Advanced Science and Technology Letters
Vol.60 (Mobile and Wireless 2014), pp.53-57
http://dx.doi.org/10.14257/astl.2014.60.14

Voice Annotation Technique for Reading-disabled People on Mobile System

Joo Hyun Park[1], KyungHee Lee[1], JongWoo Lee[1], Soon-Bum Lim[1],

[1] Multimedia Science, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-gu, Seoul, 140-742, Korea
{park.joohyun5, khlee, bigrain, sblim}@sm.ac.kr

Abstract. The eBook annotations such as bookmarks and notes are used basically in the visual form. Reading-disabled people have much difficulty to use such annotations on digital talking books. In this paper, we suggest a voice annotation technique that can record annotations by using only voice and hearing sense. The annotation recording uses voice recognition function and time-positioning algorithm to decide the position of a voice annotation on book reading. For the Reading-disabled people, our model provides an adaptive interface which services voice and display to make them read eBook effectively. Our voice annotation system is implemented on Android platforms, and it consists of 4 modules: input, analysis, storing and output.

Keywords: Time-Positioning, Voice Annotation, Annotation Input, Reading-disabled People

1 Introduction

A reading-disabled people is a person who have difficulty in reading including blind, presbyopia, dyslexia, or another physical disability[1]. As the world is changing into an Aging Society, the numbers of people with reading disabilities are growing faster than expected. This increase has led the rising demand for 'Spoken text', giving anybody the ability to use eBook without any barriers. An annotation is a note or a comment that is added to text in a book, but to use it, reading-disabled users need a voice annotation system. Aside from just listening to text, they want to record their thoughts and opinions and re-open their annotations when necessary. However, traditional eBook viewer software, even the viewer for disabled people, is screen-based without proper regard to the needs of reading-disabled users. Besides, most eBook software for the disabled only provides a bookmarking function, so users are not able to record notes of their thoughts and opinions on a specific location where they want.

In this paper, we suggest a voice annotation system that enables the disabled-readers to record annotations anywhere they want by voice. This would improve reading environment as well as user satisfaction. In addition to advantages of traditional audio books, this system is implemented to run on smartphone, allowing the user to record the additional information and browse it anytime anywhere.

ISSN: 2287-1233 ASTL

2 Related Work

AMIS[2] is an open-source playback software program, developed by the DAISY(Digital Accessible Information System)[3] Consortium. It has many practical functions such as its own vocalization interface, searching the original content, and regulating the speaking speed, but it only offers a simple bookmark function as an additional memo function.

EasyReader[4] is a charged software that provides full support for EPUB[5]. It also supports embedded objects such as SVG images, and MathML. Users can take text or audio notes by bookmarking. Audio notes can be recorded through microphone, and users can check their recorded voice with the playback button. However, there are several weaknesses in this software. There are some physical variables, such as noise which is occurred during recording because this system uses the raw recorded media files. Note file's contents and detailed information are also hard to be modified.

eBook software for mobile has relatively intuitive and simple interface comparing to one for PC. Android-Daisy-ePub-Reader[6] is a collaborative project to create an electronic book reader that supports DAISY, but it only reads text without displaying it. Annotation functions are not provided as well.

All eBook playback software discussed above provide functions such as the ability to add notes and bookmarks, but using these functions are very limited for visually impaired users since they cannot operate the screen to decide where to locate annotations. Therefore, we focus on that enables readers to record annotations anywhere they by voice.

3 Voice Annotation System

In this section, we would discuss possible issues that must be considered to implement Voice Annotation System (V-Anno). Then, we would introduce our system using new solutions.

An issue arises when reading-disabled users record voice annotations.

- The difference between user's desired location and actual location when recording annotations.

First of all, unlike low vision readers who can use both sight and auditory sense, blind readers have to completely depend on auditory sense. So when blind users add an annotation on a specific position, where annotation is actually recorded may not correspond to the position where user intended. To solve this issue, we designed a V-Anno System that enables users to record voice annotations. Also we separated annotation input & output type into screen-based and voice-based interface. Whereas blind readers add annotations using voice, low vision or learning disabled readers utilize text screen as well as voice.

Since voice is being played even when blind readers record annotations, the position where readers want doesn't correspond to where annotation is actually

recorded. Because the user adds annotations on already read sentence, this step is especially essential for readers who are highly dependent on auditory sense. Therefore, the system decides where to record the annotation based on a sentence that includes selected word or syllable.

Analysis phase is necessary for sentence determination. We measured the number of words, and playback time in 100 sentences from 2 audio books. As a result, an average playback time for one word was 0.58(±0.87) seconds and a sentence was composed of least 3 words or up to 27 words, with an average of 9.8.

We use 10 words units to determine where to record annotations. This is based on the psychology research on memory [7][8] and the statistical fact, as previously measured.

This system determines an exact location depending on word count and playback time of a sentence that includes the annotation requested time(T_{Req}). Based on the average playback time of 10 words(5.8 seconds), if T_{Req} is included within first 5 words′ playback time, it selects previous sentence(S_{i-1}) otherwise current sentence(S_i). This equation is defined Eq 1.

$$\begin{aligned} &if\ (T_{Req} < T_{iB} + \frac{T_{iE}-T_{iB}}{W_n} * 5\ ) \\ &\quad \text{A-}indx_k = i - 1 \\ &\text{else} \\ &\quad \text{A-}indx_k = i \end{aligned} \tag{1}$$

If selected position is in an invalid playback time(IT_i), space between two sentences(S_i and S_{i+1}), current sentence(S_i) is chosen. Figure 1 shows a diagram of input time.

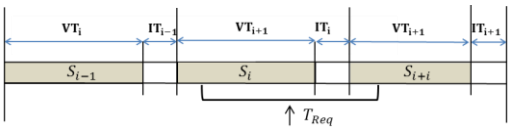

Fig. 1 A diagram of Annotated time

4 Implementation

This System is android-based[9], and used Google Speech API for voice recognition, Google Korean TTS Engine for voice synthesis. We experimented the system on Android Emulator and Samsung Galaxy Tab[10]. Figure 2 shows sample V-Anno screens for experiment.

To evaluate the success of the proposed voice annotation implementation, two small tests were conducted. 4 people participated in our test. In the first experiment, each participants record 15 annotations. Then, they checked the locations where annotations were created. As a result, 58 annotations were successfully located on the right positions.

Fig. 2 V-Anno Screens

5. Conclusion

In this research, we suggest V-Anno System that enables reading-disabled users to record voice annotations using voice and auditory sense. This system is divided into 4 modules: input, analysis, storing, and output.

We made contributions for voice annotation position determination method by adopting time positioning algorithm for our annotation model. When annotation requested in the sentence, this technique determine the annotating sentence considering the first 5 words' playback time. Our model is expected to enable the users with reading disabilities to record and navigate annotations just using auditory sense. Also, built on current smart phone, V-Anno can enrich reading environment and quality of life regardless of time and space. We expect that V-Anno can help to educate and train people with reading disabilities, and provide an environment conducive to education, stability and job creation for them.

Acknowledgments. This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2014.

References

1. Ministry of Culture Sports and Tourism: Reading Culture Promotion Law., Republic of Korea (2009)
2. AMIS, http://www.daisy.org/amis.
3. DAISY, http://www.DAISY.org/
4. EasyReader, http://www.yourdolphin.com/pro-ductdetail.asp?id=9
5. EPUB, http://idpf.org/
6. Android Daisy Epub Reader, http://code.google.com/p/android-daisy-epub-reader

7. Miller G.A.: The Magical Number Seven, plus or minus two:Some limits on our capacity for processing information. *Psychological Review.*Vol63(2), 81—97 (1956)
8. Kwang-Hee Han, etc. Cognitive Secience in Korean. Hakjisa Publisher, Seoul (2000)
9. Android, www.deveolpers.android.com.
10. Samsung Galaxy Tab, http://www.samsung.com

