Integrating Digital Pens in Breast Imaging for Instant Knowledge Acquisition

Daniel Sonntag, Markus Weber
German Research Center for AI
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany

Matthias Hammon, Alexander Cavallaro
Imaging Science Institute, University Hospital Erlangen
Maximiliansplatz 1
91054 Erlangen, Germany

Abstract

Future radiology practices assume that the radiology reports should be uniform, comprehensive, and easily managed. This means that reports must be "readable" to humans and machines alike. In order to improve reporting practices in breast imaging, we allow the radiologist to write structured reports with a special pen on paper with an invisible dot pattern. In this way, we provide a knowledge acquisition system for printed mammography patient forms for the combined work with printed and digital documents. In this domain, printed documents cannot be easily replaced by computer systems because they contain free-form sketches and textual annotations, and the acceptance of traditional PC reporting tools is rather low among the doctors. This is due to the fact that current electronic reporting systems significantly add to the amount of time it takes to complete the reports. We describe our real-time digital paper application and focus on the use case study of our deployed application. We think that our results motivate the design and implementation of intuitive pen-based user interfaces for the medical reporting process and similar knowledge work domains. Our system imposes only minimal overhead on traditional form-filling processes and provides for a direct, ontology-based structuring of the user input for semantic search and retrieval applications, as well as other applied artificial intelligence scenarios which involve manual form-based data acquisition.

Introduction

Most current user interface technologies in the medical radiology domain take advantage of the inherent advantages paper provides over digital documents. Making medical diagnosis with paper documents is intuitive and smoothly integrates with reading the written diagnostic comments at a later stage when it comes to the patient treatment process. Therefore, many radiology practices have used paper reporting over the last 20 years or more. However, this situation is not optimal in the digital world of database patient records which have many advantages over current filing systems when it comes to search and navigation in complete patient repositories called radiology information systems. In fact, modern hospital processes require a digitisation of patient reports. Until now, there is no solution available which potentially combines the virtues of paper reporting in a paper-centric practice and the real-time digitisation of the paper contents—although there exist many empirical studies that domain experts prefer paper to digital media when it comes to reporting processes. A digital pen-based interface should enable radiologists to create high-quality patient reports more efficiently and in parallel to their patient examination task.

The current practice in hospitals is that a radiologist's dictated or written patient report is transcribed by hospital staff and sent back to the radiologist for approval. (Speech recognition is used more and more to reduce transcription tasks.) The turn-around time is 2-30 hours and the process inefficient. In order to improve the general medical scenario, we implemented a first prototype for radiology findings which several unique features: (1) a real-time digitisation into PDF documents of both text and graphical contents such as sketches; (2) real-time hand-writing recognition and real-time feedback on the recognition results on a computer screen; and (3) the mapping of the transcribed contents into concepts of several medical ontologies. We evaluated the first prototype in 2011 at the university hospital in Erlangen, Germany. After its technical improvements, we conducted a second more formal evaluation in 2012, in the form of a clinical trial with real radiologists in the radiology environment and with real patients.

In our improved scenario implementation and evaluation presented in this paper, we use a pen-based interface and a new "real-time interactive" paper writing modality. We extend the interactive paper for the patient reporting approach with an even more specific approach for mammography (the examination of the human breast). The goal of breast imaging is the early detection of breast cancer and involves x-ray mammography, magnetic resonance imaging (MRI), as well as a physical examination.

The design should therefore not only integrate physical documents and the doctor's (physical) examination task into an artificial intelligence (AI) application, but also the direct feedback about the recognised annotations to avoid turn-over times and additional staff. In contrast to our approach, traditional word and graphic processors require a keyboard and the mouse for writing and sketching. Even advanced speech recognition engines for clinical reporting cannot provide a good alternative. First, the free-form transcriptions do not directly correspond to medical concepts of a certain vo-

cabulary; second, the description of graphical annotations is by far more complex and prone to misunderstandings than a sketch-based annotation. In our scenario, the direct, fast, and flexible digital pen solution is an optimal extension of the analog paper reporting process which has been highly optimised by trial-and-error over the last 50 years. In this paper, we present our prototype and compare its performance and robustness to existing (electronic and AI based) data entry solutions that exist today.

Related Work

Primary data collection for clinical reports is largely done on paper with electronic database entry later. Especially the adoption of real-time data entry systems (on desktop computers) has not resulted in significant gains in data accuracy or efficiency. (Cole et al. 2006) proposed the first comparative study of pen-based data input and other (mobile) electronic data entry systems. The lack of availability of real-time accuracy checks is one of the main reasons why digital pen systems have not yet been used in the radiology domain (Marks 2004). It is a new concept which extends other attempts to improving stylus interaction for electronic medical forms (Seneviratne and Plimmer 2010).

Only recently, a variety of approaches have been investigated to enable an infrastructure for real-time pen-driven digital services: cameras, pen tablets (www.wacom.com), ultrasonic positioning, RFID antennas, bar code readers, or Anoto's technology (www.anoto.com). The Anoto technology, which we use, is particularly interesting because it is based on regular paper and the recording is precise and reliable. In order to become interactive, documents are Anoto-enabled at print time by augmenting the paper with a special Anoto dot pattern. In addition, for example *iGesture* (Signer, Kurmann, and Norrie 2007) can be used to recognise any pen-based gestures and to translate them into the corresponding digital operations. For the recognition of the contents of the form's text fields, and primitive sketch gestures, either the commercial Vision Objects (Knerr and Augustin 1998) or the Microsoft handwriting recognition engines (Pittman 2007) can be used.

Many digital writing solutions that specialise in healthcare with Anoto (http://anoto.com/healthcare-cases-3.aspx) are available, but these systems are "one-way" and do not use a special terminology for the terms to be recognised or support any gestures. In our scenario, we developed a prototype with these extensions, which is able to process the input in real-time and to give immediate feedback to the user. While using the interactive paper, we address the knowledge acquisition bottleneck problem for image contents in the context of medical findings/structured reporting. A *structured report* (Hall 2009) is a relatively new report generation technique that permits the use of pre-determined data elements or formats for semantic-based indexing of image report elements. In other related work, (Feng, Viard-Gaudin, and Sun 2009) for example, the input modality of choice is a tablet PC. While a tablet PC supports handwritten strokes, writing on it does not feel the same as writing on normal paper. Likewise, the physical paper serves as a certificate.

Scenario Implementation

The digital pen annotation framework is available at the patient finding workstation and the examination room. The radiologists finish their mammography reports at the patient finding station where they can inspect the results of the digital pen process. With the radiologist's signature, a formal report is generated according to the mammography annotations. The sketches the expert has drawn are also included in the final digital report. Anoto's digital pen was originally designed to digitise handwritten text on normal paper and uses a patented dot pattern on a very fine grid that is printed with carbon ink on conventional paper forms. We use the highest resolution dot pattern (to be printed with at least 600 dpi) to guarantee that the free-form sketches can be digitised with the correct boundaries. To use the high resolution dot pattern, the Bluetooth receiver is installed at the finding station; this ensures an almost perfect wireless connection. Please note that we use the digital pen in the streaming mode to ensure that the radiologist can inspect the results on screen at any time; our special Anoto pen research extension accommodates a Bluetooth sender protocol to transmit pen positions and stroke information to the nearby host computer at the finding station and interpret them in real-time.

In the medical finding process, standards play a major role. In complex medical database systems, a common ground of terms and structures is absolutely necessary. For annotations, we reuse existing reference ontologies and terminologies. For anatomical annotations, we use the Foundational Model of Anatomy (FMA) ontology (Mejino, Rubin, and Brinkley 2008). To express features of the visual manifestation of a particular anatomical entity or disease of the current image, we use fragments of RadLex (Langlotz 2006). Diseases are formalised using the International Classification of Diseases (ICD-10) (Möller et al. 2010). In any case, the system maps the handwriting recognition (HWR) output to one ontological instance. Images can be segmented into *regions of interest* (ROI). Each of these regions can be annotated independently with anatomical concepts (e.g., "lymph node"), with information about the visual manifestation of the anatomical concept (e.g., "enlarged"), and with a disease category using ICD-10 classes (e.g., "Nodular lymphoma" or "Lymphoblastic"). However, any combination of anatomical, visual, and disease annotations is allowed and multiple annotations of the same region are possible to complete the form.

Digital Pen Architecture

The pen architecture is split into the domain-independent Touch & Write system (Liwicki et al. 2012) and the application level. In Touch & Write, we have conceptualised and implemented a software development kit (SDK) for handling touch and pen interactions on any digital device while using pure pen interaction on paper. The SDK is divided into two components: the Touch & Write Core and the application specific part (see figure 1). The core part always runs on the interaction computer (laptop or desktop) as a service and handles the input devices (in this scenario the Anoto pen). The SDK contains state-of-the-art algorithms for

Radiology Pen Application

Application Layer

Interpretation Layer

- -

Touch&Write Core

Event Manager

Handwriting Recognition   Pen Gesture Recognition

Online Mode Detection

Ink Collector

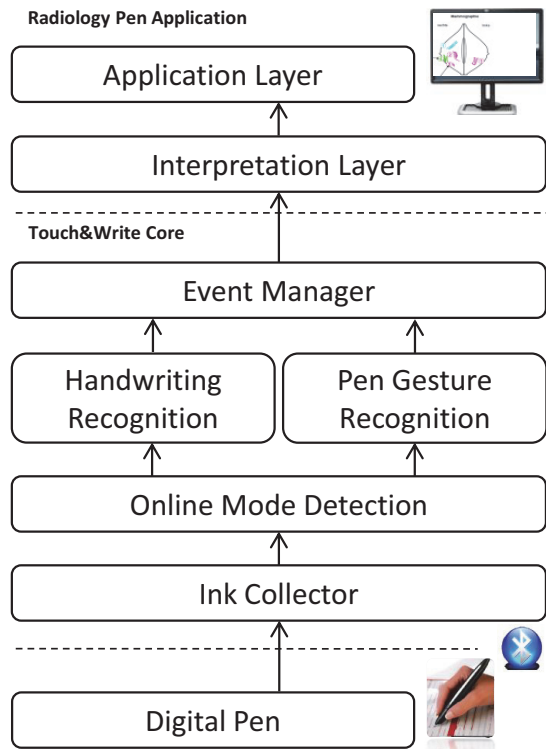- -

Digital Pen

Figure 1: Architecture of the pen interaction application.

analysing handwritten text, pen gestures, and shapes. Furthermore, it implements state-of-the-art algorithms in mode detection (Weber et al. 2011).

First, the *Digital Pen* component establishes a remote connection with the pen device via Bluetooth. Then it receives information on which page of the form the user is writing and its specific position at this page in real-time. This information is collected in the *Ink Collector* until the user stops interacting with the paper form. For the collection of the pen data, a stable connection is sufficient. The Anoto pen uses the Bluetooth connection for the transmission of the online data. Furthermore, it has an internal storage, to cache the position information, until the transmission can be completed. Here is a potential bottleneck, which could cause a delay in the interaction—a too great distance of the pen to the Bluetooth dongle could interrupt the connection. Because of the caching mechanism, no data get lost and can be collected when the connection is stable again.

Second, the *Online Mode Detection* component is triggered. Mode detection is the task of automatically detecting the mode of online handwritten strokes. Instead of forcing the user to switch manually between writing, drawing, and gesture mode, a mode-detection system should be able to guess the user's intention based on the strokes themselves. The mode detection of the Touch& Write Core distinguishes between handwriting, shapes drawing, and gesture which triggers the further analysis of the pen data. To classify the input, a number of features such as compactness, eccentric-

ity, closure, and so forth, are calculated (Weber et al. 2011). These features are used in a multi-classification system to detect the classes of handwritten information, shape drawings, or pen gestures. The system reaches an accuracy rate of nearly 98%.

Third, depending on the results of the mode detection either the *Handwriting Recognition* or the *Gesture Recognition* is used to analyse the collected stroke information. For the handwriting recognition and the shape detection the Vision Objects MyScript Engine[1] is used. The pen gestures are recognized using the iGesture framework (Signer, Kurmann, and Norrie 2007) with its extended version of the Rubine algorithm. The result of the analysis distributed via the *Event Manager* component. Both the iGesture framework and the Vision Objects engine are capable of providing immediate results, the user receives the results of the analysis and feedback on screen in less than a second.

Fourth, the application has to register at the *Event Manager* component in order to receive the pen events. There is a general distinction between the so-called low-level events and high-level events. Low-level events include raw data being processed like positions of the pen. High-level events contain the results of the analysis component (e.g., handwriting recognition results, detected shapes, or recognised gestures.)

On the application level the events are handled by the *Interpretation Layer*, where the meaning of the detected syntactic handwritten text and pen gestures is analysed depending on the position in the paper form. Finally, the application layer provides the visual feedback depending on the interpretation of the events, the real-time visualisation of the sketches, gestures, and handwritten annotations.

As in (Hammond and Paulson 2011) and (Steimle, Brdiczka, and Mühlhäuser 2009), we differentiate between a conceptual and a syntactic gesture level. On the gesture level, we define the set of domain-independent gestures performed by the (medical) user. Besides the hand-writing, these low-level strokes include circles, rectangles, and other drawn strokes. It is important to note that our recognisers assign domain-ignorant labels to those gestures. This allows us to use commercial and domain-independent software packages for those recognitions. On the conceptual level, a domain-specific meaning and a domain-specific label is assigned to these gestures. In our specific mammography form context, the position on the paper defines the interpretation of the low-level gesture. In fact, the domain-specific interpretations of the primitive gestures (examples are shown in figure 2) provides for domain practicality and the reduction of the cognitive load of the user (detailed in the usability evaluation, the digital paper can be filled out with primitive, self-explaining gestures).

The second "operational" interactive paper form for structured mammography reports (see figure 3) spans over two full pages and its division into different areas is a bit more complicated as illustrated in this paper. The interpretation example (see figure 2, bottom) shows different interpretations of a circle in free text areas, free-form sketch areas,

---

[1]http://www.visionobjects.com/. Last seen on 10/20/2012.

Gestures:

Interpretations:

Free Text Area: character "o/O", or "0" according to the text field grammar

Sketch Area: position of specific area or coordinate

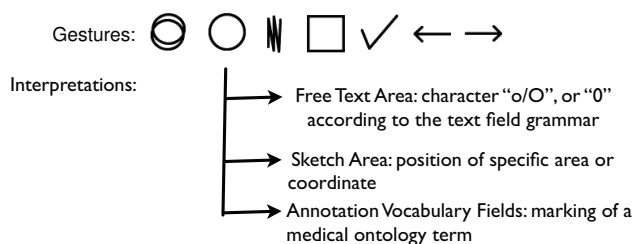Annotation Vocabulary Fields: marking of a medical ontology term

Figure 2: Gesture set and three dominant interpretations

and the predefined annotation vocabulary fields. In our first implementation (Mammo1, 2011), we did not take much advantage of predefined interpretation grammars and tried to recognise all gestures in a big free text area. The current Mammo2 design (2012, evaluated here) accounts for many of these unnecessary complications for the recognition engine. It takes full advantage of the separation of the form into dedicated areas with dedicated text and gesture interpretation grammars.

Evaluation

The following five preparation questions for improving the radiologist's interaction with the computer of the patient finding station arise:

- How is the workflow of the clinician?

- What kind of information (i.e., free-form text, attributes, sketches) is relevant for his daily reporting tasks?

- At what stage of the medical workflow should reported information items be controlled (by the clinician)?

- Can we embed the new intelligent user interface into the clinicians workflow while examining the patients?

- Can we produce a complete and valid digital form of the patient report with one intelligent user interface?

Four different data input devices were tested: the physical paper used at the hospital, our Mammo Digital Paper (AI-based), the iSoft PC mammography reporting tool (2012 version)[2], and an automatic speech recognition and reporting tool (Nuance Dragon Medical, 2012 version, AI-based)[3]. We are mostly interested in a formal evaluation of ease-of-use and accuracy so that we do not disrupt the workflow of the clinician. Additional features: (1) Multiple Sketch Annotations: the structured form eases the task of finding appropriate annotations (from FMA, ICD-10, or RadLex); some yes/no or multiple choice questions complete the finding process. Multiple colours can be selected for multiple tissue manifestations. (2) Annotation Selection and Correction: the user is able to use multiple gestures, e.g., underline or scratch out a concept in the free text fields. Then he or she has the possibility to select a more specific term (displayed on the computer screen) or refine/correct a potential

recognition error. This makes the paper interaction really interactive and multimodal. We also use the iGesture framework to select the colours on a virtual colour palette printed on the physical forms (in colour); the user can circle a new paint-pot to get this colour's ink to sketch and annotate in a specific colour.

Evaluating AI-based and Traditional Methods

In the formal clinical evaluation study, we observed two senior radiologists in the mammography scenario with real patients. Additional seven radiologists were able to test the application apart from the daily routine. These experts also controlled the creation of the accuracy evaluation. A usability engineer was present at the patient finding workstation (host) while the doctor engages in the patient examination task (without visibility) and data entry task (with visibility).

Data input using a usual paper form with and without a digital pen was used. So each doctor had to perform the form-filling process twice. This ensures minimal change to the daily routine and the possibility to observe the doctor in the daily examination routine. The input forms (paper and Mammo Digital Paper) had the same content and layout. Two radiologists with experience in breast imaging participated in this experiment. Each reader included 18 consecutive patients during clinical routine performing the two data input methods (resulting in 36 fully-specified patient records with a total of 3780 annotation fields whereby 765 have been used. Sparsity = 0.202). The usual paper form served as reference standard for data collection. After the workday every reader evaluated the documentation results. Breast cancer diagnosis included MRI imaging. Additional seven radiologists were able to test the application apart from the daily routine. Standard usability forms (questionnaires) were filled out in order to identify objective key features and a comparison to other data entry systems the radiology team was familiar with.

The form consists of 2 sections: (1) MRI imaging including different attributes for the characterisation of lesions as well as numbers for BI-RADS classification; (2) assessment of the results in free text form. The number of data entry errors was determined by comparing the results of the different methods.

Evaluation Results

The results are shown in table 1. We highlighted the new digital pen features we implemented in Mammo Digital Pen. As can be seen, the new digital pen system features of immediate validation, offline validation, real-time recognition of text, online correction of recognition errors, real-time capture to structured database, and forward capture to database (with the help of a transcriber), which have previously been reserved for PC and/or ASR systems, can now be done with digital pens employing automatic stroke interpretation. In addition, the real-time recognition of gestures and using the digital source document as a certificate (the captured signature can be officially used) are unique features of the Mammo Digital Paper system.

In many specific reporting tasks such as radiological reporting, dictation (preferably with modern ASR systems) is

Free-
form
Sketch
Area

Klinischer Befund

rechts   links

Links
Verdichtung: **yes**
Knoten:
Größe in cm: **315**
Lokalisation (cm v. Mamille/h): **2**
BI-RADS: 3

**Freeform
Text Area
(as in structured
CT reports)**

Links
Herdbefund
Form: rund / oval / **lobuliert** / **irregulär**
Rand: glatt / irregulär / spikuliert
Enhancement: homogen / heterogen / **Rand-E.**
                         dunkle int. Sept. / anreich. int. Sept.
                         zentr. Enhancement

nicht raumford. Enhancement
Verteilung: fokal / linear / duktal / segmental
                         regional / multipel / **diffus**

intern. Enhancement: homogen / heterogen
                         punktiert / gruppiert / retikulär

Symmetrie: symmetrisch / **asymmetrisch**

assoziierte Befunde: **cell Lymp**

**Annotation
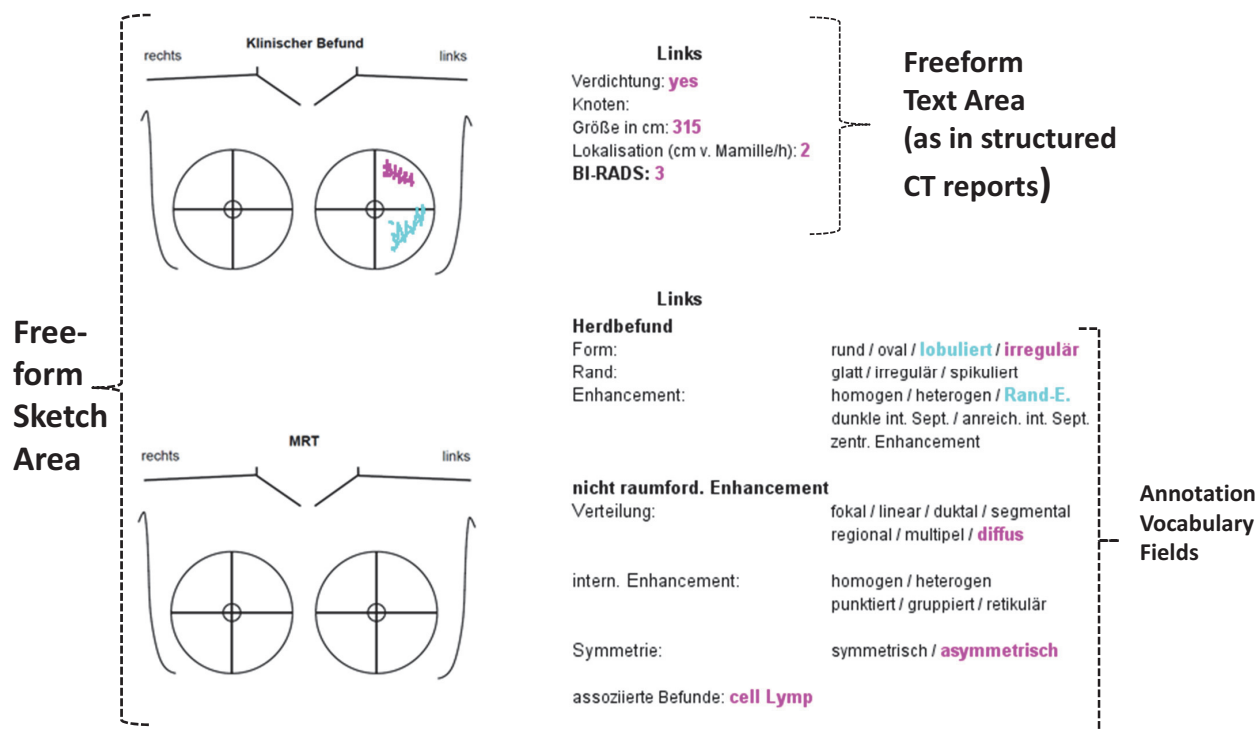Vocabulary
Fields**

MRT

rechts   links

Figure 3: Realtime interactive paper screen display for structured mammography reports

performed. Though, in the department we evaluated, paper based data collection dominates during breast imaging because many digital devices are immobile and too unwieldy. Nevertheless, flexibility is crucial in this clinical setup. The data entry system should be completely mobile in order to work with it in different situations such as taking the patient's medical history during the ultrasound examination or during the mammography reporting. The usage of the usual paper form enables quick and very comfortable data input and provides a high user satisfaction. This is partly due to the fact that because of the resemblance to the source paper forms, no additional training hours were needed. All radiologists noted that flexibility during the data input promotes a good doctor-patient relationship what is crucial for patients' satisfaction and recovery (no distraction from primary task; no distraction from patient). The user distraction from primary task is one of the main issues with any clinical PC reporting software.

Conclusion and Future Work

We presented a digital pen-based interface for mammography forms and focussed on a evaluation of normal paper and digital paper which also included a comparison to PC reporting and a potential speech recognition system. All digital data input devices improve the quality and consistency of mammography reports: the direct digitisation avoids the data transcription task of professional typists. The radiologist team was in general very supportive to test the new digital paper form. According to their comments, it can be said that most of them feel that digital documentation with desktop PC computers (without AI support) is in many respects a step backward. The results of the clinical evaluation confirm this on the measures of ease-of-use/user distraction and efficiency. The results presented here may differ with other, more integrative desktop PC or ASR reporting software. The possibility to reduce real-time recognition errors and logic errors as the data is being collected has great potential to increase the data quality of such reports over the long run. There's also great potential for reasoning algorithms and ontology-based deduction. With automatic concept checks of medical terms for example, educators may find interactive papers for mammography can help trainees learn the important elements of reports and encourage the proper use of special radiology terms. Future work includes the recognition of the meaning of the handwritten strokes in the sketch areas on the conceptual, medical level. (Feng, Viard-Gaudin, and Sun 2009), for example, already recognise more complex sketches, but their system also does not support sketches with medical meaning and handwritten textual medical annotations.

Table 1: Comparision of data entry. Key features for data collection/verification (upper part) and ease-of-use (lower part)

System features	Paper	Mammo Digital Paper	PC (iSoft)	ASR (Nuance)
Pen-on-paper interface	x	x		
Immediate validations		**x**	x	x
Offline validation (of digital content)		x	x	x
Realtime recognition (text)		**x**	x	x
Realtime recognition (gestures)		**x**		
Online correction of recognition errors		**x**	x	
Real-time capture to structured database		**x**	x	
Forward capture to database		**x**	x	x
Source Document (Certificate)	x	**x**	(x)	
Digital Source Document (Certificate)		**x**		
Training hours before effective usage	10	10	30	35
No user distraction from primary task	x	x		
No distraction from patient	x	x		(x)
Average time to complete one predefined Radlex entry	3sec	3sec	5sec	2sec

References

Cole, E. B.; Pisano, E. D.; Clary, G. J.; Zeng, D.; Koomen, M.; Kuzmiak, C. M.; Seo, B. K.; Lee, Y.; and Pavic, D. 2006. A comparative study of mobile electronic data entry systems for clinical trials data collection. *I. J. Medical Informatics* 722–729.

Feng, G.; Viard-Gaudin, C.; and Sun, Z. 2009. On-line hand-drawn electric circuit diagram recognition using 2d dynamic programming. *Pattern Recognition* 42(12):3215 – 3223. New Frontiers in Handwriting Recognition.

Hall, F. M. 2009. The radiology report of the future. *Radiology* 251(2):313–316.

Hammond, T., and Paulson, B. 2011. Recognizing sketched multistroke primitives. *ACM Trans. Interact. Intell. Syst.* 1(1):4:1–4:34.

Knerr, S., and Augustin, E. 1998. A neural network-hidden Markov model hybrid for cursive word recognition. In *Int. Conf. on Pattern Recognition*, volume 2, 1518–1520.

Langlotz, C. P. 2006. Radlex: A new method for indexing online educational materials. *RadioGraphics* 26:1595–1597.

Liwicki, M.; Weber, M.; Zimmermann, T.; and Dengel, A. 2012. Seamless integration of handwriting recognition into pen-enabled displays for fast user interaction. In *Document Analysis Systems DAS 2012 10th IAPR International Workshop on Document Analysis Systems*, 364 – 368.

Marks, R. G. 2004. Validating electronic source data in clinical trials. *Controlled clinical trials* 25(5):437–446.

Mejino, J. L.; Rubin, D. L.; and Brinkley, J. F. 2008. FMA-RadLex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. In *Proc. of AMIA Symposium*, 465–469.

Möller, M.; Ernst, P.; Dengel, A.; and Sonntag, D. 2010. Representing the international classification of diseases version 10 in OWL. In *Proc. of the International Conference on Knowledge Engineering and Ontology Development (KEOD)*.

Pittman, J. A. 2007. Handwriting Recognition: Tablet PC Text Input. *Computer* 40(9):49–54.

Seneviratne, N., and Plimmer, B. 2010. Improving stylus interaction for eMedical forms. In *Proceedings of the 22nd OZCHI Conference*, 280–287. New York, NY, USA: ACM.

Signer, B.; Kurmann, U.; and Norrie, M. 2007. iGesture: A General Gesture Recognition Framework. In *Proceedings of the Ninth International Conference on Document Analysis and Recognition - Volume 02*, ICDAR '07, 954–958. Washington, DC, USA: IEEE Computer Society.

Steimle, J.; Brdiczka, O.; and Mühlhäuser, M. 2009. Coscribe: Integrating paper and digital documents for collaborative learning. *IEEE Transactions on Learning Technologies*.

Weber, M.; Liwicki, M.; Schelske, Y. T. H.; Schoelzel, C.; Strauß, F.; and Dengel, A. 2011. MCS for Online Mode Detection: Evaluation on Pen-Enabled Multi-touch Interfaces. In *ICDAR*, 957–961.