Development of Multimodal e-Slide Presentation to Support Self-Learning for the Visually Impaired

Rustam Asnawi and Wan Fatimah Wan Ahmad

Abstract—Currently electronic slide (e-slide) is one of the most common styles in educational presentation. Unfortunately, the utilization of e-slide for the visually impaired is uncommon since they are unable to see the content of such e-slides which are usually composed of text, images and animation. This paper proposes a model for presenting e-slide in multimodal presentation i.e. using conventional slide concurrent with voicing, in both languages Malay and English. At the design level, live multimedia presentation concept is used, while at the implementation level several components are used. The text content of each slide is extracted using COM component, Microsoft Speech API for voicing the text in English language and the text in Malay language is voiced using dictionary approach. To support the accessibility, an auditory user interface is provided as an additional feature. A prototype of such model named as VSlide has been developed and introduced.

Keywords—presentation, self-learning, slide, visually impaired

I. INTRODUCTION

ONE of the codes of ethics in the Association of Computing Machinery (ACM, 1992) states that: "In a fair society, all individuals would have an equal opportunity to participate in, or benefit from, the use of computer resources regardless of race, sex, disability, national origin or other such similar factor" (also cited by [1]). The statement explains that everyone in this world should have the opportunity to use Computing Machines, which regardless of who they are, what they are capable of or not on utilizing computer resources. This also applies to the blinds or visual impaired person, they should have equal opportunity to utilize computer resources. It has been reported that in this world people with visual impairment are reaching up to 135 million [2] compared to the world population, which is approximately 6 billion people. According to [3] and also cited by [4], based on a National Eye Survey conducted in 1996, in Malaysia about 2.44% of the populations have low vision and 0.29% are blind. Exactly, after more than fourteen years, today the population of people with visual impairment has increased. This is a significant number, which means, their needs for computer resources in order to support them for a better life should be considered seriously.

Meanwhile, currently almost all universities in Malaysia are using electronic slide (e-slide) application as the teaching-learning tool. The most popular e-slide applications are

Rustam Asnawi is with Computer and Information Sciences Department, Universiti Teknologi Petronas, Bandar Sri Iskandar 31750, Tronoh, Perak, Malaysia. phone: 0175663357; (e-mail: rustam@uny.ac.id).

Wan Fatimah Wan Ahmad is an Associate Professor in Computer and Information Sciences Department, Universiti Teknologi Petronas, Bandar Sri Iskandar 31750, Tronoh, Perak, Malaysia. (e-mail: fatimhd@petronas.com.my).

Microsoft Power Point in Windows platform and Open Office Presentation in Linux platform. It is very possible that in the future schools also will use the same tool, whereby teachers will use e-slide presentation instead of chalk board, whiteboard and overhead projector. When that happens, normal students would not encounter any problem, but it could have a serious impact for students who have visual impairment or blind. For persons with visual impairment, they often have difficulties to look and examine e-slides that contain various animations and colour combinations. As for the blinds, their problem is even greater since they can neither see nor able to read any of the slide content.

This paper is proposing a model for presenting such e-slide in multimodal presentation. An e-slide is presented either in conventional slide show or in auditory slide presentation in either Malay language or English language. By our concept, any e-slide files (such as PPT) that are composed of text material can be voiced in either Malay or English language. In the design, live multimedia presentation concept is used. In the implementation, the text content in each slide is extracted using COM component. The main concept of voicing such slide is done by using text-to-voice, where voicing of text in English language is accomplished by utilizing Microsoft Speech API and for Malay language dictionary approach is used. Additionally, to support accessibility, an auditory user interface features is available. As a result, a prototype of such model has been developed (named as VSlide).

II. RESEARCH MOTIVATION

According to [5] and also cited by [6], in entire the worldwide there are more than 30 million presentations in one day using PowerPoint (e-slide), and also the PowerPoint has been installed in around 250 million computer. Consequently, there are many high-quality educational materials stored in e-slide file format available and can be easily accessed either via online (internet) or offline (class room material).. Unfortunately, in their present file format, these materials cannot be benefited by the students with visual impairment.

The recent version of Microsoft Windows operating system comes in handy with a provision of the Microsoft Narrator application. Microsoft Narrator is able read all programs' menu on the screen by focusing on one program at a time. The Microsoft Narrator can guide the blind to use this system, but as Windows platform is in English language, Microsoft Narrator is also in English language. In this situation, the user must have an understanding of English language first before they can use the system independently. Besides, the Microsoft Narrator is rather confusing when it is used to voice an e-slide since it will voice all texts that appear on the Power Point's menu and its slide show. In other words, voicing e-slide is

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

different from voicing texts on screen. In e-slide voicing, there is a (implicit) temporal constraint between presenting the slide and generating the associated voice. Therefore, voicing e-slide needs synchronization between the slide show and its text's voice.

According to [4], based on an observation of the application of assistive courseware for visually impaired learners, it is very important to use audio while developing such assistive courseware. The observation result indicated that 100% of the participants preferred to utilize the audio. Therefore, we propose an idea to voice any e-slide content in order to assist the visually impaired students when examining such e-slides.

III. SYSTEM ARCHITECTURE

The overall system architecture is as illustrated as Figure 1. The core component to digest the e-slide file is COM (*Component Object Model*) server (descendants of the *Ole Server*) [7-8]. By this component an e-slide file can be extracted into its elements including all text elements. Moreover, many important properties and functions of the e-slide can be obtained and modified programmatically. Some of the important properties and functions are: coordinate position of slides on screen (top, left position), size (width and length), name of slide name, index of slide index, and slide navigation functions (first, previous, next, last).
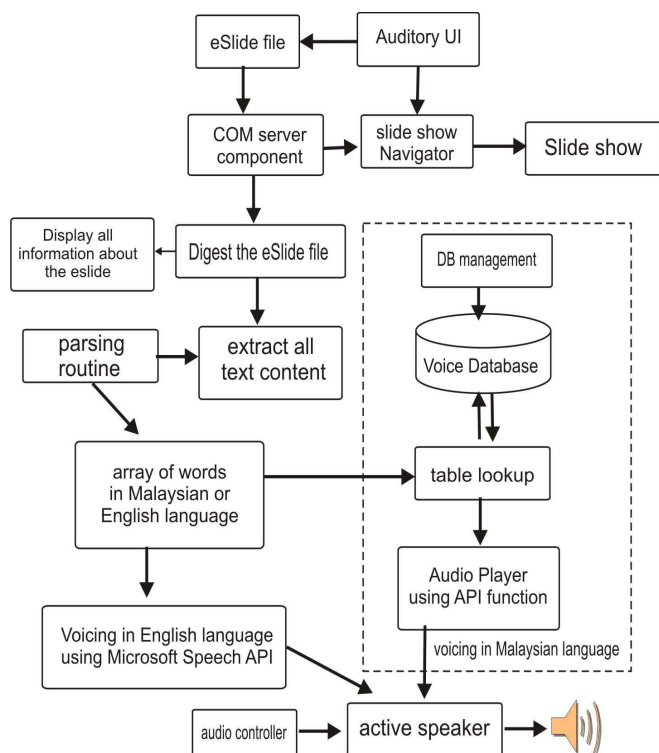

Fig. 1 System Architecture of VSlide

A piece of slide may consist of some type of shapes, and a shape may have a text frame. Extracting text frame is performed for all slides that exist in an e-slide file. The output

of this step is an array of sentences, thus all sentences need to be parsed into an array of words.

A. Auditory user interface

As we know, a visual impaired person lacks clarity of vision to see any object accurately. Consequently it is anticipated that they will meet some problem and difficulties to operate an application program if they have to use mouse and keyboard as the input devices. However, accessibility to the application can be improved by enhancing the user interface with multimodal user interface such as auditory button or audio annotation [9] combined with mouse and keyboard. In this model, we use the combination of clicked mouse, hotkey and voicing of a button's function when the mouse moves over it.

B. Parsing Routine

In Malay language, commonly in a sentence each word that has meaning is separated by a white space. Based on this assumption, a conventional parsing algorithm to parse a sentence into an array of words is developed. Below is the algorithm.

```
function Parsing
(input parameter: string StringInput;
 Output parameter: StringOut array of
string;
 Output function: integer)

 Variable declaration and initialization
 k <- 1, i <- 0, s <- StringInput;
 for j:=1 to length of StringInput do
 begin
   if (s[j] = ' ') then // encounter a
specific character
   begin
     StringOut[i] := copy(s,k,j-k);
     i:=i+1;
     k := j+1;
   end
 end;
 s := copy(s,k,j-k);
 if s <> '' then
 begin
   StringOut[i] := s;
   i:=i+1;
 end;
 result := i;
end;
```

The above function is called at each step of sentence parsing. As a result an array of words is stored into a global array.

C. Text to Voice

The next step is to voice all content of such global array. Consider Fig.1, VSlide provides two choices of languages i.e. Malay and English. For the *text to voice* in Malay language, we use dictionary approach to perform this task. Using a table lookup routine, a word that comes from the global array is

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

matched in the Malay words database. If the word is available in the dictionary then a matching voice file in the database will be looked up. The voice file is a stored audio file of the related word. Thus the relation between a Malay word and its voice file is one-to-one relation. Developing the dictionary and its voice is managed by the database management routine. Since, the aim of this project is to develop a prototype of such system (showing the main functions and specification of such system), the content of the dictionary is the common and popular words that are usually used by Malaysian blind school. However, the database management routine provides a feature for adding new content of the dictionary. Finally, API function is exploited for playing the voice file.

Complementary, for the *text to voice* in English language we use Microsoft Speech API [10] as the core component. Using this component, building the routine for the text to voice in English language is very simple, it involves just a few lines of codes.

D. Using LMP Applications Concept

The design of the VSlide model is inspired and approached by the *live multimedia presentation* (LMP) application concept. Based on the authors' earlier works [11-14] a concept for developing live multimedia presentation application has been introduced.

There are four basic concepts of LMP application development which are:

1) Live-integrated controller
2) Two groups of user (presenter and target user),
3) View separation, and
4) Temporal and spatial constraints consideration.

Since e-slide is a *metamedia* [14], it should has equal treatment with other multimedia element (video, audio, text, image and animation). In other words, in terms of multimedia presentation, while presenting all multimedia elements the e-slide should have similar controller functionalities as other multimedia element controllers. For example, it should have a particular-independent controller such as start button for starting a presentation, next button for next slide, etc. Moreover, a special output stream should be provided for this *metamedia* data.

There are two groups of user:

1. Teacher, a teacher is a person who is responsible to deliver all content of the material to the student (person with visual impairment). This user group has the responsibility to deliver all material content to the student. For preparation, this user acts as an administrator who must prepare and settle all materials and VSlide parameter.
2. Visual impairment student, a visually impaired user is the student who needs to understand all content of the presentation material. This user group has the ability to use either auditory or haptic user interface.

Following model development of the multimodal e-slide presentation, a trial implementation was conducted at a disability school in Batu Gajah, Perak, Malaysia to verify the model. The next section will describe the result of the implementation. i.e. a prototype application.
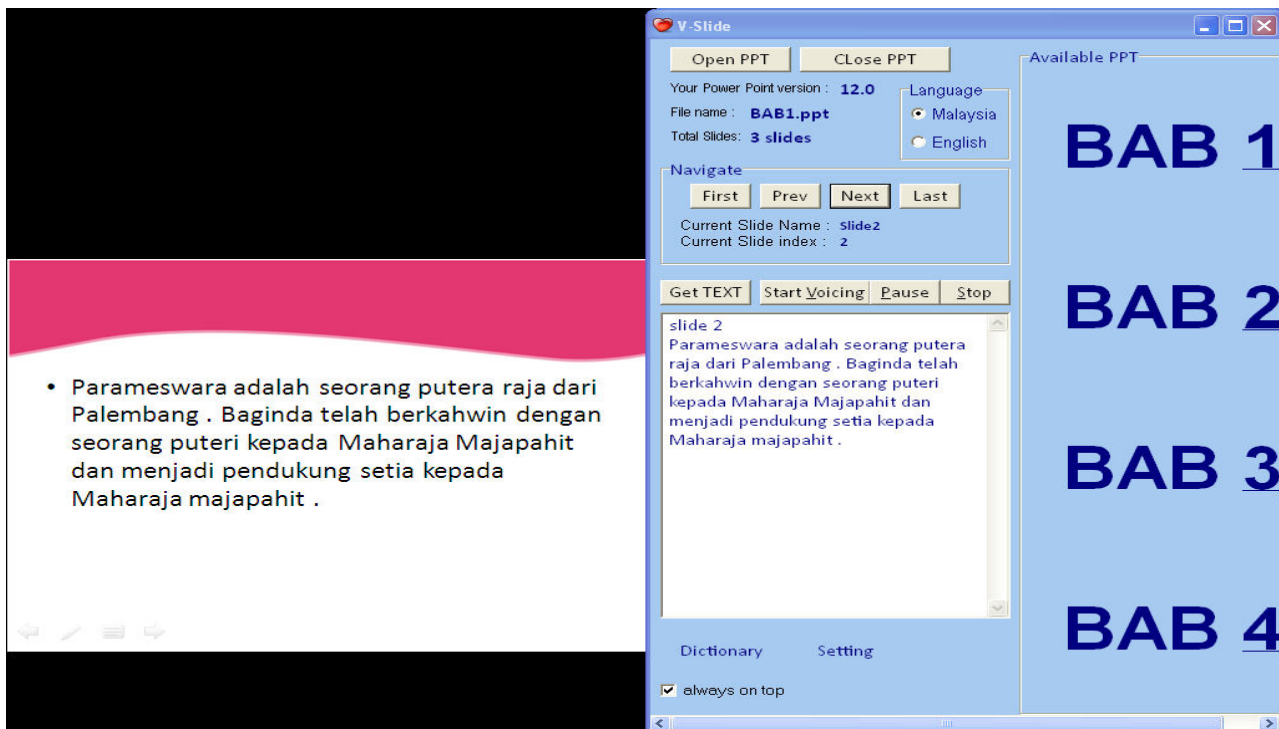

Fig. 2 Screenshot of VSlide

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

IV. RESULT AND DISCUSSION

A. The Prototype

In order to fulfill the requirement of the LMP application concept, a view separation style is applied. Consider Figure 2, the whole screen is divided into two equal parts, the left part is provided to display the slide show, and the right one is the control panel of the user interface. In this case, four different e-slide files are available for the student. For each button an extra auditory user interface is added. When the user moves the mouse cursor on top of such button, a voice about such button's function will be generated automatically.

For an example, if the user moves the cursor over "BAB 1" button, then the computer will generate "bab 1" voice automatically. So, the user can directly click the left button of the mouse to open BAB1.ppt file. Alternatively, the user also can open the PowerPoint file by pressing a hotkey (combination keys alt+1). Once opened, the PowerPoint file will be displayed in slide show mode on the left part of the screen. By default, when the file first open it will show the first slide and its text content, concurrently with the voice of the text.

To continue on to the next slide, the user just need to right click the mouse's button (or click the next button). In the case of Fig. 3, voicing of slides are in Malay language (default). To change to English language, just click the English menu items inside the radio box. The OpenPPT button menu is provided when the user needs an extra e-slide file that does not exist in the four main buttons
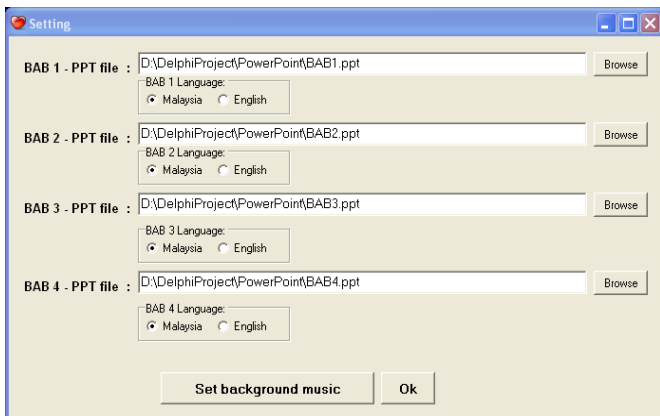

Fig. 3 The VSlide setting window is aimed to set a PowerPoint file associated with a particular button and its language

B. Voice Database Management

In the prototype, the voice database of Malay language is rather poor and needs to be improved. VSlide provides a special menu for adding, deleting, editing and modifying the database content in a simple way. Figure 4 depicts the user interface for modifying the voice database (Figure 4 can be obtained by clicking the dictionary button in Figure 3).

By this interface, any user (with normal vision) can update and modify the content of the dictionary including new voice recording. In the context of visual impairment school, this task can be accomplished by the teacher or someone who has the responsibility for teaching learning process in that school.

C. Background music

As a preliminary research, we gathered some requirement and information related to such model. An interview with teacher and student in visually impaired school was conducted. One of the requirements is they prefer to utilize background music rather than merely voicing text. To attain this task, an additional feature to play back music (in mp3 or wav format) should be developed. This music player should have an independent controller, but it must be integrated with VSlide. An audio processing such as bass booster and treble enhancer is provided to improve the performance of the player.
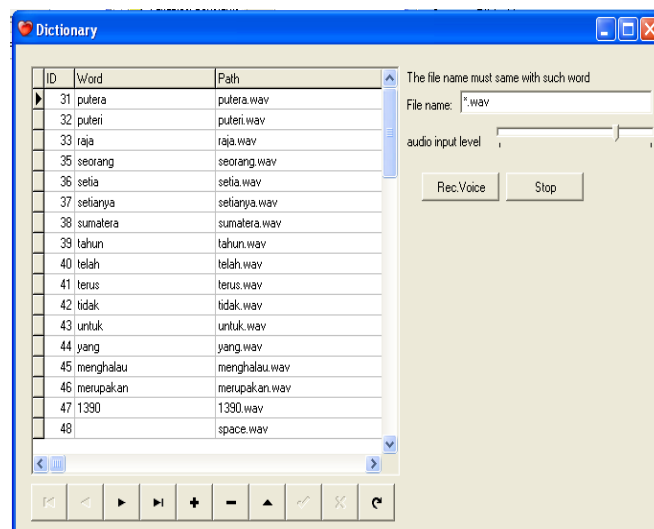

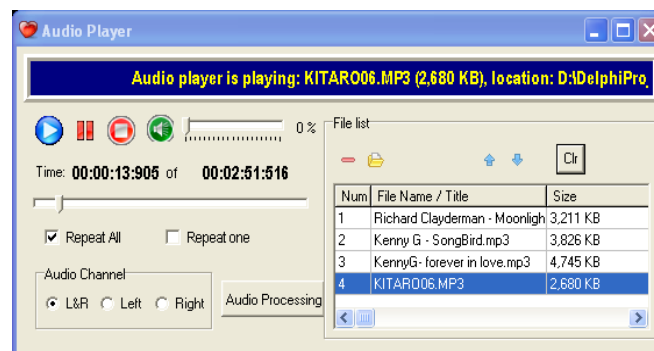Fig. 4 VSlide voice database using dictionary-like approach

Fig. 5 VSlide audio player for playing background music in mp3 files

D. Testing in Visual Impairment School

A field trial of this prototype has been carried out involving potential users at the visually impaired school "Pendidikan Khas" at Sekolah Kebangsaan Sultan Yussuf, in Malaysia. As a result, generally the teacher and the visually impaired students have accepted the VSlide as a tool to support their self learning.

V. CONCLUSION AND FUTURE WORKS

This paper describes the development of a multimodal e-slide presentation for assisting students with visual impairment

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

in their self-learning program. A prototype has been successfully developed to embody our model using rapid application programming (Delphi). The prototype has been tested in a visual impairment school in Perak, Malaysia. The prototype has been well received as most of the academic staffs in that school. Also they said that the system enable to encourage their self learning.

As a matter of fact, the VSlide is not limited to those with visual impairment only. Since it enables voicing of any text on an e-slide, thus VSlide can be utilized by normal people to support their self-learning by listening to the audio instead of reading the texts on the slide.

For future work, we are suggesting for improvement of the voice dictionary using other better algorithm of voice generating. In this prototype version of VSlide all sentences, regardless of the type of punctuation marks, are voiced in the same tone and rhythm (flat). This should be improved to generate special tone and rhythm when meeting special punctuation marks such as point, comma, exclamation mark and question mark.

REFERENCES

[1] W. Mike, "Learning Through Multimedia: Automatic Speech Recognition Enabling Accessibility and Interaction," in *Proceedings of ED-MEDIA 2006: World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 2006, pp. 2965-2976.

[2] N. A.Bradley and M. D. Dunlop, "An Experimental Investigation into Wayfinding Directions for Visually Impaired People," *Journal Personal and Ubiquitous Computing,* 2005.

[3] M. Zainal, "Prevalence of blindness and low vision in Malaysian population: results from the National Eye Survey 1996," *British Journal of Ophthalmology, ,* 2002.

[4] N. Aziz, "Assistive Courseware for Visually-Impaired," in *LNCS 5857, IVIC Proceeding* Kuala Lumpur, 2009.

[5] Michael Alley and K. A. Neeley, "Discovering the Power of PowerPoint: Rethinking the Design of Presentation Slides from a Skillful User's Perspective," in *Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition*, 2005.

[6] April Savoy*, et al.*, "Information retention from PowerPointTM and traditional lectures," *Journal Computer & Education, Elsevier,* vol. 52, pp. 858-867, 2009.

[7] Microsoft. (2011, July, 29th). *COM: Component Object Model Technologies*. Available: http://www.microsoft.com/com/default.mspx

[8] Microsoft. (2011, July, 29th). *Component Object Model (COM)*. Available: http://msdn.microsoft.com/en-us/library/ms680573%28v=VS.85%29.aspx

[9] Hilko Donker*, et al.*, "The Design of Auditory User Interfaces for Blind Users," in *Proceedings of the second Nordic conference on Human-computer interaction* 2002.

[10] Microsoft. (2011, July, 29th). *Speech Software Development Kit 5.1*. Available: http://www.microsoft.com/download/en/details.aspx?id=10121#overview

[11] Rustam Asnawi*, et al.*, "Design Architecture for IM-Player as a Tool for supporting Visual Education Presentation," in *Lecture Notes in Computer Science-IVIC Kuala Lumpur, Malaysia*, 2009, pp. 78-89.

[12] Rustam Asnawi*, et al.*, "Development of Multipurpose Integrated Multimedia Presentation System Based on the CCM Algorithm," in *Proceedings of International Conference on Computer and Automation Engineering*, Singapore, 2010, pp. 716-720.

[13] Rustam Asnawi*, et al.*, "Unifying Multimedia Player and Presentation System with Utilizing Dual-Display Feature," in *Proceedings of The 4th International Symposium on Information Technology* Kuala Lumpur, Malaysia, 2010, pp. 70-75.

[14] Rustam Asnawi*, et al.*, "Formalization and Verification of a Live Multimedia Presentation Model," *International Journal of Computer Applications,* vol. 20(3), pp. 32-41, 2011.

Rustam Asnawi,received his bachelor's and master's degrees in Computer System and Informatic from Electrical Engineering, Gadjah Mada University, Indonesia in 1995 and 2004. He is a senior lecturer in Electrical Engineering department, Yogyakarta State University, Indonesia. Currently he is a PhD student in Computer and Information Sciences Department, Universiti Teknologi PETRONAS, Malaysia. His interest researches are multimedia application and presentation system.

Wan Fatimah Wan Ahmad received her B.A and M.A. degrees in Mathematics from California State University, Long Beach, California U.S.A. in 1985 and 1987. She also obtained Dip.Ed. from Universiti Sains Malaysia in 1992. She completed her PhD in Information System from Universiti Kebangsaan Malaysia in 2004. She is currently an Associate Professor in the Information Technology/Information System program at UTP. She was a lecturer at Universiti Sains Malaysia, Tronoh before joining UTP. Her main interests are in the areas of mathematics education, educational technology, e-learning, mobile learning and multimedia.