

The applicants visits and submit the scanned documents on the banking website. The web server collects all these documents and stores in document imaging database. Then the collected documents are prepared according to the banking terms and conditions [9]. Then again the prepared documents are scanned for storage. Before storing the scanned documents, the documents are indexed for easy querying and retrieval processes. The indexing can be done manual and automatic indexing technique. Finally the indexed scanned documents are stored in the document imaging databases. By sending appropriate query to document imaging database we can view and retrieve the documents for processing. The process flow of document imaging process is shown in Figure 3.

3. Related Work

Loan processing system discusses about an abstract decision-making model that evaluates a loan file in the same way that a normal human process. Risk management approval system is used to predict the future and used to avoid the risks as much as possible. Loan selection emphasizes on the identification and management of risk among groups of loans applied [11]. The credit evaluation system checks the applicability of one of the new integrated model on a sample data taken from banks. Prior relationships in bank loan analyses the importance of retail customer's banking relationships for loan defaults using a distinctive, broad dataset of over one million loans by banks. Single loan opening system discusses the development of a comprehensive dataset of over one million loans by saving banks [13]. Single loan initiation system discusses the developments of a wide-ranging and flexible loan originating system that can be readily and seamlessly interfaced with a core banking system to cater to the banks end to end transaction processing.

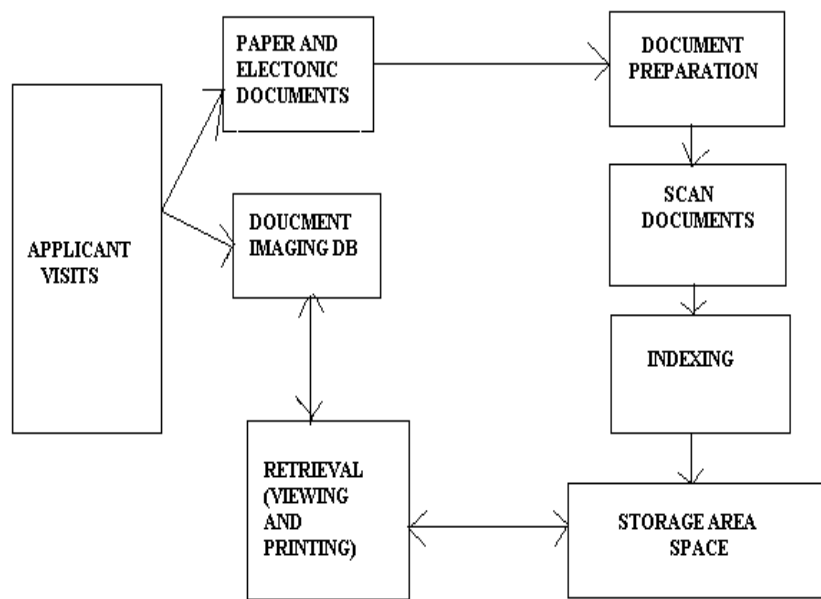


Figure 3. Document Imaging Process Flow

Mind tree uses the business process modeling technology that promises to automate the corporate loan originating process, while retaining partisanship in decision making in the loan approval process suggest solutions that could improve the customer experience in the approving process loan assessment suggests in reviewing the existing literature on decision support system for loan assessment in financial companies [15]. The loan approval process model focuses on the modeling of business process that a banking society follows, during the actions of a loan request.

4. Proposed Work

There has been many research works in the past three years in implementing web-service based bank application systems. This paper focuses on the optimization of the approval process a bank conducts during the application process of small business loans. An ideal system should have the following. Have a structured workflow for automatic routing of applications for different loan products, support multiple loan products and loan types, support, different origination channels, easily interface with different origination channels, easily interface with different external, provide internal credit inspection, credit calculations and conformity examination, preserve all the request details and provide status updates as and when required, support document imaging and archiving for various loan application related documents. The proposed web-based application, E-LAP will be used by the applicants and loan officers over the internet. The loan applicants shall be able to apply for a home loan and track responses and action items. Whereas the loan team at the bank shall be able to view, respond, and process loan applications. The admin user will be able to upload masters in to the application.

The city master will be uploaded from the CSV (comma separated values) format file. The CSV has one column (city). No duplicate records in this file. The checklist of documents required per city for home loan processing. The CSV file has three columns that is city, name of the document, number of copies. Such that this will be received from loan department in a CSV format and will be periodically uploaded by the E-LAP admin user. The E-LAP database used to hold all the E-LAP data including the uploaded data loan applications, processing status, uploaded documents. The web server is used to run the bank website, which contains the E-LAP link. The mail server used for providing acknowledgement to the applicants after the loan is applied for and is being processed by the loan department. The system architecture of E-LAP system is shown in Figure 4.

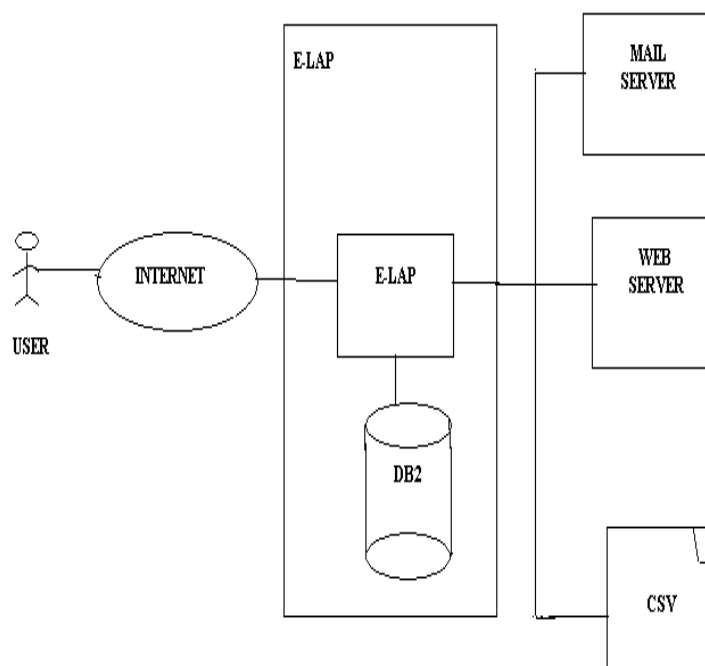


Figure 4. E-LAP System Architectre

5. Conclusion

A loan approval system that is seamlessly integrated with the core banking solution is becoming a critical requirement for banks. Thus to develop such a system that optimizes the time and space complexities of the existing system in the banking sector is highly essential. This system also uses digital imaging technology to reduce the delays and inefficiencies associated with paper documents. The home loan processing complexity is toned down for purpose of project timelines. Automatic loan EMI or interest calculations are not required; the developer will treat them as data entry fields. Most of the concerns needs to store and query huge amount of information in electronic format for retrieving effectively and easily. Thus document image processing plays a vital role in various applications like banking and financial institutions.

References

- [1] Hedjam R, Cheriet M. *Ground-Truth Estimation in Multispectral Representation Space: Application to Degraded Document Image Binarization*. Document Analysis and Recognition (ICDAR), 2013 12th International Conference. 190-194.
- [2] Obafemi-Ajayi T, Agam G. Character-Based Automated Human Perception Quality Assessment in Document Images. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*. 2012; 42(3): 584,595.
- [3] Heatwole E. Processing document images on the telco network. *Communications Magazine, IEEE*. 1993; 31(1): 40-44.
- [4] Wu XD, Street RA, Weisfield R, Ready S, Nelson S. *Page sized a-Si:H 2-dimensional array as imaging devices*. Solid-State and Integrated Circuit Technology, 4th International Conference on. 1995: 724,726.
- [5] Nunnagoppula G, Deepak KS, Rai HGN, Krishna PR, Vesdapunt N. *Automatic blur detection in mobile captured document images: Towards quality check in mobile based document imaging applications*. Image Information Processing (ICIIP), IEEE Second International Conference on. 2013: 299-304.
- [6] Yun Lin, Seales WB. *Opaque document imaging: building images of inaccessible texts*. Computer Vision, ICCV 2005. Tenth IEEE International Conference on. 2005; 1: 662-669.
- [7] Lemmi F, Mulato M, Ho J, Lau R, Lu JP, Street RA, Palma F. Active matrix of amorphous silicon multijunction color sensors for document imaging. *Applied Physics Letters*. 2001; 78(10): 1334-1336.

- [8] Barbu E, Heroux P, Adam S, Trupin E. Clustering document images using a bag of symbols representation. *Document Analysis and Recognition. Proceedings. Eighth International Conference on*. 2005; 2: 1216-1220.
- [9] Bolan Su, Shuangxuan Tian, Shijian Lu, Thien Anh Dinh, Chew Lim Tan. *Self Learning Classification for Degraded Document Images by Sparse Representation*. Document Analysis and Recognition (ICDAR), 12th International Conference on. 2013: 155-159.
- [10] Shijian Lu, Linlin Li, Tan CL. Document Image Retrieval through Word Shape Coding. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2008; 30(11): 1913-1918.
- [11] Bolan Su, Shijian Lu, Chew Lim Tan. Robust Document Image Binarization Technique for Degraded Document Images. *Image Processing, IEEE Transactions on*. 2013; 22(4): 1408-1417.
- [12] Trahanias, PE, Venetsanopoulos AN. Vector directional filters-a new class of multichannel image processing filters. *Image Processing, IEEE Transactions on*. 1993; 2(4): 528-534.
- [13] Prajapati HB, Vij SK. *Analytical study of parallel and distributed image processing*. Image Information Processing (ICIIP), 2011 International Conference on. 2011: 1-6.
- [14] Chien SA. *Automated synthesis of image processing procedures for a large-scale image database*. Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference. 1994; 3: 796-800.
- [15] Yi Liang, Yingyuan Xiao, Jing Huang. *An Efficient Image Processing Method Based on Web Services for Mobile Devices*. Image and Signal Processing, 2009. CISP '09. 2nd International Congress on. 2009: 1-5.
- [16] Casey RG, Wong KY. Document Analysis Systems and Techniques, Image Analysis Applications. *Image Analysis Applications*. 1990: 1-35.
- [17] Castleman KR. Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall, Inc. 1979.
- [18] CC Chang, DC Lin s. A Spatial DataRepresentation: an Adaptive 2D-H string. *Pattern Recognition Letters* 17. 1996: 175-185.
- [19] Sonka Milan, Hlavac, Roger Boyle. Image Processing Analysis and Machine Vision. Brooks/Cole Thomson Learning. 1999.
- [20] T Young, Gerbrands. Fundamentals of Image Processing. Paper Back. 2007.
- [21] O Gorman L, Kasturi R. Document Image Analysis Systems. *Computer*. July 1992; 25: 5-8.
- [22] Rangachar Kasturi1, Lawrence, O'gorman2. Document image analysis: A primer. S⁺adhan⁺ a; 2002: 3-22.