Study of Text Reading Method Using Important Words on Web Sites for visual impaired

Takumi Misawa^{a,*}, Yuto Tanaka^a, Yoshihisa Nakatoh^a

^aKyushu Institute of Technology

*Corresponding Author: o349429t@mail.kyutech.jp

Abstract

Generally, visually-impaired individuals obtain information from the Internet; they use a kind of software called reading software, which reads out the character information from HTML files of PC. However, it takes a long time to obtain the desired information from the conventional text-reading system compared to healthy people because Web sites have been created basically for those with good eyesight.

Therefore, in order to improve the efficiency of gathering information, we have proposed a text-reading system which users can intuitively understand the Web site by conveying important keywords in a shorter time.

This paper discusses the reading method as understanding the important words in a shorter time in order to evaluate the effectiveness. As a result, we found that to improve the hearing capacity in a short time by using the reading method superimposing the word to be played again later, based on the words that are played. In addition to the above, it was found that there is a further improvement by the words first played from a female voice, and then from a male voice to help settle-in the words.

Keywords: visual impairment, text to speech, Internet.

1. Introduction

Currently, visually-impaired individuals are generally using reading software when using the Internet. This is software for transmitting information to the user by reading out the information of the Web site displayed on the screen by speech synthesis.(1) Without this, the visually-impaired cannot operation of PC. However, it takes a long time until the visually impaired person can get to the information they

want compared to healthy people, because Web sites have been created basically for those with good eyesight.

In this system process of information acquisition from the Internet consists of five stages.

- ① Displaying the search site.
- ② Searching the Web site related to want information.
- ③ Clicking the Web site likely desired information from the search results.
- ④ Determining whether or not the Web site there has the desired information.
- ⑤ Checking the contents of the desired information.

If you do not want the information in the process of ④, return to ② or ③. Therefore, this study focused on the information acquisition process ④ in order to improve the efficiency of gathering information for the visually impaired. We thought that in order to find the desired information in Web sites from the search result in short time, the visually impaired are told not only the title of the Web site but also important words like the representative of the Web site, to help intuitively grasp the Web site. In this paper, we discuss the reading method as understanding the important words in a short time to evaluate the effectiveness.

2. Internet browsing support overview

The flow of the processing of the Internet-browsing support system is shown in Fig. 1. The system consists of four units. Its four configurations include the extraction unit that extracts sentences from the Web site, the morphological analysis unit that divides the obtained sentences into words, the selection unit for extracting important words from the obtained words for the speech

Fig. 1. Block diagram of the support system for visual-impaired individuals

synthesizer to output the important words. The details for each processing system are listed below.

2.1 Extraction of sentence

This process obtains the HTML data from the Web site and divides the information into sentences. First, it removes the HTML tags from the HTML data, then divides the sentences separated by a ".", "!", such as punctuation marks and exclamation points.

2.2 Morphological analysis

The chosen sentences are morphologically analyzed and divided word by word. At this time, the words have no meaning since particles and auxiliary verbs are excluded, and nouns, verbs and adjectives are selected only for words that have meaning.

2.3 Extraction of important words

The selection of key words in the sentences, use words in the document "d" represented by the following formula "t" Severity "tf * idf (t) (2) " as an indicator.

$$w(t, d) = tf(t, d) * idf(t) \quad (1)$$

$$idf(t, d) = \log \frac{N}{df(t)} \quad (2)$$

However, "tf (t, d)" is the word frequency of word "t", "df (t)" is the number of documents in which the word "t" appears, and "N" is the total number of documents in the text.

The word frequency "tf (t, d)" value as the word of the

subject is contained in a large amount in texts; we think that has been written for more information about the word. On the other hand, we think that the words "df (t)" values that are used in the document are of higher importance, even though there are more of the words that are used in a small document. Thus, a word representing the document is derived.

2.4 The method of reading important words

The method of reading important words selected by speech synthesis is shown in Fig. 2. Previously realizing the word and understanding it in a short time by superimposing the word to be played again later, based on the words that are played. Fig2 shows the six kinds of superimposed methods (a) 0%, (b) 25%, (c) 50%, (d) 75% and (e) 100%. Playback time in Fig2 is a value at the time of the 400 phoneme per minute. This value is speed of a typical read speech (3). There is a difference of 2.56 sec when compared to (a) and (e). In addition to this, the word preceding the set male voice is later followed by the same word from a female voice (or vice versa). In the evaluation, the methods are assessed.

Fig. 2. The reading method of important words

3. Evaluation

3.1 Experimental content

In the experiment, six word outputs in five types of audio reproduction methods in Fig2 only using male voice were compared, and an evaluation was made based on how many the ten individuals could target or could hear. This not only, methods were also investigated the case preceding the set male voice is later followed by the same word from a female voice (or vice versa). For the six word outputs, the man was set in because of the "Magical Number 4" (4); that is the extent to which things in fours can be stored in short-term memory. We were using actual news sites for the important word extraction from the Web sites. It shows some examples in Table. 1. Also, when reading aloud, a speech synthesis, in which one PC-Talker from the reading software used in Japan, was used.

3.2 Experimental result

The result of each reading method is shown in Fig. 3 by male voice only. This figure shows whether were able to hear how many units in six. As a result, an average 2.46 units is when superimposed 0%, an average 2.90 units is when the overlaid 25%, an average 2.65 units is when the overlaid 50%, an average 2.25 units is when the overlaid 75%, it becomes the average 1.91 units is when the overlaid 100%, it can be seen 25% overlaid reproduction method is most dominant.

According to the survey results, there were many opinions that individuals were able to hear, but 25% was

the strongest; it is considered that it is important to have the continuity of the speech. However, since the phoneme number of words that was used was a wide range of words 3-14, it was not possible to hear more words than the phoneme number.

In addition, the result of order for each word to be reproduced is shown in Fig. 4. This figure shows the percentage of match between the catch word and the reproduction word. From this, the highest was with the first, second and sixth words of the correct rate of the word. On the other hand, the third, fourth and fifth words are low. Therefore, it is considered that in order to understand more Web sites by placing a high importance on words around or near the exit-start playback.

Furthermore, the result of each reading method is shown in Fig. 5 by male voice and female voice. On the basis of Fig3, which is a low evaluations 75% and 100% were omitted. Similar to Fig3 this figure shows whether

Fig. 3. Number of words understood in each

Table. 1. Example of Important Words

		Important Words					
		1	2	3	4	5	6
Web Site	1	リニアモーターカー	タカラトミー	JR 東海	電磁石	中央リニア新幹線	時速 500 キロメートル
	2	常用労働者	パートタイマー	飲食サービス業	建設業	アルバイト	労働時間
	3	パブロ・サンドバル	レッドソックス	SNS	いいね	出場停止処分	メジャーリーガー
	4	はやぶさ2	打ち上げ	200日	イトカワ	C型小惑星	機器類
	5	永井大	中越典子	披露宴	8月2日	挙式	渡辺謙
	6	セキュリティー	マイナンバー	社会保障	サイバー攻撃	政官民	厚生労働相
	7	YASKAWA	安川電機みらい館	ロボット村	津田純嗣会長	モートマン	北九州市
	8	Windows 10	Windows 8	タブレット端末	無料アップグレード	タスクバー右端	リリース後 1 年間
	9	梅雨入り	雨雲	九州北部	九州南部	福岡管区気象台	くもり
	10	自転車運転者講習	ルール違反	安全運転義務違反	危険運転	改正道路交通法	反則金

(a) 3 from 1

(b) 6 from 4

Fig. 4. Correct-answer rate for each playback order

were able to hear how many units in six. As a result, an average 2.64 units is when superimposed 0%, an average 3.22 units is when the overlaid 25%, it becomes the average 2.75 units is when the overlaid 50%, it can be seen 25% overlaid reproduction method is most dominant. It was improved of the listening capability than when the male voice only.

Next, the result of order for each word to be reproduced is shown in Fig. 6. This figure shows the percentage of match between the catch word and the reproduction word. From this, the highest was with the second, fifth and sixth words of the correct rate of the word. Especially fifth words were improved of the listening capability than when the male voice only. On the other hand, the third and fourth words are low.

4. Conclusions

In this paper, the hearing of the important words, increased listening ability by superimposing the words to be

Fig. 5. Number of words understood in each reproduction method

Fig. 6. Correct-answer rate for each playback order

played later, and the word to be reproduced first is observed. In the future, it is expected to carry out the evaluation experiments in cases of changing the speed and size of the spoken word.

References

- (1) Manabi Miyagi, Tetsuya Watanabe, Kazunori Minatani, and Hideji Nakgaoka: " A survey of mobile phones, personal computers, and the Internet users with visual impairments" , The 8th Meeting of the Japanese Society for Low-vision Research and Rehabilitation , 2007.
- (2) Karen Sparck Jones: " A statistical interpretation of term specificity and its application in retrieval " , Journal of Documentation , 28 , 11-21 , 1972
- (3) Mitome Yukio: "Applications and a future prospect of speech synthesis" , THE JOURNAL OF THE ACOUSTICAL SOCIETY OF JAPAN , 49(12), 875-880, 1993-12-01
- (4) Nelson Cowan: " The Magical Number 4 in Short-term Memory: A reconsideration of mental storage capacity " , BEHAVIORAL AND BRAIN SCIENCES , 24 , 87-185 , 2000