Yang, J. C. et al. (Eds.) (2018). Proceedings of the 26th International Conference on Computers in Education. Philippines: Asia-Pacific Society for Computers in Education

# Transferring Learning Footprints Across Versions within E-Book Reader

**Christopher Ching-Yuan YANG[a*], Gokhan AKCAPINAR[b, c], Brendan FLANAGAN[b] & Hiroaki OGATA[b]**

[a] *Graduate School of Informatics, Kyoto University, Japan*
[b] *Academic Center for Computing and Media Studies, Kyoto University, Japan*
[c] *Department of Computer Education & Instructional Technology, Hacettepe University, Turkey*
*yang.yuan.57e@st.kyoto-u.ac.jp

**Abstract:** E-book readers support users to create learning footprints in many forms, it also allows instructors to update their documents in the e-book reading system. However, e-book users often face problems when trying to find the learning footprints they made in a new version document. In this paper, we deal with how to determine similar pages for learning footprints transferring across versions within an e-book reader in a coursework scenario. We propose an algorithm include Diff method and Transfer method to transfer learning footprints across versions based on page image similarity comparison, and we also try to find an optimal threshold for similar pages determination. After that, we prepare some data for testing and then analyze how well can this algorithm address the problem and what is the optimal threshold for similar pages determination, and the analysis results will be presented.

**Keywords:** E-book reader, learning footprint, structural similarity

## 1. Introduction

Despite e-book readers allow instructors to update their learning material freely, furthermore, users can also make their learning footprints in any place in the interface of an e-book reader. There is a problem exist within e-book reader. In a coursework, when instructor update the learning material, there is a new version of learning material will show up. At that time, we may have to automatically transfer learning footprints made by users since some of the page old contents may be changed, and the location of learning footprints may also not correct (Yang, Flanagan, & Ogata, 2018). This is similar to annotation reflowing or annotation repositioning issue, such techniques have also been considered in several researches (Bargeron, & Moscovich, 2003; Golovchinsky, & Denoue, 2002). According to that, this paper will briefly summarize those issues to learning footprint transferring problem.

### *1.1 E-Book Reader and Learning Footprint Transferring Problem*

This paper uses the BookRoll digital learning material reading system, users can create marker, memo and bookmark on BookRoll (Ogata et al., 2017). Currently, learning material contents can be uploaded to BookRoll in PDF format, and this paper will focus on slide-based learning materials.

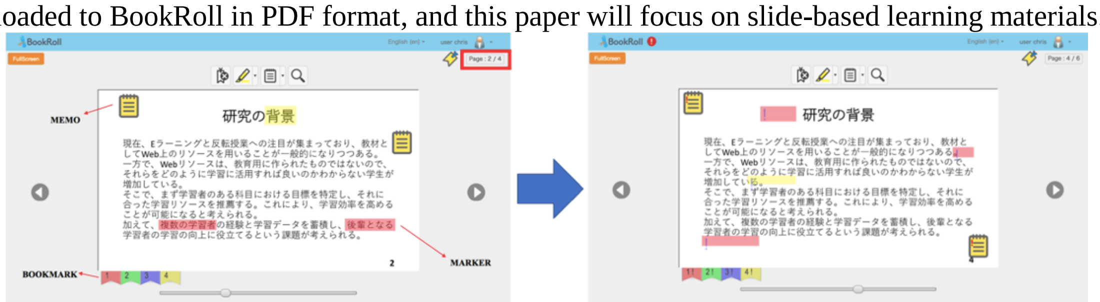


*Figure 1.* Learning footprints transferring problem in two different versions.

These learning footprints need to be transferred automatically from source page in old version learning material to its target page in new version learning material or that will become a reading problem for the readers since learning footprints are still being connected to the old version learning material as shown in Figure 1. To address this problem, in this paper we try to find similar pages in two different version learning materials based on the proposed algorithm.

## 2. Method

In this section we demonstrate how we find similar pages based on page image similarity comparison and transfer learning footprints across versions by using some data prepared by the authors as our learning footprints in a learning material. In this paper, we prepared 100 learning footprints, include 40 markers, 40 memos, and 20 bookmarks in a 20-pages slide-based learning material for the testing of learning footprints transferring algorithm.

### *2.1 Algorithm*

In the proposed algorithm, we will first execute Diff method then be able to compare all the page image similarities between two different versions by implementing structural similarity (SSIM) index which has been used in heterogeneous document image retrieval for many years (Shin, & Doermann, 2006). After that, learning footprints will be either transferred to the target page in new version learning material or removed from the current learning material by Transfer method. Although we can find similar pages, it is still hard to define why are they similar, so we analyze the optimal threshold determining similar pages in two different version learning materials to make sure the threshold will not be set too low or too high.

### *2.2 Evaluation*

In the optimal threshold analysis, if the image structural similarity between two pages is higher than the threshold, all the learning footprints on the source page will be transferred to the target page with the closest image structural similarity, otherwise if none of them is higher than the threshold, then they will all be removed. We run the algorithm for 20 times, testing the threshold value from 0 to 1, divided by 0.05 in each time to compute the $F_1$ score based on confusion matrix. The $F_1$ score and threshold can be seen in Figure 2.

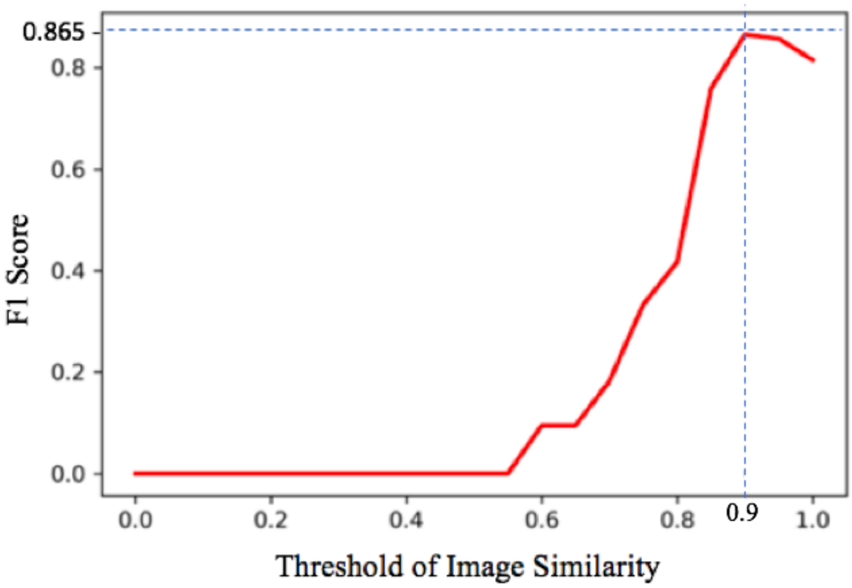


*Figure 2.* The analysis for deciding optimal threshold value.

## 3. Results

As shown in Figure 2, when the optimal threshold value is set to 0.9, we can get the highest $F_1$ score 0.865, and the confusion matrix result of image similarity 0.9 is shown in Table 1. In this confusion matrix we define *Transfer* class as learning footprints that need to be transferred, *Non-Transfer* class as learning footprints that should not be transferred, respectively. According to this confusion matrix, 81 out of 100 learning footprints are transferred correctly, but 19 learning footprints are transferred incorrectly. To understand these errors, we did error analysis and found that there are two types of error in it, 4 of them were transferred incorrectly and 15 of them were removed incorrectly, we also found that the reason that caused this error is due to the change of image position or text position.

Table 1
*Confusion Matrix for Optimal Threshold*

| | **Gold-Standard** | |
|---|---|---|
| **Algorithm** | Transfer | Non-Transfer |
| Transfer | 61 (TP) | 0 (FP) |
| Non-Transfer | 19 (FN) | 20 (TN) |

## 4. Conclusions

In this paper, a method for learning footprints transferring was presented and evaluated. This method can detect pages higher than 0.9 image similarity between two different versions, but for other pages with lower than the optimal threshold value, we do not know if we can transfer some of the learning footprints on it. Meanwhile, the $F_1$ score obtained from analysis result is still not perfect enough, 19 of them will still be processed incorrectly. In the future, we will investigate better analysis method in different analysis level such as text content comparison between two different versions to eliminate these errors.


## Acknowledgements

This research was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (S) Grant Number 16H06304.


## References


Bargeron, D., & Moscovich, T. (2003, April). Reflowing digital ink annotations. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 385-393). ACM.

Golovchinsky, G., & Denoue, L. (2002, October). Moving markup: repositioning freeform annotations. In *Proceedings of the 15th annual ACM symposium on User interface software and technology* (pp. 21-30). ACM.

Ogata, H., Oi, M., Mohri, K., Okubo, F., Shimada, A., Yamada, M., ... & Hirokawa, S. (2017). Learning Analytics for E-Book- Based Educational Big Data in Higher Education. In *Smart Sensors at the IoT Frontier* (pp. 327-350). Springer, Cham.

Shin, C. K., & Doermann, D. S. (2006, June). Document Image Retrieval Based on Layout Structural Similarity. In *IPCV* (pp. 606-612).

Yang, C. C. Y., Flanagan, B., & Ogata, H. (2018). Connecting Learning Footprints Across Versions within E-Book Reader. In *The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018* (pp. 1-4).