Intelligent Document Writer (iDocWriter)

Conference Paper · February 2016

CITATIONS
0

READS
424

3 authors, including:

Dilshan De Silva
Sri Lanka Institute of Information Technology
16 PUBLICATIONS **50** CITATIONS
SEE PROFILE

Udara Samaratunge
Sri Lanka Institute of Information Technology
20 PUBLICATIONS **24** CITATIONS
SEE PROFILE

Intelligent Document Writer

(iDocWriter)

D. I. De Silva[1], S. A. U. S. Samaratunge[2]
Faculty of Computing
Department of Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
[1]dilshan.i@sliit.lk, [2]udara.s@sliit.lk

K. Srirankan[3]
Faculty of Computing
Department of Software Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
[3]it12134236@my.sliit.lk

***Abstract*—With the advancement in technology now it has come to an era where a vast majority of people opting to use computers for their day to-day life activities. However, in terms of document writing most of the people struggle to create appropriate documents according to defined formats in a shorter period of time. In addition, people with less information technology literacy find it extremely difficult to create documents by themselves using existing software. The proposed system acts as a solution to this problem. It allows users to operate most of the features that are helpful in creating documents through their voice. In addition, it will automatically create documents based on a user given title, identify almost all the punctuation issues in a document and correct them automatically, suggest related words while a user is speaking or typing, search for references based on the user's input and list and cite them accordingly, allow users to easily and quickly place the images from the internet in appropriate places in the document file, and will read the documents according to the user's desires. Through this application users will be able to create document easily, quickly and conveniently. To check the accuracy level of recognizing the pronunciation accents of the Sri Lankan people, 50 people who resided in different locations were chosen randomly. For that sample, the system managed to achieve an accuracy level of 75 percent. Correcting grammar mistake feature achieved an 80 percent accuracy. Reference management feature was tested against 150 different documents and achieved 70 percent accuracy. The feature which was used for easy handling of images was also tested was accuracy using 50 different documents and it managed to achieve 85 percent accuracy.**

Keywords—Speech recognition; reference management; grammar checker

I. INTRODUCTION

In the modern world a vast majority of people make use of computers to carry on their day-to-day life activities. However, in terms document writing, most of the people struggle to create appropriate documents according to defined formats in a shorter period of time. In addition, people with less information technology (IT) literacy find it extremely difficult to create documents by themselves using existing software. Thus, they are left with only two options. That is either to get the help from another person or to control the computer through their voice. But the accuracy level of existing voice based systems are not in a satisfied status. At times users are required to give step by step instructions to accomplish a particular task which can be very annoying and most of these systems are not specifically focused towards document writing. The available few document creating applications also have some drawbacks. While narrating (creating a document), users have to say each and every word separately. They are not intelligent enough to determine spaces, paragraphs, punctuation marks, etc. Also users have to handle all the other document related functionalities such as saving, printing, aligning, etc. without their voice.

As a solution for these problems Intelligent Document Writer was proposed. It allows users to operate most of the features that are helpful in creating documents through their voice. In addition, it will automatically create documents based on a user given title, identify almost all the punctuation issues in a document and correct them automatically, suggest related words while a user is speaking or typing, search for references based on the user's input and list and cite them accordingly, allow users to easily and quickly place the images from the internet in appropriate places in the document file, and will read the documents according to the user's desires. The application is expected to be faster, user friendly and efficient than the current voice based document creating applications

II. LITERATURE REVIEW

Document writing is one of the most on growing needs for business people and students. They are spending a lot of time in creating the documents, letters, reports etc.

Out of the available document creating applications, Microsoft Word and Libre office are considered as two of the best applications that can be used with the Windows operating systems and Linux operating system respectively [1]. However, these to applications also have their drawbacks. Libre Office does not include the grammar dictation, auto error correction and synonymous suggestion [2] [3].

At present there exists some voice based report creating software. One of the popular applications among them is "Dragon NaturallySpeaking" [4]. The cost of this software is considerably high. . Some drawbacks of this software are the application control commands used are too cumbersome for

978-1-5090-0082-1/16/$31.00 ©2016 IEEE

regular use [5] and sometimes there is a possibility of it unexpectedly losing the connection with microphone.

TalkTyper is another document creating application available. It doesn't understand different pronunciation accents even when user spoke clearly, it trip up on some words and it can only be used simpler stuff and shorter spoken content [6].

Another voice based software used to create documents is Windows Speech Recognition [7]. The accuracy level of this software gets dipped when long texts are dictated. This too has a difficulty of detecting some pronunciation accents and it does not support all the aspects of document writing.

In his "The Road Ahead" book, Bill Gates mentions that Automatic Speech Recognition (ASR) is the most important feature in future [8]. Mainly there are two types of ASR. One is Direct Voice Input (DVI). DVI based applications are mainly focused on predefined voice commands where the voice commands will be ended in single word. Other one is Large Vocabulary Continuous Speech (LVCSR). LVCSR based systems are used in form filling and document creation. While DVI systems are dealing with small amount vocabularies LVCSR systems are dealing with large sometimes huge amount of data. Generally LVCSR systems are configured to transcribe thousands of words. LVCSR system mostly runs as real time applications and they are expected to perform faster.

III. PROPOSED SYSTEM

The proposed system consists of the following features:

A. Easy document creating through voice

System will automatically place the punctuations at the desired places, identify the grammar mistakes and correct them accordingly, identify places where spaces are needed, and also user will be able to edit the word document (font size, font family, and margin) through the voice commands. Also users will be able to create entire documents by just providing the title.

B. Best user experience

If user needs to check the document, system will read back the document based on the user request. This feature can be used for several other scenarios as well. For example, a parent can write or open an existing document which contains a story and can ask the application to dictate it to the child using a desired voice (preferably a child's voice) and speed. User can adjust the reading speed based on his/her preference.

IV. METHODOLOGY

This section describes the methods and techniques that the research team used to come up with the proposed system. System will observe the users' voice and convert it into text using Google speech API. That text will then be tested for grammar and punctuation mistakes. Identification of the places to put spaces, breaking of paragraphs, and suggesting of appropriate words once a certain text is typed will also happen in parallel. Then final text will be placed in a Microsoft word.

Research team is using iterative design methodology to carry out the project from beginning to end. Iterative design is a design methodology [9] based on a cyclic process of prototyping, testing, analyzing and refining a product or process. Based on the results of testing the most recent iteration of a design, changes and refinements are made.

The functionalities of the main features of the proposed application are listed below.

A. Implementation of document writer

In order to incorporate the voice recognition with the system, Google voice API has been used. From Android mobile application with the help of the Google voice, user's voice will be converted into text and will be transferred from mobile to desktop application via Wi-Fi. Using the destination IP address android mobile application can identify where to transfer the converted text. Once the text reaches the destination, desktop application triggers the voice recognition method and place the text in the document applications.

B. Implementation of punctuation analyzer

Before placing the text in the Microsoft word, it will go through two tests: grammar dictation and punctuation analyzer. These tests are performed using some libraries which detect the grammar mistakes in the given text and correct them accordingly. After that punctuation analyzer will go through the text and will try to identify where dots, comma and quotations marks should be placed.

C. Implementation of smart doc creator

To automatically a generate report, just user needs to give the title of the document. If the user wants to give specific formatting details, then he/she can customize the default formatting details. Also user can generate this report in three different formats PDF, DOC, and HTML.

D. Implementation of image handler

When the user wants to insert an image, user can search for that image from the proposed android mobile application under "Get an image" function. From this function user will get some sample images that relates to his/her search. Simply by clicking that image user can insert the image into the word document. Image formatting can be done from the desktop application.

V. RESULTS AND DISCUSSION

Recognizing the pronunciation accents of the Sri Lankan people was one of the major challenges that were faced by the development team. To check the accuracy level of this feature, 50 people who resided in different parts of Sri Lanka were randomly chosen. For that sample, the system managed to achieve an accuracy level of 75%. However, the system failed to generate the intended results in crowded places due to the noise in the environment.

Correcting grammar mistake achieved an 80% of accuracy. Reference management feature was tested against 150 different documents and achieved 70% accuracy. The feature which was used for easy handling of images was also tested for accuracy using 50 different documents and it manage to achieve an 85% accuracy.

978-1-5090-0082-1/16/$31.00 ©2016 IEEE

Features	iDocWriter	MS Word	Wordpad	Notepad	Notepad++
Create Document through voice.	✓	✓	✓	✗	✗
Managing font setting through voice.	✓	✗	✗	–	–
Identifies punctuation errors and correct them automatically.	✓	✗	✗	✗	✗
Suggesting regular words.	✓	✗	✗	✗	✗
Read document.	✓	✗	✗	✗	✗
Searching reference.	✓	✗	✗	✗	✗
Detecting Plagiarism.	✓	✗	✗	✗	✗
Creating smart documents automatically.	✓	✗	✗	✗	✗
Search images from internet.	✓	✗	✗	–	–

Fig. 1. A comparison of the features of the existing applications and the proposed application

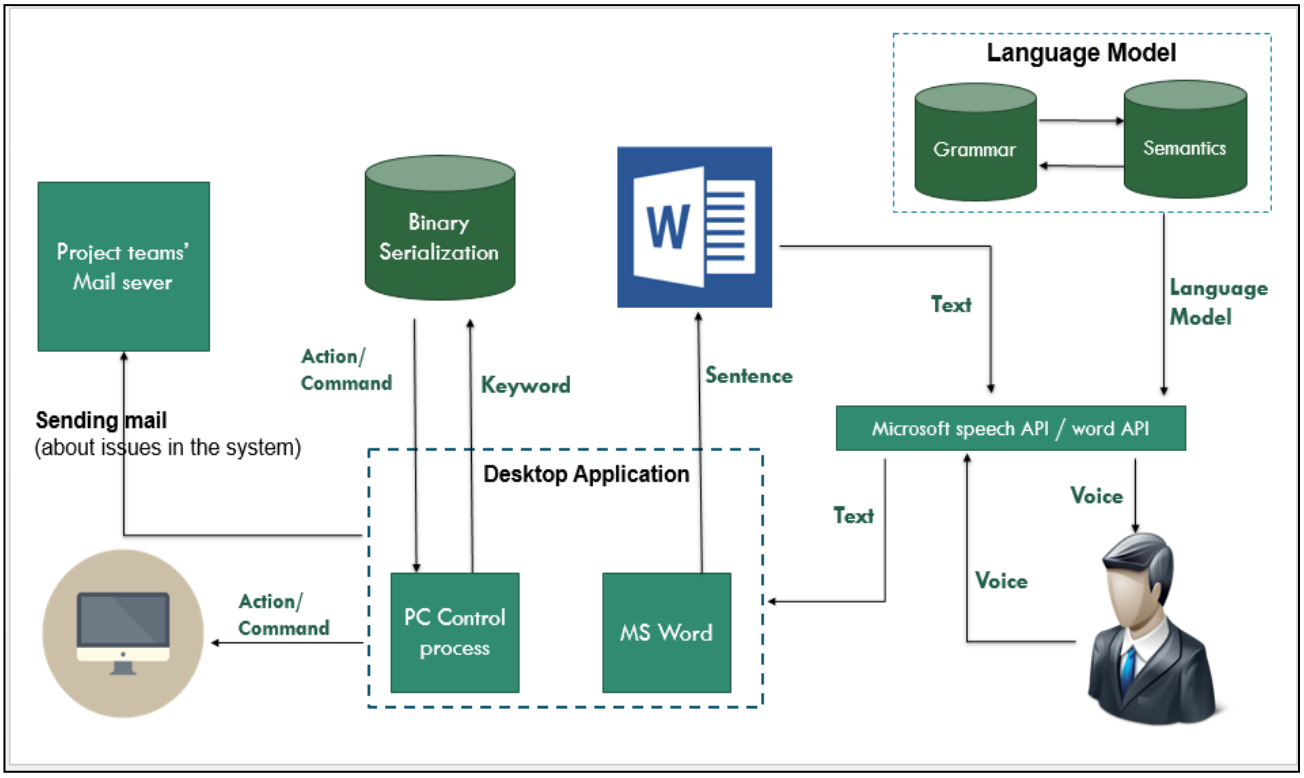

Fig 2. System Diagram

978-1-5090-0082-1/16/$31.00 ©2016 IEEE

VI. CONCLUSION

The paper proposes a fully automated document creating tool. Using the application even people with less IT literacy will be able to create document by themselves using their voice. In addition, it will automatically create sample documents based on a user given title, identify almost all the punctuation issues in a document and correct them automatically, suggest related words while a user is speaking or typing, search for references based on the user's input and list and cite them accordingly, allow users to easily and quickly place the images from the internet in appropriate places in the document file, and will read the documents according to the user's desires.

Apart from normal users, visually challenged people will find the reading back the document feature very helpful. Ability to adjust the reciting speed will help users who are not that much fluent in English the most. Automatic report creating feature will help users to get the brief idea regarding a certain topic in a quick same of time, before attending a meeting, presentation or workshop.

The proposed Mobile Application currently works only on android phones. In the future it can be developed to work with IOS and Windows phones as well. Also the proposed application can be further extended to recognize the user's voice even in crowded places.

REFERENCES

[1] G. Tiwari, M. Pandey, and M. Shrestha. "Text Prompted Remote Speaker Authentication". 2011. Presentation.

[2] P. Rohit. "What are the Advantages or disadvantages of Libre office compared to Microsoft Office?" in Quora Online Discussion Forum, 10 Feb. 2015. [Online]. Available: http://www.quora.com.

[3] E. Mendelson "Microsoft Word 2013", 10 Feb. 2015. [Online]. Available: http://www.pcmag.com/article2/0,2817,2416138,00.asp

[4] Nuance.com, "Dragon NaturallySpeaking - World's Best-Selling Speech Recognition Software | Nuance". Web. 28 Feb. 2015.

[5] C. Null, "Dragon NaturallySpeaking 13 review", 28 Feb. 2015. [Online]. Available: http://www.pcworld.com/article/2458363/

[6] A. Gragg, "A Review of the Web Application", 13 Feb. 2015. [Online]. Available: http://www.geekbraindump.com/2014/08/08/

[7] Windows.microsoft.com, "What can I do with Speech Recognition? - Windows Help". Web. 16 Jan. 2015.

[8] R. K. Moore, "Automatic Speech Recognition", [Online]. Available: http://www.kornai.com/MatLing/

[9] Mais.umich.edu, "ITS | Methodologies > Iterative Design & Development Framework". 2013. Web. 5 Feb. 2015.

978-1-5090-0082-1/16/$31.00 ©2016 IEEE