Smartphone Sketches for Instant Knowledge Acquisition in Breast Imaging

Alexander Prange and Daniel Sonntag
German Research Center for AI
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany

Abstract
To improve reporting practices in breast imaging, we allow the radiologist to write structured reports with a pen on a smartphone. In this way, we provide a knowledge acquisition system for digital mammography. 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. 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. We focus on the implementation of the smartphone interface based on the design in Sonntag et al. (2013) and the employed algorithms for beautification and on-the-fly-corrections.

Introduction and Related Work
In order to improve the general medical reporting scenario, we implemented a prototype for radiology findings with several unique features compared to other radiology reporting tools: (1) a real-time digitisation into PDF documents of both text and graphical contents such as sketches; (2) real-time hand-writing/gesture recognition and real-time feedback on the recognition results on the smartphone screen; and (3) the mapping of the transcribed contents into concepts of several medical ontologies.

This work is a direct extension of the work presented in Sonntag et al. (2013) where we used digital pens on normal paper. 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. This is a new concept and has been extended by other attempts to improving stylus interaction for electronic medical forms (Seneviratne and Plimmer 2010). Our demo is the first of its kind that uses a fully digital solution in this application domain: the user can freely write and draw directly on a smartphone screen into different recognition areas for text selection, text recognition and free-form sketch recognition (shapes and labels on images).

Scenario Implementation
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. 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”, “oval”, “unscharf/diffuse”, “isoechogen”, which are predefined annotation fields to be encircled), and with a disease category using ICD-10 classes (e.g., “Nodular lymphoma” or “Lymphoblastic, see short cite Sonntag et al. (2013) for more information). However, any combination of anatomical, visual, and disease annotations is allowed and multiple annotations of the same region are possible to complete the form (figure 1).

In contrast to our approach, which extends the digital pen implementation for normal paper (Sonntag et al. 2013) to digital pens on smartphone touch screens, 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 misunderstanding than a sketch-based annotation. In our context, the easy to use, powerful and flexible pen-based mobile application solution is a technological advancement of the analog paper reporting process which has been highly optimised by trial-and-error over the last 50 years (Sonntag et al. 2013). The main difference and improvement of this process is that with physical paper, the
The turn-around time is 2-30 hours (manual transcription, corrections, final approval). The design of our mobile application should therefore not only adopt the benefits of physical documents, but also extend it to a point where direct feedback about the recognised (ontological) annotations is given, resulting in lower turn-over times and no need for additional staff.

**Recognition Process**

The pen gestures on the smartphone are recognised using the Vision Objects framework which uses an extended version of the widely used single and multi stroke algorithm. Both the iGesture framework (Signer, Kurmann, and Norrie 2007) and the Vision Objects MyScript engines are capable of providing immediate results, the user receives the results of the analysis and feedback on screen in less than 300ms. Figures 1 and 3 illustrate a current combination of written and hand-drawn annotations. The Android API for touch input is event based. Every event contains several coordinates captured since the last one, including a pressure value for every point. We store these coordinates together with the pressure values and combine them into strokes. Whenever we receive a pen-down event, all following events until the next pen-up event identify a stroke event. As soon as the user stops writing for more than half a second, these strokes are bundled into a multiple-stroke sketch and send through the handwriting and gesture recognition engines.

**Annotation Types**

In our context the user needs to annotate certain areas of an image. These annotations then have to be digitalised for further usage. It is important to know which area of the image was annotated with text. Therefore we implemented a simple heuristic that determines which strokes form a valid annotation: We calculate a rectangular bounding box around each text figure and a circular area around circles. For lines we make areas around the two end points. If there is an input configuration where a circle, a line, and a text stroke intersect in a valid way, we assume that it is an annotation (figures 2 and 3). The marked areas and the annotated text are then stored as a complex ontological annotation.
Implementation
We implemented our smartphone-based digital pen solution as a prototypical Android application, currently running on a Samsung Galaxy Note 3 smartphone which uses Wacom’s advanced digital pen sensor system in combination with the S-Pen stylus to provide accurate pen-based user input. Currently, we are using the Android 4.3 version, but it should be straightforward to downgrade or upgrade to other versions, since we do not use any platform dependent features.

For handwriting recognition and gesture recognition we use the MyScript framework from VisionObjects (www.visionobjects.com/en/myscript), which provides us with the ability to recognise text and simple gestures in real-time. We improved the recognition accuracy by limiting the set of recognised forms to circles, rectangles and lines, as well as by adding a dictionary of medical terms (RadLex (Langlotz 2006) and ICD-10 (Möller et al. 2010)).

User Interface
The physical paper version of the mammography form consists of two pages in letter format. For the digital version (figure 1), we created a raster graphics image and divided it into several parts (MRI, sonography, etc.) in order to fit the screen while maintaining readability. For the navigation between the different parts, we use a tab based approach. Nowadays most radiologists are familiar with smartphones and are used to common touch gestures for interacting with mobile applications. However, the differentiation between pen-based move gestures and actual drawn sketches is quite difficult because even simple drag gestures can be readily interpreted as lines for example. Therefore we decided to use a mixture of independent and exclusive finger gestures and handwritten stylus input gestures. Within the displayed application surface on the mobile screen, the user is able to move around (dragging and zooming in/out by using a pinch gesture or the corresponding buttons). The actual handwriting input is obtained from the smartphone stylus. Direct feedback about the recognition results in form of pop-up notifications at the bottom of the screen is provided to the user. A demo video can be found at http://ul.to/006g5jhi.

Beautification
To improve the handwriting experience of the users, we use variable stroke widths depending on the supplied pressure at any given point. From each Android touch input event we extract not only the coordinates but also the supplied pressures. These touch points are then connected with straight lines to form line segments. The thickness and color is calculated using the pressure values. Unfortunately, the Android API is only able to draw lines with a fixed width, and in some cases (especially at the end of strokes) this leads to a rather rigid step formation. See figure 4, the red dots indicate the original touch points delivered by the Android OS.

To avoid a rather rigid step formation, we interpolate values between all those pairs of touch points where the difference in pressure and distance reaches a certain threshold (figure 5). See algorithm 1. The appearance of strokes could be further improved by using cubic Bezier curves for interpolation, yielding in smoother lines.

\begin{algorithm}
\textbf{input} : Two points \((x_1, y_1), (x_2, y_2)\) and their pressure values \((p_1, p_2)\)
\textbf{output}: an interpolated line between those two points
\begin{align*}
m &\leftarrow \frac{y_2-y_1}{x_2-x_1} \\
b &\leftarrow y_1 - m \times x_1 \\
d &\leftarrow \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} \\
\text{steps} &\leftarrow \left\lfloor \frac{p_1-p_2}{pStep} \right\rfloor \\
pStep &\leftarrow \frac{p_1-p_2}{\text{steps}} \\
dStep &\leftarrow \frac{d}{\text{steps}} \\
\text{if } x_1 < x_2 \text{ then} \\
\quad xi &\leftarrow x_1 + dStep \\
\quad \text{while } xi \leq x_2 \text{ do} \\
\quad \quad yi &\leftarrow m \times xi + b \\
\quad \quad \text{drawPoint}(xi, yi, p_1) \\
\quad \quad p_1 &\leftarrow p_1 - pStep \\
\quad \quad xi &\leftarrow xi + dStep \\
\quad \text{end} \\
\text{else} \\
\quad xi &\leftarrow x_1 - dStep \\
\quad \text{while } xi > x_2 \text{ do} \\
\quad \quad yi &\leftarrow m \times xi + b \\
\quad \quad \text{drawPoint}(xi, yi, p_1) \\
\quad \quad p_1 &\leftarrow p_1 - pStep \\
\quad \quad xi &\leftarrow xi - dStep \\
\quad \text{end} \\
\end{align*}
\end{algorithm}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{stroke_without_interpolation.png}
\caption{Stroke without Interpolation}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{stroke_with_interpolation.png}
\caption{Stroke with Interpolation}
\end{figure}
On-the-fly corrections

Our digital approach has a huge advantage over both the traditional pen and paper input and the digital pen version (Sonntag et al. 2013) on normal paper: we can implement on-the-fly corrections on the digital smartphone screen. The user is able to modify and delete previously drawn sketches in place. Our prototype recognises several intuitive correction gestures: crosses, lines, or to scratch something out (figure 6). This works as follows: whenever a recognition result returns a line, a cross or a zig-zag, we check whether this gesture intersects with any previously drawn figure or text. Depending on the coverage ratio between the potential correction gesture and the multiple stroke image, either the underlying figure is removed or the gesture accepted as a valid additional stroke (e.g., a line). It is also possible to only remove individual strokes of an previously drawn sketch or single characters and words. In the latter case, the recognition process for the modified group of strokes is triggered again to analyse the changes (includes a possible reinterpretation of the whole multi-stroke sketch). See algorithm 2. The lack of availability of real-time accuracy checks and on-the-fly corrections is one of the main reasons why digital pen systems have not yet been used in the radiology domain (Marks 2004). We hypothesise a large increase in the acceptability of our mobile digital reporting system with on-the-fly corrections for validating electronic source data in clinical trials.

**Algorithm 2:** On-the-fly corrections with practical gestures

```plaintext
input : The current sketch s to be analysed and the set of previously drawn sketches S
result ← analyse(s)
if result ∈ correction gestures then
    foreach sketch ∈ S do
        if s intersects sketch then
            if intersection area / sketch area ≥ threshold then
                l remove sketch from screen
            end
        end
    end
if s is valid correction gesture then
    l remove s from screen
else
    l draw s
end
```

**Conclusion and Future Work**

We presented a digital pen-based interface for mammography forms where users can directly write and sketch onto the smartphone screen. The resulting digital data input device can improve the quality and consistency of mammography reports because the direct digitisation avoids the data transcription task and turn-over times of analog paper reporting.

The possibility to reduce real-time recognition errors (which the digital pen version could not provide) and logic errors as the data are being collected has great potential to increase the data quality of such reports over the long run.

Educators may find these digital reporting devices can help trainees learn the important elements of reports and encourage the proper use of radiology terms. Future research will however focus on a clinical evaluation of the comparison of digital pen and smartphone and the implementation of new beautification and correction methods.

**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 European EIT ICT Labs in the Medical Cyber-Physical Systems activity.

**References**


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)*.

