A Proposal of Writing-Life Log and Its Implementation Using a Retrieval-Based Camera-Pen

Koichi KISE [a], Riki KUDO [a] , Masakazu IWAMURA [a], Seiichi UHICDA [b] and Shinichiro OMACHI [c]
[a] *Dept. of Computer Science and Intelligent Systems,*
Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuencho, Naka-ku, Sakai, Osaka 599-8531, JAPAN
[b] *Department of Advanced Information Technology,*
Faculty of Information Science and Electrical Engineering, Kyushu University
744, Motooka, Nishi-ku, Fukuoka 819-0395, JAPAN
[c] *Department of Electrical and Communication Engineering,*
Graduate School of Engineering, Tohoku University
6-6, Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi 980-8579, JAPAN

{kise@, kudo@m., masa@}cs.osakafu-u.ac.jp, uchida@ait.kyushu-u.ac.jp, machi@ecei.tohoku.ac.jp

Abstract. This paper proposes a notion called writing-life log and a camera-pen for implementing the notion. The writing-life log is to log all of user's writing activities for capturing the generation of information by the user. We focus here on logging user's handwriting activities. For this purpose, we employ a camera-pen developed by ourselves. As compared to the Anoto pen, it has an advantage that it works on ordinary paper documents and is capable of locating recovered digital ink onto the corresponding electronic document. In order to make the pen more realistic, we introduce the recognition of pen-up and down states as well as global repositioning of recovered handwriting by matching it to physical ink.

1. Introduction

The flow of information between a person and an environment is important to know about the person. Once we are successful to capture it, we can use this source of information for personalized recommendation, estimation of the person's preference, the level of knowledge about topics the person is interested in, etc. Possible modalities of the information can be classified into verbal and non-verbal; there is no doubt that the most important and conveying most information is verbal --- the natural language, which is represented in various media such as speech and text. Among them, the text is fundamental and transform/record a vital portion of information.

In order to capture the flow of information represented by text, we have launched the project for logging the flow between a user and his/her environment. The log of the flow to the user is called "reading-life log" and that from the user is named "writing-life log". The reading-life log is to tap the user's information taking activities by "reading" and record all of the information taken by the user. The writing-life log, on the other hand, is to capture the flow from the user by recording all text produced by the user. It is almost needless to say but the text is mainly produced by either typing or handwriting. In this paper, we focus on logging handwritten text.

Nowadays we have various devices to record our handwriting such as tablets, ultrasonic- and camera-pens. However they are not perfect and there exist some prerequisites to use these devices. Among them, a camera-pen called Anoto pen is known as a flexible and easy-to-use device for recording handwritings. However, it still has a prerequisite: it is necessary to use special paper to record handwritings in addition to the special camera-pen.

In this paper we propose an implementation of writing-life log by using a camera-pen that does not require special paper. Our pen has a capability of associating handwritings to the document the user is writing on with the help of document image retrieval. This means that if the user writes something on a document which is stored in the document image database, the newly added handwriting is recorded onto the document as a digital ink. We have already proposed a basic prototype of this pen (Chikano et al., 2012). However this pen has a limitation that it cannot deal with handwritings with multiple strokes since there is no capability to recognize pen-up and down states. In addition, there is a room for improving the accuracy of recovered handwritings. The camera-pen proposed in this paper is an extension of the basic prototype by adding the following capabilities: (1) Recognition of pen-up and down states by using a micro switch for detecting a writing pressure, (2) global repositioning of recovered handwritings for improving the accuracy by matching digital ink to physical ink.

The organization of this paper is as follows. In Section 2, we briefly overview related work of writing-life log and camera-pens to locate our trial in this context. Section 3 is devoted to propose a notion of "writing-life log". In Section 4, the improvements of our camera-pen are described. Section 5 is to show results of fundamental experiments of the improved camera-pen. In Section 6 we conclude this research and show future work to realize a user friendly writing-life log.

2. Related Work

One of the earliest research trials related to writing-life log is (Arai et al., 1997). They implemented a function that establishes a link from a printed word to electronic contents. This implies that the user can record his/her interests on words and recover them by using the proposed pen device. In this sense, it captures a part of activity recognition about writing, but unfortunately it is not capable of covering the recovery of handwritings.

Handwriting recovery has been implemented by using various ways. Even if we limit cases to a combination of paper and a pen, we have ultrasonic pens, the Anoto pen and other camera-pens (Munich et al., 2002, Yasuda et al.,2008, Seok et al., 2008).

Although ultrasonic pens have already been available for many years, they are not dominant since they require a station from which the distance to the pen is measured and thus it is mandatory to fix the paper to the station. Another important limitation is that there is no way to find a document on which the pen works. Thus the task of associating the digital ink to an electronic document is the task for users.

This problem has been solved by the Anoto pen. By using unique distribution of fine dots on the paper, it is capable of sensing not a relative position of the pen on the paper but the absolute position on the whole space a part of which is assigned to the paper. This means that if it has already known the relationship between contents of a document and the sheet of paper, the digital ink can be easily associated with the contents. On the other hand, a serious drawback of the Anoto pen is that it requires the special paper with fine dots; it is not possible to capture the handwriting on a normal sheet of paper.

In order to solve the problem we have proposed a camera-pen that functions on normal paper and still able to associate the recovered digital ink to the contents of a document, if the document has already been stored in the database (DB). These functions are enabled by using the following set of technologies:

(1) Paper fingerprint and tracking: When the pen is moving on a sheet of paper, the pen tracks its movement by using paper fingerprint --- the microscopic structure of the paper surface which is unique to recognize the surface the pen has already seen. This allows us to know the position in the sheet of paper.

(2) Document image retrieval: If the contents (foreground) of a document have already been stored in the DB, the pen is capable of associating the digital ink recovered by the function (1) with the document in the DB. In other words, the recovered digital ink can be placed onto the electronic version (PDF) of the document the user is writing on. The technology of document image retrieval is scalable and can handle more than 10 million page images in real-time with the retrieval accuracy of more than 99% (Takeda et al. 2011). In addition, it is not unnatural to assume that the digital version of a printed document is available. From these points we consider that there is no problem on the document image retrieval side.

The remaining problem to be solved for realizing writing-life log is twofold: (1) the usability of the camera-pen, and (2) the accuracy of recovered handwriting. The problem (1) is that the previous pen assumes that the pen-tip is always on a sheet of paper and thus it is not possible to distinguish pen-up and down states. This poses a problem on the usability. The problem (2) is that the global position of a recovered digital ink is determined based only on the result of document image retrieval. Although the page level retrieval accuracy is high enough, it is not always possible for the document image retrieval to estimate the accurate position within the page. This paper proposes a method to solve these problems.

3. Writing-Life Log

Before going into the details of the improvement of the camera-pen, we describe a general notion of writing-life log. We are doing the research on logging a person’s reading and writing activities: the former is called reading-life log, which records all the test a person has read, and the latter is called writing-life log, which is to record all text and other materials written by a person. In this paper we focus on the writing-life log.

We consider that the writing-life log is a kind of human activity recognition with a special focus on the activity “writing”. The writing is an activity for outputting from a person the information represented mainly as a natural language text. In addition some simple figures such as underlines, circles and arrows are employed for representing information as annotations to a document. The writing-life log is an attempt to tap the information flow from a person to his/her environment (printed documents, computer systems, or other persons, etc.) to record text and graphics produced by a person.

As for writing text, we have two means: typing and handwriting. Typing is relatively easy to capture by using software such as a key logger. On the other hand, handwriting requires a device to digitize the user’s handwriting. Nevertheless it is important to capture the handwriting because when we work on printed documents or taking notes, handwriting is often more natural than to write by typing. Another advantage of handwriting is that it can also produce figures without changing the device to use.

Once we have a writing-life log of a person, we are able to use it in many different ways. One obvious application is to infer a person’s preference and interests. If a person puts an underline to a specific word, for example, this tells us that the person is interested in it. Another, a bit complicated application scenario is for education. Suppose students have their own devices for logging their writing activities while studying. The movement of the pen, written text and figures include fruitful information about a student. By analyzing the log it would be possible to assess parts the student is not good at.

4. Retrieval-Based Camera-Pen

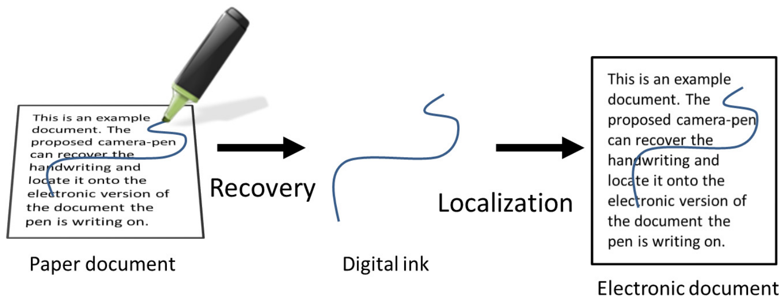

Figure 1. Functions necessary for a camera-pen.

For the implementation of writing-life log, it is necessary to have two functions as shown in Fig.1. The process of recovery is to extract the handwriting as a digital ink on a certain coordinate system. The process of localization is to transfer the recovered digital ink to on the coordinate system of the corresponding electronic document, if it is available. Both of these functions enable us to correlate the handwritings to the contents of a document.

Figure 2 shows the appearance of the camera-pen. The camera is currently mounted at the middle of the body looking downward to the surface of paper. In the current implementation the image is transferred to a computer through USB; but in the future captured images are either transferred through a wireless connection or recorded on the memory of a pen. The ultraviolet light is employed to stabilize the extraction of paper fingerprint. The micro switch is to distinguish pen-up and down states physically based on the pressure at the pen-tip. Although this signal is similarly transferred to a computer via USB, a solution using either wireless or memory can also be applied.

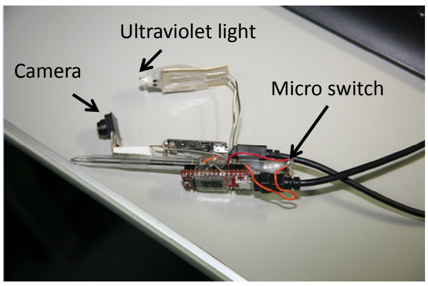

Figure 2. Camera-pen prototype.

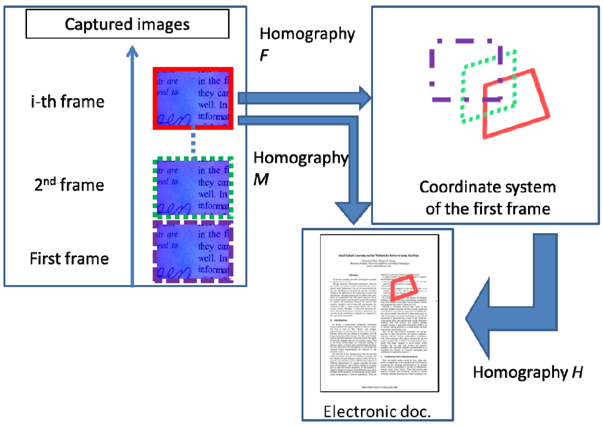

Figure 3. Processing Steps.

The overall processing flow is shown in Fig. 3, which is the same as the one proposed in (Chikano, et al., 2012). Let us briefly summarize it. We have two processing paths from captured images to a corresponding electronic document. First we explain the lower flow. By using a document image retrieval called LLAH (Takeda et al., 2011), it is possible to estimate the homography from each frame to the electronic document, which is represented as M. Next, let us explain the upper flow based on the homographies F and H. In our method all succeeding frames are perspectively transformed into the coordinate system of the first frame to unify the information. This transformation of i-th frame to the first frame is represented by the homography F, which is obtained by the frame-wise matching based on the LK tracking (Lucas & Kanade, 1981). Since the resultant coordinate system after applying F and H is also that of the electronic document, $M=FH$ holds. This indicates that we can calculate H by $H = F^{-1}M$ at every time we receive a new frame. This value is sometimes wrong because of the inaccurately estimated M and F. Thus the final estimation of the homography H is done by using all calculated H's to minimize the impact of these inaccuracies.

We have newly introduced two improvements: (1) recognition of pen-up and down states, and (2) global repositioning of a recovered handwriting for better localization. The pen-up and down recognition is quite straight-forward. As we have already seen, it is implemented physically with a micro switch. As compared to a software implementation, it is more direct and reliable. The improvement of localization needs some explanations. As mentioned above, even with a careful estimation of the homography H, it sometimes fails due to inaccurately estimated F and M. In order to solve this problem, we focus here on the physical handwriting captured in the images. The digital ink should have the same shape and position as the physical handwriting. Thus by matching the recovered digital ink to the physical handwriting, we can globally reposition the digital ink. As shown in Fig. 4, it is possible for us to find a position of digital ink that matches better to the physical ink. In the current implementation, it is achieved by moving digital ink locally around the original position.

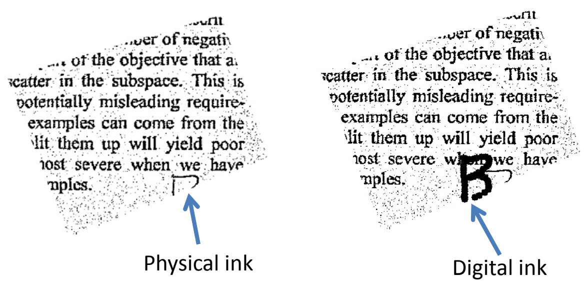

Figure 4. The use of physical ink for repositioning of digital ink.

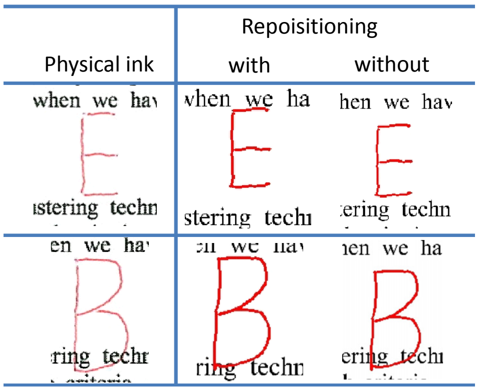

Figure 5. Experimental results.

5. Fundamental Experiments

In order to evaluate the improvements, we tested our method in the following way. We wrote capital letters with multiple strokes, A, B, D, E, on a document three times for each with heights of two to four lines, and compare them to the physical ink. Figure 5 illustrates the results. As shown in this figure, the pen-up and down recognition works well. The same holds for repositioning. It is often better to apply repositioning to have a better global position. However for characters with smaller heights, repositioning sometimes made the position worse due to erroneous matching between physical and digital ink. This is because the current method does not distinguish printed and handwritten parts for matching. This problem can be solved as follows. We are able to estimate the handwritten physical ink by, for example, subtracting the document image generated from PDF from the captured image. Applying the matching only with the resultant handwritten ink, a better position can be estimated. Further adjustment of size of digital ink can lead the repositioning even better.

6. Conclusion

In this paper we have proposed the notion of writing-life log and a way to realize this notion by using a camera-pen. Our camera-pen is unique since it requires no special paper to recover and localize the handwritings to an electronic document. The future work is to stabilize more the positioning of digital ink as well as to evaluate the method with larger experiments.

References

Chikano, M., Kise, K., Iwamura, M., Uchida, S., & Omachi, S. (2012). Recovery and localization of handwritings by a camera-pen based tracking and document image retrieval, Pattern Recognition Lett., in press, http://dx.doi.org/10.1016/j.patrec.2012.10.003

Arai, T., Aust, D., & Hudson, S.E. (1997). Paperlink: A technique for hyperlinking from real paper to electronic content. Proc. of the SIGCHI conference on Human factors in computing systems, pp.327–334.

Munich, M. E., & Perona, P. (2002). Visual input for pen-based computers, IEEE Trans. PAMI, 24 (3), pp.313–328.

Yasuda, K., Muramatsu, D., & Matsumoto, T. (2008). Visual-based online signature verification by pen tip tracking, Int'l Conf. on Computational Intelligence for Modelling, Control and Automation pp.175–180.

Seok, J. H., Levasseur, S., Kim, K., & Kim, J. H. (2008). Tracing handwriting on paper document under video camera, Proc. of ICFHR2008.

Takeda, K., Kise, K., & Iwamura, M. (2011). Real-Time Document Image Retrieval for a 10 Million Pages Database with a Memory Efficient and Stability Improved LLAH, Proc. ICDAR2011, pp.1054–1058.

Lucas, B.D., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision, Proc. of IJCAI, pp.674-679.