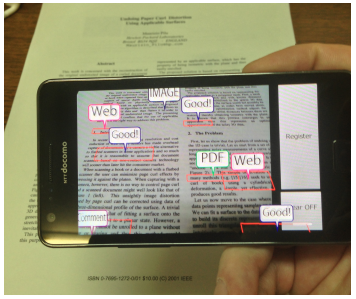


**Figure 1:** The first prototype. On the left are two buttons, one for registration of new annotations, one for removal.

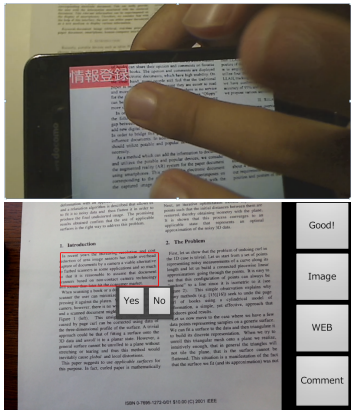


**Figure 2:** The user selects a position for an annotation and can choose the type of annotation (also first prototype).

# Annotate Me – Supporting Active Reading using Real-Time Document Image Retrieval On Mobile Devices

**Kai Kunze**
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
firstname.lastname@gmail.com

**Katsuma Tanaka**
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
Katsuma@m.cs.osakafu-u.ac.jp

**Masakazu Iwamura**
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
masa@cs.osakafu-u.ac.jp

**Koichi Kise**
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
kise@cs.osakafu-u.ac.jp



http://dx.doi.org/10.1145/2494091.2494165

## Abstract
We present a novel system to support active reading. Utilizing a mobile device the user can add digital annotations to physical documents like papers and posters. We present first prototype implementations of the mobile phone interface with and without server support. Server support let's you share annotations with your friends. We discuss findings from an initial user evaluation and present an improved prototype. We believe annotating paper using document image retrieval is a promising technology for active reading support.

## Author Keywords
active reading, document image Retrieval, augmented reality, handheld, interactive paper

## ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces Graphical User Interfaces;; H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

## Introduction
Active reading –reading a text while working with it through highlights and annotations– is a critical knowledge acquisition task. There has been substantial research in better understanding and enhancing it [14, 10]. Paper has a high degree of affordance [13, 14].

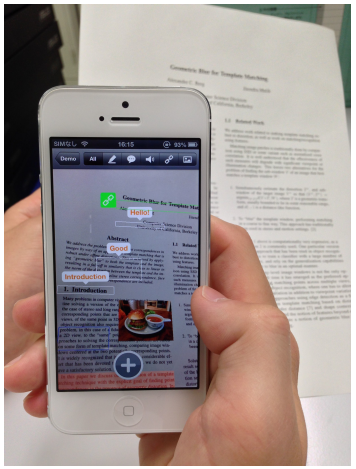

**Figure 3:** The main screen of the second prototype. It now supports landscape mode and easy annotation entry.

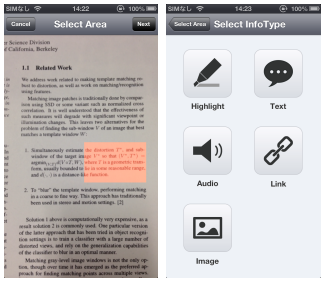


**Figure 4:** Second Prototype: The user selects a position for an annotation and can then chose the type of annotation (highlight, text, audio, url or image)

It can be easily annotated, it can be spread out and navigation between content is simple. The spacial information of the annotations on paper helps us to easily find things and memorize facts easier.

Therefore, we wonder: is there a way to combine the high degree of affordance of paper with the advantages of digital technology (easy replication, search-ability etc.)?

In this work we contribute towards this goal: (1) We present a system to register and annotate text (documents) in real time using mobile devices.The system can work in stand-alone mode and with server support. In stand-alone mode we can retrieve pre-registered digital versions of a document or register a un-known documents locally on a device. The server version lets you retrieve registered documents over a service and share annotations to text documents with your friends and colleagues. (2) We build annotation libraries for Android and iOS, that enable the system developer to register a new digital document and to add highlights, free text, audio, urls and images to the printed version of the document using a mobile device. (3) As a proof of concept, we show our first mobile annotation prototypes and improvements based on user evaluations.

## Our Approach

We augment paper directly with the help of digital technology. With this approach, we hope to retain all positive features of paper and enhance it accordingly to overcome the problems presented beforehand. There are several ways to augment a paper sheet, most paper-based augmented reality systems use QR codes or other passive markers [1]. This method requires specially prepared documents with markers embedded into them. We decided to use a document image retrieval method based on "Locally Likely Arrangement Hashing" (LLAH) [11]. The main reasons for using LLAH are explained in the following. LLAH is very fast. The feature extraction method can be implemented on todays smart phone devices and run in perceived realtime. Smaller databases ( a couple of pages) can be searched on a high end smart phone as well. For larger document databases we need a client server infrastructure. Still the retrieval algorithm is fast. For example, retrieval from a database with 100 Million pages (around 440 thousand books) takes around 178 ms on a single server core.The correct retrieval rate is around 99% [5]. The method works under perspective distortion, uneven lighting, and non-linear deformation.

## Implementation

*System Libraries*

We currently have working application level libraries for iOS, Android and Windows. As mentioned before, the application libraries can be used in stand-alone mode. This means the developer can add and register documents and the applications do not need to access the network. The user can also register new documents manually. However, the number of documents and performance is limited by the processing power of the local machine ( e.g. a couple of pages on a slow mobile device). The other supported mode is client-server, where multiple clients can access the database to register new documents or retrieve them.

*Mobile Applications*

As a proof of concept, we implemented mobile clients for the Android and iOS platforms, using our system libraries and iterative design to improve the user interfaces.

*User Evaluation*

We performed several user tests with early prototypes. The initial prototype shown in Fig. 1 and 2 enabled the

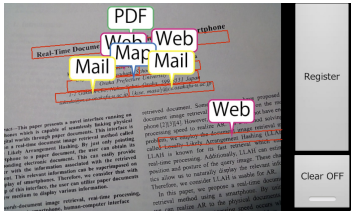


**Figure 5:** The main screen of the first prototype.

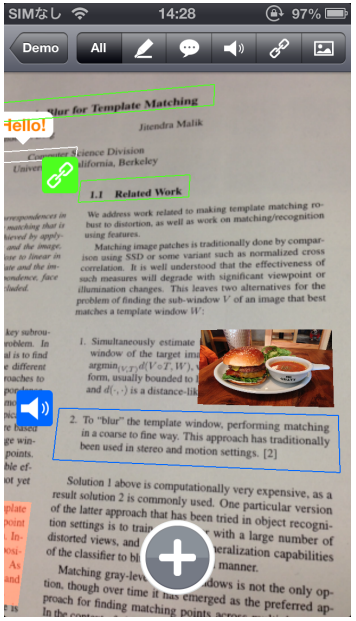


**Figure 6:** The main screen of the second prototype with UI improvements.

user to add a text region highlight, text comment, images and urls. While observing how 5 users (age between 24 and 53) interacted with the system and interviewing 2 HCI experts, we identified and improved on the following problems. The old and the improved interfaces can be seen in Fig. 5 and 6. The initial prototype just supported landscape mode and it was impossible to use only one hand to annotate a document. We added portrait mode and an easy reachable "+" button, so easy one-hand operation is now possible ( Fig. 3 and 4).

Another problem users encountered is that the screen can be cluttered with annotations of different types, so it's harder to click on them or see them. Thus, we introduce buttons to filter the different annotation types (from left to right in the upper corner: all annotations, highlight, free text, audio, links and images). Regarding the redesign, all users agreed that it is better and more straightforward to use. Two new user requests are adding search and allowing to copy pen annotations from the paper document [4].

*Server Components + Web Interface*
If we run our system in client server mode a larger number of documents can be supported ( We used 100 Million pages for demonstration purposes with an average retrieval time of 178 ms). We also show a Web interface based on Ruby on Rails, allowing the user to see all documents he annotated and access his annotations as well as the public annotations of his friends (see Fig. 7 and 8 for an example).

## Related Work

We build on our work showing that realtime document retrieval can be implemented on todays smartphones [15, 11]. Compared to our previous work, LLAH has not been used for augmenting paper documents before.

There is already some work trying to support active reading on mobile devices using annotations [4, 6, 9, 12]. The works we are aware of use SIFT or similar features, for indexing and retrieval and thus can deal only with a small number of pages (compared to the 100 Million pages for LLAH). Also the accuracy is usually lower and processing speed higher. Our goal is not to build a single system or support a single application, but to create an open platform for active reading support also available for other researchers.

Researchers already explore augmented reality (AR) to support knowledge acquisition mostly in education and work environments [8, 3].

Other work focuses on augmenting paper with AR annotations with different application goals: interfaces for multimedia control or retrieval of travel information [2]. Additionally, researchers in AR tend to use QR codes or similar optical markers [1, 7].

## Conclusion and Future Work

Combining paper documents and digital systems using document image retrieval have some very interesting properties for active reading support. We can benefit from all of the advantages of paper without having some of its limitations. We showed an initial prototype for annotating text documents. As a next step, we want to further improve the prototypes, before we evaluate their usability for active reading tasks in realistic scenarios.

We want to make the presented application and the LLAH based document image retrieval system available to the community. Interested researchers can already try LLAH


**Figure 7:** The web interface showing documents annotated by the user.

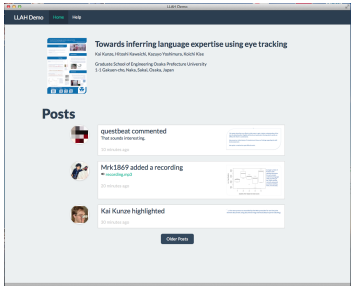


**Figure 8:** The detail web view for one of the documents, a poster, showing the user's and his friends' annotations.

for the desktop registering their own documents `http://imlab.jp/LLAH/`.

As a next step we plan on offering an online service for document matching and an iOS and Android library, as well as an improved version of the mobile applications described in this paper.

## Acknowledgements

This work was supported in part by the CREST project from Japan Society for the Promotion of Science (JSPS). Ideas for this work in progress were gathered during the Dagstuhl Seminar 12492 Human Activity Recognition in Smart Environments.

## References

[1] Chen, N.-S., Teng, D. C.-E., Lee, C.-H., and Kinshuk. Augmenting paper-based reading activity with direct access to digital materials and scaffolded questioning. *Comput. Educ.* (2011), 1705–1715.

[2] Erol, B., Antúnez, E., and Hull, J. J. Hotpaper: multimedia interaction with paper using mobile phones. In *Proc of MM 2008*, MM '08 (2008), 399–408.

[3] Graham, J., and Hull, J. Video paper: a paper-based interface for skimming and watching video. In *In Proc. of ICCE. 2002* (2002), 214 – 215.

[4] Guimbretière, F. Paper augmented digital documents. In *Proc. of UIST*, ACM (2003), 51–60.

[5] Iwamura, M., Nakai, T., and Kise, K. Improvement of retrieval speed and required amount of memory for geometric hashing by combining local invariants. *Proc. of BMVC2007* (2007), 1010–1019.

[6] Liao, C., Liu, Q., Liew, B., and Wilcox, L. Pacer: fine-grained interactive paper via camera-touch hybrid gestures on a cell phone. In *Proceedings of CHI 2010* (2010), 2441–2450.

[7] Liu, Q., Liao, C., Wilcox, L., Dunnigan, A., and Liew, B. Embedded media markers: marks on paper that signify associated media. In *Proc of*, IUI '10 (2010), 149–158.

[8] Mackay, W. E., and laure Fayard, A. Designing interactive paper: Lessons from three augmented reality projects. In *In Proc. of IWAR98*, A K Peters, Ltd (1999), 81–90.

[9] Mackay, W. E., Pothier, G., Letondal, C., Bøegh, K., and Sørensen, H. E. The missing link: augmenting biology laboratory notebooks. In *Proc. of UIST*, ACM (2002), 41–50.

[10] Marshall, C. C. Saving and using encountered information: Implications for electronic periodicals. In *In Proc. of CHI 05*, ACM Press (2005), 111–120.

[11] Nakai, T., Kise, K., and Iwamura, M. Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval. *In Proc. of DAS 2006 3872* (Feb. 2006), 541–552.

[12] Pietrzak, T., Malacria, S., and Lecolinet, É. S-notebook: augmenting mobile devices with interactive paper for data management. In *Proc. of AVI*, ACM (2012), 733–736.

[13] Piper, A. M., and Hollan, J. D. Tabletop displays for small group study: affordances of paper and digital materials. In *Proceedings of CHI 2009*, CHI '09 (2009).

[14] Schilit, B. N., Golovchinsky, G., and Price, M. N. Beyond paper: supporting active reading with free form digital ink annotations. In *In Proc. of CHI 1998* (1998), 249–256.

[15] Takeda, K., Kise, K., and Iwamura, M. Real-time document image retrieval on a smartphone. *in Proc. of IAPR* (2012), 225–229.