

## Effect of Enhancing Learning through Annotation Similarity and Recommendation System

Jae-Kyung Kim<sup>1</sup>, Won-Sung Sohn<sup>1\*</sup>, Kyung Hur<sup>1</sup> and YangSun Lee<sup>2</sup>

<sup>1</sup>*Dept. of Computer Education, Gyeongin National University of Education,  
Incheon, Korea*

<sup>2</sup>*Dept. of Computer Engineering, Seokyeong University,  
Seoul, Korea*

{kimjk, sohnws\*, khur} @ginue.ac.kr<sup>1</sup>, yslee@skuniv.ac.kr<sup>2</sup>

### **Abstract**

*With the use of smart mobile devices, learning opportunities have been greatly increased in education. Especially, digital textbooks are the new primary digital instructional resource for elementary and middle schools. The use of electronic textbooks in a mobile learning environment assists collaborative learning and increases access to educational resources. Learners are now able to share their bookmarks, notes and highlights both on and off campus. These annotation features can lead to enhanced learning effect for students. Digital textbooks still face several hurdles as the value of using electronic textbooks for students remains unclear. Major reasons limiting the uptake of digital textbooks include usability and design issues. One of the most important features in education is interaction between resources and students such as annotating or highlighting. To enhance value of digital textbooks, we designed a prototype note-taking interface reflecting students learning pattern and self-directed learning. The interface and functionality were defined in order to support touch-based annotation input and cooperative learning. It has been implemented a touch-based device for smart education environment.*

**Keywords:** Annotation, Collaborative learning, Smart education

### **1. Introduction**

With the increasing adoption of tablets in education, the tablet pc would allow students to use digital textbooks instead of the traditional printed copies [1, 2]. For students, the textbooks may no longer need sturdy backpacks and arms in the near future. Many studies and articles discuss about the advantages of electronic textbooks are obvious.

The main advantage of electronic textbooks is that students can access textbooks anywhere, anytime, without the need to print or carry physical books. Electronic textbooks are usually cheaper than the classic paper books, because large institutions such as universities may license digital content on more favorable pricing terms. When paper textbooks became outdated, students need to buy a new book, while digital content can be constantly updated for a small additional cost. In addition, by incorporating traditional textbook content with the interactive capabilities of the Internet, digital textbooks can better prepare learners for learner-centered paradigm. For example, in tablets reading environments with proper social

---

\* Corresponding Author

media features, it will allow students to share bookmarks and comments with one another as shown in Figure 1.

In early 2012, U.S Government challenged the nation's schools to adopt electronic textbooks within five years. In late 2010, Japan Government, MEXT (the Ministry of Internal Affairs and Communications) proposed deployment of digital textbooks to all elementary and junior high school students until 2015, in the "Haraguchi – Vision" [3]. In South Korea, the digital textbook program was announced by the Education Ministry of South Korea on March 8, 2007. The digital textbook is currently being tested in several primary schools and will be distributed free to every school nation-wide by 2013 [4].

However, the use of digital textbooks in education could be limited if it is not designed properly, and some students are not receptive to the idea of using them [5, 6]. Students like to annotate and highlight, but doing them on touch device requires more time and effort. These user experience and usability effect on the learning process. Poor design and functionality may cause negative effects in learning attitude [7, 8, 9].

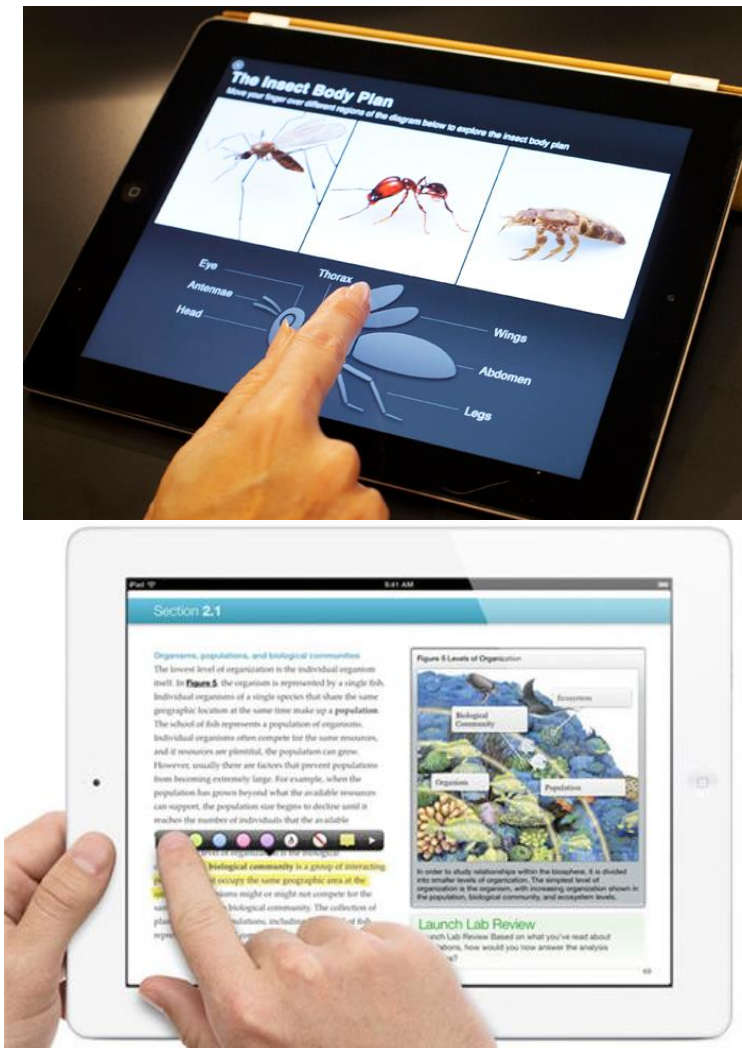


Figure 1. Interactive features of digital textbooks

Therefore, we designed and implemented the knowledge recommendation interface based on the proposing annotation-contents model. Our system provides note-taking style interface for students to annotate their work. Thus, annotated information in digital textbooks is analyzed and stored in a database according to the model. The similarity algorithms can find other annotations created by students who have top grade or similar learning style. Thus, the student can access other annotation related to his or her learning style. The main idea of the system is to get more information from annotation about our students, and to put students at the center of the learning system

In this paper, we propose annotation-contents model and algorithm in Section 2, implemented system in Section 3, evaluate it in Section 4. Then we conclude in Section 5.

## 2. Related Works

Many electronic book applications on tablets and readers support annotation functionality such as a note-taking and highlighting. For instance, the Kindle application allows a user to create highlights and take notes in reading materials the same as on a Kindle device. Any annotations a user made in a book or personal documents are automatically synchronized across devices. Amazon saves notes and highlights so that they become a part of user data called “Popular Highlights”, a collection of highlights by other readers of the same book that can be viewed on both Kindle and the Amazon Kindle website as shown in Figure 2.

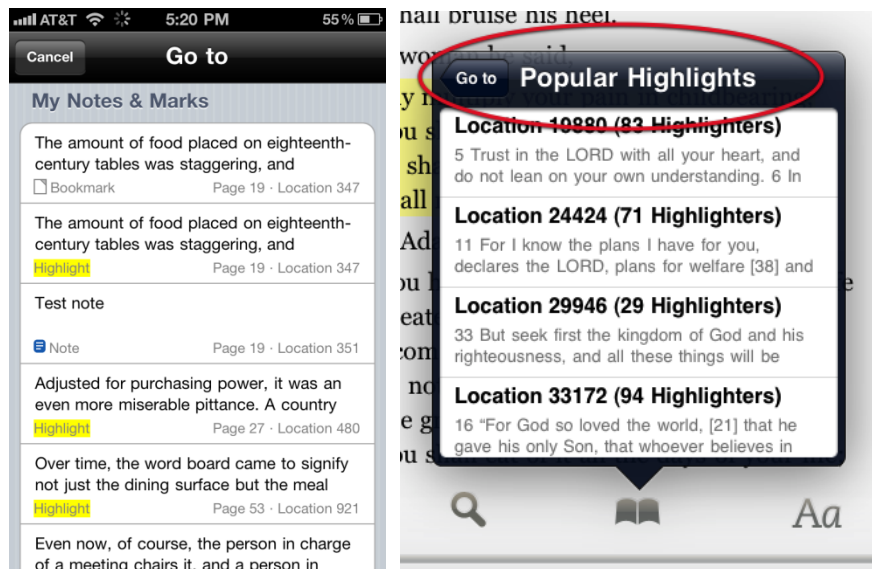


Figure 2. Highlights feature of Kindle application

Thus, a user can read the same notes on the other Kindle devices. It also allows a user to highlight images, charts and figures within print replica textbooks on iPad, in addition to text. A user can view all available pages which are bookmarked, highlighted, or annotated. After a user has located the section in the list, tap that item to jump to the page on which the note or mark is found [10].

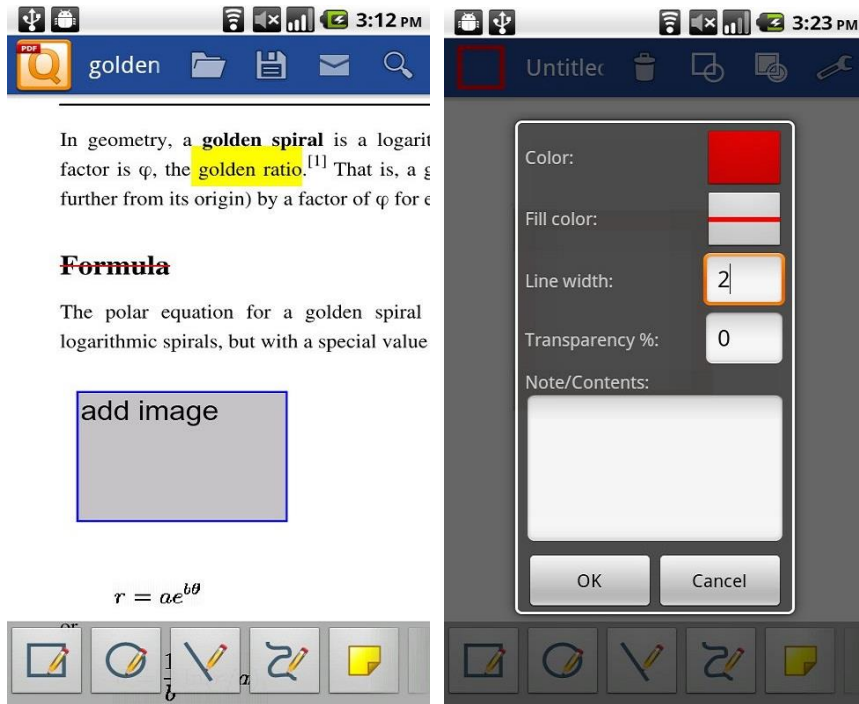


Figure 3. Note interface of qPDF reader

PDF documents are still popular book format in the market. Many PDF readers support viewing and adding comments to Adobe PDF files and fill blank forms on mobile phone or tablet computer. qPDF Notes is an Android PDF reader and markup tool to annotate PDF documents, fill forms as shown in Figure 3 [11].

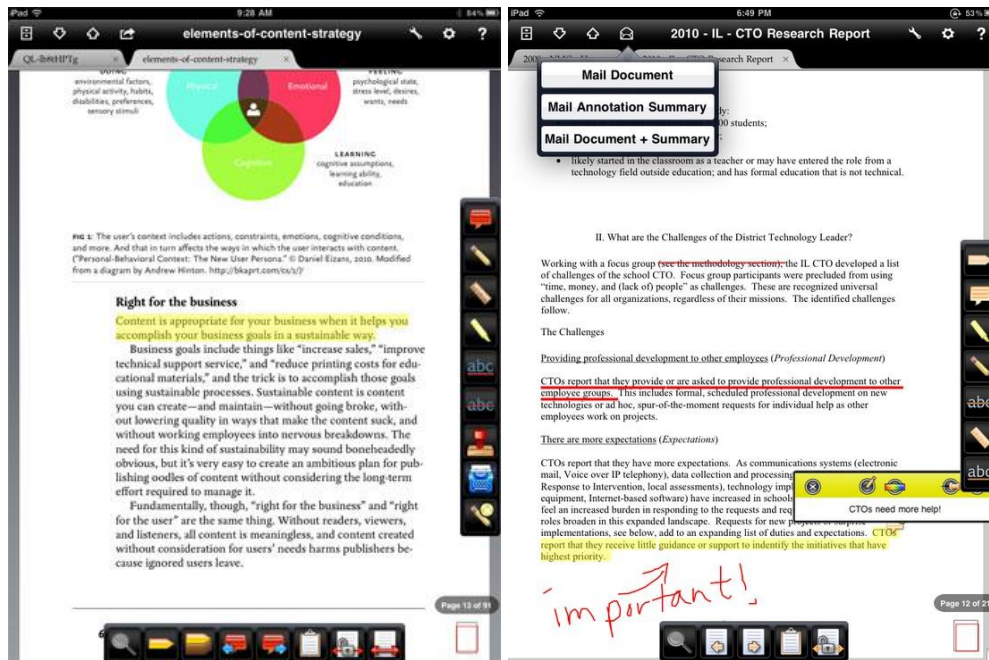


Figure 4. Annotation features of iAnnotate

Another PDF reader application is iAnnotate PDF application for the iPad. iAnnotate PDF allows a user to choose from the pen, highlighter, typewriter, stamps, notes, underline, strikethrough, photo, voice recording, and date stamp functions. It also assembles user's own set of custom highlighter colors, draws and imports new stamps such as a signature. A user can manage, store, and annotate PDF files on their iPad. However, it doesn't synchronize annotation data to a web server. Instead, it allows a user to e-mail created annotations in a text format as shown in Figure 4 [12].

## 2. Annotation-content model and similarity algorithm

In this section, annotation model and similarity algorithm is developed for the interactive and collaborative learning system. The system needs to collect annotation activities for self-directed learning. Thus, the students' activity data is stored according to the model. The proposing model is defined based on the relation of annotation such as note and digital textbook content.

### 2.1. Annotation model

In this paper we defined relations that occur between touch area and educational resources. According to this definition, an area beautification technique is applied to create accurate annotation information. Moreover, an annotation can be stored as reusable information in learner's profile database.

In tablets, students make their own annotation by touching their finger on the screen. They reserve an area as a bookmark or note-taking space. They write then a *comment*, *question*, *answer* to other annotation and tags if they want to put. Basically, annotation can be consisted of *area*, *text*, *types* and *tags* in this model.

All annotations are anchored to digital textbook content such as text, image or page itself. In the low-level, they are floating on the screen by coordinates. However, in the abstract level, they are anchored information to book, chapter, section, or text phrase somehow. Therefore, the area is consisted of *book*, *chapter*, *section*, *HTM path* and *coordinates* on the screen. Formula (1) shows the proposed definition of annotation model.

$$\text{Annotation} = \{ \text{Area}, \text{Text}, \text{Type}, \text{Tags} \}$$

$$\text{Area} = \{ \text{Book}, \text{Chapter}, \text{Section}, \text{HTML\_Path}, \text{Coordinates} \} \quad (1)$$

$$\text{Type} = \{ \text{Memo}, \text{Question}, \text{Answer} \}$$

### 2.2. Annotation-content based learner's profile

According to the annotation model, we propose the learner's profile. It also includes fundamental information such as quiz score or attendance, but it is not sufficient to reflect students' learning behavior. The key characteristic of the model is using dynamic information from students. They make, read, and share their annotation over semesters. The annotation is rich and valuable knowledge in education and all these dynamic activities need to be collected.

Therefore, we defined the annotation-content based learner's profile as shown in Table 1. It mainly organizes and manages information occurring among annotation-content-student. The data is then extracted and stored in database based on XML schema

**Table 1. Annotation-content based learner's profile**

Category	item	Description	
Basic Info.	Name	Basic profile	
	Sex		
	Grade/Classroom		
Test Report	Mid/Final Term	Mid/Final Score	
	Quiz	Quiz score	
Learning Progress	Textbook Progress	Textbook reading progress.	
Annotation Activities	Creating	Creating frequency	The number of created annotations
		Annotation types	Ratio of created annotation types
		Anchored content	Semantic location of annotations
		Notebook creating frequency	The number of created notebooks
	Reading	Accessing frequency	Annotation accessing frequency
		Reading types	Ratio of accessed annotation types
		Notebook reading frequency	The number of accessed notebooks
Foxonomy	Tag information	Tag information in each annotations	
	Tag frequency	Tag frequency in a tag cloud	
	Tag access frequency	The number of hit tags	
Attendance	Login info	Login connection information	
	Connected time	Total connected time to system	

### 2.3 Similarity Algorithm for Annotation Recommendation

Active sharing of student knowledge is an essential element of effective and meaningful way of learner-centered teaching. Thus, students can access a public notebook, a collection of user annotations, by other learners in the proposed system. The notebooks are displayed on the bookshelf categorized by courses and chapters.

The notebooks in the same category have different weight values of similarity and recommendation calculated by the proposing algorithm. The similarity weight value of a

notebook is determined by a vector inner-product of HTML structure and student annotation score.

The format of digital textbooks on the system use ePub3, which is basically based on HTML5. Thus, the structure of textbooks is explicitly divided into chapter, section, paragraph, etc. Every structural element has a generated unique id, and an annotation can be attached to one of the element.

$$Relevance(a, b) = \sum_{i=1}^n N_{a,i} \times N_{b,i} \quad (2)$$

If no student( $s_n$ ) annotation activity occurs on structural element  $i$ , the value,  $v$  would be zero. Scores of  $v$  for simple comment or memo, question and answer are 1, 2, and 3 respectively. Accumulated annotations on the same  $i$  have higher score. Table 2 shows the example of annotation-content values stored in database on a web server.  $N$  is normalized value for each  $v$  on  $i_n$

**Table 2. Student learning pattern based on annotation-content relation**

	$i_1$		$i_2$		$i_3$		$i_4$	
	$v$	$N$	$v$	$N$	$v$	$N$	$v$	$N$
$s_1$	0	0.00	1	0.41	2	0.82	1	0.41
$s_2$	3	0.42	2	0.28	6	0.85	1	0.14
$s_3$	2	0.53	0	0.00	3	0.80	1	0.27
$s_4$	0	0.00	2	0.55	0	0.00	3	0.83

The relevance result value between two students can be deducted from the formula (2). The relevance value is sum of multiplying normalized value of each element  $i$  and student  $a, b$ .

Students' activities are stored in the learner's profile such as questions, answers and comments they wrote. To support cooperative learning, we analyzed the activities and defined the following formula to search students who have similar learning styles and interests. The similarity between two annotations( $i, j$ ) is calculated by sum of normalized factor R whereas user and area stands for  $u$  and  $i$  respectively

$$Similarity(i, j) = \frac{\sum(R_{ui} - \bar{R}_i)(R_{uj} - \bar{R}_j)}{\sqrt{\sum(R_{ui} - \bar{R}_i)^2} \sqrt{\sum(R_{uj} - \bar{R}_j)^2}} \quad (3)$$

Thus, students who have similar learning styles and patterns can be grouped and the group members can share and discuss about their knowledge on our system via annotations. For instance, if a student created a question annotation in the digital textbook, then our system recommend it to other group members to solve it together. In addition, the students can scan



other members' similar questions and answers to gather additional information about their annotation.

### 3. Implementation

According to the annotation definition and similarity formula, the system was implemented on iPad on iOS 5.1. Digital textbooks on the system follows ePub3 standard based on HTML5.

Figure 5 shows the annotation user interface allowing users to draw annotation area around a target object such as paragraphs, images, tables and etc. When a touch-up event occurs at the new area, the annotation module then recognizes target area information such as *Book*, *Chapter*, *Section*, *HTML\_Path*, *Coordinates*, defined in formula 1. It prompts then the input menu for text comment, note type and tags associated at the new annotation. Users select type of annotation such as *Memo*, *Question* or *Answer* and fill additional information out (*Text*, *Tags*).

The inaccuracy result that occurs during the interpretation stage of touch input and objects within the digital textbooks needs to be resolved since it is a big obstacle when accurately selecting the anchored area to annotate. To do this, the proposed system recognizes and decides the annotation area by semantic structure of digital textbooks as shown in equation (1).

When students draw an area on the screen as shown in Figure 5(left), the system tracks object information such as Course, Chapter, Section, Path and Coords defined in equation (1) within the area, and it determines the best candidate object on the page. Next, students select type of annotation such as Memo, Question or Answer and fill out other information such as Text or Tags as shown in Figure 5(right).



Figure 5. Annotation creation interface(left) and notebook page(right)



Detailed steps to process annotation are described in [13]. In this paper, we focus on the model and similarity function.

Students can browse their annotations as a list on the proposed system as well as other's comments on them. Students may touch 'Find Similar Questions and Answers' button to see other's annotation based on the proposed similarity equation (2). For instance, the student created a question annotation about 'Asking for directions in English' as shown in Figure 6(left), and other's annotations about the same subject were searched as shown in Figure 6(right).

Swiping through the page with two fingers switches to the notebook page, collection of annotations. All created annotations are gathered here as a form of learner's personal notebook which stores their own summaries of annotations.

The "note-bookshelf" as shown in Figure 7 lets students browses public notebooks shared by other learners. Once students see something that interests them, students can tap a book icon to read the annotations.

As students pile their notebooks in the shelf, there can be lots of different kinds of notebooks available. Even though the note-bookshelf allow students organize them by course category, it would be hard to find something useful. Therefore, the system executes the proposed algorithm to sort the notebooks by learning pattern similarity. Students whose annotation activity is similar to the current login user in application have high similarity scores. The notebook with the highest score in a subject category would be located to the most left side of the row.

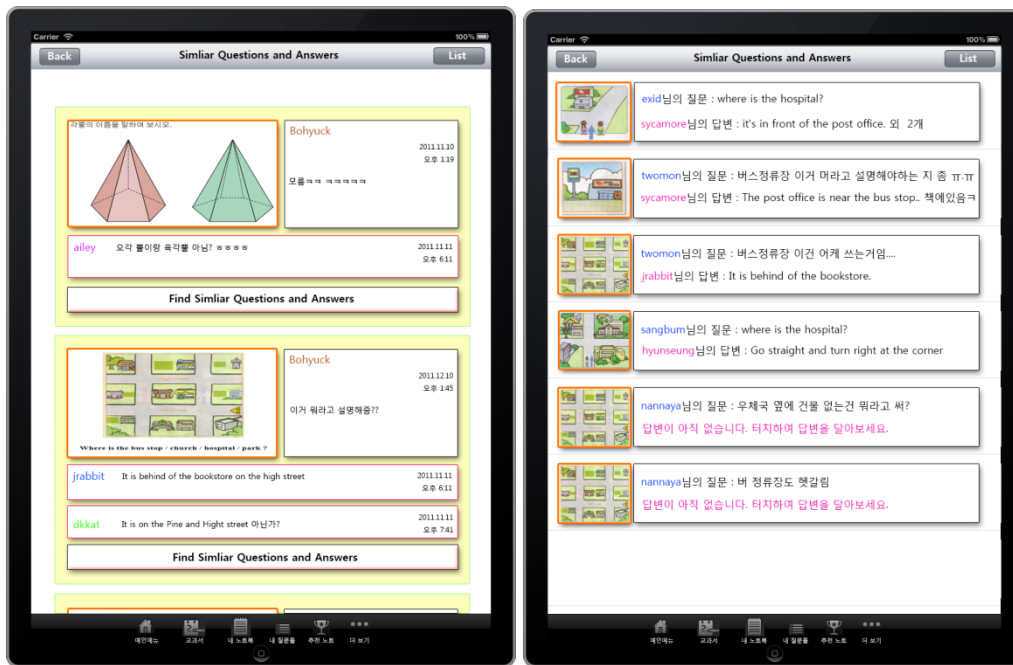
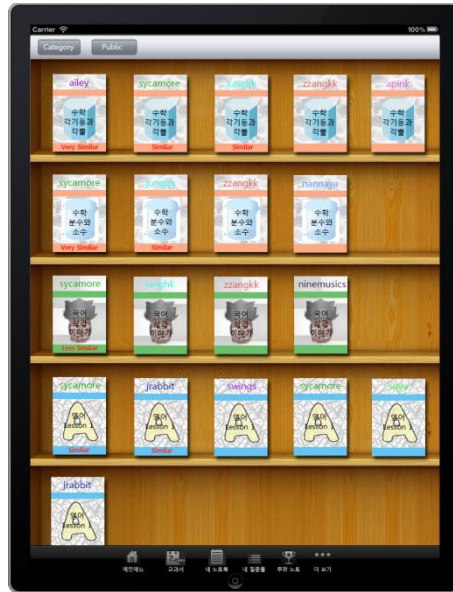


Figure 6. Similar question and answer list

As a result, notebooks are sorted by the scores in each subject categories in descending order. Students can access to the most similar notebooks as well as the most not similar ones. By comparing their own work with others, they would find information they had missed, or learn how others study for the subject. Therefore, the notebook-self is interface where learners rely on each other to achieve the same goal.



**Figure 7. Public notebooks in the ‘note-bookshelf’ ordered by similarity weight**

If students haven’t created annotations enough, they are not able to identify the similar notebook yet. As collaborative filtering methods recommend items based on users’ past preferences, new users will need to rate sufficient number of items to enable the system to capture their preferences accurately and thus provides reliable recommendations. However, looking on the bright side, it encouraged students to do more annotation activities, which proved us that the fact we mentioned in the introduction section were correct, and the usability feature of the application was designed properly.

#### 4. Evaluation

In this evaluation,  $G_1$  represents the experimental group using the proposed application, and  $G_2$  represents the control group using the same application without similarity and recommendation functionality.  $G_2$  still can annotate and share the notebooks, but they are not sorted by the formulas. 40 elementary students (22 male and 20 female, the 6th grade student) participated as evaluators. Enough time were given to be familiar with the application before the evaluation. Our hypothesis is that  $G_1$  will improve student’s knowledge more than  $G_2$  after using the application.

First, we took the pre-test to measure students’ basic knowledge on the course material.  $G_1$  and  $G_2$  scored average of 71.82 and 73.01, standard deviation of 4.51 and 4.34 respectively. The difference of test results between  $G_1$  and  $G_2$  is not significant ( $p = .79$ ). This ensures that those two groups have a similar level of basic knowledge on the material.

**Table 3. A comparison between the post-test of two groups**

	N	Avg.	SD	df	t	p
$G_1$	21	92.01	9.89	15.73	1.012	.000
$G_2$	21	84.71	10.21	16.84		

After the pre-test, one week was given to both groups for using the application to read, annotate and share them on digital textbooks.

Then, we took the post-test to measure how the proposed functionality affected them.  $G_1$  and  $G_2$  scored average of 92.01 and 84.71 respectively. The significant difference was found ( $p = .00$ ) as shown in Table 3. The proposed application improved students' knowledge more than digital textbooks with basic annotation functionalities. Therefore, we found that the use of the proposed notebook application had a positive effect on problem solving skills.

## 5. Summary

The prototype note-taking interface reflecting students learning pattern was developed to enhance value of digital textbooks. The interface was designed touch-based annotation input and cooperative learning. Students allowed creating their own work by using the proposed note-taking interface. The annotations and textbook content were synchronized to a web server based on the proposed structure. The proposed similarity algorithms can find other annotations created by students who have top grade or similar learning style. Thus, the student can access other annotation related to user's learning style. The main idea of the system is to get more information from annotation about our students, and to put students at the center of the learning system. It deployed on a touch-based device and evaluated to measure its effectiveness. As a result, learner's knowledge was improved by sharing notebooks recommended by the similarity module. In addition, learners were satisfied with creating their own content by using the proposed note-taking interface. Thus, the digital textbooks offer more learning opportunities and facilitate the improvement of learning achievement.

## Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2012-0006751)

This paper is a revised and expanded version of a paper entitled Strategies based on Learning Style Similarity for Smart Education presented at Multimedia, Computer Graphics and Broadcasting, December 16-19, 2012 at Korea Woman Training Center, Kangwondo, Korea.

## References

- [1] P. Nagarajan and G. W. Jiji, International Journal of u- and e-Service, Science and Technology, vol. 3, no. 4, (2010), pp. 37-48.
- [2] O. Boyinbode, A. Bagula and D. Ngambi, International Journal of u - and e - Service, Science and Technology, vol. 4, no. 3, (2011), pp. 11-18.
- [3] Y. Taizan, S. Bhang, H. Kurokami and S. Kwon, International Journal for Educational Media and Technology, vol. 6, no. 1, (2012), pp. 85-93.
- [4] Digital textbook, [http://en.wikipedia.org/wiki/Digital\\_textbook](http://en.wikipedia.org/wiki/Digital_textbook), (2013).
- [5] R. C. Meurant, International Journal of u - and e - Service, Science and Technology, vol. 3, no. 4, (2010), pp. 49-62.
- [6] A. Balla, International Journal of Advanced Science and Technology, vol. 6, no. 1, (2009), pp. 1-14.
- [7] I. A. Ovsianikov, M. A. Arbib and T. H. McNeill, International Journal of Human-Computer Study, vol. 50, no. 4, (1999), pp. 329-362.
- [8] C. C. Marshall, The 35th Annual GSLIS Clinic: Successes & Failures of Digital Libraries, (1998) University of Illinois at Urbana-Champaign.

- [9] P. Lam, J. Lam and C. McNaught, "How usable are eBooks in an mLearning environment?", *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 20, no. 1, (2010), pp. 6-20.
- [10] Amazon Kindle, <http://www.amazon.com/gp/help/customer/display.html?nodeId=200298480>.
- [11] qPDF, <http://alternativeto.net/software/qpdf-notes-full-version/about/>.
- [12] iAnnotate PDF, <https://itunes.apple.com/us/app/iannotate-pdf/id363998953?mt=8>.
- [13] J. K. Kim, W. S. Sohn, T. H. Kim and Y. S. Lee, The 1st International Conference on Advanced Information Technology and Sensor Application, (2012) Daejun, Korea.

## Authors



### Jae-Kyung Kim

He received the B.S. degree in Statistical Computing/Chemistry from Dankook University in 2000 and the M.S and PhD degrees in Computer Science from Yonsei University in 2002 and 2007. He is currently a senior researcher at Smart Education Research Center at Gyeongin National University of Education. His research interests include smart education, human-computer interaction, annotation, and electronic textbooks.



### Won-Sung Sohn

He received the B.S. and M.S. degrees in Computer Engineering from Dongkuk University in 1998 and 2000 and the PhD degree in Computer Science from Yonsei University in 2004. From 2004 to 2006 he was a postdoctoral associate in the Computational Design Laboratory at Carnegie Mellon University. He is currently a professor at Department of Computer Education, Gyeongin National University of Education. His research interests include educational design research, human-computer interaction and computer education.



### Kyeong Hur

He is currently an Associate Professor in the Department of Computer Education at Gyeongin National University of Education, Republic of Korea. He was senior researcher with Samsung Advanced Institute of Technology (SAIT), Korea from September 2004 to August 2005. He received a M.S. and Ph.D. in Department of Electronics and Computer Engineering from Korea University, Seoul, Korea, in 2000 and 2004, respectively. His research interests include; computer network designs, next generation Internet, Internet QoS, and future All-IP networks.



### YangSun Lee

He received his B.S., M.S. and ph.D degrees in Department of Computer Engineering from Dongguk University, He is currently a professor at Department of Computer Engineering, Seokyeong University. He is also a member of board of directors of Smart Developer Association, Korea Multimedia Society and Korea Information Processing Society. His research interests include smart system solutions, programming languages, and embedded systems.