
OCR Interfaces for Visually Impaired

TOPIC ASSIGNMENT 2

TOPIC PROPOSAL

Author:

Sachin FERNANDES

Graduate 8

Undergraduate Team 2

Instructor:

Dr. Robert PASTEL

March 4, 2016

List of Figures

1	Case study of Desktop application	2
2	Sign with Braille	3
3	Case study of KNFB reader	4
4	Mock up of simpler interface with only camera display	5
5	Representation of hands free Google Cardboard	6

Contents

List of Figures	i
1 Abstract	1
2 Problems with Current Interfaces	1
3 Real World Analysis	3
4 New Design Proposal	4
4.1 Phase 1 - Mobile Application	5
4.2 Phase 2 - Google Cardboard Implementation	6
5 Current Implementation	7
6 Conclusion	7
Appendices	8
Appendix A Email questionnaire	8

1 Abstract

The advancement of technology has caused a digitization revolution that allows most documents to be stored electronically. Even with this boom in digitization a majority of our books and reading material are still accessed in physical paper form. This content is largely unavailable to visually impaired people. To provide access to this paper content I want to propose a system that can scan text on the paper and read it out to the visually impaired person. The interface of such an application is of utmost importance because of the way in which the application will be used. Most user interfaces today are geared towards non-visually impaired people. This makes it hard to understand what a good visually impaired friendly interface is. New accessibility features in mobile phones help simplify this process but there is still a lot of work to be done in this area.

My idea is to implement a new way in which a visually impaired person can interact with an unfamiliar environment and be able to read important information as and when required by them. Similar applications exist but have very complex interfaces and try to do many other things as well. I want to simplify these interfaces and focus them towards a specific task rather than making them do many tasks. Simplification and minimization of the current interfaces is the primary goal of this proposal.

I also want to go one step further and implement a wearable version of the application that will do away with the need for an interface all together. Having no interface and using the application as if it were a natural extension of the body is the goal of wearable implementation of the application.

2 Problems with Current Interfaces

Current interfaces in the OCR (Optical Character Recognition) domain are extremely cluttered and are very complex to use even for non-visually impaired users. The problem with current interfaces is one that plagues a lot of today's technology i.e. they have too many features and try to do everything. Most current technology in this area targets users who are proficient in the use of phones and are technologically inclined. This quickly becomes an issue when targeting people who are not used to using phones at an advanced level. As a side note, current technologies are also relatively expensive and most mobile OCR applications cost around a 100 USD to purchase. This seems very expensive to me and I want to release my mobile application for free.

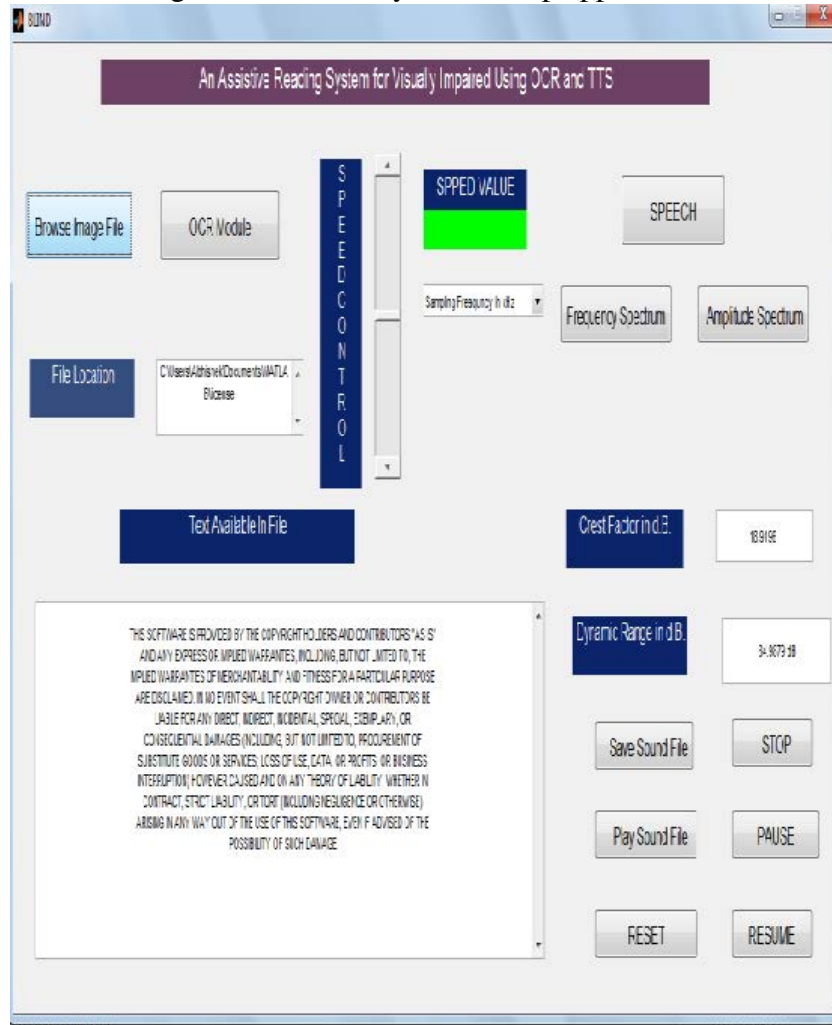
An additional issue that I have noticed with the current interfaces is the need to constantly take pictures in order for it to be read out the text. This can become tiring for the user if they need to read a lot of documents. The constant picture-taking and then waiting for the text to be read can waste a lot of the user's time.

Desktop applications are also used for OCR but are generally very complicated to use and are not portable. Figure 1 shows an interface for a desktop application that was developed as an OCR application for the blind [3]. I could not find a better image for the application interface. Figure 1 was the only image of the interface that the paper had. It is quite clear from the image that the interface would take some time to understand. It would also be quite difficult to use on a daily basis.

With the explosion in phones today, almost everyone has a mobile device. The desktop tech-

nology for this purpose has become obsolete. Phones with better cameras and faster processing facilitate easier development of mobile OCR applications. This is the main reason I will stick to a mobile application interface and not implement a desktop version.

Figure 1: Case study of Desktop application



A physical world example of design in this area is the braille on the walls outside most rooms. Figure 2 shows a sign with a room number and braille under it. From the image it is clear that it would be very easy to miss this braille lettering. How does a visually impaired person in an unfamiliar environment know where exactly these braille letters are placed. They would need to feel around the walls until they reached a sign with these braille letters. Even then there is a chance of missing the letters. They would need to feel for every door until they find the room number they are looking for. My application could help with this by reading out the numbers as they pass each numbered sign. They would not need to go around feeling for the braille letters but simple look in different directions till the application reads out the room number they are looking for.

Figure 2: Sign with Braille



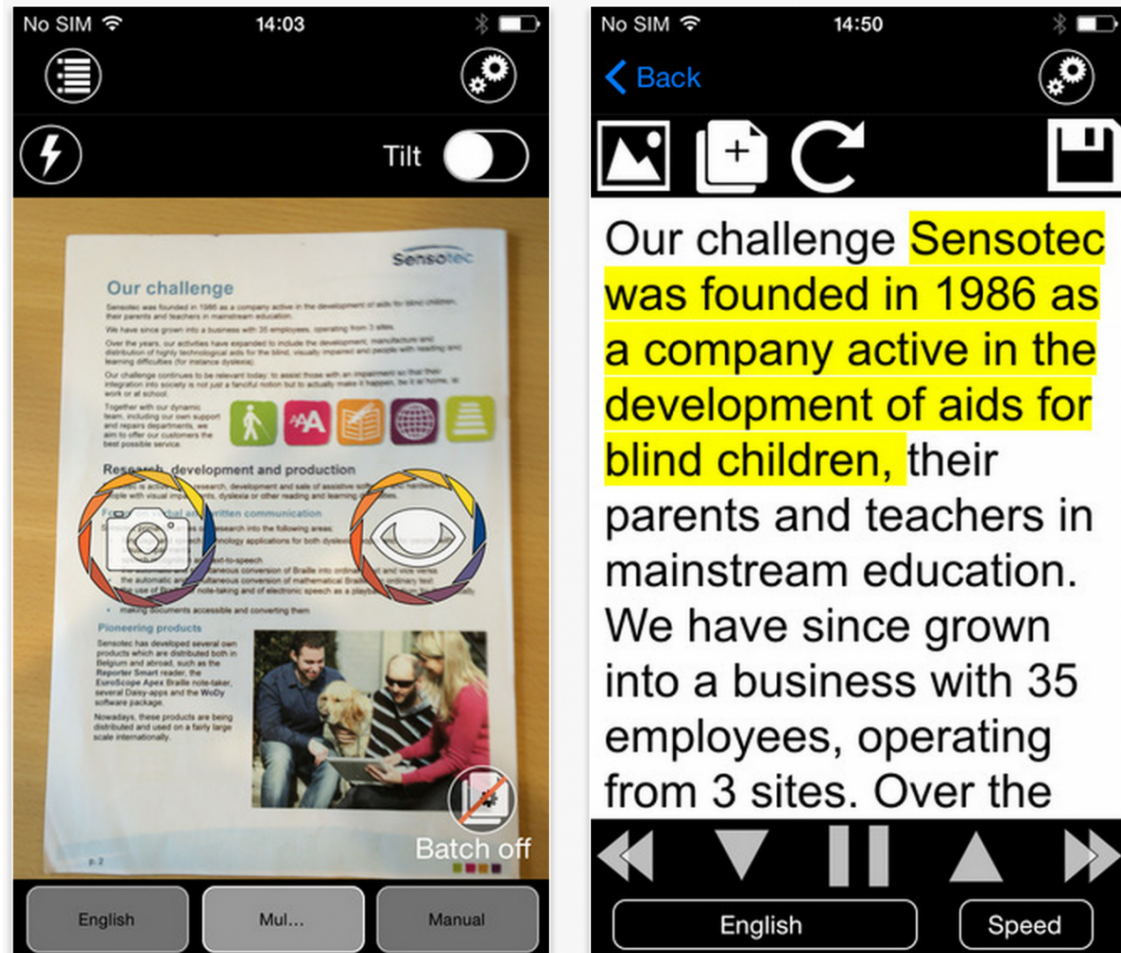
3 Real World Analysis

Figure 3 shows a screen shot of the KNFB reader [2] which is currently used by some visually impaired people. The image on the left shows the camera screen along with many buttons for flash, taking a picture, getting information about the current position, tilt etc. It can be seen from this example that there are too many buttons and too many functions. The application is not focused on what it is trying to achieve. The screen on the right shows the converted text screen. This screen has more buttons than the previous screen. It also has many sub-menus that go into details about speed, saving a file and uploading to Dropbox. This screen also has buttons on the bottom to pause and continue the reading along with skip sentences functions. All these extra functionalities deviate from the main task the user is trying to perform.

The interface takes a long time to learn and it is easy to make a mistake and lose yourself in the interface. Although the OCR capabilities of the application are very good, my worry is that it takes a long learning curve to start using and can also easily frustrate the user. The application makes good use of inbuilt accessibility features but the clutter of all the buttons can often lead to confusion.

I conducted an email interview with a visually impaired CS major named Kevin in order to get more insights about how he uses his phone and applications. I wanted to learn what his idea or concept of an application was. The conversation with Kevin helped me understand how I could improve my application. The email questionnaire has been included in the appendix section A of the report. I have learned that he primarily uses the KNFB application to scan one document at a time. As previously mentioned this can get very tiring. Additionally, mis-clicks are very easy to make due to the close proximity of the buttons in the interface.

Figure 3: Case study of KNFB reader



4 New Design Proposal

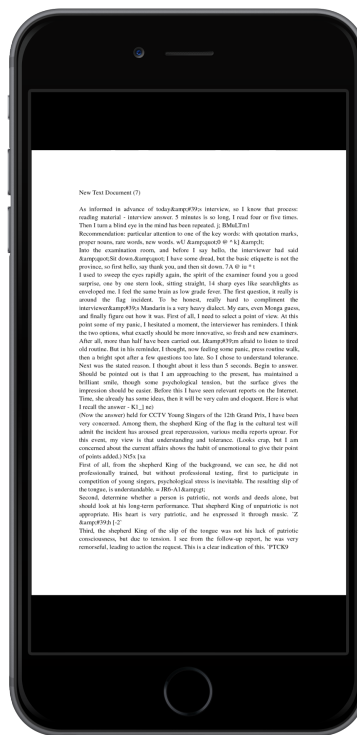
My design proposal is to have the simplest interface possible for the visually impaired. The design will make the application straight forward to use and will have a negligible learning curve. The simple design will also help with getting visually impaired children used to using the application from a young age and not feel intimidated by the complexity of the application itself. On the other

end of the spectrum, it can also help elderly people who have visual disabilities learn how to use this application without having to learn about other complex features.

The new application design has no unnecessary buttons and it is focused on doing one task well. The interface has a single camera screen that continuously scans for text in front of the camera. This will eliminate the need for the user to constantly take a picture in order to get a document read. The idea will extend into a Google cardboard interface that will require no physical interaction by the user leaving the user's hands free to perform other tasks. The two implementations are called Phase 1 and Phase 2. They are described in detail in the following subsections.

4.1 Phase 1 - Mobile Application

Figure 4: Mock up of simpler interface with only camera display



Phase 1 of the process is to have a mobile application with just the camera screen and feedback to the user. The feedback will be in terms of both voice feedback and haptic vibrations. The haptic feedback can be useful in noisy environments or in case the volume is turned off. The main idea of the interface is to automate all the tasks that would usually be performed manually by a user. The camera will auto detect the text without the need for a click button and will automatically give feedback to the user about the current position of the camera.

Once the text is processed the application will automatically start reading the text without the need to go to the reading page. I believe that this automation process is what makes the interface better to use. The lesser options a user is provided with, the lesser mistakes they will make. A mistake in the case of visually impaired users could prove hard to recover from. For example, if the user clicks on a wrong button and gets lost inside the application, they might need to exit the application and return to the home page to try again. The constant feedback with automation and the simple interface could avoid these mistakes and make the application seamless.

Additionally, I want to incorporate natural motions into the application. Instead of using buttons to perform an action, a user could swipe left or right to decrease or increase the speed of reading. A swipe with two fingers could skip the current sentence or go back to a previous sentence. I want to be able to avoid buttons as much as possible. This no button mechanism limits the number of mistakes a user could make. By actively having users perform an action i.e. swiping right or left the user knows what the outcome of an interaction will be, which may not be as clear with a button press. A long tap on the screen could indicate pause and play of the reading of text. The natural actions that the application will use makes it easier to learn and remember the functions the application can perform. Once the application can perform well in the mobile environment I would like to do away with the mobile interface all together and further simplify the application by implementing it on the Google Cardboard [4].

4.2 Phase 2 - Google Cardboard Implementation

Figure 5: Representation of hands free Google Cardboard



Phase 2 of the interface will be implemented using the Google cardboard. The phone will be placed inside the Google cardboard with the camera facing outwards. The camera will then constantly process it's environment and try to figure out if there is anything in the field of view of the camera that needs to be read. If it finds something worth reading it will start the reading process.

The Google cardboard also has a singular button on the side of the glass. This button will be the pause and play button for the reading in the application. Once the application starts reading the text it has detected, the user can pause/stop the sound but pressing this physical button. Similarly, the user can resume playing using the same button. The extreme simplicity of this version allows the user to walk around with the cardboard not having to worry about pushing buttons on a mobile screen. Using phone buttons while walking around could be dangerous and could distract them from concentrating on finding their way.

The idea with the Google cardboard is to create a seamless experience for the user with the application adding to their daily experience and not hindering it. The hands free aspect of this implementation allows the user to perform other tasks with their hands like flip a page or adjust the position of a book. The hands free aspect is also very important from a safety point of view as the user might need to carry a walking stick or may have a guide dog with them. In such cases the user requires their hands to be free to navigate using the stick or guide dog.

5 Current Proposal Implementation

My idea is in the development stage but I have been able to get good results from my current application. I have uploaded a video that shows how the application currently functions. You point the camera at some text and if it detects any text it will try and read it out to the user. The video link can be found in the references [1]. In the first video, the application detects text on a piece of paper and starts reading it out. The second video shows how the application detects an emergency sign and reads it out to the user. This is how the Google cardboard version of the application would function. Although my application is in the preliminary stage of development, the interface is relatively effective and gets the job done.

Eventually, I would like the application to learn what it is looking at. This will be done using cloud based neural network technology. For example, if many visually impaired people are pointing at EXIT signs but the application cannot recognize it, eventually the neural network engine will teach the application what an EXIT sign looks like. This will greatly help the intelligence level of the application and further simplification of the interface can be done based on what the application has learned.

6 Conclusion

I believe that this new interface and application could be an improvement over the current technologies that are present for use today. The fact that the new application is free will also give access to many people who cannot afford expensive alternatives. I have been getting in touch with many disability centers to get feedback about the idea and how I can better improve it. It is essential to get the user interface correct in this application if the idea is to succeed.

Appendix A Email questionnaire

1. Do you use a mobile device often? If yes what do you primarily use it for?

I use an iPhone with VoiceOver accessibility a constantly. I check email, read documents, manage contacts, use map and GPS apps, social media including Facebook, Twitter, LinkedIn... I browse the web, I use the KNFB reader to read printed documents, I use Voice Stream to convert text documents to speech.

2. Do you use an android or IOS device?

IOS exclusively. Android has a long way to go to catch up to accessibility that is available in IOS.

3. Have you used the camera functions of your mobile device?

Yes primarily to OCR documents.

4. Which apps are you familiar with on your device?

I have approx 40 apps of which I use 25 a lot. I mentioned some above.

5. Are the apps simple or complex to use (do they have a large learning curve)?

No more learning curve than would be required for anyone else, as long as the Apps are VoiceOver accessible.

6. What is the one thing that an app must have for you to be able to use it?

VoiceOver accessibility built-in.for example voice commands, accessibility features?

7. Have you used an app that can convert images to speech before?

Yes KNFB Reader.

8. If you were using a live video stream, text to speech application how would you use it?
Basically what would you point it at?

Any items that I would like information about.

9. How would you like the user interface of such an application to be.

The App needs to follow VoiceOver and IOS accessibility Standards.

10. Any feedback you would like to provide would be extremely valuable.

You don't need additional feedback instructions as long as you follow VoiceOver IOS accessibility standards.

References

[1] Sachin Fernandes, The Story Teller Youtube reference videos.

<https://www.youtube.com/playlist?list=PLTyBgCiwUMoTOlyTQoVWjmTmHBngf0gkB>

[2] KNFB reader

<http://knfbreader.com/>

[3] *An Assistive Reading System for Visually Impaired using OCR and TTS*, PEC University of Technology, International Journal of Computer Applications

[4] Google Cardboard by Google

<https://www.google.com/get/cardboard/>