Share Me – A Digital Annotation Sharing Service for Paper Documents with Multiple Clients Support

Kazuma Tanaka, Motoi Iwata, Kai Kunze, Masakazu Iwamura, Koishi Kise
Dept. of Computer Science and Intelligent Systems
Osaka Prefecture University
{*tanaka,kunze*}*@m.cs.osakafu-u.ac.jp* {*motoi,iwamura,kise*}*@m.cs.osakafu-u.ac.jp*

***Abstract*—In this paper we describe a novel annotation service which is capable of seamlessly linking physical and digital worlds through paper documents. Our service uses a real-time document image retrieval method called Locally Likely Arrangement Hashing (LLAH) for providing information associated with the document. By using this service digital annotations can be added to physical documents and shared with friends via mobile devices. We present the prototype implementation, and provide a discussion covering the technical details of the system.**

I. Introduction

Electronic books (E-books) have been increasingly popular in recent years, with many applications appearing for PCs, smartphones, and other mobile devices. E-books have some advantages over traditional documents For this characteristics, various services are available; for example, you can add highlights or comments to e-books using mobile devices, and share the annotations with your friends.

However, even if e-books continue to spread widely, traditional paper has still a high level of affordance and is preferred by a fast number of users compared to digital devices.

So we wonder if it is possible to have the best of the digital and physical world. To achieve this goal we recognize physical documents and link digital information to the documents. We choose real-time document image retrieval method called Locally Likely Arrangement Hashing (LLAH) [8], [5].

The mobile devices like tablet PCs or smartphones have become widespread recently and they have enough ability to display information using AR. The contributions of the paper are given as follows: (1) we present a novel annotation system combining LLAH together with augmented reality to add highlight and text to paper documents. (2) By using this system digital annotations can be added to physical documents and be shared with friends. This system also has a web interface allowing users to see documents and annotations online via web browser in PCs or smartphones. (3) We evaluate the system over user and performance tests on a commodity laptop and smart phone.

We believe this system is a promising technology for advanced document use. In this paper, first we describe an overview of LLAH, and the next we explain about our proposed system. We also present the results from system performance test and remaining problems to be solved in the future.

II. Related Work

This work is complementary efforts to use document image retrieval techniques to replicate the affordances of paper [7]. Kunze et.al. focus on how to use document image retrieval to support active reading[3]. In this paper, we give a detailed system description. In addition, the system vastly improved in processing and tracking, described later.

So far we are unaware of any document annotation service that is available for researchers to use. Our work is based on using realtime document retrieval on todays smartphones [8], [5]. However, so far, document image retrieval has not been used to augment paper documents.

There is some work using SIFT or other similar computer vision features to augment paper documents [2]. However, the problem of using SIFT features etc. is that the feature calculation is quite expensive. Therefore these systems can only deal with a relative small number of pages (10 – 100 compared to the 100 Million pages for our approach). Also the accuracy is usually lower and processing speed higher.

978-1-4799-2190-4/13 $26.00 © 2013 IEEE
DOI 10.1109/ACPR.2013.182

There are also other paper augmentation systems using mostly QR codes or other specially designed optical marker on the paper [1]. Of course, they have the limitation that the methods only work with special paper/markers.

III. Approach

Locally Likely Arrangement Hashing (LLAH) [5] is a method for real-time document image retrieval. In this method, a query image is transformed into a collection of feature points, then features are calculated from arrangements of these points. Calculated features are used for accessing to document image database to retrieve images by voting. LLAH is characterized by following features. **Fast and Accurate** The past experiments confirmed that LLAH has the accuracy of 99% and the processing time of 50[ms] with a 20 million pages database [8], [5]. Therefore, LLAH is enough fast and accurate to retrieve images in real-time. **Robust** In this method, calculated features consist of geometric invariants which are stable under perspective distortion. Therefore, LLAH is robust to variation of camera angles. **Works even with a low-resolution image** Not only an entire image, but also a partial image can be used as a query image. Owing to this characteristic, LLAH works even with a low-resolution image captured by a cheap web camera.

IV. Implementation

In the following we will discuss our novel annotation system using real-time document image retrieval with LLAH. By using this system, users can add digital annotations to physical documents, and can share them with other users. The system is based on client-server architecture and consists of two backend servers and a mobile client application. It also has a web interface, where users can access to all the documents and annotations using a web browsers. In the rest of this section, we explain the technical details of this system.

A. System Library

We develop an application level LLAH library with multi-platform support. This library provides core functions of LLAH: feature extraction, feature matching, optimized hash table, and database serialization. In our proposed system it is used in the LLAH server for document image retrieval.

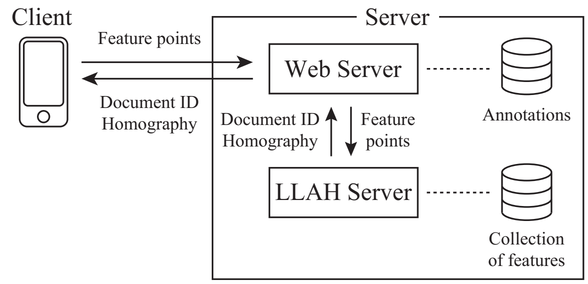

Figure 1. System Architecture

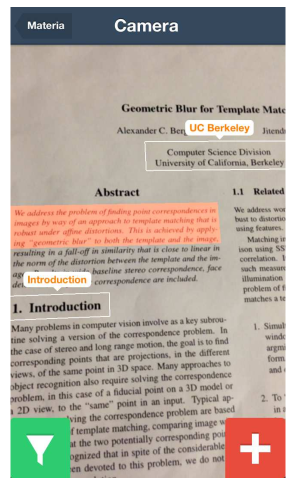

Figure 2. Mobile Application

B. Architecture

Figure 1 depicts the system architecture. The system has two servers: the first server is a web server bridging clients and the second server. The second server is a LLAH server performing document image retrieval using LLAH.

Each of the servers have individual databases. The database of the web server stores registered annotations for web interface, and the database of the LLAH server stores collections of features used for document image retrieval.

The web server and clients communicate using WebSocket protocol. WebSocket uses an independent TCP-based protocol after the handshake through HTTP. By using this protocol there are fewer communication loss and fewer affection to other connections than using only HTTP. Therefore, it is suitable for live content with real-time data exchange.

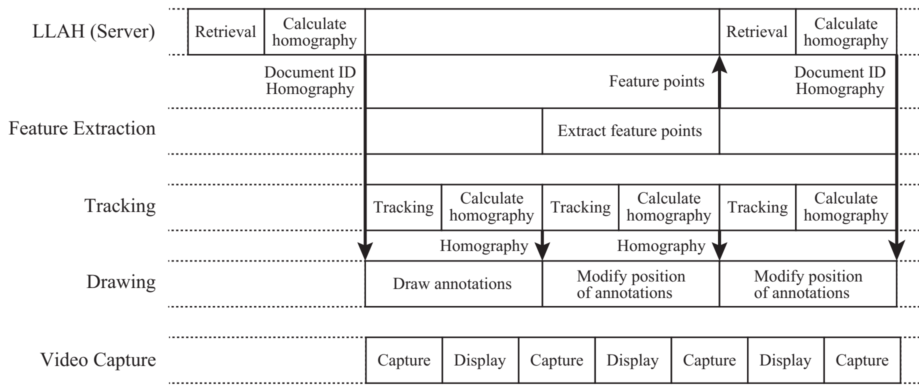

Figure 3. Multithread Processing to improve overall performance

C. *Mobile Application*

We implemented a mobile client application for the iOS platform. Figure 2 shows a screenshot from prototype. In this implementation, a Twitter account is used for user authentication. This can be changed in the future implementation. The features of this implementation are described in the following sections.

1) AR with LLAH: Owing to the characteristics of LLAH, it is possible to estimate the actual coordinate of a query image on original document. This enables the application to display annotations on its screen as if they were on the document.

Annotations are displayed using Augmented Reality (AR), and each annotation has an individual color for convenience. Users can tap the annotation to get more information or filter them by tapping the filter button on the bottom left.

2) Multithread Processing: In order to improve processing speed, we employ multithreaded processing in this system. As shown in figure 3, we use four threads: feature extraction, tracking, drawing and video capture. Annotations can be drawn seamlessly by performing a heavy task like feature extraction in background thread.

3) Tracking: Due to the limitation of the capacity of the server, clients cannot send feature points continuously. Therefore coordinates of annotations on screen are not updated frequently. For solving this, we use optical flow tracking to compensate for homographies while it is not provided by the server. In prototype Lucas-Kanade [4] method is used for optical flow estimation.

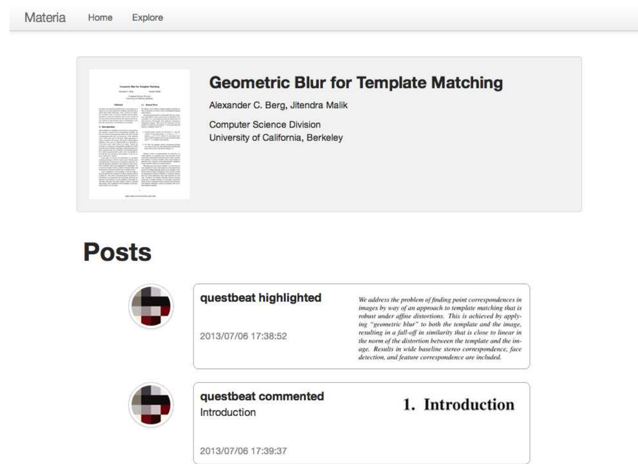

Figure 4. Web Interface

D. *Web Interface*

This system has a web interface based on Tornado (Python web framework). This allows users to see all documents and its annotations from anywhere via general web browsers in PCs or smartphones. Figure 4 is a screenshot from the prototype.

V. Evaluation

In this section we present the results from system performance test and a small user test. We measures the average processing time of 1000 times retrieval with a 4 pages document database while a user use the application on each of the pages.

In this experiment we used Apple MacBook Air (1.8GHz Intel Core i7, 4GB DDR3) as the server and Apple iPhone 5 (1.3GHz Apple A6, 1GB DDR2) as

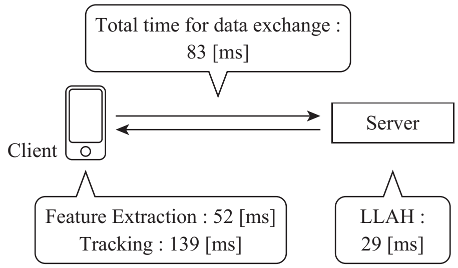

Figure 5. Results from System Performance Test

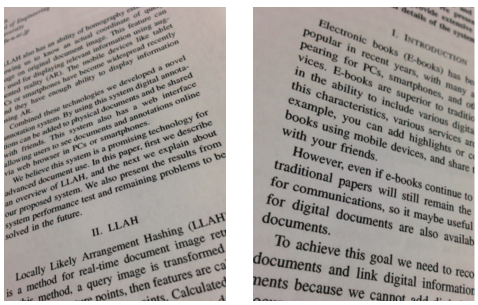

Figure 6. Examples of correctly retrieved query images

the client. Both of them are placed on the same local area network.

Figure 5 shows the mean processing time. Feature extraction and tracking on the client respectively takes 52[ms] and 139[ms]. LLAH on the server takes 29[ms]. The total time for data exchange takes 83[ms].

Figure 6 shows examples of correctly retrieved query images which tended to include many words. On the other hand, figure 7 shows the examples of the queries that were not retrieved. LLAH requires many characters to be captured to realize the stable retrieval, therefore, the images having less text cannot be retrieved.

We tested the proposed system with one ui expert and 3 users, annotating documents. The system works as advertised. The expert liked the functionality of adding highlights and seeing annotations from other users.

Still the tracking can be improved as quick movements make the augmentation lag a bit. There are also still a lot of UI issues to be solved, how to integrate annotations from a large number of users and how to select a particular user or a specific annotation type.

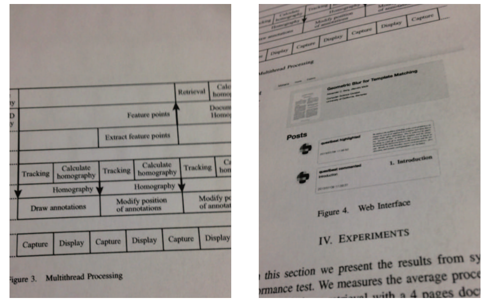

Figure 7. Examples of not correctly retrieved query images

VI. Conclusion and Future Work

In this paper, we introduced a novel annotation system, allowing users to use document image retrieval with LLAH from mobile application easily.The experimental results confirmed that this system suits minimum requirements for practical use. As a next step we plan on offering this system. Therefore, further improvements to the mobile application and more feedback from users are required.

References

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

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

[3] K. Kunze, K. Tanaka, M. Iwamura, and K. Kise. Annotate me supporting active reading using real- time document image retrieval on mobile devices. *in Adj. Proc. of Ubicomp*, pages 225–229, 2013.

[4] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81*, pages 674–679, 1981.

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

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

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