Pen and Paper-based Interaction with the Semantic Desktop

Marcus Liwicki, Kinga Schumacher and Andreas Dengel
German Research Center for AI (DFKI GmbH)
Knowledge Management Department, Kaiserslautern, Germany
{Firstname.Lastname}@dfki.de

Nadir Weibel, Beat Signer, and Moira C. Norrie
Institute for Information Systems, ETH Zurich, Switzerland
{weibel,signer,norrie}@inf.ethz.ch

Abstract

In this paper we propose a system which recognizes and interprets the semantics of handwritten annotations on printed documents. The semantic information will be sent to the Semantic Desktop, the personal Semantic Web on the desktop computer, which supports users in their information management. This allows a seamless integration of interactive paper into the individual knowledge work. The current implementation of the proposed system works with OpenOffice documents printed on Anoto paper. However, our system generally works on any kind of interactive paper documents and the support of other document formats is planned in near future.

1 Introduction

The *paperless office*, i. e., the philosophy of working with minimal paper and converting everything into digital documents, was predicted more than thirty years ago. However, there is still an increasing amount of paper used by humans in their everyday work. Using paper is motivated by several issues. First, the person is not bound to a specific device or to a specific medium. Second, paper is portable, allowing for making notes anytime and anywhere. Furthermore, among other issues, writing on paper is more natural to most persons making the pen their preferred writing instrument for tasks such as brainstorming, collaborative work or reviewing documents [9].

While using a pen on paper makes it easier and faster for the user to attach their thoughts to the document, the interpretation of the notes is often more complex compared to the situation where digital comments are made. To overcome this drawback, several document analysis approaches have been proposed which automatically process human-made notes and transform them into digital format.

In workflows like reviewing, where the annotations have a meaning for the marked text, the problem of mapping the paper to the digital counterpart arises. A variety of approaches have been investigated to enable this kind of paper-driven digital services. They use cameras, Wacom Graphics Tablets[1], ultrasonic positioning, RFID antennas, bar code readers, or Anoto's Digital Pen and Paper technology[2]. The Anoto technology is particularly interesting because it is based on regular paper and the recording of the interactions is precise and reliable.

In this paper we propose the Semantic eInk system, which automatically processes handwritten annotations on printed documents and interprets the semantic information of these annotations. This information will be expressed through formal semantics using the individual's vocabulary, and integrated into the personal knowledge base, the Semantic Desktop. The integration makes this knowledge searchable, reusable, sharable and gives a context for its interpretation. Semantic eInk extends the Semantic Desktop with a new input modality, interactive paper. Thus it supports personal knowledge work on paper.

The rest of this paper is organized as follows. First, Section 2 presents the information management tool for which the Semantic eInk system is developed. Next, the interactive paper framework on which our prototype builds is described in Section 3. The Semantic eInk system is then proposed in Section 4. Finally, Section 5 summarizes the current state of the development and gives an outlook for future work.

2 The NEPOMUK Social Semantic Desktop

The Semantic Desktop [2, 3] is a means for personal knowledge management; it builds the personal Semantic

[1] http://www.wacom.com
[2] http://www.anoto.com

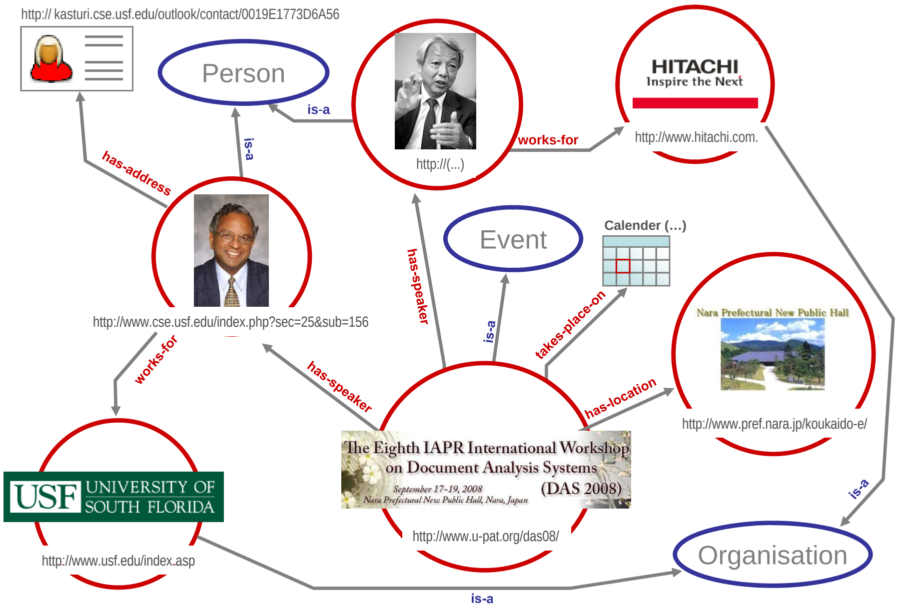


Figure 1. PIMO extract: example representation of the event "DAS 2008" and the keynote speakers

Web on desktop computers. The consistent application of Semantic Web standards such as the Resource Description Framework[3] (RDF) and RDF Schema[4] (RDFS) provides the identification of digital resources, i. e., text documents, e-mails, contacts, multimedia files, by unique URIs, across application borders. In contrast to current limitations in file and application based information management, the user is able to create his or her own classification system which reflects the way of thinking: it consists of projects, people, events, topics, locations etc. Furthermore, the Semantic Desktop enables the user to annotate, classify and relate all resources, expressing his or her view in a *Personal Information Model* (PIMO) [8]. Figure 1 illustrates an extract of a PIMO which represents part of the information about the event "DAS 2008" and the keynote speakers of this conference. The figure shows some *ontological concepts* (classes like "Event" and instances like "USF"), which are related to the DAS conference and semantically describes the kind of relations, e. g., "take-place-on".

Several systems supporting personal information management already exist, e. g., Haystack at MIT [6], D-Bin by SEMEDIA [11], IRIS by SRI [1] and Gnowsis at DFKI [7]. These systems integrate different kinds of applications, mostly calender, email, notes and provide different functionalities. A full-featured Semantic Desktop involves many data sources, indexes the text and metadata of all documents and categorizes them. Together with the collected facts about the instances, e. g., contact information and membership of a person, a critical amount of information is available to the user. The Semantic Desktop offers effective access to this personal memory since the information is not only personalized, facilitating navigational search, but also machine readable. This allows the development of machine processes, learning methods which assists the user with their every-day activities and the extension and maintenance of the PIMO.

Thus, the Semantic Desktop supports individual knowledge workers in their information management. The Social Semantic Desktop NEPOMUK[5] extends the described personal aspects of the information management with social aspects. The desktop is expanded to a collaboration environment in order to facilitate collaborative knowledge work. NEPOMUK supports knowledge exchange and sharing, distributed search and hence distributed collaboration. These capabilities enable the user to conduct information management across social and organizational relations [4].

3 Interactive Paper

In order to retrieve written annotations from paper documents, our system builds on the interactive paper (iPaper) framework [5]. The iPaper framework enables the capture of user actions on paper and the activation of corresponding digital services or access to digital resources. In order to

[3] http://www.w3.org/RDF/
[4] http://www.w3.org/TR/rdf-schema/

[5] NEPOMUK (Networked Environment for Personalized, Ontology-based Management of Unified Knowledge) is funded by the European Union IST (Grant FP6-027705). For more details see the project homepage http://nepomuk.semanticdesktop.org/xwiki/bin/view/Main1/

become interactive, documents are Anoto-enabled at print time by augmenting the paper with a special Anoto dot pattern. The Anoto pattern represents an absolute positioning system based on (x,y) coordinates that can be read by special digital pens such as the Magicomm G303[6]. After decoding the pattern, the pen transmits the (x,y) positions to a computer through a wireless Bluetooth connection.

The iPaper framework interprets the paper-based user actions and activates the corresponding digital actions. We use the iGesture framework [10] to recognize any pen-based gestures and to translate them into the corresponding digital operations. For the recognition of text, either the MyScript Builder from Vision Objects[7] or the Microsoft handwriting recognition engine[8] can be used.

Even though the digital pen retrieves the position of the written annotation from paper, this information is usually not sufficient to access the digital content to which an annotation refers. In order to associate the recognized pen-based annotation with a specific word, sentence, or section within the digital document, a precise mapping between the digital content and its printed version has to be performed. Since a lot of information about the content of the printed document is available at publishing time, a special publishing plug-in (iPublish) performs an analysis of the document during the printing process and stores the information about the physical positions on paper for any required structural element (sections, paragraphs, words, characters). To store this information in a effective way, we use a dynamic approach that analyzes a paginated version of the document and use a dynamic storage scheme to its structural elements, refining the granularity of the stored information on demand [12].

4 Aim of the System

The system proposed in this paper allows a seamless integration of interactive paper technology into personal knowledge work. To be more specific, the workflow of printing a document, annotating it while reading, and integrating the new information into the personal knowledge base will be supported by an automated interpretation of user annotations. Therefore, the documents are printed onto the Anoto paper and annotations are made with the digital pen. Then, a set of gestures and handwritten text is recognized and finally, the information is sent to the NEPOMUK system.

It is allowed to use three kinds of annotation (see Fig. 2).

1. The user can make a comment at any place in the document (The topmost handwritten text in Fig. 2). This text is recognized by the handwriting recognition engine and stored as a comment in the knowledge base.

2. The user can mark a text passage with right angle strokes ("⌈" "⌉") and write a comment or a word representing an ontological concept in in the knowledge base ("Title" in Fig. 2). For processing the annotation, the handwriting is first recognized. In the case of a comment, the comment and its relation to the marked text is stored in the knowledge base. In the case that the written words represent an ontological concept in the knowledge base, we need to detect the corresponding property, its range of values and the concrete value, e. g., the marked phrase as string or as an instance of the class person. The particular steps of this process are dependent on the kind of recognized ontological concept. The identified property, its value and possibly required instances are stored in the knowledge base. For the latter case, we plan to allow the user to define shortcuts for the most frequent ontological concept that they store (e. g., "T" for "Title")

3. The user can create a side mark and write a comment which will be associated with a specific part of the document. In this case, a new property for the document is inserted in the knowledge base, where the property value is the marked text.

5 Project State and Future Work

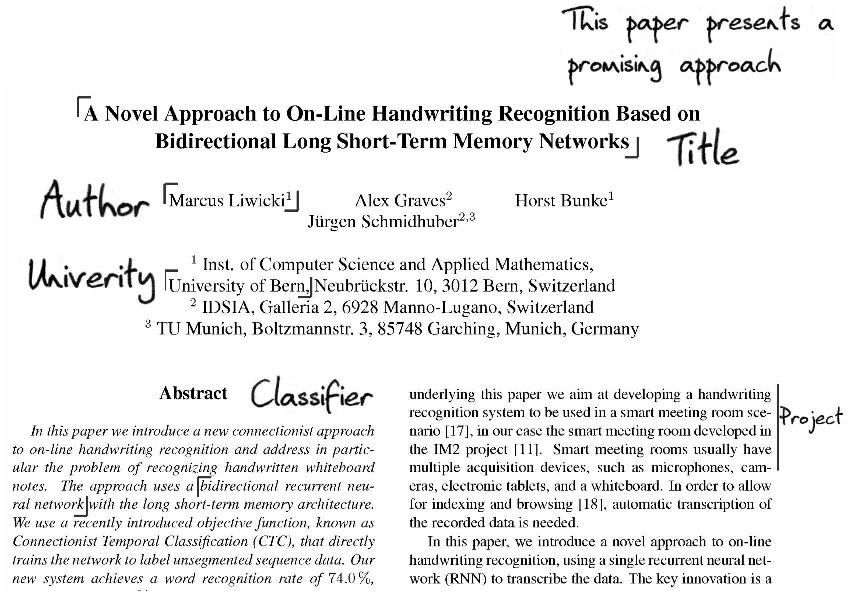
The current prototype of the Semantic eInk system builds on the iPaper framework and the publishing approach presented in Section 3 and is designed to work with OpenOffice documents. An interactive paper version of an OpenOffice document can be published via the iPublish plug-in. During the publishing process, an XPS file is generated and analyzed for structural information. In this way, the document is ready to be annotated on paper. In order to support this paper-based interaction, the information about the Anoto pattern used and the location of the OpenOffice documents are published to iPaper.

After annotating a published paper, the stroke information is sent to the application, where the identification and the recognition of the gestures is realized. The next step will be the interpretation of the recognized text and the corresponding update commands sent to the NEPOMUK server. There the PIMO will be updated according to the message and the information will be available in all other tools of the Semantic Desktop. We plan to implement the possibility to define shortcuts for ontological concepts frequently used in the annotations.

In the future we plan to extend the functionality of Semantic eInk to support other document types such as PDF and Microsoft Word.

In order to evaluate the usability of Semantic eInk we will perform user studies where participants are asked to an-

[6]http://www.magicomm.co.uk
[7]http://www.visionobjects.com
[8]http://www.microsoft.com/windowsxp/tabletpc/default.mspx



A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks

Marcus Liwicki[1] Alex Graves[2] Horst Bunke[1]
Jürgen Schmidhuber[2,3]

[1] Inst. of Computer Science and Applied Mathematics,
University of Bern, Neubrückstr. 10, 3012 Bern, Switzerland
[2] IDSIA, Galleria 2, 6928 Manno-Lugano, Switzerland
[3] TU Munich, Boltzmannstr. 3, 85748 Garching, Munich, Germany

Abstract

In this paper we introduce a new connectionist approach to on-line handwriting recognition and address in particular the problem of recognizing handwritten whiteboard notes. The approach uses a bidirectional recurrent neural network with the long short-term memory architecture. We use a recently introduced objective function, known as Connectionist Temporal Classification (CTC), that directly trains the network to label unsegmented sequence data. Our new system achieves a word recognition rate of 74.0 %,

underlying this paper we aim at developing a handwriting recognition system to be used in a smart meeting room scenario [17], in our case the smart meeting room developed in the IM2 project [11]. Smart meeting rooms usually have multiple acquisition devices, such as microphones, cameras, electronic tablets, and a whiteboard. In order to allow for indexing and browsing [18], automatic transcription of the recorded data is needed.

In this paper, we introduce a novel approach to on-line handwriting recognition, using a single recurrent neural network (RNN) to transcribe the data. The key innovation is a

Figure 2. Annotations on the document which will be interpreted by Semantic eInk

notate papers and to perform tasks in the NEPOMUK system with the help of their annotations. Afterwards, they have to fill out a questionnaire. From these studies we can infer the usefulness of paper-based annotations for the Semantic Desktop. Furthermore, future research directions will arise from answers in the questionnaire.

References

[1] A. Cheyer, J. Park, and R. Giuli. Iris: Integrate. relate. infer. share. *1st Workshop on The Semantic Desktop. 4th International Semantic Web Conference*, page 15, Nov 2005.

[2] S. Decker and M. R. Frank. The networked semantic desktop. In *WWW Workshop on Application Design, Development and Implementation Issues in the Semantic Web*, 2004.

[3] A. R. Dengel. Knowledge technologies for the social semantic desktop. In *In Proceedings of the Second International Conference on Knowledge Science, Engineering and Management*, 2007.

[4] T. Groza, S. Handschuh, K. Moeller, G. Grimnes, L. Sauermann, E. Minack, C. Mesnage, M. Jazayeri, G. Reif, and R. Gudjonsdottir. The NEPOMUK project - on the way to the social semantic desktop. In T. Pellegrini and S. Schaffert, editors, *Proceedings of I-Semantics' 07*, pages 201–211. JUCS, 2007.

[5] M. C. Norrie, B. Signer, and N. Weibel. General framework for the rapid development of interactive paper applications. In *Proc. of CoPADD 2006, 1st Int. Workshop on Collaborating over Paper and Digital Documents*, pages 9–12, 2006.

[6] D. Quan, D. Huynh, and R. Karger. Haystack: A platform for authoring end user semantic web applications. In *Proceedings of the 2nd International Semantic Web Conference*, pages 738–753, 2003.

[7] L. Sauermann, G. A. Grimnes, M. Kiesel, C. Fluit, H. Maus, D. Heim, D. Nadeem, B. Horak, and A. Dengel. Semantic desktop 2.0: The Gnowsis experience. In *Proc. of the ISWC Conference*, Nov 2006.

[8] L. Sauermann, L. van Elst, and A. Dengel. PIMO – a framework for representing personal information models. In *Proceedings of I-SEMANTICS 2007*, pages 270–277, 2007.

[9] A. J. Sellen and R. H. R. Harper. *The Myth of the Paperless Office*. MIT Press, 2001.

[10] B. Signer, U. Kurmann, and M. C. Norrie. iGesture: a general gesture recognition framework. In *Proc. of ICDAR 2007, 9th Int. Conference on Document Analysis and Recognition*, pages 954–958, 2007.

[11] G. Tummarello, C. Morbidoni, P. Puliti, and F. Piazza. The DBin semantic web platform: an overview. In *Workshop on The Semantic Computing Initiative at WWW2005*, 2005.

[12] N. Weibel, M. C. Norrie, and B. Signer. A model for mapping between printed and digital document instances. In *Proc. of DocEng 2007, 7th ACM Symposium on Document Engineering*, pages 19–28, 2007.