Pen-Based Retrieval in Handwritten Documents

Sascha Schimke[1], Claus Vielhauer[1]

[1] ITI Research Group on Multimedia and Security,
Universitätsplatz 2, 39106 Magdeburg, Germany
{sascha.schimke, claus.vielhauer}@iti.cs.uni-magdeburg.de
http://wwwiti.cs.uni-magdeburg.de/iti_amsl/

Abstract. This paper describes techniques for searching in handwritten and handdrawn documents. We assume that more and more handwritten documents will accrue, as pen-based computers like PDA or TabletPC get increasingly popular. In order to manage the anticipated amount of handwritten documents, powerful abilities for searching within the documents are needed. The techniques, which we propose, are well researched in the field of bioinformatics, where they are used for finding parts within sequences of amino acids or genes. This paper documents a work in progress research activity.

1 Introduction

While pen-based computers like PDA or TabletPC get popular, a large amount of handwritten documents will be produced. Often it's easier and less disturbing in a meeting to minute using a pen instead of typing with a keyboard. And sometimes for rapid drawing or sketching the pen is the only possible opportunity. To handle and manage a large amount of such documents, consisting of handwritten and handdrawn parts, the ability for retrieval is absolutely necessary.

By document retrieval, we mean the following: having a set D of documents $\{d_1, d_2, d_3, \dots, d_n\}$ and using a query word q, the result of the retrieval is a list D' of documents $\{d'_1, d'_2, \dots, d'_m\}$ out of D, where the retrieval query q is contained one or more times within the documents of D'. For each document of D', even the position(s) of the occurrence(s) of the query q is available. In the case of pen-based retrieval in handwritten and –drawn documents, the query q as well as the elements of D is a result of writing or drawing process. (See Fig. 1)

From the users' point of view, the retrieval task can be performed in two manners: a) by writing or drawing the query or b) by manually selecting one occurrence of the query within a document.

The handwritten documents in our work are no scanned images of sheets of paper, as used in the field of off-line handwriting processing. We confine ourselves to the field of so called on-line data. This means, that the pen-movement data are acquired directly during the writing process, using special hardware. The resulting data are sequences of sampled pen tip positions and pressure values: $x(t)$, $y(t)$, $p(t)$. From these directly measurable data, further data can be derived, e.g. velocities in direction of x- or y-axis or the track velocity ($v_x(t)$, $v_y(t)$, $v(t)$).

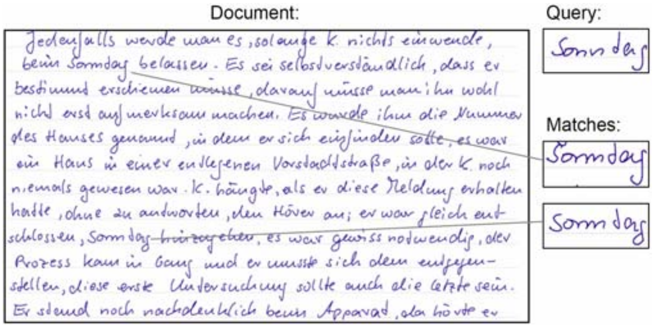

Fig. 1. Illustration of the search process for a query word “Sonntag” and two marked matches in a text from Kafka (“Der Prozeß”)

The most intuitive idea would be to perform a textual recognition for the documents of *D* and for the query *q*. Using the resulting textual data, a simple string search function could be used, as is available in every word processor. The problem is that the textual recognition is never absolute perfect and therefore a simple search would fail in most cases. Another problem with using textual recognition is the case, where no text at all exists to be recognized, namely when using handdrawn images or sketches instead of words. To solve these problems, we try to use a kind of *direct handwriting matching* instead of textual recognition, as described in section 3.

2 Related Works

A large amount of related work has been done before by other researchers. But most of them had different approaches or different goals for their pen-based retrieval.

Srihari et al. presented a search engine for handwritten documents [8]. Contrary to our approach, their documents were acquired off-line. Due to the nature of off-line handwriting data, their processing steps are more complex then ours. A similar approach is for example presented in [3] for historical documents.

In [5] Landay and Davis describe a set of experiments for shared note taking. The participants of these experiments used PDAs and paper-based digitizer devices to write notes and other documents. Later, management of the accumulated documents should be possible, for example browsing and searching.

Besides this text oriented retrieval in pen based data, there are several approaches for pen aided retrieval in image databases. In [7] a system is described, which extracts shapes from images and makes a comparison of these shapes with pen drawn query inputs. A similar idea is described in [1]. Here the pen is used for retrieval in ClipArt data, i.e. vector graphic files.

An approach, which is slightly similar to our work, was described in [9]. The goal of the authors is the same as ours – to search in on-line captured handwritten documents – but they use different features and different matching algorithms. We expect to reduce the computational complexity of their system by using our own approach.

3 On-line Comparison for Retrieval

Our idea is to use an algorithm for finding substring in noisy data. Similar algorithms are used in genetics and bioinformatics for example to find sequences of amino acids or genes. The problem is that often there is no 100%-match between the query sequence and a part of the complete sequence. Instead, only a certain similarity can be ascertained.

3.1 Approximate String Searching

The goal of *approximate string searching* is to find those substrings (approximate matches) $m_1, m_2, \ldots, m_n$ of a document d, so that the distance of these substrings and a query string q is smaller than a threshold τ. Here the distance is defined as the *edit distance* [4]. The approximate string searching problem is to find those positions of d, where the approximate matches $m_1, m_2, \ldots, m_n$ end.
May the query string q consist of k characters $q[1], q[2], q[3], \ldots, q[k]$. Furthermore, may the document d consist of l characters $d[1], d[2], d[3], \ldots, d[l]$. The approximate string searching problem is equal to find those values j, so that $D(k, j)$ is smaller than the threshold τ:

$$D(i,j) = \begin{cases} 0 & \text{if } i = 0, \\ D(i-1,0)+1 & \text{if } i > 0 \text{ and } j = 0, \\ \min \begin{Bmatrix} D(i,j-1)+1 \\ D(i-1,j)+1 \\ D(i-1,j-1)+\delta(i,j) \end{Bmatrix} & \text{else.} \end{cases}$$

with $0 \leq i \leq k$ and $0 \leq j \leq l$. The character-wise distance δ is defined as follows:

$$\delta(i,j) = \begin{cases} 0 & \text{if } q[i] = d[j], \\ 1 & \text{else.} \end{cases}$$

The asymptotic computational complexity of this recursion D is $O(k \cdot l)$, but the required memory can be reduced to $O(k)$ by smart implementations.

3.2 Textual Recognition

For retrieval in handwritten documents, we interpret each document as a long sequence of elements. These sequences are the basis for estimation of local similarity regarding a (shorter) query sequence q. One problem is to define the elements of the sequences, i.e. the alphabet. Using text recognition algorithms, the alphabet would be the set of characters and signs, which are the output of the algorithm and which form the words, sentences and the whole text. The similarity search is used to overcome the problem of misrecognised characters.

To limit the computational complexity, which is implied by using text recognition algorithms, in our first experiments we use more rudimental features from handwritten documents. Two kinds of features are described in the following two subsections.

3.3 Stroke Direction Features

In [2] a method is described for converting a line drawing into a sequence of stroke directions. The idea of Freeman is to have a very short and compressible description of line drawings, but we try to adapt his method for substring search in handwritten documents.

The background of Freeman's approach is to code a line drawing using a chain of strokes with a certain direction, as can be seen in Fig. 2. Following the quantized strokes and coding every direction with a certain symbol, we obtain a chain of direction codes. The granularity of the resulting code string is influenced by the size of the quantizing grid.

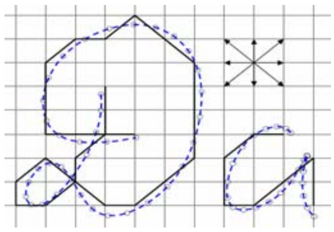

Fig. 2. A writing sample (dashed line) and the corresponding sequence of stroke directions (solid line), using a square grid quantization

3.4 Biometric Features

Another idea of coding handwritten inputs as strings is presented in [6]. Here, from the handwriting signals ($x(t)$, $y(t)$, $p(t)$, $v_x(t)$, $v_y(t)$, $v(t)$) the local extrema are extracted, i.e. points of minimal or maximal x-/y-value, pressure or velocity. That way for a given pen-based input a string is created, consisting of successive local extrema of the signals. The original purpose of this approach was biometric authentication using handwritten signatures.

3.5 Fusion of Single-Feature Results

Having different searching results for a query q within a document d by using different base features (and/or different parameters), derived from the original pen-based input, it is necessary to make a kind of fusion of these different results.

For example, by using the direction coded string (3.3) we could get two matches m_1 and m_2 (positions in the original document d). By using the minima and maxima for creating a string (3.4) to describe the pen movement we could get three matches m_3, m_4 and m_5, where m_2 and m_5 are identical. Two trivial solutions to handle this situation are the conjunction and the disjunction. In conjunction, only those matches, which are identical for both kinds of feature strings, are taken as final matching result. Disjunction means, that all the matches are taken as the result. The choice of conjunction or disjunction affects the recall and precision rates of the retrieval. Using the disjunction approach, the recall will presumably increase, but in the same moment, the precision will decline.

4 Test Settings

For testing of our system, we collect handwriting data using TabletPCs and special pens. In our case the formers ones have a spatial resolution of 1,000 points/cm, temporal resolution of 100 Hz and can distinguish 1,024 degrees of pressure. The latter devices are pens, which are equipped with an optical sensor to read a special pattern at paper, which allows them to get precise position information [10]. The spatial resolution is about 280 points/cm, sampling rate is up to 50 Hz and 128 degrees of pressure can be distinguished. The documents, acquired with these devices, will be used to obtain recall and precision results for our handwriting retrieval system.

Acknowledgements

This work has been partly supported by the EU Network of Excellence SIMILAR (FP6−507609). The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Union.

References

1. M. Fonseca, B. Barroso, P. Ribeiro, J. Jorge, "Sketch-based retrieval of ClipArt drawings", *Proceedings of the working conference on Advanced visual interfaces*, 2004, pp. 429–432.
2. H. Freeman, "Computer Processing of Line-Drawing Images", *Computer Surveys*, Vol. 6, No. 1, March 1974, pp. 57–97.
3. V. Govindaraju, H. Xue, "Fast Handwriting Recognition for Indexing Historical Documents", *Proceedings of First International Workshop on Document Image Analysis for Libraries (DIAL'04)*, 2004, pp. 314–320.
4. D. Gusfield, *Algorithms on Strings, Trees, and Sequences*, Cambridge University Press, ISBN 0-521585-19-8, 1997.
5. J. A. Landay, R. C. Davis, "Making sharing pervasive: Ubiquitous computing for shared note taking", *IBM Systems Journal*, Vol. 38, No. 4, 1999, pp. 531–550.
6. S. Schimke, C. Vielhauer, J. Dittmann, "Using Adapted Levenshtein Distance for On-Line Signature Authentication", *Proceedings of the 17th International Conference on Pattern Recognition*, 2004.
7. L. Schomaker, E. de Leau, L. Vuurpijl, "Using Pen-Based Outlines for Object-Based Annotation and Image-Based Queries", *Lecture Notes in Computer Science*, Vol. 1614, 1999, pp. 585–592.
8. S. Srihari, C. Huang, H. Srinivasan, "A Search Engine for Handwritten Documents", *Proceedings of SPIE-IS&T Electronic Imaging*, 2005, pp. 66–75.
9. K. Sun, J. Wang, "Similarity based matching method for handwriting retrieval", *Proceedings of SPIE 2003 – Document Recognition and Retrieval X*, 2003, pp. 156–163.
10. Anoto Group AB, http://www.anotofunctionality.com/, 2005.