PREFERENCES OF HANDWRITING RECOGNITION ON MOBILE INFORMATION SYSTEMS IN MEDICINE
Improving Handwriting Algorithm on the Basis of Real-life Usability Research

Andreas Holzinger, Martin Schlögl
Institute of Medical Informatics, Statistics and Documentation, Research Unit HCI4MED, Medical University Graz Graz, Austria

Bernhard Peischl
Institute of Software Technology, Graz University of Technology, Graz, Austria

Matjaz Debevc
Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Austria

Keywords: Handwriting Recognition, Mobile Computer, Human-computer Interaction, Usability, Real-life, Health Care.

Abstract: Streamlining data acquisition in mobile health care in order to increase accuracy and efficiency can only benefit the patient. The company FERK-Systems has been providing health care information systems for various German medical services for many years. The design and development of a compatible front-end system for handwriting recognition, particularly for use in ambulances was clearly needed. While handwriting recognition has been a classical topic of computer science for many years, many problems still need to be solved. In this paper, we report on the study and resulting improvements achieved by the adaptation of an existing handwriting algorithm, based on experiences made during medical rescue missions. By improving accuracy and error correction the performance of an available handwriting recognition algorithm was increased. However, the end user studies showed that the virtual keyboard is still the overall preferred method compared to handwriting, especially among participants with a computer usage of more than 30 hours a week. This is possibly due to the wide availability of the QUERTY/QUERTZ keyboard.

1 INTRODUCTION AND MOTIVATION FOR RESEARCH

In cases of emergency, rapid patient information collection is very important. This information is most often collected by first aiders (first responders) and paramedics (e.g. Red Cross). Prompt and accurately recorded and well communicated vital patient data can make the difference between life and death (Holzman, 1999), (Anantharaman & Han, 2001).

The data acquisition should have as little disruptive effect on the workflow of the emergency responders (rescue staff) as possible. A possible solution for data input can be an mobile application on a lightweight handheld device (Baumgart, 2005), (Chittaro, Zuliani & Carchetti, 2007).

Due to the fact that emergencies are usually within difficult physical situations, special attention to the design of information technology for emergencies has to be taken into consideration (Klann et al., 2008). A key issue of any such information system is the acquisition of textual information. However, extensive text entry on mobile devices is principally to be avoided and a simple and easy to use interface, in accordance with the proverb: less is more, is a supreme necessity (Holzinger & Errath, 2007).

Holzinger A., Schlögl M., Peischl B. and Debevc M. (2010).
PREFERENCES OF HANDWRITING RECOGNITION ON MOBILE INFORMATION SYSTEMS IN MEDICINE - Improving Handwriting Algorithm on the Basis of Real-life Usability Research.
In *Proceedings of the International Conference on e-Business*, pages 14-21
DOI: 10.5220/0002979900140021
Copyright © SciTePress

The basic evidence that entering data onto a mobile device via a stylus is slower, more erroneous and less satisfactory for end users than entering data via a QWERTZ (de) or QUERTY (us) keyboard has been demonstrated in some studies (Haller et al., 2009), although, on the other hand the use of a stylus is much faster and more accurate than using finger touch (Holzinger et al., 2008b). A specific study for "Ambulance Run Reporting" shows good results for acquiring text with a virtual keyboard, while acquiring text by the application of handwriting recognition showed some serious usability problems (Chittaro et al., 2007). Motivated by this previous work, we focus in this work on handwriting recognition and on how to improve its usability – in case of need, also by adaptation of existing handwriting algorithms. Consequently, in this paper we report on real-life experiences and on some improvements achieved by the adaptation of an existing handwriting engine.

2 BACKGROUND

A big difficulty of handwriting recognition is that handwritten characters are variable on an individual basis and that these characters are usually separated into alphabets, numerals, and symbols, despite the different characters of the language itself. Although handwriting recognition will benefit in future from improved adaptive and context-sensitive algorithms, improving the user experience of novice end users with the respective technology is possibly the most important factor in enhancing user acceptance (MacKenzie & Chang, 1999). This is even more important in medical or health care contexts, where the difficulty is in the environmental conditions, e.g. if the person is on the move or in a hurry (Holzinger et al., 2008a). Whereas the first problem might be solved by the training modus opportunities, in order to adapt the system to the individual handwriting style, the second problem is only solvable by an extremely robust and usable system. Especially in the health care domain, good end user acceptance and usability can only be obtained by providing simple operation (good user guidance), very short response times and low error rates (Holzinger, Geierhofer & Searle, 2006).

Basically, there are several methods for handwriting recognition; these belong basically to two distinct families of classification:

I) Structured and Rule Based Methods

Because of the fuzzy nature of human handwriting, it makes sense to adapt the well known fuzzy logic technique for this purpose (Gader et al., 1997). Rather than evaluating the two values as in digital logic, fuzzy terms admit to degrees of membership in multiple sets so that fuzzy rules may have a continuous, rather than stepwise, range of truth of possibility. Therefore non-identical handwritten numerals, from same or different users, can be approximated using fuzzy logic for fast and robust handwriting recognition (Shi & Li, 2006).

II) Statistical Methods

a) Hidden Markov Modeling (HMM)

The attractiveness of HMM for various pattern recognition tasks is mainly due to their clear and reliable statistical framework. Many efficient algorithms for parameter estimation and model evaluation exist, which is an important prerequisite for their practical implementation for real-life applications (Plotz & Fink, 2009). The methods using HMM (Marti & Bunke, 2002), are based on the arcs of skeleton graphs of the words to be recognized and an algorithm applied to the skeleton graph of a word extracts the edges in a particular order, which is transformed into a 10-dimensional feature vector. Each of these features represent information about the location of an edge relative to four reference lines, the curvature and the degree of the nodes incident to the considered edge. Training of the HMM is done by use of the Baum-Welch algorithm, while the Viterbi algorithm is used for recognition (Bunke, Roth & Schukattalamazzini, 1995), (Xue & Govindaraju, 2006).

b) Neural Networks

The methods based on Neural Networks were driven by the emergence of portable, pen based computers. A typical approach is to combine an artificial neural network (ANN), as a character classifier, with a context-driven search over segmentation and word recognition hypotheses (Yaeger, Webb & Lyon, 1998).

However, handwriting recognition not only consists of the recognition itself; the data must undergo some preprocessing:

(I) Reduce noise;
(II) Normalization, and
(III) Segmentation.

The last step, the segmentation phase, segments the input into single characters (Plamondon & Srihari, 2000). Writing discrete characters requires no segmentation; this is done by the users themselves (Tappert, Suen & Wakahara, 1990).

Another way to improve recognition is to decrease the set of possible alternatives, such as to restrict the set to accepting only lower case letters or digits (Frankish, Hull & Morgan, 1995).

3 RELATED WORK

To date only a few studies considered handwriting recognition on mobile devices and very few in the health care domain.

A very early work by Citrin et al. report very general on the usage of a pen on a flat surface of a LCD unit (scribing and tapping). They reported that with the maximum rate of 100 selections of direction per second for pen, scribing may produce strokes with the speed of 300 (100×3) bps. However, no more results were found (Citrin et al., 1993).

MacKenzie showed that the recognition accuracy for a set containing upper and lower case letters was lower than for a set containing just lower case letters (MacKenzie et al., 1994).

Chittaro evaluated a system for recording data on a system during a running ambulance drive, having first responders as participants. Text entry via virtual keyboard and handwriting recognition (MS Transcriber – Calligrapher) were also performed. Text entering by handwriting was considered very laborious and difficult by the users (Mean 3.8, Var 6.6), while entering text by use of the virtual keyboard was quite easy (Mean 7.2, Var 1.8). (0=Hard, 9=Easy). Furthermore, they emphasized the bad usability of entering text by using handwriting recognition. Most words were wrongly recognized and there were enormous problems in correcting those wrongly recognized words (Chittaro et al., 2007).

4 METHODS AND MATERIALS

The aim of our study was to increase the performance of available handwriting recognition by improving accuracy and error correction following solid usability engineering methods (Holzinger, 2005).

We focused on separate character recognition, since the correction of a single letter, at the moment of false recognition, can be made more naturally, and efficiently, than attempting to correct or delete a single letter within a recognised word.

Due to limited space, there could be some problems inputting long words. Therefore, only one character at a time can be written and recognized.

4.1 Experimental Device

The device used for the prototype was an Asus MyPad A626 PDA (Personal Digital Assistant).

This device is equipped with an anti-glare touch screen display. For typing on the touch screen, a stylus is used.

Table 1 contains the technical specifications of this device.

Table 1: Specifications of the PDA ASUS MyPal A626.

CPU	Marvell XScale, 312MHz
Operating System	MS Windows® Mobile™ 6
Memory	256MB Flash ROM and 64 MB SDRAM
Display	3.5" Brilliant TFT LCD 65k full-colours, anti-glare 16-bit display QVGA, 240x320 px touch screen
Weight	158g
Physical dimensions	117 mm x 70.8 mm x 15.7cm

4.2 Dialog Design

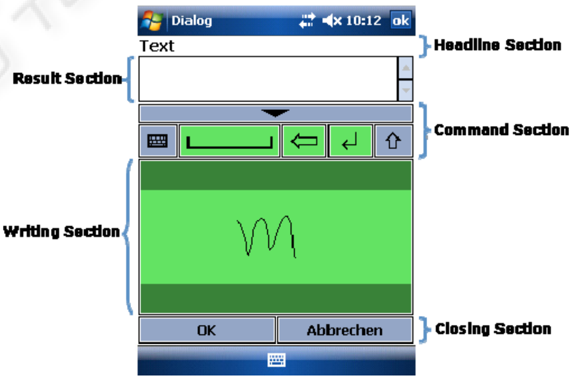

Figure 2: Design of the handwriting dialog.

The light green area within the writing sections defines the optimal size for handwritten lowercase characters of 80 points (Phatware, 2002).

4.3 Handwriting Recognition

We used the SDK of the handwriting recognition engine Calligrapher (in MS Windows® Mobile Transcriber) in the version 6.0 (Phatware, 2002).

This SDK makes it possible to define single character recognition. We can handle the results and a custom timeout (after which time the recognition starts) can be defined.

4.3.1 Adaptive Timeout

A handwritten character consists of one or more strokes. The recognition starts after the character is finished. The system has to await a timeout before starting recognition because the system doesn't know whether the character consists of just one or more strokes.

A stroke is defined as the writing from pen down to pen up (Tappert et al., 1990).

Because of the different writing speeds of each user, this timeout has to be calculated for each user. Therefore, the system stores the last ten times which elapse between two strokes.

$$T = \frac{\sum_{i=1}^{11} s(i)}{11} * \frac{X}{100}$$

Figure 3: Calculation timeout T [sec].

Figure 2 shows how the timeout is calculated every time a timeout is requested. s(1) is the last calculated average time between strokes, s(2)..s(11) are the last ten stored times between strokes. X is a factor, in this experimental setting X is 200. The result T is the timeout in seconds.

4.3.2 Correction Intervention

Calligrapher SDK 6.0 doesn't adapt recognition on users' handwriting because of the use of static Fuzzy-Neuronal Nets (Strenge, 2005).

There are problems with some user's style of writing letters – the user writes a letter (e.g. an "a") but the recognition engine recognizes another letter (e.g. figure 4).

Figure 4: Written "a" but not recognized as "a", instead as "ir".

A recognition result is a list of possible characters and its weight (maximum 5 entries). Every time the same letter is wrongly recognized for a user (as in Figure 2), the lists returned by the recognition are similar.

These lists (characters and its weight) with its representing letter are stored. Each of them is called schema.

During writing, the recognition result will be compared to the stored schemas as follows. (Example in Figure 5)

For each stored schema:

Characters from the result list and the list of the schema are compared. If the result list consists of 2 or 3 characters, at least 2 have to match to the stored schemas lists characters. (2 of 2, 2 of 3). If there are 4 or 5 characters in the result list, at least 3 have to match (3 of 4, 3 of 5). This means, the resulting list is validated to the list of the schema.

If the list is valid according to the list of the schema, the average deviation between these matching characters is calculated.

Inputted list			**Stored Schema's list**		**Deviation**
u	52		k	41	4
n	47		M	35	0
k	37		n	31	16
M	35		h	26	
A	22		m	22	
Validity: 3/5 → VALID					
Average Deviation (16+0+4)/3:					**6**

Figure 5: Example of a list comparison.

The representing letter of the schema with the lowest average deviation will be put in first place of the recognition result.

4.3.3 Calibration

The calibration is designed to collect user specific data for each letter. This data contains weights, which present every character explicitly. Also, schemas of wrongly recognized letters (Chapter 3.3.2) are collected.

The system prompts the user to input a letter.

If the result list of the recognition has the prompted letter in first place, the weight will be stored for this letter. In the calibration phase, at least 2 weights will be stored for each letter.

If not, the result list will be stored as a schema with the prompted letter as a representing letter. In the calibration phase, a maximum of 10 schemas for each letter is stored.

This calibration is done once for each user. A continuous calibration is also done during writing in the handwriting recognition dialog, saving weights and schemas for correctly recognized letters (but not for deleted letters)

4.3.4 Other Interventions on Recognition Results

To avoid side effects, the intervention described in Chapter 3.3.2 is only made when the weight of a recognized letter is less than the average weight for this letter (average of the weights for this letter collected by calibration).

Other interventions are made to avoid potential. problems with highly confusable pairs such as “r” and “v” (Frankish et al., 1995). (I) While writing a word, only letters and punctuation marks are valid, recognized results. (II) Just deleted letters (with BACKSPACE) are not valid, recognized results for the next recognition (III) Special handling for “O” and “0” as first letter of a word or number

4.4 Experiment

The real life environment is mostly a seat in an ambulance car. To avoid negative effects on ambulance responder’s work, the experiment is done in their recess in the ambulance service rooms, simulating the circumstances (sitting in a car) by doing the experiment sitting on a chair, holding the PDA in their hand, without laying down the elbows on e.g. an armrest. (Kjeldskov et al., 2004) shows that simulating environments gives almost the same results.

Figure 6: Participants during experiments in real life.

Participants were people who work as ambulance officers (professionals, volunteers and former civilian service). No previous experience with mobile computers was required.

They were asked to fill out a background questionnaire to obtain data about their age, education and use of computers.

The prototype for the experiment is divided into two parts, one for virtual keyboard based text input, and the other for handwriting recognition input. Within these two parts, the users have the opportunity to become familiar with the input methods. After that, the user has to input a given text to the experimental dialog (for measuring the accuracy). Due to measuring the accuracy, text entry is done as text copy (MacKenzie & Soukoreff, 2002). This text consists of 13 German words (94 characters without spaces, 106 with spaces). After the keyboard based experimental dialog, the calibration of the handwriting is done.

Speed in wpm, words per minute (Lewis, 1999, MacKenzie et al., 1994) and the accuracy of the handwriting recognition are measured and calculated. At the end, a feedback questionnaire is filled out by the user. Some questions are based on the study of Chittaro (Chittaro et al., 2007).

5 RESULTS

5.1 Participants

The participants of the experiment were professional (9) and volunteer (8) first responders of the Austrian Red Cross, one student of medicine and three others (because everyone could be a volunteer first responder).

10 are experienced on a PDA or a mobile phone with touch screen, while 11 have no experience with touch screens.

Their ages ranged from 20 to 85 years. Two elderly people (68 and 85 years) were chosen because they had never before used a QWERT keyboard or a PC.

The average use of a PC is 12.3 years, using a PC 31 hours per week. 11 participants use a PC ≤ 30 hours a week, while 10 participants use a PC for more than 30 hours.

One of the 21 participants was left-handed.

5.2 Accuracy

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
99.1	6.28	100	11.5	99.06	1.44

Figure 6: Accuracy inputting text with virtual keyboard [%]; all participants, participants ≤ 30 hours and above.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
89.25	34.3	91.43	30.20	88.00	37.34

Figure 7: Recognition accuracy [%] of handwriting recognition; all participants, participants ≤ 30 hours and above with interventions.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
84.66	57.6	86.99	79.15	83.33	38.21

Figure 8: Recognition accuracy [%] of handwriting recognition; all participants, participants ≤ 30 hours and above without interventions.

The participants using a PC ≤ 30 hours a week include the two elderly people.

The 85 year old participant has an accuracy of 89.2% for inputting text with the virtual keyboard and a recognition accuracy of 80.1% with interventions and 65.6% without interventions.

The 68 year old participant had an accuracy of 100% for inputting text with the virtual keyboard and a recognition accuracy of 95% with interventions and 90.8% without interventions.

The 85 years old participant has an accuracy of 89.2% for inputting text with the virtual keyboard and a recognition accuracy of 80.1% with interventions and 65.6% without interventions.

The 68 year old participant has an accuracy of 100% for inputting text with the virtual keyboard and a recognition accuracy of 95% with interventions and 90.8% without interventions.

5.3 Speed

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
13.17	27.7	12.88	29.46	13.43	18.29

Figure 9: Words per minute virtual keyboard; all participants, participants ≤ 30 hours and above.

Overall		≤ 30 weekly usage		> 30 weekly usage	
Mean	Var	Mean	Var	Mean	Var
8.44	4.59	8.11	5.37	8.71	1.95

Figure 10: Words per minute handwriting recognition; all participants, participants ≤ 30 hours and above.

Participants using a PC ≤ 30 hours a week include two elderly people.

The 85 year old participant wrote 2.87 wpm with the keyboard and 2.82 wpm with handwriting recognition. The 68 year old participant wrote 4.88 wpm with the keyboard and 4.17 wpm with handwriting recognition.

5.4 User Questionnaire

Overall	**Mean**	**Var**
Keyboard		
Inputting Data (+4=easy, -4=difficult)	3.0	2.6
Correction of wrong inputted data (+4=easy, -4=difficult)	4.0	2.8
Handwriting		
Inputting Data (+4=easy, -4=difficult)	2.0	4.9
Correction of wrongly input/recognized data (+4=easy, -4=difficult)	3.0	1.4
Did the recognition slow down your writing (+4=no, -4=yes)	0.5	9.2
I would prefer (+4=handwriting, -4=keyboard)	-2.5	7.9
Basic Information		
Use of colour is (+4=useful, -4=useless)	2.0	2.5
The handwriting recognition positively surprised me (+4=yes, -4=no)	2.5	7.1
Characters on the PDA are easy to read (+4=yes, 4=no)	4.0	3.9

Figure 11: Overall results of user questionnaire.

Weekly computer usage [hours]	**Mean (<=30)**	**Var (<=30)**	**Mean (>30)**	**Var (>30)**
Keyboard				
Inputting Data	3.5	1.4	3.0	3.7
Correction of wrongly input data	4.0	1.7	3.5	4.5
Handwriting				
Inputting Data	2.5	4.2	0.5	3.8
Correction of wrongly input/recognized data	4.0	0.5	2.5	1.8
Did the recognition slow down your writing	2.5	8.4	-0.5	8.2
I would prefer	-1.5	2.0	-2.5	4.9
Basic Information				
Use of colour	2.0	2.5	1.5	2.6
The handwriting recognition positively surprised me	4.0	7.7	0.0	4.7

Characters on the PDA are easy to read	4.0	4.6	4.0	3.2

Figure 12: Results of user questionnaire for weekly usage of computer ≤ 30 hours and above.

6 CONCLUSIONS

Entering text with the virtual keyboard (Mean 3.0, Var 2.6) was easier for the participants than with handwriting (Mean 2.0, Var 4.9). However, compared to the study of (Chittaro et al., 2007), we could reach an significant improvement by inputting data with handwriting. Interestingly, inputting data by handwriting recognition was rated easier by participants who use computers less than or equal to 30 hours a week than by participants with extensively more use (Mean 2.5; Var 4.2; against Mean 0.5, Var 3.8 of virtual keyboard). Also, the correction on the handwriting recognition dialog was rated easier (Mean 4.0, Var 0.5; against Mean 2.5, Var 1.8; of virtual keyboard). Participants with a computer usage of more than 30 hours a week preferred the virtual keyboard (Mean -2.5, Var 4.9) more than the other participants (Mean -1.5, Var 2.0). This could be a result of hardly any handwriting during work and much more typing text on classical keyboards (QWERTZ or QUERTY). Consequently, the two elderly participants were included in this study, in order to obtain data regarding participants who never used any computer or handheld device. The elderly participants were the only ones who provided a complete preference to the handwriting recognition in contrast to the virtual keyboard. This is also clearly visible in the results for these participants, although both groups have quite comparable results in wpm for the virtual keyboard and the handwriting text input.

This is an interesting result; however, it is not of practical relevance, since there are hardly any people left – at least amongst people able to volunteer as a first responder – without experience on computer keyboards. Today, from elementary school on, children get used to work with computers by using the QWERTZ or QUERTY keyboard.

Nevertheless, our interventions on the basis of the results of the handwriting recognition, finally paid off in an significant improvement on the recognition accuracy (over all participants a better accuracy of Mean +4.39%, Var 9.54).

These interventions can also be useful for the improvement of other handwriting recognition engines, due to the fact that our interventions were only made on the results of the engine, achieving better accuracy. The use of a handwriting recognition engine with a higher accuracy than e.g. Calligrapher, in combination with our demonstrated interventions, may even improve the overall accuracy. Our methods on operating on the results of the handwriting recognition engine operate context independent. Using a dictionary to add the likelihood of upcoming characters may improve the accuracy in that part of the problem regarding confusable pairs, such as “r” and “v”. Because of typing in characters one by one, a word completion feature could be added to handwriting recognition too. This also would increase the writing speed.

ACKNOWLEDGEMENTS

This study was performed with support of FERK-Systems. We cordially thank the engineering team for their continued and effective industrial support of this work. The research was partially funded by the Austrian Research Promotion Agency (FFG) within one „Innovationsscheck Österreich“.

REFERENCES

Anantharaman, V. & Han, L. S. (2001) Hospital and emergency ambulance link: using IT to enhance emergency pre-hospital care. *International Journal of Medical Informatics,* 61, 2-3, 147-161.

Baumgart, D. C. (2005) Personal digital assistants in health care: experienced clinicians in the palm of your hand? *The Lancet,* 366, 9492, 1210-1222.

Bunke, H., Roth, M. & Schukattalamazzini, E. G. (1995) Off-Line Cursive Handwriting Recognition Using Hidden Markov-Models. *Pattern Recognition,* 28, 9, 1399-1413.

Chittaro, L., Zuliani, F. & Carchetti, E. (2007) Mobile Devices in Emergency Medical Services: User Evaluation of a PDA-based Interface for Ambulance Run Reporting. In: Löffler, J. & Klann, M. (Eds.) *Mobile Response*. Berlin, Heidelberg, New York, Springer, 19-28.

Citrin, W., Halbert, D., Hewitt, C., Meyrowitz, N. & Shneiderman, B. (1993) Potentials and limitations of pen-based computers. *Proceedings of the 1993 ACM conference on Computer science.* Indianapolis, Indiana, United States, ACM.

Frankish, C., Hull, R. & Morgan, P. (1995) Recognition accuracy and user acceptance of pen interfaces. *Conference on Human Factors in Computing Systems.* Denver, Colorado, United States, ACM Press/Addison-Wesley Publishing Co.

Gader, P. D., Keller, J. M., Krishnapuram, R., Chiang, J. H. & Mohamed, M. A. (1997) Neural and fuzzy methods in handwriting recognition. *Computer,* 30, 2, 79-86.

Haller, G., Haller, D. M., Courvoisier, D. S. & Lovis, C. (2009) Handheld vs. Laptop Computers for Electronic Data Collection in Clinical Research: A Crossover Randomized Trial. *Journal of the American Medical Informatics Association,* 16, 5, 651-659.

Holzinger, A. (2005) Usability Engineering for Software Developers. *Communications of the ACM,* 48, 1, 71-74.

Holzinger, A. & Errath, M. (2007) Mobile computer Web-application design in medicine: some research based guidelines. *Universal Access in the Information Society International Journal,* 6, 1, 31-41.

Holzinger, A., Geierhofer, R. & Searle, G. (2006) Biometrical Signatures in Practice: A challenge for improving Human-Computer Interaction in Clinical Workflows. In: Heinecke, A. M. & Paul, H. (Eds.) *Mensch & Computer: Mensch und Computer im Strukturwandel*. München, Oldenbourg, 339-347.

Holzinger, A., Hoeller, M., Bloice, M. & Urlesberger, B. (2008a). *Typical Problems with developing mobile applications for health care: Some lessons learned from developing user-centered mobile applications in a hospital environment*. International Conference on E-Business (ICE-B 2008), Porto (PT), IEEE, 235-240.

Holzinger, A., Höller, M., Schedlbauer, M. & Urlesberger, B. (2008b). *An Investigation of Finger versus Stylus Input in Medical Scenarios*. ITI 2008: 30th International Conference on Information Technology Interfaces, June, 23-26, 2008, Cavtat, Dubrovnik, IEEE, 433-438.

Holzman, T. G. (1999) Computer-human interface solutions for emergency medical care. *interactions,* 6, 3, 13-24.

Kjeldskov, J., Skov, M. B., Als, B. S. & Høegh, R. T. (2004) Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in Field. *Mobile HumanComputer Interactoin - MobileHCI2005*. Springer Berlin / Heidelberg, 529-535.

Klann, M., Malizia, A., Chittaro, L., Cuevas, I. A. & Levialdi, S. (2008) HCI for emergencies. *CHI '08 extended abstracts on Human factors in computing systems.* Florence, Italy, ACM.

Lewis, J. R. (1999). *Input Rates and User Preference for three small-screen input methods: Standard Keyboard, Predictive Keyboard and Handwriting*. Human Factors and Ergonomics Society.

MacKenzie, I. S. & Chang, L. (1999) A performance comparison of two handwriting recognizers. *Interacting with Computers,* 11, 3, 283-297.

MacKenzie, I. S., Nonneke, B., Riddersma, S., McQueen, C. & Meltz, M. (1994) Alphanumeric entry on pen-based computers. *International Journal of Human-Computer Studies,* 41, 5.

MacKenzie, I. S. & Soukoreff, R. W. (2002) Text Entry for Mobile Computing: Models and Methods, Theory and Practice. *HUMAN-COMPUTER-INTERACTION,* 17, 2, 147–198.

Marti, U. V. & Bunke, H. (2002) Using a statistical language model to improve the performance of an HMM-based cursive handwriting recognition systems. *Hidden Markov models: applications in computer vision*. World Scientific Publishing Co., Inc., 65-90.

Phatware (2002) Calligrapher SDK 6.0 Developer's Manual.

Plamondon, R. & Srihari, S. N. (2000) On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22, 1, 63-84.

Plotz, T. & Fink, G. A. (2009) Markov models for offline handwriting recognition: a survey. *International Journal on Document Analysis and Recognition,* 12, 4, 269-298.

Shi, B. & Li, G. (2006) VLSI Neural Fuzzy Classifier for Handwriting recognition. Patent US 7,146,037.

Strenge, M. (2005) Konzepte und Toolkits zur Handschrifterkennung.

Tappert, C. C., Suen, C. Y. & Wakahara, T. (1990) The State of the Art in On-Line Handwriting Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12, 8, 787-808.

Xue, H. H. & Govindaraju, V. (2006) Hidden Markov models combining discrete symbols and continuous attributes in handwriting recognition. *Ieee Transactions on Pattern Analysis and Machine Intelligence,* 28, 3, 458-462.

Yaeger, L. S., Webb, B. J. & Lyon, R. F. (1998) Combining neural networks and context-driven search for on-line, printed handwriting recognition in the Newton. *Neural Networks: Tricks of the Trade*. Berlin, Springer-Verlag Berlin, 275-298.