

Touch Scan-n-Search: A Touchscreen Interface To Retrieve Online Versions of Scanned Documents

Fabrice Matulic

Ricoh Software R&D Group

1-1-17 Koishikawa, Bunkyo-ku, Tokyo, Japan

{lastname}.dm.{firstname}@nts.ricoh.co.jp

ABSTRACT

The system described in this paper attempts to tackle the problem of finding online content based on paper documents through an intuitive touchscreen interface designed for modern scanners and multifunction printers. Touch Scan-n-Search allows the user to select elements of a scanned document (e.g. a newspaper article) and to seamlessly connect to common web search services in order to retrieve the online version of the document along with related content. This is achieved by automatically extracting keyphrases from text elements in the document (obtained by OCR) and creating “tappable” GUI widgets to allow the user to control and fine-tune the search requests. The retrieved content can then be printed, sent, or used to compose new documents.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – GUI, Input devices and strategies, Interaction Styles; I.7.5 [Document and Text Processing]: Document Capture – Document Analysis; H.4.3 [Information Systems Applications]: Communications Applications – Information browsers

General Terms

Performance, Design, Human Factors

Keywords

GUI, scanned document, online news retrieval, keyword extraction

1. INTRODUCTION

Many documents nowadays are available both in digital and paper form. In the news publishing industry particularly, readers have been able to access the contents of most major newspapers, magazines and journals on publishers' web sites for several years now. Producing hardcopies of digital content can be simply achieved through printing, however retrieving relevant online media based on paper documents is not as straightforward. While many retrieval techniques have been developed using scanned

document images as input, few have addressed this particular task from a usability perspective. Most text-based retrieval systems require users to type their queries with a keyboard for lack of an integrated segmentation and OCR interface.

The main goal of this work is to provide users with simple, intuitive tools to steer the retrieval process in order to quickly find online versions of scanned documents as well as related content. Touch Scan-n-Search is intended as an additional module for SmartPublisher [1] so that it can benefit from its integrated content manipulation model and its drag-and-drop interface. Moreover, the retrieved material can be immediately re-used for document creation and printing, with the advantage of working with perfect digital data instead of error-ridden OCR text.

2. DESCRIPTION OF THE APPLICATION

2.1 Interface Overview

As with SmartPublisher, scanned documents are first acquired, and previewed in the Scanner Module, where page layouts can be analysed to segment them into separate regions (OCR is performed in the background to obtain machine-readable text for text zones). As an alternative source, the Media Browser can be utilised to process document images captured with a digital camera (see [1] for further details).

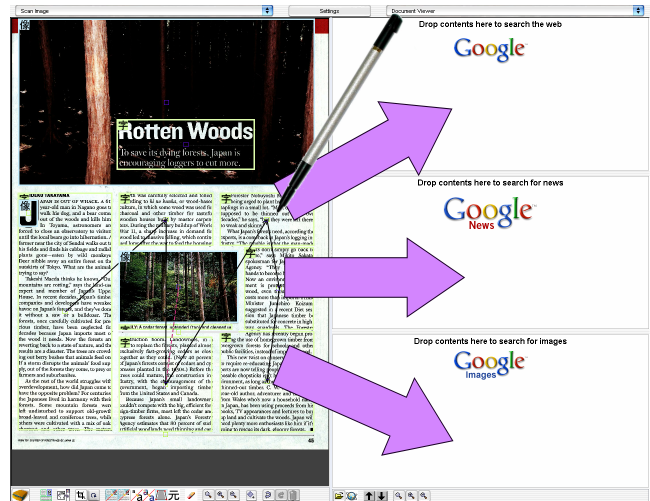


Figure 1. GUI with the Scanner Module on the left and the Document Viewer on the right. Document Elements can be dragged and dropped onto specific zones to perform various kinds of searches.

The Touch Scan-n-Search feature builds on top of the Document Viewer of SmartPublisher, where searches are performed in the embedded web browser. Document Element Objects obtained in the Scan Module or the Media Browser can be dragged over to the Document Viewer and dropped onto special areas, each of which triggers a different type of Internet search. For this prototype, Google™ Web Search, Google™ News and Google™ Image Search were used (Figure 1) but other configurations are of course possible depending on the application needs and the available search services.

2.2 Keyword Extraction

The keywords to perform the searches are automatically extracted from the content objects when a drag motion is detected. If the text boxes are part of a chain, the user needs only drag one box to perform a search based on the combined text contained in the whole sequence.

The keyword extraction algorithm itself, while not being the main focus of the application, needs to be accurate enough to output meaningful keywords for the searches to give good results. The current prototype uses a customised version of KEA, an open-source keyphrase indexing package written in Java [2]. In free indexing mode (i.e. without any controlled vocabulary), KEA needs to be trained on a set of text data before it can be put to practical use. Tests have shown that about 30 newspaper and magazine articles with manually assigned keywords were enough for the purpose of the application. Keyword extraction using KEA or other similar engines has the advantage of being fairly robust to noisy OCR text, as misrecognised words typically appear only once or twice at most and are therefore not picked up by the extractor (that is, if the text is long enough). Moreover, searching using keywords obtained from whole articles rather than just titles or headlines enables the user to not only retrieve originals but also related documents that are more relevant.

To enhance the search capabilities beyond keyword-based retrieval, CBIR-approaches can be considered if such search services are available. In that case, the user would be able to perform both text searches and queries by image content, depending on the type of the object dragged onto the search zones of the Document Viewer.

2.3 Touch-Controlled Search Interface

When content objects are dropped on a search zone, the Document Viewer switches to web browsing mode with the appropriate search interface. After a short time, the browser window updates and the search results are displayed. Depending on the specificity of the keywords, the search engine will be able to find very similar web pages/news/images and the original article if it is available. For better accuracy, the number of keywords to be extracted can be increased. With 10 keywords instead of 5, for example, the original piece of news should appear on the top of the hit results if it exists. However, the number of related pages will very likely decrease (possibly returning only pages with the original article).

What makes the search interface dynamic and touchscreen-specific is the toolbar at the top which includes toggle buttons representing each extracted keyphrase. At any time, the user can check or uncheck keyphrases that form the search query by

tapping on the respective buttons. The browser automatically updates the search results after a new combination of keywords is set by activating/deactivating the toggle buttons (Figure 2).

The keyword switch toolbar gives users more control to fine-tune search queries so that they can better locate online originals and related content, this without ever using a keyboard. In the future, the interface might also include a handwriting recognition engine to let users add new keywords that were not extracted or were not contained in the original document.

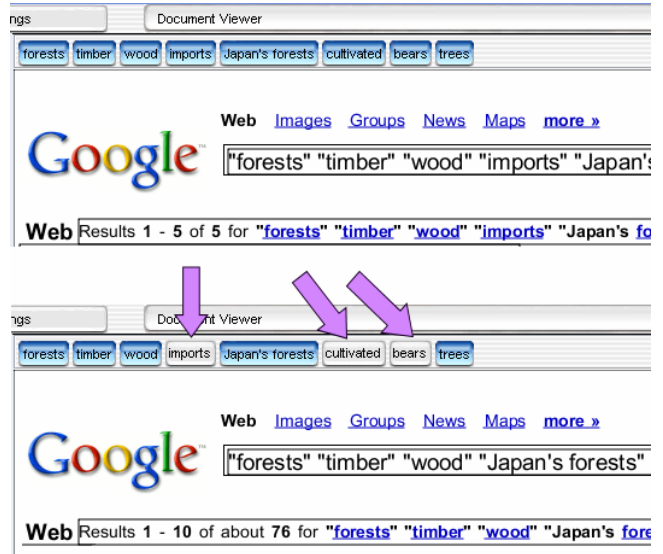


Figure 2. Keyword buttons can be toggled on and off to refine or broaden search queries.

2.4 Using the Search Results

The retrieved web pages can be used for several purposes: they can be, of course, printed (with a significantly higher quality compared to image-based printouts), sent by email or serve as input material for the creation of new documents. By combining high-resolution image elements from scanned documents and 100% accurate digital text with font and style information retrieved from the web, high-quality copies of news articles and magazines can be obtained.

3. CONCLUSION

Touch Scan-n-Search is a compromise between all-automatic but imperfect document retrieval solutions and systems that require a great deal of user intervention. With a few touches, the user can “link” his/her paper documents to digital content available on the web using common search services. While the interface was designed for scanners and MFPs, it could also possibly be integrated into next-generation mobile phones with high-resolution cameras and touch controls such as the Apple iPhone™.

4. REFERENCES

- [1] F. Matulic. SmartPublisher: Document Creation On Pen-Based Systems Via Document Element Reuse. *Proceedings of DocEng 2006, Amsterdam, The Netherlands*
- [2] KEA, keyphrase extraction algorithm: <http://www.nzdl.org/Kea>