

An AI Intelligent system for Blind People

THRINAI BATCHU
Department of Computer
Science
University of Alabama in
Huntsville
tb0055@uah.edu

ABSTRACT

Communication is one of the most common and important thing in each and every one's life, knowingly or unknowingly we will manage our most of the life with the help of the communication. A very important domain of communication in our office world is the business letter domain. There are certain standard forms of business letters are evolved by using a logical structure and a layout structure that is captured in the international standard using Office Document Architecture.

In this paper we will discuss how to present printed documents to blind people who cannot perceive the information provided in the printed documents with the help of the layout structure of a document by using some ODA standards and ALV. Also we will deal with some philosophical aspects about how to design perfect system for blind people in order to withstand in real world.

Keywords

Layout structure, logical structure, ODA (Office Documents Architecture), OCR (Optical character Recognition), ALV (Automatic Reading and Understanding), Expectation driven analysis, Text skimming, Document analysis, Document understanding.

Copyright © CS530. Summer, 2015. Huntsville, Alabama, USA.
All Rights Reserved

1. Introduction

In present technical world there are so many changes in usage of computer applications especially in office world, computer application based tools are essential in every company. These applications are playing a vital role for communication among humans like within the company as well as to maintain the external contacts of the company. Using this computer technology there are a plenty of benefits for present business world i.e., information is available online which can be accessed at any arbitrary number of different places and information exchange can be done immediately[1].

For blind people there are many improvements in their communication facilities. Some of the examples are Integration of a computer with braille output device; speech synthesizer enables blind people to have access to information provided by electronic mail boxes.

With an optical scanner and optical character recognition software even paper bound information can be perceived without seeing it[1].

Nowadays most of the office information in large extent is still paper bound. Suppose a paper documents contain a different kinds of information with different layout structures, often some of them doesn't contain explicit information. In such situations it is difficult for blind people to perceive information that doesn't specify explicitly. In this situation just usage of scanners and OCR-software is not suggestible

because they are used for reading purpose but not for understanding. In order to obtain feature like reading the document along with understanding we can use ALV (Automatic Reading and Understanding).

ALV is a system has been implemented that allows which transforms the printed letters into an ODA based representation which can consider both logical and the layout structure of a document [1].

2. ODA Architecture

ODA enables compound documents containing character text, raster graphic images and geometric graphics to be encoded and interchanged electronically between systems conforming to the standard. ODA can interchange the electronic documents such a way that they can be printed or imaged and also can edit by the receiver [3].

ODA document model divides the content of a document into three primary categories. They are logical information, layout information and content information.

ODA model has three further components. They are: Generic structures (both logical and layout), styles and document profile.

2.1 Logical Information:

Logical information is the information in a document that is independent of page layout. It also defines relationships between the components of the document content and structuring of the content information in terms of hierarchy and order. For example, structure content of a book into hierarchy of chapters to sections and sections to paragraphs.

2.2 Layout Information

Layout information in ODA is concerned with size, positioning and other image related properties of the content. Organization of layout information is called “layout structure”. These layout structures has a hierarchy of positioning components called page sets, pages, frames and blocks.

2.3 Content Information

This is used directly by the human, which is perceivable and used to convey the meaning. It also includes geometric shapes like circuit symbols.

In content information all the information provided is not considered as a real content. For example, some parts of the code like new line, space, upper and lower case. are not the real visible marks. The ODA standard specifies three types of content: character text, geometric graphics and raster graphics[3].

2.4 Generic Structures

Generic structures are sets of rules within the document itself that can be thought of as defining the "class" of the document. The ODA model enables this type of information to be held in the document itself as generic logical and layout structures. This information may potentially be used by an editing system to ensure that the document always conforms to its own "class". This is an extremely powerful feature of ODA[3].

2.5 Document Profile

This is the final aspect of the ODA model in which information concerns in document as a whole. It contains management information like title, author name and also it contains technical information like logical or layout structures or both. These are the key architecture aspects of ODA.

3. Automatic Reading and Understanding

This is a system which has been implemented that allows transforming a printed letter into an ODA based representation. We have chosen ODA standard because it can deal with both logical and layout structure of a document.

Firstly, the given business letters are scanned and perform several steps for analysis then the text is recognized. Obtained results may contain some gaps, word alternatives. after this step in the subsequent phase it extracts the important predefined information which is in our “message type model”.

In this approach basically the input data is fragmentary texts instead of having one text analysis module, predictor controls a set of different and partially alternative substantiators.

There are three working phases of a predictor they are start, discrimination and instantiation phase. Start phase is solved by applying specialized substantiators for classifying a business letter into message types. Twofold discrimination is performed to select appropriate expectations with in the message type hypotheses. A coarse discrimination reduces a number of message type alternatives, lastly a fine discrimination chooses one expectation within one or few previously selected message types. Now according to expectation selected substantiators are activated. Several rules are applied for both verification of the substantiator results and error recovery if the results are insufficient[4].

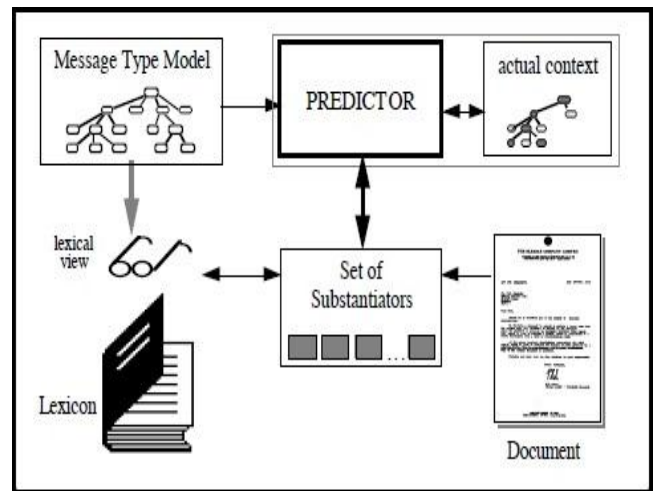


Figure.1. Concept of Text Skimming in ALV

Lexicon:

This contains full forms and stem forms of words. Full forms are used for text recognition.

Here some semantic information is given in the form of clustered words like names of customers or all words often contained in specific logical objects. This forms some groups from whole lexicon and each and every group is called a “lexical view”.

Substantiators:

Substantiators are the analysis components which are operating on text in the document. They are used to cover different analysis techniques.

Statistical based information retrieval methods, keyword and pattern matching techniques and island parsing approaches. Here each substantiator is appropriate for a special task of analysis, according to text structure of logical objects and quality of extracted information. During analysis all substantiators use the partitioned lexicon to get necessary word information.

4. Document Analysis Procedure

The document analysis procedure gives the details about how to give printed letter as input which is then converted into an electronic version in which all the logical entities are maintained. All these logical entities can be considered explicitly by using Office Document Architecture (ODA).

The information structure for a business letters can be classified into three types:

1. Document structure
2. Message type
3. Domain knowledge

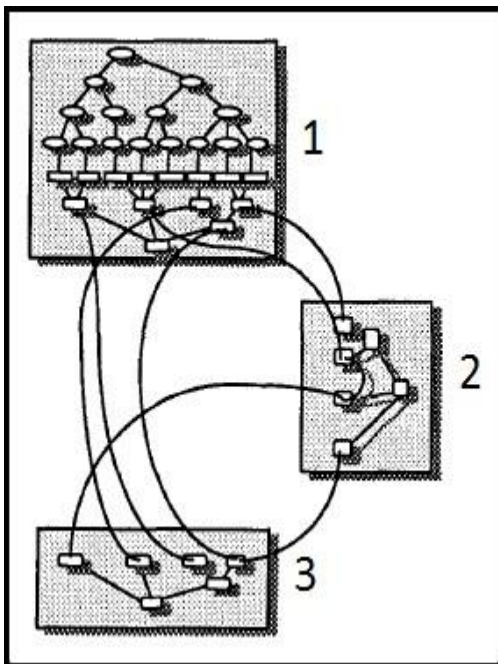


Figure.2.Information structure of a business letter

The above are some parts of knowledge which are useful for understanding business letter. Knowledge of these types has to be modeled in order to understand business documents by using a computer program.

4.1. Document Structure

Document structure contains the information of document content as its own i.e., regardless of the environmental information. The content is further bifurcated into content portion, text portion. Using this information we can obtain two different hierarchies namely the layout and logical structure of a particular document or business letter.

For example, Layout objects can be blocks, pages, words, lines and Logical objects are recipient, sender and signature and so on. Using the connection with content portion we get a mapping of the layout structure onto a logical structure which is known as logical labeling. In ALV logical labeling is done using only geometric information i.e., information of some layout object in relation with others. Which means that hypothesis for a logical label for content portion can be generated before an OCR procedure has started at all[6]. The approach we discussed before is almost similar to the way human perceive a letter which may not be useful for blind people. Now we have to find the existing technique to compensate the above deficiency to some degree.

4.2. Message Type

This message type knowledge can be used to understand logical layout of a business letter. For example, for gifts via courier we will have a separate column for the gift message.

Apart from the above approach we have some different approaches like content and the layout of the letter.

4.3. Domain knowledge

Domain knowledge is some kind of background information of the subject of the letter which can be useful in understanding a letter before reading it.

Using all this we want to capture all knowledge about processes or workflow in a company or office like the temporal precedence relation between an inquiry, an offer, an order and a bill. If two letters are inconsistent according to this relationship they are unlikely to fit together into a correspondence. However, if a new document fits in such a chain this gives supporting argument to classify a letter as a bill.

5. An Integrated Approach and Technology-Assessment

Here we are dealing with how to integrate the technical things based on philosophical aspects and how we will do the technology assessment and what are the benefits of the technical assessment.

In this new field of research a broad knowledge about methods and techniques was composed also we have started to imitate human cognitive competence like sensory perception, problem solving, speech recognition and pattern recognition. scientific community is goes on increasing knowledge and searching for appropriate application to fulfill requirements.

The document analysis concept for paperbound information is based on AI techniques. Document analysis will have a wonderful impact on the future design of information systems, especially on information systems for blind people. The idea of developing a communication tool is used to enable blind people to work in an office environment. This can change a blind people life

and also make them to compete in this present world.

The Artificial intelligence is used to imitate human cognitive competence using machine processes. This research field is closely connected with anthropological disciplines like philosophy, neurobiology or neurophysiology, linguistics[5]. Therefore, AI projects are interdisciplinary approach with are used to design new communication and information technology as user friendly as possible. Evaluation of impacts caused by intelligent information system is made easy with the help of the analysis of interfering framework of different.

So now we came to know that the AI techniques are very much useful for visually challenged people in terms of social and functional integration along with this we should also concentrate on the risks for blind users.

Now we can analyze the possibility of developing the system which is friendly towards the human being, society and nature by integration of technology assessment in development phase itself. On coming new technologies are very complex and dynamic which are often intransparent for the user. We can know from analysis that which effects the system has on the self-experience and on self-esteem of the blind user, either the computer is a communication partner or if it dominates the user by information abstraction and information selection done by intelligent information system[1]. So now we came to know that we have to design a system which has to be personality developing work structure as well and a sociotechnical system design along with compensating the knowledge representation. Our main aim to develop a system which is integrates blind people into the working world by providing real world workspace. The

reason why we are discussing about all these philosophical aspects

Disciplines of science	Ethical Aspects	Anthropology aspects
	Compatibility with 1. Social 2. Human 3. Nature related demands	1. Effect on self assesment 2. Cognition & visual impairment 3. Knowledge representation
	Intelligent communication Tools	
	1. Integration by building a real work space 2. Communication partner	1. Work organization supporting personality development 2. Sociotechnical System Design
	Aspects of Rehabilitation	Human Aspects

figure.3. Chances and risks in aspects: Ethics, anthropology, rehabilitation and human aspects.

Being aware of the ambiguity of intelligent information system, we have achieved an integrated approach of AI and technology assessment[7]. Here, technology assessment is defined as systematic analysis, in which all the impacts and complications, direct or indirect, real or potential, present and future of technology are defined, evaluated and measured as well as cause effect relationships identified[8]. We should not misunderstand the conceptual decision to assess impacts caused by intelligent systems as “technology arrestment”. This approach will be used to achieve higher performance of the designed system and greater acceptance by visually impaired users.

6. Conclusion

Here we are dealing with the problem of integration of document analysis tools into an office workspace for blind people. Main idea of this paper in artificial intelligence point of view is that the paper document used for office communication can be understood for the blind people by arranging them in a way of their layout structure corresponding to a commonly agreed logical structure.

Basically if the data is only in ASCII or textual representation the data will be destroyed. Therefore, In order to overcome this drawback we have to follow other form of representation i.e., something which maintains the logical entities based on some representation like ODA standard for business letters.

Going beyond the single document like business letter the overall goal is to exploit the technologies and integrate them into an intelligent office communication system. This system also allows random access to paper bound information which can be used to compensate the perceptual deficiency of blind people. Also in this the person can control the workflow in the office environment. For example, ALV technique using ALV the system can use the previous knowledge in order to understand the newly reading document. Finally, developing the plan based, goal oriented dynamic query answering mechanisms that enable the system in order to cope up with more complex queries for the document database.

7. Result

Result for our paper is that providing a good communication system for blind people in order to perform their regular office work without any help of other people. This can be done with help of technical things like ALV using ODA representation as well as designing a system such a way that the system should be implemented to achieve all philosophical aspects such that it shouldn't lead to any risks for blind people.

8. Future Work

Future work in this research needed is that how can to help the blind people in domains other than business letters and make their life easy in each and every situation. Some of the examples are system to read books, helping them to do their regular activities, guiding them while doing something new, taking them to new places using gps integrated with some sensors in order to reach the destination safely.

Acknowledgements

I am grateful to prof. Dan Rochowaik who wrote his valuable comments and suggestions on abstract of this paper.

References

- [1]. Thomas Kieninger, N. K. (1994). *An intelligent information system for blind people — AI technology and philosophical aspects*. Springer Berlin Heidelberg. ISBN: 978-3-540-58476-6
- [2]. 8613, I. (1988). *Information processing - text and office systems - office document architecture*. ISO.
- [3]. Robinson, P. J. (1990). Office Document Architecture. *ELEDIS JOURNAL*.
- [4]. Rainer Bleisinger, K.-p. G. (1994). Text Skimming as a Part in paper Document Understanding. *DFKI Technical Memos*, 14.

- [5]. VDI Report. Kfintliche Intelligenz. Leitvorstellungen und Verantwortbarkeit, 1993.
- [6]. A. Dengel, R. B. (1992). *7CODA: The Paper Interface to ODA*. Deutsches Forschungszentrum fiir Kfintliche Intelligenz.
- [7]. W. Rammert. Wie KI und TA einander ns In KI, volume Heft 3, pages 11-16. 1993
- [8]. Bartocha, B. (1973). *An instrument for goal formulation and the selection of problems*. New York: M. Centron and B. Bartocha editors.