

E-Commu-Book: An Assistive Technology for Users with Speech Impairments

Ka Ho Wong, Wai Kim Leung and Helen Meng

Stanley Ho Big Data Decision Analytics Research Centre,
The Chinese University of Hong Kong,
Hong Kong SAR, China

khwong@se.cuhk.edu.hk, wkleung@se.cuhk.edu.hk, hmmeng@se.cuhk.edu.hk

Abstract

We are developing a voice-enabled, mobile and customizable electronic communication book (e-Commu-Book) as an aid for people with speech impairment. When a user touches an image icon on the e-Commu-Book, the caption under the image icon will be read out by speech synthesis. The caregivers of the users can help customize contents to fit the users' needs. The record of a user's behavior can be uploaded to the cloud for future analysis so as to improve the user experience of the e-Commu-Book.

Index Terms: assistive technology, web technology, speech impairment, big data

1. Introduction

Users with speech impairment may have difficulty in their daily life. Speech therapists attempt to use assistive tools to help these patients communicate more effectively. The "communication book" [1] is a traditional hardcopy image book that has been used as an assistive tool. Users can express their thinking by pointing to the indexed vocabulary images in the book. However, traditional communication books cannot be personalized to fit the users' needs, e.g. adding the photos of a user's family members to make it easier to refer to them. By taking advantage of the advancements in web technologies, mobility and hand-held devices, we can develop an electronic communication book (or e-Commu-Book) that have enhanced capabilities compared to the traditional communication book. The e-Commu-Book provides personalized content, fluid displays and layouts, together with synthetic speech outputs. The e-Commu-Book also logs the users' behavior for analysis in order to improve the user interface.

In the next section, we shall discuss the development platform of the e-Commu-Book. In Section 3, we present the user interface. In Section 4, the design of web services and the cloud support for big data will be discussed. Finally, we present our conclusions in Section 5.

2. Cross-Platform Development

The e-Commu-Book is browser-based and there is no need for native plugins, so it is compatible with a variety of devices, such as desktop PCs, tablets, iPads and smartphones. The challenge of cross-platform development includes the compatibility of the browsers' API. For example, `getUserMedia` is used to access the device's camera. It is supported by Android 5.x WebView and Firefox browser but not in iOS Safari 8.1 or

Internet Explorer 11.¹ In a cross-platform environment, it is difficult to maintain platform-specific native plugins. This can be solved by using an open source Cordova platform [2].

Cordova incorporates different system functions (e.g. storage access) from a variety of platforms into a set of JavaScript API. Any applications built on Cordova can be easily converted into different platforms without changing any codes (see Figure 1). Developers can use cross-platform languages such as HTML 5 or JavaScript to develop Cordova applications.

3. Capacity for Personalization

3.1. User Interface (UI)

The UI of the e-Commu-Book is shown in Figure 2. The left side is a menu of "actions". When an action category is selected, all the action items are shown on the right in terms of labelled image icons. Some subjects may have difficulty manipulating fine-level finger controls. Hence, most of our controls only need a single touch, a drag-and-drop action with a single finger only. Also, an image icon will be automatically enlarged when a user touches it (Figure 2).

On the e-Commu-Book, users can re-order the categories and images by drag-and-drop actions. They can also delete a category or an image by dropping it to the recycle bin, or by clicking the deletion button (Figure 3). Caregivers can help the users to add their own words and items. For example, a user may want to add his family members to the book. The caregiver can take a photo of a family member and write the name corresponding to the photo (Figure 4). This name is synthesized by our homegrown Cantonese text-to-speech engine and the wave file is automatically generated in the cloud and downloaded to the user's client device.

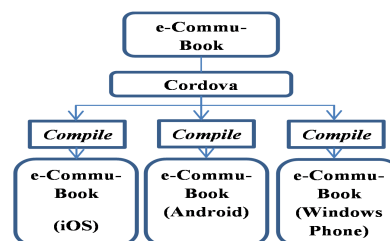


Figure 1: How to convert a generic Cordova application to a platform-specific application.

¹ "Can I Use": <http://caniuse.com/#search=getUserMedia>

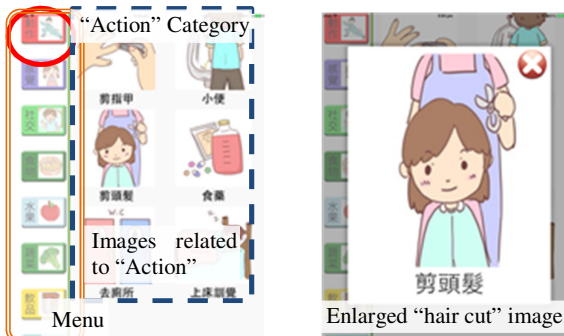


Figure 2: When user touches the “hair cut”(剪頭髮) image (left), the image will be enlarged (right).

3.2. User Profile and User Behavior

The user profile is designed to store personalized content, layout preferences and user behavior. The user profile aims to allow users to synchronize the content across different devices. The record of user behavior will be uploaded to our cloud for future analysis. The record includes any events about touch, drag-and-drop, actions selection, etc.

4. Web Services and Database

4.1. RESTful APIs

RESTful is a web service in which the software architecture follows a REST (REpresentational State Transfer) style [3]. A client uses a HTTP request to access our RESTful web services and retrieves the required data in JSON (JavaScript Object Notation) format.

JSON [4] is a lightweight data-interchange format. It is easy for both humans and machines to understand. WebKit¹ is an open source layout engine software component for rendering web pages which is used by Google’s Chrome, Apple’s Safari, etc. Webkit supports JSON directly. The e-Commu-Book is using HTTP POST/GET methods to exchange data among clients and servers in JSON format. A client can upload/restore the user profile and download new synthesized speech from the server.

4.2. Big Data Database

The development of the e-Commu-Book will be a continuous process that involves participation of the user base. New functionalities will be added as needed and this involves updating the database schema. Also the system will log user actions for analyzing user behavior.

We have chosen MongoDB to be the database engine. MongoDB is a free, open-source, cross-platform, document-oriented, NoSQL² database. NoSQL can deal with the challenges of large volumes of data, frequently changing database schemas, and multiple servers in the cloud. In MongoDB, the database schema can be dynamically edited to respond to the continuous development. Multiple servers can increase the stability of the entire suite of web services. MongoDB provides a set of functionalities for a multiple-server environment,

¹ WebKit: <http://www.webkit.org/>

² NoSQL: <http://www.mongodb.com/nosql-explained>



Figure 3: Users can delete an image icon by touching the cross button (left) or by drag-and-drop into the recycle bin (right).

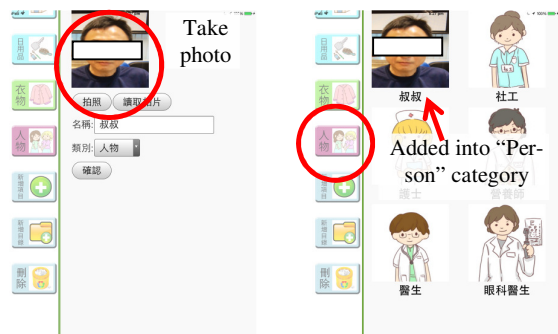


Figure 4: Caregivers can take the subject’s uncle photo (left) and add into his personalized e-Commu-Book (right).

e.g., replication, load balancing, batch processing data and aggregation operation (e.g. MapReduce).

5. Conclusions

We have developed an assistive technology, the e-Commu-Book, to help users with speech impairment. The personalized user interface allows users to adjust the layout. The record of user behavior can support behavior analysis in the future. More functions such as downloading and sharing new categories, will be implemented. A demonstration of the e-Commu-Book video can be found in:

<http://www.se.cuhk.edu.hk/~khwong/ecommu-book/ECB.mp4>

6. Acknowledgements

This project is partially supported by a grant from the HKSAR Government’s Innovation and Technology Support Program Public Sector Trial Scheme (reference no. ITT/006/13GP).

7. References

- [1] Coordinating Committee in Speech Therapist, Communication Book, the Hong Kong Hospital Authority.
- [2] The Apache Software Foundation, "Apache Cordova," October 2012. [Online]. Available: <https://cordova.apache.org/>. [Accessed 2015].
- [3] R. T. Fielding, "Architectural Styles and The Design of Network-based Software Architectures," PhD Dissertation. Department of Information and Computer Science, Irvine, 2000.
- [4] *ECMAScript Language Specification*, ECMA-265, June 2011.