Design of an Assistive System of Online Video Lecture for Visually Impaired Students

Dong-Yeon Park
Dept. of IT Engineering
Sookmyung Women's University
Seoul, Republic of Korea
yeon0729@sookmyung.ac.kr

Soon-Bum Lim
Dept. of IT Engineering
Sookmyung Women's University
Seoul, Republic of Korea
sblim@sm.ac.kr

Debbie Ko
Dept. of Software Engineering
Bethlehem University
Bethlehem, Palestine
hongheeko@bethlehem.edu

***Abstract*— We propose an assistive system to improve understanding when visually impaired students take online video lectures. The system generates a commentary text by dividing the content of the video into static and dynamic elements. Typically, static elements include screen data slides, and dynamic elements include the instructor's reference and emphasis action. When the commentary is provided as a voice, the proposed method will contribute to an increase in the understanding of visually impaired students.**

Keywords—visually impaired student, online learning, video lecture

I. Introduction

To prevent the spread of covid-19, lots of educational institutions have introduced a non-face-to-face lecture system[1]. However, the accessibility of the non-face-to-face lecture system is not properly guaranteed, so the digital weak is suffering from many inconveniences[2]. In particular, most of the non-face-to-face lecture systems are based on visual elements, but visually impaired students cannot use vision. Typically, the understanding level of the lecture contents of the visually impaired student decreases because it's impossible to know what contents appear on the screen of the video lecture[3,4]. Therefore, in this paper, we propose an assistive system that automatically provides voice commentary for video lecture screens for visually impaired students.

II. Design of Speech Commentary Assistive System

We propose a system that explains the elements on the video screen by voice. These elements can be largely divided into static and dynamic elements. The static element refers to a screen on which lecture materials are displayed. A typical example is a 'Microsoft PowerPoint Slide' format. Dynamic elements refer to emphasized actions used by instructors to help or supplement content understanding during class. The system automatically generates commentary by dividing the video content into static and dynamic elements. Finally, the system converts commentary text files into voice files using the TTS engine and merged them into the original video lecture.

III. Static Screen Data Commentary

First of all, the system explains the screen data that continuously appears in a common form in the video. There are various types of screen data for video lectures, but in this study, we limit the category of screen data to slide type. The system analyzes the elements of the screen data and provides a commentary suitable for the characteristics of each element. Representative elements: text, picture, table, graph, formula. (text) The system reads text from top to bottom and from left to right in the same order as the normal reading. (picture) The system checks whether the picture is captioned or not, and if there is no caption, explains the content of the picture in one or two sentences. (table) The system reflects the structural matrix information of the table and explains it in units of cells. (graph) The system reads the title of the graph, the title of each axis, and the unit. (formula) The system reads the formula according to the four-step rule of the previous study[5].

IV. Using the Template

There are various types of emphasis actions in the dynamic elements of video lectures. As a result of analyzing 80 lecture videos of 8 different types, it could be divided into 4 types as follows: movement of a cursor, note action, screen display action, referent pronoun. The movement of the cursor is a method of emphasizing a target object as a movement near the object. Note action generally means handwriting. It is characterized by drawing free forms of text and pictures using digital pen tools to supplement and explain the contents of the lecture. Screen display actions refer to and emphasize objects using digital pen tools, such as note actions. However, it is characterized by its relatively regular shape, such as underlines, circles, and squares. The referent pronoun is using the indicator pronoun to refer to the target object. However, it is not used alone but is usually used with other emphasis actions. In this way, identifying and explaining the characteristics of each emphasized action and the subject of reference can contribute to improving the understanding of visually impaired students.

Conclusion

In this paper, we proposed a video lecture learning assistive system that can be used by visually impaired students. This system allows visually impaired students to learn independently by explaining the contents of the screen. This study can find significance in guaranteeing the infringed right to study for visually impaired students. In addition, if the proposed system is actually implemented, it will be able to contribute greatly to improving the class understanding of visually impaired students.

References

[1] Adnan, M., & Anwar, K, "Online Learning amid the COVID-19 Pandemic: Students' Perspectives," Online Submission. vol. 2, no. 1, pp. 45-51, 2020.

[2] Bae, Y.I., & Shin, H.L., "The Shadow of the Non-face-to-face Era," Digital alienation Issue & Diagnosis, pp. 1-25, 2021.

[3] Yoo, J.Y., & Oh, C.W., "A Study on Corona19 Experience among Visually Impaired People as Disaster Vulnerability Group," The Korean Journal of Visual Impairment, vol. 36, no. 3, pp. 1-22, 2020.

[4] Park, D.Y., Jo, S.R., & Lim, S.B., "Requirement Analysis of Visually Impaired Students in the Non-Face-to-Face Lecture Environment," Proceedings of the 17th International Conference on MITA, 2021.

[5] Hwang, J.S., & Lim, S.B., "A Study on Phased Reading Techniques of Mathematical Expression in the Digital Talking Book," Journal of Korea Multimedia Society, vol. 17, no. 8, pp. 1025-1032, 2014.

This research was supported by the Step 4 BK21 plus program through the National Research Foundation (NRF) funded by the Ministry of Education of Korea.