Integrating Digital Pens in Breast Imaging for Instant Knowledge Acquisition

Daniel Sonntag, Markus Weber, Alexander Cavallaro, Matthias Hammon

■ *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.*

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 again 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 digitization of patient reports. Until now, there is no solution available that potentially combines the virtues of paper reporting in a paper-centered practice and the real-time digitization 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 turnaround 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 with several unique features: (1) a real-time digitization into PDF

  ISSN 0738-4602

documents of both text and graphical contents such as sketches; (2) real-time handwriting or gesture 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 article, 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 physical examination, X-ray mammography, ultrasound, as well as magnetic resonance imaging (MRI).

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 recognized annotations to avoid turnover 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 vocabulary; 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 optimized by trial-and-error over the last 50 years. In this article, we present our prototype and compare its performance and robustness to (electronic and AI based) data-entry solutions that exist today.

Background

The THESEUS Radspeech project (Sonntag 2013) focuses on the knowledge-acquisition bottleneck of medical images: we cannot easily acquire the necessary image semantics of medical images that ought to be used in the software application as it is hidden in the heads of the medical experts. One of the selected scenarios aims for improved image and patient data search in the context of patients that suffer from cancer. During the course of diagnosis and continual treatment, image data is produced several times using different modalities. As a result, the image data consist of many medical images in different formats, which additionally need to be associated with the corresponding patient data, and especially in the mammography case, with a physical examination. In addition, in the current mammography practice the reporting results are often inaccurate because prespecified standardized terminologies for anatomy, disease, or special image characteristics are seldom used.

That's why we base our development on a special mammography form where radiologists can encircle special characteristics and add free text commands that can only be filled in with medical terms of a prespecified medical vocabulary or free-form sketches. Thus, we are concerned with a semantic tagging task, which is the primary reporting task of the radiologist. According to the results of the evaluation, we advocate the usage of both digital and physical artifacts. From an ecological perspective, the key for providing a successful interface lies in the seamless integration of artificial intelligence into the distributed cognition task of the user — to draw sketches and annotate radiology terms while the patient is present or the doctor skims the radiology pictures. Regarding the fact that the user and the intelligent system should have a collaborative goal (Sonntag 2012) to allow for a dedicated integration of AI technology, we followed the Cognitive Work Analysis (CWA) approach, which evaluates first the system already in place and then develops recommendations for future design (Vicente 1999). The resulting evaluation is then based on the analysis of the system's behavior in the actual medical application context. Cognitive task analysis (CTA) is also related as it underlines that advances in technology have increased, not decreased, mental work in working environments while doing a specific work-related task (Militello and Hutton 1998).

For example, all radiology doctors in our evaluation confirm that they can better focus on the patient while writing with a pen instead of using a computer. We should therefore support the digital interaction with printed documents to report in a digital form and to allow a radiologist to control the recognition results on the fly. The potential of a digital version is the possibility of a real-time processing (and sharing) of the handwriting and the recognition results. The advantage of a digital pen version over a tablet PC is for example the high tracking performance for capturing the pen annotations without restricting the natural haptic interaction with a physical pen or the need for additional tracking devices.

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 digital 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 digital pen systems have not yet been used in the radiology domain (Marks 2004). It is a new concept that 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 made compatible with Anoto 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 recognize 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 or the Microsoft handwriting recognition engines (Pittman 2007) can be used.

Many digital writing solutions that specialize in health care with Anoto (see Anoto's Industry Case Studies online) are available, but these systems are "one-way" (that is, the results can only be inspected after the completion of the form, thus no interaction is possible) and do not use a special terminology for the terms to be recognized or support any gestures. In our scenario, we developed a prototype with these extensions that 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 predetermined data elements or formats for semantic-based indexing of image report elements. In other related work, for example, the input modality of choice is a tablet PC (Feng, Viard-Gaudin, and Sun 2009). While a tablet PC supports handwritten strokes, writing on it does not feel the same as writing on normal paper. Another difference is that 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 (see figure 1). Anoto's digital pen was originally designed to digitize 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 digitized 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 a new continuous streaming mode to ensure that the radiologist can inspect the results on screen at any time; our special Anoto pen research extension accommodates a special 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 formalized 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 (for example, "lymph node"), with information about the visual manifestation of the anatomical concept (for example, "enlarged," "oval," "unscharf/diffuse," "isoechogen," which are predefined annotation fields to be encircled), and with a disease category using ICD-10 classes (for example, "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 (Dengel, Liwicki, and Weber 2012) and the application level. In Touch & Write, we have conceptualized 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

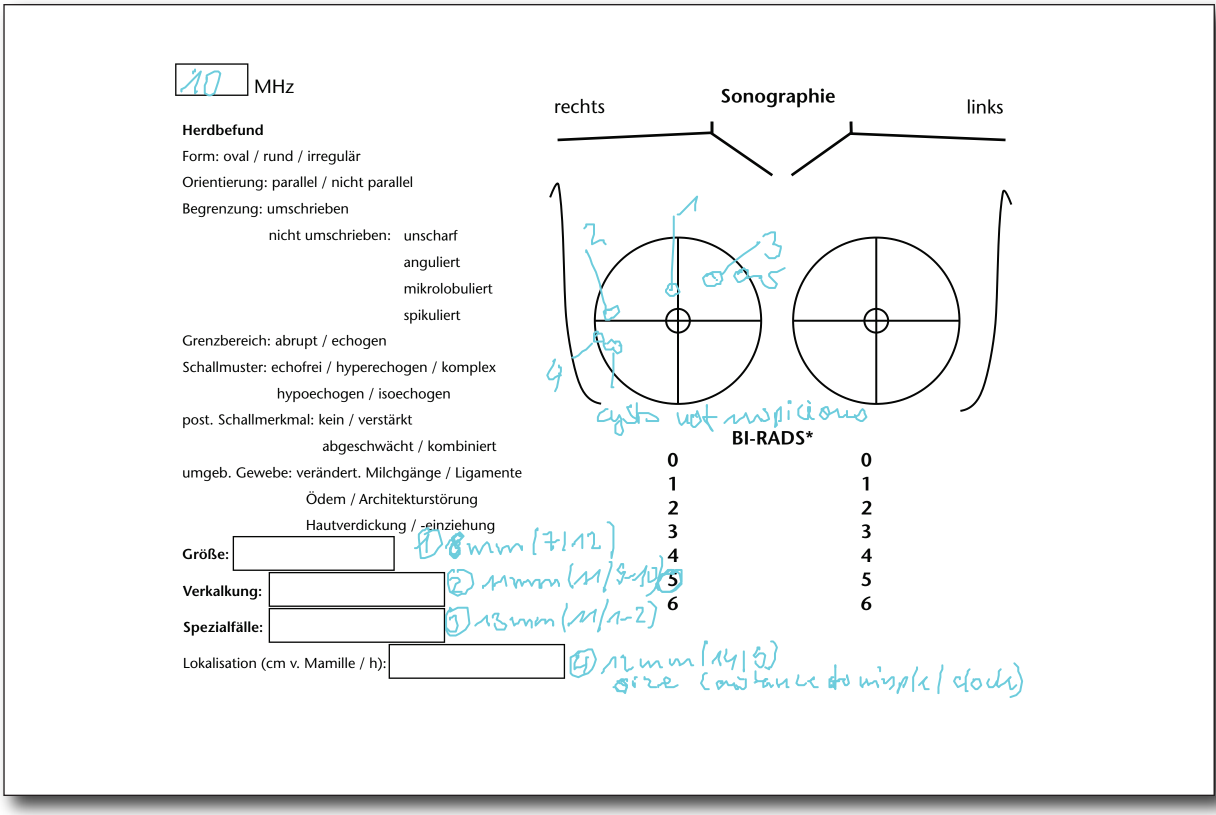

Figure 1. Hand-Drawn Sketches and Free Text Annotations.

Core and the application-specific part (see figure 2). 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 analyzing handwritten text, pen gestures, and shapes. Shape drawings are sketches using simple geometric primitives, such as ellipses, rectangles, or circles. The shape detection is capable of extracting the parameters for representing the primitives such as the centroid and radius for a circle. By contrast, pen gestures trigger a predefined action when performed in a certain area of the form, such as selecting or deselecting a term, or changing the ink color. Furthermore, the SDK implements state-of-the-art algorithms in mode detection (Weber et al. 2011), which we will, due to their importance, describe in greater detail.

First, the Digital Pen establishes a remote connection with the pen device through 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 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 gestures that trigger the further analysis of the pen data. To classify the input, a number of features such as compactness, eccentricity, closure, and so forth, are calculated. These features are used in a multiclassification and voting system to detect the classes of handwritten information, shape drawings, or pen gestures. The system reaches a recognition rate of nearly 98 percent — see also Weber et al. (2011). Mode detection is essential for any further automatic analysis of the pen data and the correct association of the sequential information in the Interpretation Layer. In fact, online mode detection provides for domain practicality and the reduction of the cognitive load of the user.

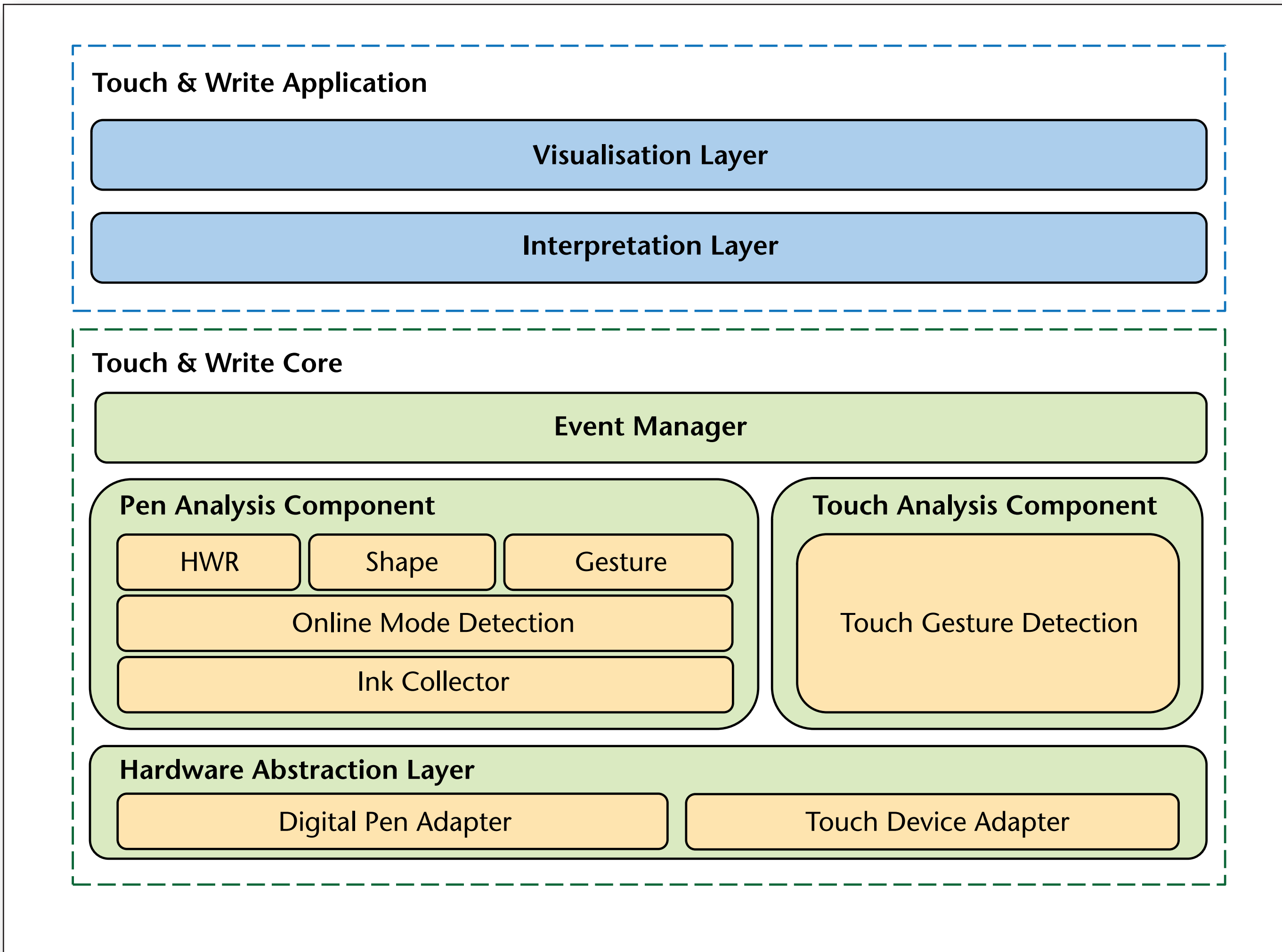

Figure 2. Architecture of the Pen Interaction Application.

Third, depending on the results of the mode detection either the Handwriting Recognition or the Gesture Recognition is used to analyze 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), which uses an extended version of the widely used single- and multistroke algorithm presented in Rubine (1991). The result of the analysis is distributed through 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. Figure 1 illustrates a current combination of written and hand-drawn annotations.

The application has to register at the Event Manager 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 (for example, handwriting recognition results, detected shapes, or recognized 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 analyzed 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 visualization 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 handwriting, these low-level strokes include circles, rectangles, and other drawn strokes. It is important to note that our recognizers assign

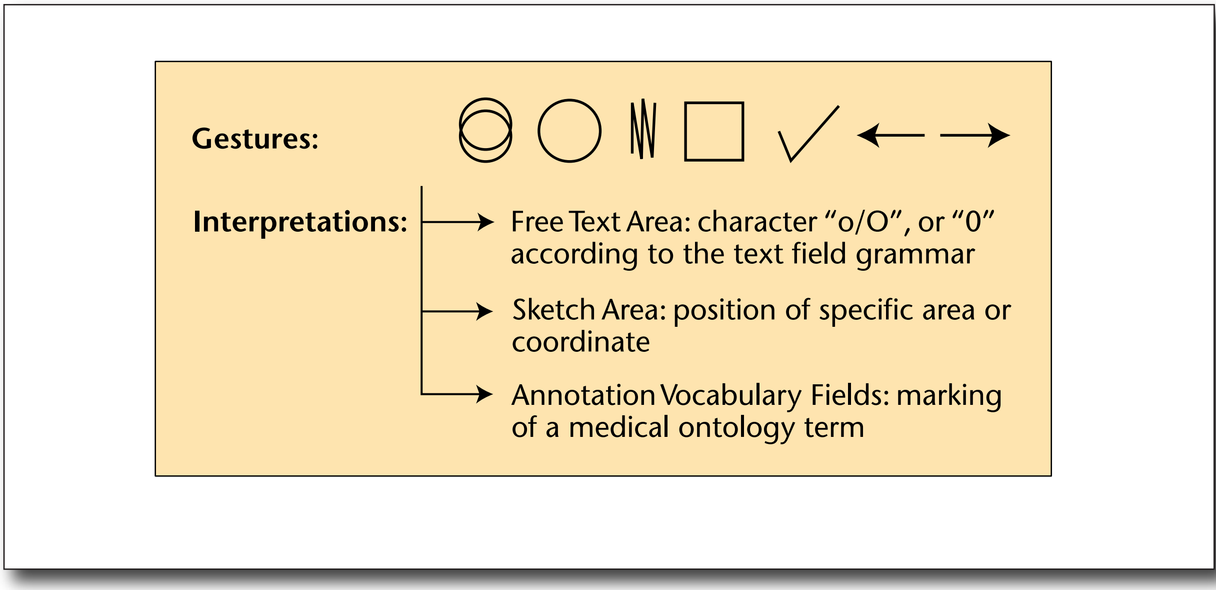

Figure 3. Real-Time Interactive Paper Screen Display for Structured Mammography Reports.

domain-ignorant labels to those gestures. This allows us to use commercial and domain-independent software packages for the recognition of primitives on the syntactic level. On the conceptual level, a domain-specific meaning and a domain-specific label are assigned to these gestures (see next section). In our specific mammography form context, the position on the paper defines the interpretation of the low-level gesture.

The resulting screen feedback of the 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 article. The interpretation example (see figure 4, bottom) shows different interpretations of a circle in free text areas, free-form sketch areas, and the predefined annotation vocabulary fields. In the mammography form implementation of 2011, we did not take much advantage of predefined interpretation grammars and tried to recognize all gestures in a big free text area. The current design of 2012/2013, which is 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.

Pen Events and Online Mode Detection

Pen events are declared by basic notations introduced here. A sample

$$\vec{s}_i = (x_i, y_i, t_i)$$

is recorded by a pen device where (x_i, y_i) is a point in two-dimensional space and t_i is the recorded time stamp. A stroke is a sequence S of samples,

$$S = \{\vec{s}_i \mid i \in [0, n-1], t_i < t_{i+1}\}$$

where n is the number of recorded samples. A sequence of strokes is indicated by

$$D = \{S_i \mid i \in [0, m-1]\},$$

where m is the number of strokes. The area A covered by the sequence of strokes D is defined as the area of the bounding box that results from a sequence of strokes.

For a low-level pen event, the following raw data are provided by our new recognizer API taking the pen's global screen coordinates and force as input: *pen id* (a unique ID for the pen device); *(x, y)* (the relative *x, y* screen coordinate and time stamp [sample s_i]); *force* (normalized pen pressure measured by the pen device); *velocity* (*x, y*) (the velocity in *x* and *y* direction); *acceleration* (the acceleration); and *type* (the type indicates whether the event is a pen down, pen move, or pen up event).

High-level events contain the results of the pen analysis component (for example, handwriting recognition results, detected shapes, or recognized gestures). An event for the handwriting contains strokes (the sequence of strokes D on which the analysis is being applied); bounding box (a rectangle that defines the area A of the strokes D); and results (the handwriting recognition results, that is, a list of words and their alternatives in combination with confidence values).

Shape events are composed of *strokes* (the sequence of strokes D on which the analysis is being applied); and *shapes* (the list of the detected shapes and their parameters).

Currently, the shape detection detects circles,

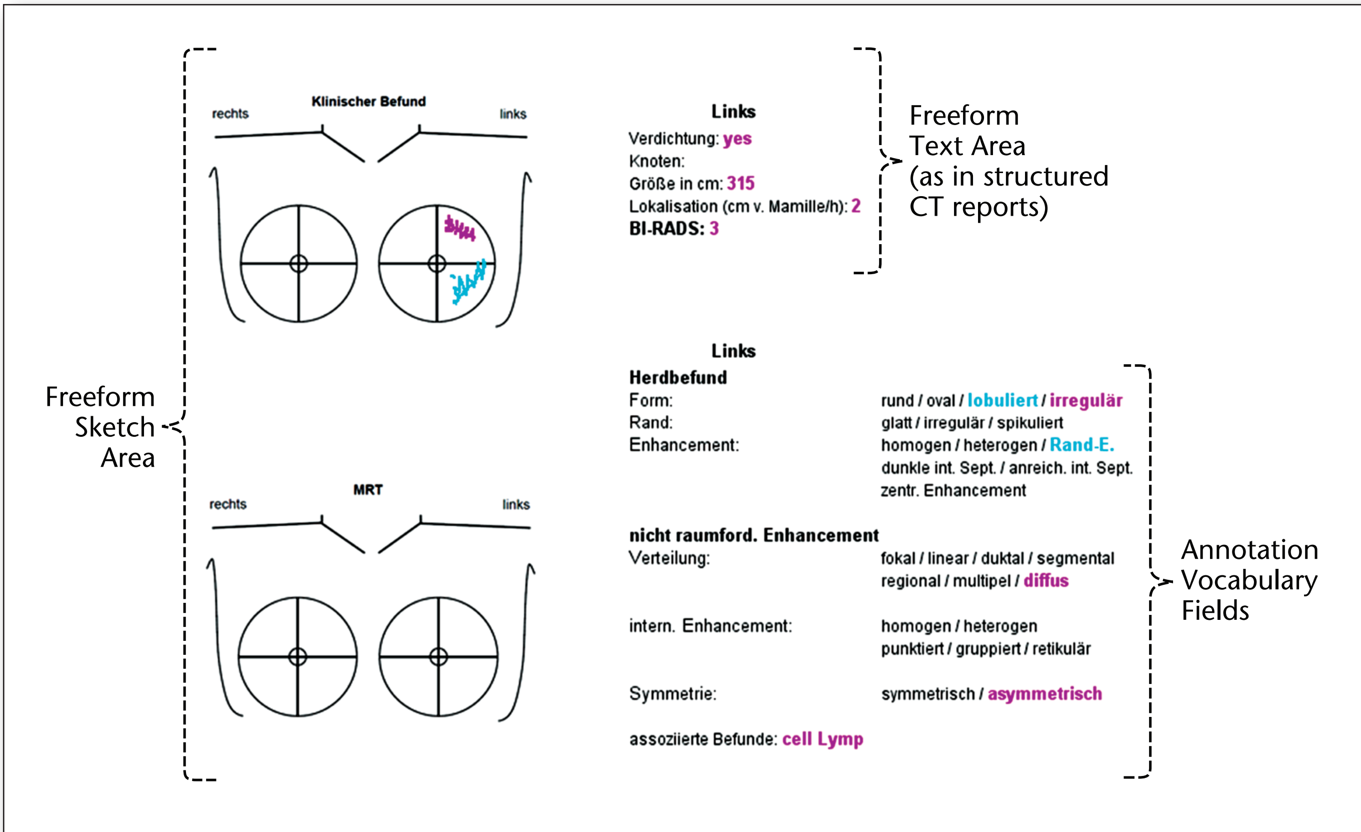

Figure 4. Gesture Set and Three Dominant Interpretations.

ellipses, triangles, and quadrangles. If none of these geometries are detected with high confidence, a polygon is approximated. Gesture events are composed of *strokes* (the sequence of strokes D on which the analysis is being applied), *gesture type* (the type of the gesture), and *confidence* (the aggregated gesture confidence value).

On the application level, the meaning of the detected syntactic event, for example, handwritten symbols and pen gestures, is analyzed according to the position in the paper form and domain-specific recognition grammars for areas. The usage of predefined (medical) stroke and text grammars are exclusively specified on the Application Layer. Finally, the visualization layer provides the visual feedback depending on the interpretation of the events, the real-time visualization of the sketches, gestures, and handwritten annotations.

Of course, the interaction with the text and sketch based interface should be intuitive, and a manual switch between drawing, handwriting, or gestures modes must be avoided. Thus it becomes necessary to distinguish between such different modes automatically. In the mammography form, we distinguish between three major modes: handwriting, drawing, and gesture mode. Our online mode detection is based on the method proposed by Willems, Rossignol, and Vuurpijl (2005). We will introduce basic notations for the mode detection. In addition to strokes, the centroid μ is defined as

$$\vec{\mu} = \frac{1}{n}\sum_{i=0}^{n-1}\vec{s_i}$$

where n is the number of samples used for the classification, the mean radius μ_r (standard deviation) as

$$\mu_r = \frac{1}{n}\sum_{i=0}^{n-1}\| \vec{s_i} - \vec{\mu} \|,$$

and the angle φ_{si} as

$$\varphi_{s_i} = cos^{-1}\left(\frac{(s_i - s_{i-1})\cdot(s_{i+1} - s_i)}{\| s_i - s_{i-1} \|\| s_{i+1} - s_i \|}\right).$$

Figure 5a shows an example of a recorded area together with its bounding box and the calculated centroid; figure 5b shows the extracted angle, which is only available from online detections. Table 1 contains a listing of the online features used for the classification of the mode. As long as the user is writing or drawing (continuous stylus input according to a time threshold), the strokes are recorded in a cache. As a result, the feature values are calculated for stroke

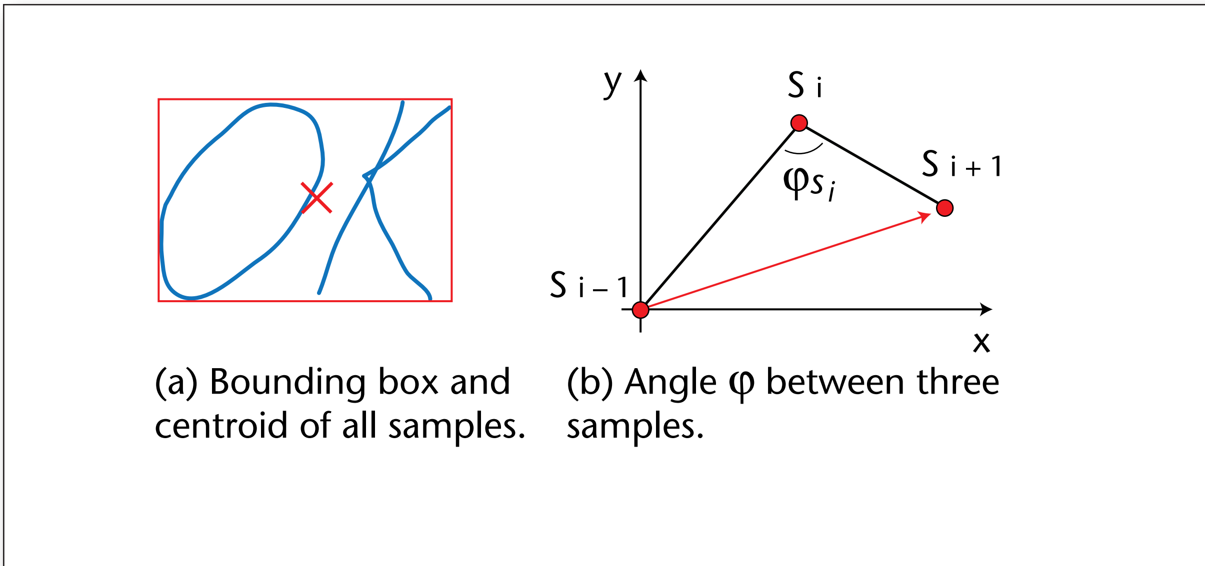

Figure 5. A Recorded Area Together with Its Bounding Box and the Calculated Centroid.

(a) Bounding box and centroid of all samples. (b) Angle φ between three samples.

sequences *D* by calculating individual stroke information and summing up the results. Whenever the detection is triggered, the feature vectors are computed from the cached data and the classification is performed in real time.

To classify the input, a number of features such as stroke length, area, compactness, curvature, and so forth, are calculated. Each mode of pen interaction, such as drawing, handwriting, or gestures, has its characteristics. Thus the extracted features should represent them and make the modes separable. For example, handwritten text tends to be very twisted and compact; hence the values of compactness and curvature are quite high in comparison to drawing mode where more primitive shapes are prevalent. Many distinctive features are based on the online feature angle φ.

Evaluation

The following five preparation questions for improving the radiologist's interaction with the computer of the patient finding station arise: (1) How is the workflow of the clinician; can we save time or do we try to digitize at no additional costs? (2) What kind of information (that is, free-form text, attributes, and sketches) is most relevant for his or her daily reporting tasks? (3) At what stage of the medical workflow should reported information items be controlled (by the clinician)? (4) Can we embed the new intelligent user interface into the clinician's workflow while examining the patients? (5) Can we produce a complete and valid digital form of the patient report with one intelligent user interface featuring automatic mode detection?

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 (according to the CWA/CTA procedures). Additional test features of Mammo Digital Pen are the following: (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 colors can be selected for multiple tissue manifestations. (2) Annotation Selection and Correction: the user is able to use multiple gestures, for example, 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 colors on a virtual color palette printed on the physical forms (in color); the user can circle a new paint-pot to get this color's ink to sketch and annotate in a specific color.

Evaluating AI-Based and Traditional Methods

In the formal clinical evaluation study, we observed two senior radiologists with experience in breast imaging in the mammography scenario with real patients. Additional seven radiologists were able to

ID	Feature	Description	Note
0	Number of Strokes	N	
1	Length	$\lambda = \sum_{i=0}^{n-1} \| vecs_i - vecs_{i+1} \|$	s_i denotes a sample.
2	Area	A	
3	Perimeter Length	λ_c	Length of the path around the convex hull.
4	Compactness	$c = \frac{\lambda_c^2}{A}$	
5	Eccentricity	$e = \sqrt{1 - \frac{b^2}{a^2}}$	a and b denote the length of the major or minor axis of the convex hull, respectively.
6	Principal Axes	$e_r = \frac{b}{a}$	
7	Circular Variance	$v_c = \frac{1}{n\mu_r^2} \sum_{i=0}^{n} (\| \vec{s}_i - \vec{\mu} \| - \mu_r)^2$	μ_r denotes the mean distance of the samples to the centroid μ_r.
8	Rectangularity	$r = \frac{A}{ab}$	
9	Closure	$c_l = \frac{\| \vec{s}_0 - \vec{s}_n \|}{\lambda}$	
10	Curvature	$\kappa = \sum_{i=1}^{n-1} \varphi_{\vec{s}_i}$	φ_{s_i} denotes the angle between the $\overline{s_{i-1}s_i}$ segments and $\overline{s_i s_{i+1}}$ at s_i.
11	Perpendicularity	$p_c = \sum_{i=1}^{n-1} \sin\left(\varphi_{\vec{s}_i}\right)^2$	
12	Signed Perpendicularity	$p_{sc} = \sum_{i=1}^{n-1} \sin\left(\varphi_{\vec{s}_i}\right)^3$	
13	Angles after Equidistant Resampling (6 line segments)	$\sin(\alpha), \cos(\alpha)$	The five angles between succeeding lines are considered to make the features scale and rotation invariant (normalization of writing speed).

Table 1. Online Features.

test the application apart from the daily routine. These experts also controlled 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 contents and similar layouts.

Each reader (senior radiologist) examined 18 consecutive patients/patient cases 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 and the seven additional radiologists evaluated the documentation results. Breast cancer diagnosis included MRI imaging. Standard usability forms (questionnaires) were filled out in order to identify objective key features and to provide a comparison to other data-entry systems the radiology team was familiar with.

The form evaluation focused on two medical sections: (1) MRI imaging including different attributes for the characterization 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 2. 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. This corresponds to the workflow of the clinician. We evaluated that a direct digitalization at no additional cost counts most. 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.

What kind of information counts most? In many specific reporting tasks such as radiological reporting, dictation (preferably with modern ASR systems) is performed. However, 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. Predefined annotation fields can be recognized at the recognition rate of the online mode detection of 98 percent. HWR and drawings vary according to the predefined grammar, where a trade-off between accuracy and coverage must be investigated in future evaluations. It cannot be said that any information of a specific mode is more important than that of another mode as this is highly case specific. In any case, real-time digitized information items should be controlled/corrected at acquisition time to avoid the data transcription/verification task of professional typists (which is also prone to error).

Can we embed the new intelligent user interface? 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. Can we produce a complete and valid digital form? The evaluation in table 2 was based on this requirement (given the correction at acquisition time), which has been justified empirically by the expert questionnaires.

Conclusion and Future Work

We presented a digital pen-based interface for mammography forms and focused on an evaluation of normal paper and digital paper, which also included a comparison to PC reporting and an automatic speech-recognition system. All digital data-input devices improve the quality and consistency of mammography reports: the direct digitization 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. Finally, after controlling the results on screen, a signature triggers a PDF report-generation process where the original user input can be seen, as well as the transcribed database entry. In

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	3 sec	3 sec	5 sec	2 sec

Table 2. Comparison of Data Entry.

Key features for data collection/verification (upper part) and ease of use (lower part).

addition, our approach provides other means to increase the data quality of future reports: with normal paper forms, logic errors can arise, for example by skipping required fields (such as the BI-RADS classification) or annotating with words that do not stem from the predefined vocabularies.

The possibility to reduce real-time recognition errors and logic errors as the data are 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. We firmly believe that a large-scale implementation of the Mammo Digital Pen technology all over the country can help improve the quality of patient care because similar cases can be found more easily and used in case-based reasoning applications toward automatic decision support. Toward this goal, the reliability of recognition concerning sketches and text labels at various positions has to be improved considerably; this detection assumes a perfect detection of different modes in fast succession. Future work includes the interpretation of the handwritten strokes in the sketch areas on the conceptual, medical level, for example, "does the form value 'round' correspond to the shape in the sketch area?"

Acknowledgements

This research has been supported in part by the THESEUS Program in the RadSpeech project, which is funded by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016 and the EIT ICT Labs in the Medical Cyber-Physical Systems activity. We would also like to thank Daniel Gröger and Marcus Liwicki for their support in the technical realization of the pen-enabled form and the radiology team of the Image Science Institute in Erlangen, Germany, for their participation in the technical evaluation.

Notes

1. See www.visionobjects.com/.

2. See www.isofthealth.com/en/Solutions/Department/Radiology.aspx.

3. See www.nuance.com/for-healthcare/capture-anywhere/radiology-solutions/index.htm.

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. *International Journal of Medical Informatics* 75(10–11): 722–729. dx.doi.org/10.1016/j.ijmedinf.2005.10.007

Dengel, A.; Liwicki, M.; and Weber, M. 2012. Touch & Write: Penabled Collaborative Intelligence. In *Knowledge Technology, Communications in Computer and Information Science,* Volume 295, 1–10. Berlin: Springer.

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): 3 215–3223.

Hall, F. M. 2009. The Radiology Report of the Future. *Radiology* 251(2): 313—316. dx.doi.org/10.1148/radiol.2512090177

Hammond, T., and Paulson, B. 2011. Recognizing Sketched Multistroke Primitives. *ACM Transactions on Interactive Intelligent Systems.* 1(1): 4:1–4:34.

Langlotz, C. P. 2006. Radlex: A New Method for Indexing Online Educational Materials. *RadioGraphics* 26 (November): 1595–1597. dx.doi.org/10.1148/rg.266065168

Marks, R. G. 2004. Validating Electronic Source Data in Clinical Trials. *Controlled Clinical Trials* 25(5):437–446. dx.doi.org/10.1016/j.cct.2004.07.001

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 *Proceedings of the American Medical Informatics Association (AMIA) Symposium*, 465–469. Washington, USA, US National Library of Medicine. PMCID: PMC2656009

Militello, L. G., and Hutton, R. J. B. 1998. Applied Cognitive Task Analysis (ACTA): A Practitioner's Toolkit For Understanding Cognitive Task Demands. *Ergonomics* 41(11): 1618—1641. dx.doi.org/10.1080/001401398186108

Möller, M.; Ernst, P.; Dengel, A.; and Sonntag, D. 2010. Representing the International Classification of Diseases Version 10 in OWL. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD)*. Lisbon, Portugal: The Institute for Systems and Technologies of Information, Control and Communication.

Pittman, J. A. 2007. Handwriting Recognition: Tablet PC Text Input. *Computer* 40(9): 49–54. dx.doi.org/10.1109/MC.2007.314

Rubine, D. 1991. Specifying Gestures by Example. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques,* SIGGRAPH '91, 329–337. New York: Association for Computing Machinery.

Seneviratne, N., and Plimmer, B. 2010. Improving Stylus Interaction For Emedical Forms. In *Proceedings of the 22nd Australian Computer-Human Interaction Conference Conference* (OZCHI), 280–287. New York: Association for Computing Machinery.

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,* ICDAR '07, Volume 2, 954–958. Los Alamitos, CA: IEEE Computer Society.

Sonntag, D. 2012. Collaborative Multimodality. *Künstliche Intelligenz* 26(2):161–168. dx.doi.org/10.1007/s13218-012-0169-4

Sonntag, D. 2013. Incremental and Interaction-Based Knowledge Acquisition for Medical Images in THESEUS. In *Integration of Practice-Oriented Knowledge Technology: Trends and Prospectives,* ed. M. Fathi, 97–108. Berlin, Heidelberg: Springer. dx.doi.org/10.1007/978-3-642-34471-8_8

Steimle, J.; Brdiczka, O.; and Mühlhäuser, M. 2009. Coscribe: Integrating Paper and Digital Documents for Collaborative Learning. *IEEE Transactions on Learning Technologies* 2(3): 174–188. dx.doi.org/ 10.1109/TLT.2009.27

Vicente, K. J. 1999. Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work. Boca Raton, FL: CRC Press LLC.

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 *Proceedings of the 14th International Conference on Document Analysis and Recognition,* ICDAR '05, 957–961. Los Alamitos, CA: IEEE Computer Society.

Willems, D.; Rossignol, S.; and Vuurpijl, L. 2005. Mode Detection in On-Line Pen Drawing and Handwriting Recognition. In *Proceedings of the Eighth International Conference on Document Analysis and Recognition,* ICDAR '05, 31–35. Los Alamitos, CA: IEEE Computer Society.

Daniel Sonntag is a project leader and senior research scientist at the Intelligent User Interface Department (IUI) at Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) and permanent member of the editorial board of the German Journal on Artificial Intelligence (KI). He received his Ph.D. degree in computer science and his M.Sc. degree in computational linguistics from the Saarland University in 2008 and 2001, respectively. He has worked in industrial research projects in natural language processing, text mining, multimodal interface design, and dialogue systems for more than 16 years. Most recently, he won the German High Tech Award with RadSpeech, a semantic dialogue system for radiologists.

Markus Weber received his M.Sc. degree in computer science from the University of Kaiserslautern, Germany, in 2010. He is currently working toward the Ph.D. degree in the Department of Augmented Vision at the German Research Center for Artificial Intelligence (DFKI) in Kaiserslautern. His research interests include pen-based interaction, sequential pattern recognition, and human computer interaction.

Alexander Cavallaro is an attending physician at the Department of Diagnostic Radiology, University Hospital Erlangen, since 1997 and responsible for the digitalization of patient data and IT-workflows in radiology. Additionally, since 2010 he is a professor for medical technologies at the Georg Simon Ohm University of Applied Science in Nuremberg. His research interests include medical imaging, knowledge representation, medical standards, and computer applications in radiology including semantics.

Matthias Hammon is in his fourth year of the diagnostic radiology residency program of the University Hospital Erlangen, Germany. He received his M.D. degree from the University of Erlangen-Nuremberg, Germany in 2010. He is currently working toward the Ph.D. degree in the Department of Diagnostic Radiology, University of Erlangen-Nuremberg. His research interests include intelligent user interfaces and knowledge acquisition in radiology, clinical research (musculoskeletal and genitourinary imaging), and magnetic resonance imaging (MRI).