DigiClip: Activating Physical Documents

Christian Decker[1)], Michael Beigl[1)], Adam Eames[2)], Uwe Kubach[3)]
[1)]*TecO, University of Karlsruhe, Germany*
[2)] *Massachusetts Institute of Technology, Cambridge, MA, USA*
[3)]*Corporate Research, SAP AG, Karlsruhe, Germany*

Abstract

Today many documents are held in parallel as electronic files and as paper based documents. Keeping the properties of both versions in sync is important to assure consistency between the physical and virtual instances of one document. This paper introduces the DigiClip system that provides a solution to automatically enforce such consistency by converting passive paper documents to active physical documents. The main component of the system is the DigiClip device, a clip with embedded computing, sensing and communication technology that is attached to a paper document. The device is able to supervise the status of the document and to communicate to a software component that supervises the electronic instance of the same document. In this way the DigiClip system enables applications to automatically coordinate properties like integrity and access restrictions and to keep track of changes both of the physical and the virtual document.

1. Introduction

Bridging the virtual and physical world is one of the major topics and concepts in Ubiquitous and Pervasive Computing. The concept refers to the fact that many physical objects relate to (virtual) information in a computer system. The concept requires that both instances – the physical and virtual – should be aware of and reactive to changes at the other instance. As a consequence changes in the virtual world – e.g. changing information in a database – lead to an immediate reaction by the physical object – e.g. a change in the status displayed on the object. Vice versa, a change of the physical object's state – e.g. a relocation of the object – is reflected by a change of the information in the virtual world.

In this paper we concentrate on bridging the virtual and physical world for *documents*. Documents are often held as copies both on paper and electronic files. A decade ago, documents were created as paper and were made available for electronic processing afterwards. Today, many documents originate as electronic instances in computer files, while copies are printed out and then filed in shelves for legal issues, reliability issues or convenience. The system we present in this paper bridges the gap between the status of electronic documents and their printed-out copies in the physical world, by converting passive paper to *active physical documents*.

There are several research efforts going on in the area of integrating digital and paper based documents and information. One research area is using Tablet Computers and transferring (the majority of) the handling metaphors of physical paper to such devices. While these systems are able to support "natural" interaction as annotation in the electronic document [8] they are still not as convenient to use, regarding issues of form-factor and runtime as paper. Upcoming Digital Paper technology could be of help here, However, major players like Gyricon and Philips failed to fulfill their 2001 promise to integrate Digital Paper in consumer products by 2003 [4], indicating that the promise of availability of this technology in large quantities in 2010 may also be too optimistic. While these efforts purely look into the use of digital media in the same way as paper is used today, other research looks more into linking physical objects and virtual information. Webstickers [6] is one example where physical objects can be used to directly access Web-pages containing associated information by utilizing Barcodes that are already printed on many objects. In the CoolTown [5] project a similar system was constructed using active beacons that are attached to objects.

The DigiClip system focuses on the automatic tracking of changes in properties, characteristics and the overall composition of documents with instances in the physical and virtual world. Our primary application domain in this paper, concentrates on document processes with one paper document corresponding to one electronic document. Such processes can be found for instance in document archives and in publishing companies. Further, we assume that the system is used in a cooperative environment, where users do not cheat the system.

The DigiClip system introduces a new type of appliance device that is able to bring both worlds in sync with the help of sensor, communication and computing technology. Problems that have to be solved to apply such functionality are subject for analysis in the next section. The section also includes a first look at the basic functionality of the DigiClip system. Applications for the DigiClip are presented in section 3, while technology – especially of the DigiClip device – is shown in section 4. In section 5 we present

Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04)
0-7695-2087-1/04 $20.00 © 2004 **IEEE**

some results of the use of DigiClip system and the paper closes with a conclusion and future work in section 6.

2. Problem Analysis and DigiClip

The DigiClip system consists of four basic elements: The physical document with an attached, digitally enhanced paper clip (Figure 1), an electronic document corresponding with the physical document, a mediator software service and an infrastructure providing the connection and communication between both document instances (Figure 3). This system focuses on physical documents that are generated through print-out from an original electronic document. The basic idea behind the DigiClip system is to enforce the same properties and handling restrictions on the electronic version of the document and the physical paper counterpart.

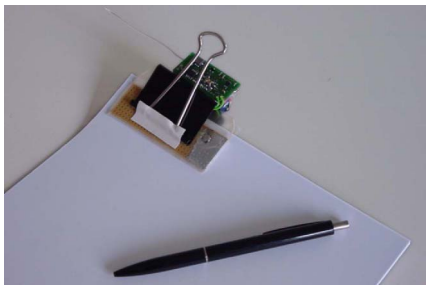

Figure 1. DigiClip clipped on some papers.

2.1. Problem Analysis

A prerequisite to enforce matching of characteristics between paper and electronic form of a document is that changes of properties of each of the document instances can be matched to the other instance. The following list shows a summary of document properties that we found important after analyzing several document handling tasks in our organization:

- An electronic document contains additional *meta-information* inside the document that are not or only partially available at the printed document. Such information typically includes author, theme and keywords, but also a complete revision history.
- *Access* to electronic documents is restricted on a per-user basis automatically through the computer system's document database or file system, while paper documents are only shielded from unauthorized access when stored in supervised places as archives.
- *Validity* of electronic documents can be bound to several activities. Examples are generation of a new document or the annulment of a process. In contrast to that, paper based validity is restricted to a predefined static property as the validity period.
- *Integrity* of a document is handled by applying rules to documents in electronic document management systems. Electronic integrity mechanisms automatically enforce the authorization or refusal to add or remove information from a document. In the physical world there isno possibility to enforce integrity of documents, apart from a manual supervision in archives.

In summary, the main difference between physical paper and virtual electronic documents is the lack of active, automatic supervision of document handling of a paper based document. The DigiClip system adds active supervision functionality to the paper document through embedding sensing, computing and communication technology in an object that is already used for paper processing: the paper clip. This DigiClip device keeps track of the paper document by supervising the states through sensors. The DigiClip mediator software running at the document file or database is responsible for supervising the states of the electronic document. To enforce synchronization between the DigiClip device and the mediator both parts exchange information wirelessly and are able to notify the user or document administrator about violation of the properties.

2.2. System Run and Potential

For the DigiClip system, an active physical document process starts with the print-out of an electronic document. After printing a DigiClip device is attached to the set of papers that typically make up the document. The DigiClip is immediately activated and constantly supervises the status of the document. At the same time, it is required to bind the physical print-out to its electronic original. Currently, this is left to the initiator of the print-out by manual input of the attached DigiClip device's unique ID for the electronic document. In this way the passive physical document is transformed to an *active physical document* that is able to electronically apply and enforce a set of properties such as:

- Validity of the document through a red LED indicating invalidity or green LED indicating validity
- Access to the electronic document indicated through a blue LED
- Updates of general properties from the electronic version, which are stored in the DigiClip's memory and readable wirelessly by small devices, e.g. PDAs
- Restrictions in use by supervising the space where it is used in and by indicating the relocation of the paper out of "authorized" places through sending a notification to the document administrator.
- Supervising integrity of the document content. E.g. the DigiClip can supervise that only additional paper sheets are allowed to be added to the document.

Vice versa, changes of the paper version of the document are automatically added as properties in the DigiClip's electronic version. E.g. adding a sheet of paper to the document results in a new entry in the document's revision history with time and place of change, in the same way moving the document from one location to another can be tracked by viewing the revision history.

3. Typical Applications

For active physical documents we see two main advantages over regular physical documents. First, we can electronically keep track of a physical copy's location. Second, we can ensure the consistency between electronic documents and their physical counterparts. In the following we describe two applications that make use of these characteristics.

3.1. Location Tracking

Since we can electronically keep track of an active physical document's location, for example the room, shelf or folder where it is located, this information can be stored as an additional attribute of the electronic original. Hence, a physical copy can be easily found by referencing the electronic original. This is an important advantage in environments where physical copies of electronic documents are mandatory, e.g. in lawyer and patent offices, and should not be misplaced.

Additionally this location tracking feature can be used in order to ensure that physical copies are only located within certain areas, for example within a certain building. We implemented a signaling mechanism that triggers an event when a document leaves a predefined area, for example the reading room of a library or the non-public part of an office building. In combination with a personal access control or tracking system the DigiClip system can be used to implement access restrictions to physical documents, since only people authorized to enter the area will have access to the documents.

3.2. Consistency

Another major problem with regular physical documents is to ensure their consistency with their electronic counterparts. Once a document is printed there is no link between a regular physical document and its electronic original. Hence changes of the electronic version have no effect on the physical copies. In order to manage consistency hereby, the occurrence of changes in the electronic original have to be communicated to the corresponding physical copy. The other way around, it is not always possible to see whether pages have been added or removed from the document after it has been printed. Therefore, physical documents are required to report back any changes in their page count to the electronic original. Consistency is a major issue in publishing companies printing paper based books and journals. Authors submit their manuscripts electronically, but copy-editors and proofreaders review and correct them manually on paper using the standardized proof-reading marks. During this process the electronic version is not allowed to be edited, otherwise the consistency is not warranted anymore. Furthermore, it should also be warranted that no page is left out.

With a signalization system based on active physical documents we achieve two improvements: First, we implemented an out-of-date signalization to all physical versions of a document as soon as the electronic original was changed. Second we were able to supervised the physical copy's integrity, at least in terms of ensuring that pages are not added to or removed from the copy without sending a notification to the electronic original.

4. Technology

This section will first analyze technological requirements resulting from the system functionality described in the previous sections and then explain the implementation of the DigiClip system in detail.

4.1. Requirements

We identified various requirements for a technology appropriate to support active physical copies of electronic documents. These are analyzed and explained in more detail below.

Location System. A location system able to keep track of a physical copy is required to implement control of access restrictions. Physical spaces can serve as a guarantor for limiting access to resources like documents since entering can be easily restricted. Hereby, a location system for physical documents enables a fine grain mapping of electronic access restrictions to physical spaces.

The electronic document contains various identification properties such as author, title etc., which enable a query for the document. A location system for physical documents enables the location of a physical counterpart and can track further the entire history of previous locations.

Sensing and computing at the physical copy. Sensing and computation technologies are required in order to control restrictions and constraints, which go beyond the capabilities of a location system. Moreover, for controlling the integrity of physical documents, sensing technology is required, which ensure the togetherness and integrity of multiple sheets of paper and reports when pages are removed or added. The technology has to be in-situ, i.e. directly at the document, to allow a fine-grained and clear control of these constraints.

Bi-directional wireless communication between the electronic original in the virtual world and the physical copy in the real world is needed to report changes on each

instance to the respective counterpart. In particular, applications concerned with consistency of documents require a confirmed communication for all notifications in order to work properly. Therefore, this communication is required to be bi-directional. Further, it is the appropriate solution for the convenient usage of physical documents.

No change of use with active physical documents. As we outlined in the introduction of this paper, physical copies are made for reasons like convenience, legal or editing issues. However, the use of technology is not allowed to weaken or even to negate these reasons when transforming regular physical copies into active ones. Importantly, users who do not want to benefit from the technology should not be affected in their documentary practices even when interacting with active physical documents.

Simple Usage. One major advantage of paper based physical copies is the simplicity of handling for human beings. In order to keep this convenience, it is required that the technology does not bother the user who wants to benefit from it. For this reason we demand exclusion of actions for technical operation or technical maintenance. The user should be unaware of the technology, unless it needs his attention. If an interaction with the physical copy is needed, e.g. setup of a new active physical copy, it should be embedded in an inartificial action performed on regular paper based copies.

4.2. DigiClip System

In the following section we describe the components of the DigiClip system in their technical details.

The DigiClip device consists of a regular paper clip augmented by computing, sensing and wireless communication capabilities. The post-hoc augmentation approach does not rely on prepared paper and can be applied to any kind of paper based physical document. As underlying hardware and firmware platform providing these capabilities, we used TecO's Smart-Its Particles[1], which were extended by a additional sensors and packed together with a AAA battery onto an off-the-shelf paper clip. Figure 2 shows a close look at DigiClip device.

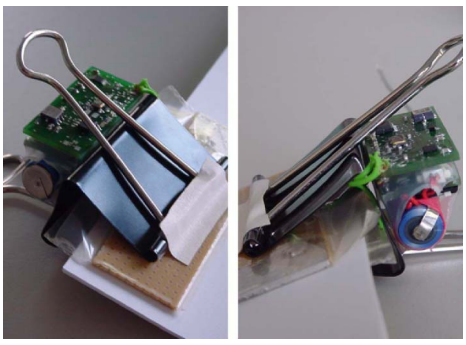

Figure 2. Close look at the DigiClip device.

In the current implementation, the DigiClip device's sensing capabilities allow the system to monitor the number of pages currently clipped and to recognize situations like clip is empty, clip is opened, paper is inserted or removed, document is moved, and document is packed in a bag. These situations are recognized on the clip itself. Different events like invalidation of the electronic original and violation of access rights can be displayed using the clip's multi-color LEDs. In order to exploit this functionality the DigiClip device provides a bi-directional wireless communication to a document management system.

Infrastructure. The link between the electronic original and the physical copy of a document relies on a network infrastructure. Electronic documents are stored in a document management system. The latter provides for each electronic document a DigiClip mediator component. The mediator recognizes changes in the electronic version of the document while the DigiClip device keeps track of the status of the physical version. Both, the mediator and the DigiClip device keep both document instances in sync by cooperatively updating the status of the respective other. Mediators extend the existing information about an electronic document by the location information and sensed situation of each assigned physical document. Mediators reside in the Document Management System which is connected via an IP network to a gateway which forwards messages from the DigiClip devices and vice versa. Using the mediators, the management system knows all assignments of electronic documents in relation to their physical copies and their respective electronic access restrictions This information can be queried by client computers via Internet. A web interface is provided for easy access documents and their associated information and to configure the whole system. The figure 3 summarizes all system components.

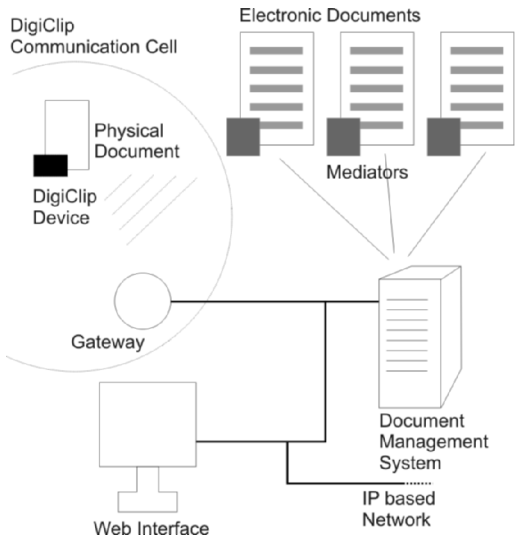

Figure 3. DigiClip System.

Location system. The infrastructure is used to establish a cell based location system with a room level granularity (Figure 3) using standard features of the Smart-Its Particle system. In each office room there are fixed installed Smart-its gateways which are aware of their own location.. The underlying Smart-Its network uses this information to provide the sending location in all communication packets[2]. Active physical documents are located from the mediator software by interpreting the received communication data from a DigiClip device. A timestamp – applied when these data arrives at the management system – is used to record the entire history of locations of each active physical document. This allows the tracking of clipped documents over time. The web interface in figure 3 illustrates the location of one document on a TecO floor plan.

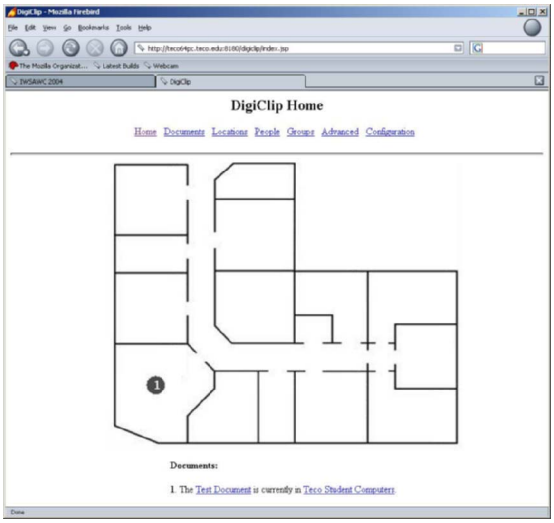

Figure 4. A Document's location is presented in a web interface

Communication and Sensing. The Smart-Its Particle platform used by the DigiClip offers a communication board enabling a wireless bi-directional communication and a generic sensor board with computation capability and various sensors. For the current implementation we mounted a light sensor, an accelerometer and a capacitive sensor.

Data from the accelerometer in cooperation with the light sensor data enabled us to detect whether the document was put into a bag or taken out of it. We implemented a first and very simple algorithm that uses a movement patterns together with a light level patterns to conclude on the situation "document packed into a bag".

The capacitive sensor on the DigiClip is able to sense the opening of the clip, and it serves for measuring of the document's page count. To do so, a stack of pages is considered as a capacitor. The sensor board implemented a circuit based on a sine wave generator and resistor for measuring this capacity. The principle can be found in [7]. The voltage across the capacitor was measured according to the number of pages between the capacitor. Since the capacity depends on the thickness of the stack the correspondence between page count and voltage is only valid for a stack of equal-think pages. This method will fail for a stack containing pages of various thickness or when using very different type of paper material. We did not see this as a limitation since in our analysis of the application domain the same kind of paper is used for all documents over years. In the figure 5 our results for the typical 80g/m² paper are presented.

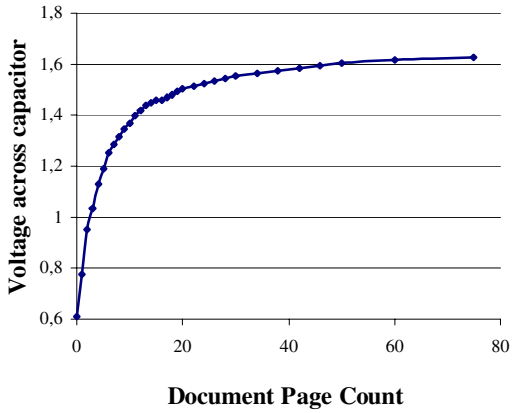

Figure 5. Measurement of pages in DigiClip

The measurement showed that for less than 20 pages in one stack, the measurement works very accurate. The adding and removal of one page could be precisely measured. With our current method beyond 20 pages the curve becomes too flat leading to inaccurate results.

Nevertheless, once this curve was initially measured and stored on the DigiClip device, it was able to recognize page count changes and to compute the actual page count of a document.

Web interface. The functionality of a DigiClip infrastructure and document management is accessible through a web interface. It serves as graphical user interface to the DigiClip mediators for each document. One can assign an electronic document to a physical copy, define an owner and a group of workers using a document. The user can use this web interface to keep track of all locations of each active physical documents and their current state. The interface provides further a convenient access for configuring the location of fixed installed gateways in the infrastructure.

5. First Results

Technical oriented results. We did some experiments to evaluate the handling of active physical documents. The room level location system works successfully, but needed initially some effort for adjustment. However, in seldom

cases where DigiClip devices couldn't be assigned precisely to one room, this led to problems in the location tracking application. More fine grained location resolution will be possible in the future based on the Particle's RF Receiver Signal Strength Indication (RSSI) and field strength regulation feature. On the sensor side, the DigiClip device was able to recognize its opening reliably without any failure. Measuring the page count was accurate in respect to the conditions we describe in 4.2 The detection ratio of situations – like the above mentioned "document packed in bag" – was not sufficient for usage in our applications. The implemented simple detection algorithm did not reach the minimum detection ratio we demanded for the usage in our applications. This indicates the need for more sensors. The underlying Smart-Its Particle communication platform fits the needs for the DigiClip application very well regarding bandwidth and reliability.

Application oriented results. The most important issue for our applications was that the DigiClip device should not negatively affect the handling with paper document. We found that the electronic augmentation of the paper clip did not influence its original functionality at all. The slightly increased size as a consequence of the capacitor's plates had also no effect on its usage. A difficulty arose from the weight of the clip. Although the electronic including the battery on the clip is less than half of the overall weight (Table 1), the DigiClip device is too heavy for only a few pages document. Such documents cannot be conveniently read without being placed on a table.

Plain Clip	32g
Battery	12g
Electronic and Plates	16g
DigiClip	60g

Table 1. Weights of Device Components

We did not have the time to look deeply into battery life time of the AAA battery supplying the DigiClip device. However, as a result of our experiences with the Chair appliance in the AwareOffice[3] an operational lifetime for about one year should also be possible for the DigiClip device since both systems share many of the hardware and software components. Also, the general behavior of the application on the DigiClip device is similar to the Chair appliance. This is mainly founded by the assumption that the DigiClip device - as the Chair appliance - needs only to measure and communicate when a physical trigger like a movement of the document occurs. Otherwise the device is powered down. The usual handling of the DigiClip device and a one-year battery lifetime does not bother the user with technical issues and therefore allows an convenient usage. The PC based components of the application are mainly built from off-the-shelf modules and technologies and worked as expected without major problems.

6. Conclusion and Future Work

For bridging the gap between electronic documents and their respective physical copy we introduced in this paper the concept of the active physical documents and its implementation as digital enhanced paper clips, called DigiClip devices. DigiClip is a system that fulfils requirements for activating physical document. Two typical example applications focusing on location tracking for applying access restriction, and on consistency between electronic original physical copy show how bridging the gap between the two instances of one document - the electronic and the physical one - can lead to new types of applications. First results indicate that this is a promising approach worth for further investigations. More sensors and more fine-grained mediation technology between physical and electronic documents will enable the system to supervise more complex settings as many physical documents per electronic document.

For the future we plan to refine the technology in order to make the DigiClip device more robust and to extend the usage of active physical documents. In a field test including lots of these clips a major effort will be spent in exploring of how various active physical documents can benefit from each other with and without the presence of an infrastructure.

Acknowledgements

The authors would like to thank all students workers at TecO for their appreciated work. In particular, we would like to thank Kay Schmidt for his help during the hardware design of the sensor board. The work is partially funded by Gottlieb Daimler and Karl Benz Foundation.

References

[1] Beigl, M., Zimmer, T., Krohn, A., Decker, C., Robinson, P.: Smart-Its - Communication and Sensing Technology for UbiComp Environments. Technical Report ISSN 1432-7864 2003/2

[2] Beigl, M., Zimmer, T., Decker, C.: A Location Model for Communicating and Processing of Context. Personal and Ubiquitous Computing Vol. 6 Issue 5-6, pp. 341-357, ISSN 1617-4909, 2002

[3] Beigl, M., Krohn, A., Zimmer, T., Decker, C., Robinson, P.: AwareCon: Situation Aware Context Communication. Proceedings of Ubicomp 2003, Oct. 12-15, Seattle, USA

[4] Ditlea, S.: The electronic paper chase. Scientific American, November 2001.

[5] Kindberg, T. et al.: People, Places, Things: Web Presence for the Real World. WMCSA 2000, December 7-8, Monterey, California, USA, pp. 19.

[6] Ljungstrand, P., Holmquist, Lars E.: WebStickers: Using Physical Objects as WWW Bookmarks. In Extended Abstracts of CHI '99, ACM Press

[7] Measurement of capacity in a RC circuit. Available Online: http://www.kfunigraz.ac.at/exp8www/Praktikum/kondensator1.htm [Accessed: 9/12/2003]

(8) Schilit, Bill N., Price, Morgan N., Golovchinsky, Gene: Digital Library Information Appliances. In Proceedings of the third ACM conference on Digital libraries, Pittsburgh, Pennsylvania, United States, 2003, pp.217-226