

# *TexT* - Text Extractor Tool for Handwritten Document Transcription and Annotation

Anders Hast<sup>1</sup>, Per Cullhed<sup>2</sup>, and Ekta Vats<sup>1</sup>(✉)

<sup>1</sup> Department of Information Technology, Uppsala University, Uppsala, Sweden  
`{anders.hast,ekta.vats}@it.uu.se`

<sup>2</sup> University Library, Uppsala University, Uppsala, Sweden  
`per.cullhed@ub.uu.se`

**Abstract.** This paper presents a framework for semi-automatic transcription of large-scale historical handwritten documents and proposes a simple user-friendly text extractor tool, *TexT* for transcription. The proposed approach provides a quick and easy transcription of text using computer assisted interactive technique. The algorithm finds multiple occurrences of the marked text on-the-fly using a word spotting system. *TexT* is also capable of performing on-the-fly annotation of handwritten text with automatic generation of ground truth labels, and dynamic adjustment and correction of user generated bounding box annotations with the word being perfectly encapsulated. The user can view the document and the found words in the original form or with background noise removed for easier visualization of transcription results. The effectiveness of *TexT* is demonstrated on an archival manuscript collection from well-known publicly available dataset.

**Keywords:** Handwritten text recognition · Transcription  
Annotation · *TexT* · Word spotting · Historical documents

## 1 Introduction

When printing was invented in the mid 15th century, a sort of transcription revolution took place all over Europe. Single handwritten texts were transformed into multiple copy books. Although this invention was crucial for the growth of knowledge, the process of writing continued well into the 20th century very much as before, with the help of pen and ink.

A similar media-revolution is taking place right now when modern technology in the form of electronic texts is revolutionizing our reading habits and our media distribution possibilities. One of the most crucial steps for science in this modern media-revolution is the ability to search within texts. Optical Character Recognition (OCR) technology [1–3] has opened up even old printed texts to modern science in an unprecedented way. In libraries, meta-data is no longer the sole entry to collections, electronic content can speak for itself and this also changes library practices. However, the large mass of handwritten texts in our

libraries and archives is still waiting to be transformed into searchable texts. The reason for this is a combination of technical and economic factors. Modern technology does not yet give us the good results of OCR technology, which nowadays can be so successfully applied to printed texts that it is a straightforward part of digitization processes world-wide.

Handwritten text recognition (HTR) [4–8] is an emerging field and can be quite successful in certain circumstances, especially when applied to an even and uniform handwriting, but rarely so for the non-homogeneous handwritten texts that fill our archives. In most cases, manual transcription is still the most common way to produce reliable electronic texts from handwritten texts, but modern technology advances and many projects try to tackle this problem. Manual transcription is typically expensive and prone to human error. The incentives to open up this material to computerized searches is high. The information in archives and library collections world-wide, represent an enormously important source to history and only relatively small parts of it is available as electronic texts.

Semi-automatic transcription of manuscripts typically requires hundreds of already transcribed pages, with thousands of examples of each word, in order to produce a useful transcription of the rest of the text. Due to the time consuming machine learning procedures involved, this is computed as off-line batch jobs overnight [7]. However, this means that if just a dozen pages exist, the transcriber is forced to complete the transcriptions without the help of HTR techniques, unless a similar handwriting style exists. An alternative approach to fast transcription of text with a low cost is using computer assisted interactive techniques.

This paper introduces a simple yet effective text extractor tool, *TexT* for transcription of historical handwritten documents. *TexT* is designed for quick document transcription with the help of user interaction where the system finds multiple occurrences of the marked text on-the-fly using a word spotting system. Other advantages of *TexT* include on-the-fly annotation of handwritten text with automatic generation of ground truth labels, adjustment and correction of user labeled bounding box annotations such that the word perfectly fits inside the rectangle. Nevertheless, the transcribed words are cleaned using filtering methods for background noise removal.

This paper is organized as follows. Sections 2, 3 and 4 discusses various transcription and annotation methods and tools available in literature, and discusses related work on handwritten text transcription. Section 5 explains the proposed text extractor tool *TexT* in detail. Section 6 demonstrate the efficacy of the proposed method with implementation details on well-known historical document dataset. Section 7 concludes the paper.

## 2 Transcription Methods and Tools

Transcriptions can be made by several different techniques, by reading and typing, typically done by one person interested in using the contents of the documents, as opposed to collective transcription where many individuals make

transcriptions using crowdsourcing techniques. HTR, and dictation, are other techniques that can be used to produce transcriptions. An example of the latter is the war-diary of Sven Blom, a Swedish volunteer in The Foreign Legion during the Great War. The diary is kept in Uppsala University Library and was transcribed by dictation [9].

Due to the labour-intensive task involved in transcriptions, crowdsourcing, a term originally coined by Jeff Howe in Wired Magazine in 2006 [10], has been a useful way of distributing transcription work to many people and therefore it sits at the core of many successful transcription projects. The *Transcribe Bentham* project at the University College of London is often mentioned as an example [11]. Like so many others, *Transcribe Bentham* is built with components from the open-source software *MediaWiki*, also used for the perhaps biggest crowdsourcing project on the planet, *Wikipedia*. *Transcribe Bentham* started in 2010 and has to this date completed approximately 43% of the whole collection [12]. They now collaborate with the READ project [13] and the application Transkribus [14], which can combine HTR with manual transcription.

There are numerous other transcription tools on the Internet. Zooniverse [15], based in Oxford, include transcription as one of their crowdsourcing tasks, among many others. The plugin Scripto [16] is one of the oldest, typically created in an environment close to the history discipline, the Roy Rosenzweig Center for History and New Media at George Mason University. It is also based on *MediaWiki* and can be used as a plugin for *Omeka*, *Wordpress* and *Drupal*. Veele Handen [17] is a Dutch application which offers crowdsourced transcriptions as a tool for archives and libraries wishing to open up their collections. They have recently included progress bars where followers and participants can monitor progress.

This feature is very similar to the Smithsonian Institution and their “Digital Volunteers” [18]. In fact, the Smithsonian Institution can be regarded as one of the pioneers in assigning tasks to volunteers. Already in 1849, soon after the founding of Smithsonian Institution, it’s first secretary, Joseph Henry, was able to initiate a network of some 150 volunteers for weather observations, all over the United States [19]. The “Smithsonian Digital Volunteers” is a very successful transcription application and their Graphics User Interface (GUI) combines a clear topical structure with progress bars and a general layout which has incorporated well-established practices used in proof-reading. The work of volunteer number one, has to be approved by a second volunteer and finally the result needs to be approved by the mother institution, wishing to publish the results on the web. Together with other activities, such as promoting projects via social networks, they have managed to achieve good results, demonstrating the importance of an attractive GUI in crowdsourcing. The topical structure facilitates for the user to find attractive tasks.

Uppsala university library is Sweden’s oldest university library and its manuscript collections consist of approximately four kilometers of handwritten material in letters, diaries, notebooks etc. The handwritten manuscript collections date back 2000 years; from BC till the 21st century. The medieval manuscripts

are plentiful and the 16th to 20th centuries are well represented with many single important collections, such as the correspondence of the Swedish King Gustav III, containing letters from, for example the French Queen Marie Antoinette and the Waller collection of 38000 manuscripts with letters from both Isaac Newton and Charles Darwin. The languages in the collection are also diverse (e.g. Swedish, Arabic, Persian etc.). However, the main languages for this project include Swedish, Latin, German, and French.

Since a few years back it has been possible to publish digitized material in the Alvin platform [20], a repository for cultural heritage materials shared among the universities in Uppsala, Lund and Göteborg, as well as other Swedish libraries and museums. However, as so often is the case, very little of the handwritten material is transcribed. The collection can therefore be accessed only through meta-data and cannot be analyzed by computational means, a problem which may only be tackled by long term and multifaceted strategic planning for producing more handwritten document transcriptions.

As a start, Alvin [20] has been adapted to allow for publishing transcriptions alongside the original manuscripts. One example of this is a transcription made from a testimony of refugees arriving to Sweden in 1945 from the concentration camp in Ravenbrück, kept at Lund University Library [21]. In this case, the transcriptions in textual electronic format (such as PDF) are a result of manual transcription and are open to Google indexing, thus making the original manuscripts searchable on the Internet. However, this is only an example, to open up more texts for use in digital humanities, a combination of HTR technology and manual crowdsourced transcriptions is probably as far as our present technologies admit. This work takes an initiative towards transcription and annotation of huge volumes of historical handwritten documents present in our university library using HTR methods such as word spotting [22].

### 3 Document Annotation Methods and Tools

Several document image ground truth annotation methods [23,24] and tools [25–32] have been suggested in literature. Problems related to ground truth design, representation and creation are discussed in [33]. However, these methods are not suitable for annotating degraded historical datasets with complex layouts [34]. For example, Pink Panther [25], TrueViz [26], PerfectDoc [27] and PixLabeler [28] work well on simple documents only and perform poorly on historical handwritten document images [35].

A highly configurable document annotation tool GEDI [29] supports multiple functionalities such as merging, splitting and ordering. Aletheia [30] is an advanced tool for accurate and cost effective ground truth generation of large collection of document images. WebGT [31] provides several semi-automatic tools for annotating degraded documents and has gained importance recently. Text Encoder and Annotator (TEA) was proposed in [32] for manuscripts annotation using semantic web technologies. However, these tools require specific system

requirements for configuration and installation. Most of these tools and methods are either not suitable for annotating historical handwritten datasets, or represent ground truths with imprecise and inaccurate bounding boxes [35].

Our previous work [34] takes into account such issues, and proposed a simple method for annotating historical handwritten text on-the-fly. This work employs this annotation method with improvements using word spotting algorithm. A detailed discussion of the annotation tools and methods is out of scope of this paper, and the reader is referred to [34] for a deeper understanding of ground truth annotation methods, and on-the-fly handwritten text annotation in general.

## 4 Related Work on Handwritten Text Transcription

Manual transcription of historical handwritten documents requires highly skilled experts, and is typically a time consuming process. Manual transcription is clearly not a feasible solution due to large amounts of data waiting to be transcribed. Fully automatic transcription using HTR techniques offers a cost-effective alternative, but often fails in delivering the required level of transcription accuracy [36]. Instead, semi-automatic or semi-supervised transcription methods have gained importance in the recent past [36–40].

The transcription method proposed in [40] uses a computer assisted and interactive HTR technique: CATTI (Computer Assisted Transcription of Text Images) for fast, accurate and low cost transcription. For an input text line image to be transcribed, an iterative interactive process is initiated between the CATTI system and the end-user. The system thus generates successively improved transcription in response to the simple user corrective feedback.

Image and language models from partially supervised data have been adapted in [38] to perform computer assisted handwritten text transcription using HMM-based text image modeling and n-gram language modeling. This method has been recently implemented in GIDOC (Gimp-based Interactive transcription of old text Documents) [41] system prototype where confidence measures are estimated using word graphs that helps users in finding transcription errors.

An active learning based handwritten text transcription method is proposed in [39] that performs a sequential line-by-line transcription of the document, and a continuously re-trained system interacts with the end-user to efficiently transcribe each line.

The performance of CATTI system [40], and the methods proposed in [38] and [39] is dependent upon accurate detection of the text lines in each document page. However, the line detection and extraction in historical handwritten document images is a challenging task, and advanced line detection techniques [42] are required.

In practical scenarios, such methods are not appropriate as a system should ideally accept a full document page as an input and generate full transcription of the words as an output. An end-to-end system for handwritten text transcription is presented in [36,37] that also uses HMM-based text image modeling with interactive computer assisted transcription. The transcription method proposed

in this work addresses these issues and introduces *TexT* for quick transcription of handwritten text using a segmentation-free word spotting algorithm [22]. The following section explains the proposed method and its advantages in detail.

## 5 *TexT* - Text Extractor Tool

This paper presents a framework for semi-automatic transcription of historical handwritten manuscripts and introduces a simple interactive text extractor tool, *TexT* for transcribing words in textual electronic format. The method is based on the idea of transcribing each unique word only once for the whole document, including annotations such as gender, geographical locations, etc. This will both speed up the tedious work of transcription and also make it less exhausting. Furthermore, an interactive approach is proposed where the system finds other occurrences of the same word on-the-fly using so-called word spotting system [22, 43]. The user simply identifies one occurrence, and while the word is being written by the user, the HTR engine finds other possible occurrences of the same word, which are shown to the user, meanwhile it continues in the background to search other pages. Further, the user helps the HTR engine in marking words that are correctly identified and correcting misclassified words. By marking these words, writing their corresponding letter sequence, and adding annotations, the HTR engine in the meanwhile processes these words and more accurately identifies them, making a better distinction between these two classes of words.

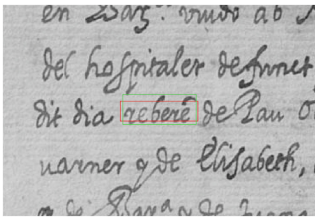
The proposed method inherits features from our previous work [34] and efficiently performs on-the-fly annotation of handwritten text with automatic generation of ground truth labels, and dynamic adjustment and correction of user annotated bounding box labels with perfect encapsulation of the text inside the rectangle. Interestingly, the transcriptions are generated such that the transcribed word contains no added noise from the background or surroundings. This is made possible by the use of two band-pass filtering approach for background noise removal [44]. This is followed by connected components extraction from the word image.

The following features are important parts of the *TexT* project planning:

- A simple yet informative, and user-friendly GUI that may attract users according to well defined topics such as botany, history, theology, diaries, etc.
- A GUI where the user can download the transcription results on-the-fly as they are distributed in the University library digital repository.
- Presence on social networks.
- A ranking system combined with a merit-report for the use of the contributor.
- A proof-reading structure with a first and a second proof-reader and a safe yet quick ingestion mechanism for the repository.
- A graphic illustration of progress for each topic.

- An administration of the application which includes active outreach to find interested audiences, close monitoring of the uploaded content and general advertising of opportunities, news and activities, including events which might give contributors extra value, such as exhibitions and shows of the original material.
- An HTR application, active only in the background, making use of the user input through machine learning and delivering better results based on the user input.

The combination of crowdsourcing and HTR is crucial and, it is believed to be one of the key factors for the *Text* project. Human interaction with AI (artificial intelligence) might be the best way to combine IT-technologies with those interested in contributing to the cultural heritage [45].



reberé

(a) Input document with user marked word (red) and system corrected bounding box (green).

(b) Clean transcribed word *reberé* with background noise removed.

**Fig. 1.** The user marks a word in the document (in the left), shown in red bounding box. The system finds the best fitting rectangle (in green) to perfectly encapsulate the word. The background noise is removed and the clean transcribed word generated is shown on the right. Figure best viewed in color. (Color figure online)

## 6 Experimental Framework and Implementation Details

This section emphasize on the overall experimental framework of *Text* along with insight on its implementation details. The proposed framework is tested on the Esposalles dataset [46], a subset of the Barcelona Historical Handwritten Marriages (BH2M) database [47]. BH2M consists of 244 books with information on 550,000 marriages registered between 15th and 19th century. The Esposalles dataset consists of historical handwritten marriages records stored in archives of Barcelona cathedral, written between 1617 and 1619 by a single writer in old Catalan. In total, there are 174 pages handwritten by a single author corresponding to volume 69, out of which 50 pages are selected from 17th century. In future, the ancient manuscripts from the Uppsala University library will be used for further experimentation.

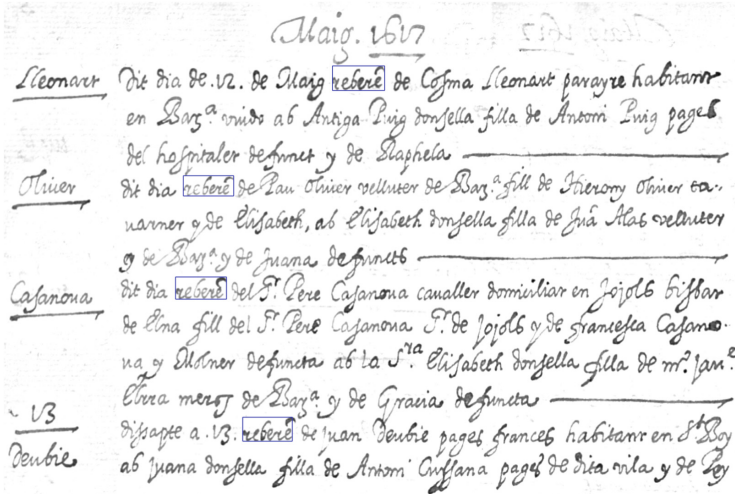


Fig. 2. The result of searching one word marked by the user (for example, reberé), represented using blue bounding box. Figure best viewed in color. (Color figure online)

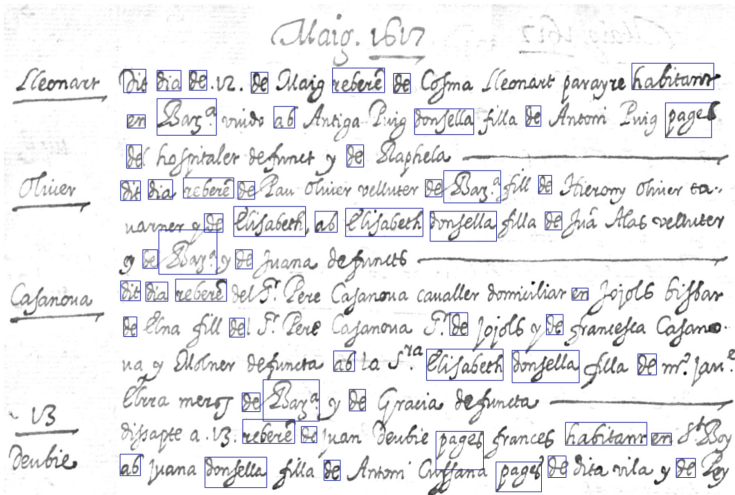
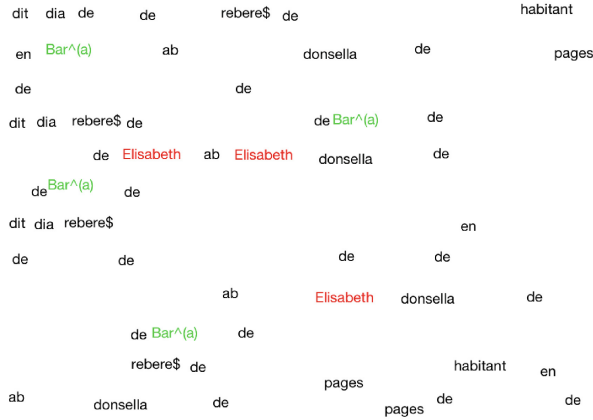


Fig. 3. The transcription can be performed in any order and in this case 11 different words have been marked, and the other occurrences are found automatically. Figure best viewed in color. (Color figure online)

The text transcription method based on word spotting is performed as follows. The system generates a document page query where the user marks a query word with a so called rubber band rectangle. The user marked red bounding box is highlighted in Fig. 1a for a sample word reberé. The system automatically finds the best fitting rectangle to perfectly encapsulate the word, as shown in





**Fig. 4.** The ongoing transcription results in words being identified in their corresponding places. In this case, the user has also annotated names and places using different colors. Figure best viewed in color. (Color figure online)

Fig. 1a using green bounding box, and extracts the word. Furthermore, the noise from the background and surroundings is efficiently removed using two band-pass filtering approach in order to make the subsequent search more reliable (see Fig. 1b).

The system starts searching for the word in the document page and the result is shown in Fig. 2. Note that only a cropped part of the document page from the dataset is shown for demonstration. The search is performed while the user inserts the transcribed text together with the annotations. Now the user can let the system learn by clicking on one or several word boxes confirming that they are correctly found. If the system find words that are misclassified, the user can inform the system by clicking a button to switch from correct to incorrect mode, and then selecting the words. While doing this, the system continues to perform word search on other document pages and update the search on the basis of information the system learns from the user (Fig. 4).

The user can select words in any order by marking them once. Figure 3 shows how 11 words have been chosen and the system finds the rest. The corresponding transcription is shown in Fig. 3. In this case, the user has annotated some words as names (highlighted in red) and others as geographical places (highlighted in green). This example of a place represents the abbreviation for the word *Barcelona*.

## 7 Conclusion and Future Work

The transcription tool *Text* presented in this paper is based on an interactive word spotting system, and lends itself to collaborative work, such as online crowdsourcing for large-scale document transcription. The proposed method can

be further improved using client-server or cloud-based solution to perform transcription without much latency. So far algorithms for word spotting [22] have been developed and a simple experimental framework is proposed to support the transcription approach presented herein.

As future work, we intend to implement a transcription framework on ancient manuscripts from Uppsala University Library that works as follows. Each user can freely mark words, annotate them and also identify words found by the search as correct or incorrect. The major part of the search will be performed on a dedicated computer that splits the work in parallel, making it possible to search even large documents in a few seconds. It can be noted that searching one word in our MATLAB implementation takes about 2s for the example shown in Fig. 2. The word spotting approach used in this work [22] efficiently performs parallel processing such that the search in a single page can be distributed into several processes, and hence making the search much faster. Different learning methods are being evaluated to improve the transcription algorithm. Deep learning techniques can be used only when several hundreds of annotated examples are available for a document, but when starting a transcription of an entirely new document, no such are usually available.

**Acknowledgment.** This work was supported by the Riksbankens Jubileumsfond (Dnr NHS14-2068:1) and the Swedish strategic research programme eSENCE.

## References

1. Mori, S., Nishida, H., Yamada, H.: Optical Character Recognition. Wiley, New York (1999)
2. Govindan, V.K., Shivaprasad, A.P.: Character recognition - a review. *Pattern Recogn.* **23**(7), 671–683 (1990)
3. Blanke, T., Bryant, M., Hedges, M.: Open source optical character recognition for historical research. *J. Doc.* **68**(5), 659–683 (2012)
4. Plamondon, R., Srihari, S.N.: Online and off-line handwriting recognition: a comprehensive survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(1), 63–84 (2000)
5. Marti, U.V., Bunke, H.: Hidden Markov Models, pp. 65–90. World Scientific Publishing Co., Inc., River Edge (2002)
6. Toselli, A.H., Vidal, E.: Handwritten text recognition results on the Bentham collection with improved classical N-gram-HMM methods. In: *Proceedings of the 3rd International Workshop on Historical Document Imaging and Processing, HIP 2015*, pp. 15–22. ACM, New York (2015)
7. Espana-Boquera, S., Castro-Bleda, M.J., Gorbe-Moya, J., Zamora-Martinez, F.: Improving offline handwritten text recognition with hybrid HMM/ANN models. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(4), 767–779 (2011)
8. Parvez, M.T., Mahmoud, S.A.: Offline Arabic handwritten text recognition: a survey. *ACM Comput. Sur.* **45**(2), 23:1–23:35 (2013)
9. <http://urn.kb.se/resolve?urn=urn:nbn:se:alvin:portal:record-12537/> (2017)
10. Howe, J.: The rise of crowdsourcing. *Wired Mag.* **14**(6), 1–4 (2006)
11. Moyle, M., Tonra, J., Wallace, V.: Manuscript transcription by crowdsourcing: transcribe bentham. *Liber Q.* **20**(3–4), 347–356 (2011)

12. <http://blogs.ucl.ac.uk/transcribe-bentham/2017/08/21/transcription-update-22-july-to-18-august-2017/> (2017)
13. <http://read.transkribus.eu/>
14. <http://transkribus.eu/Transkribus/>
15. Borne, K., Team, Z.: The zooniverse: a framework for knowledge discovery from citizen science data. In: AGU Fall Meeting Abstracts (2011)
16. <http://scripto.org/>
17. <http://velehanden.nl/> (2017)
18. <http://transcription.si.edu/> (2017)
19. <http://siarchives.si.edu/blog/smithsonian-crowdsourcing-1849/> (2017)
20. <http://www.alvin-portal.org/> (2017)
21. <http://urn.kb.se/resolve?urn=urn:nbn:se:alvin:portal:record-101351/> (2017)
22. Hast, A., Fornés, A.: A segmentation-free handwritten word spotting approach by relaxed feature matching. In: 2016 12th IAPR Workshop on Document Analysis Systems (DAS), pp. 150–155. IEEE (2016)
23. Héroux, P., Barbu, E., Adam, S., Trupin, É.: Automatic ground-truth generation for document image analysis and understanding. In: Ninth International Conference on Document Analysis and Recognition, ICDAR 2007, pp. 476–480. IEEE (2007)
24. Pletschacher, S., Antonacopoulos, A.: The page (page analysis and ground-truth elements) format framework. In: 2010 20th International Conference on Pattern Recognition (ICPR), pp. 257–260. IEEE (2010)
25. Yanikoglu, B.A., Vincent, L.: Pink panther: a complete environment for ground-truthing and benchmarking document page segmentation. *Pattern Recogn.* **31**(9), 1191–1204 (1998)
26. Kanungo, T., Lee, C.H., Czorapinski, J., Bella, I.: TRUEVIZ: a groundtruth/metadata editing and visualizing toolkit for OCR. In: Document Recognition and Retrieval VIII, vol. 4307, pp. 1–13. International Society for Optics and Photonics (2000)
27. Yacoub, S., Saxena, V., Sami, S.N.: PerfectDoc: a ground truthing environment for complex documents. In: Proceedings of the Eighth International Conference on Document Analysis and Recognition, pp. 452–456. IEEE (2005)
28. Saund, E., Lin, J., Sarkar, P.: PixLabeler: user interface for pixel-level labeling of elements in document images. In: 10th International Conference on Document Analysis and Recognition, ICDAR 2009, pp. 646–650. IEEE (2009)
29. Doermann, D., Zotkina, E., Li, H.: GEDI - a groundtruthing environment for document images. In: Ninth IAPR International Workshop on Document Analysis Systems (DAS) (2010)
30. Clausner, C., Pletschacher, S., Antonacopoulos, A.: Aletheia - an advanced document layout and text ground-truthing system for production environments. In: 2011 International Conference on Document Analysis and Recognition (ICDAR), pp. 48–52. IEEE (2011)
31. Biller, O., Asi, A., Kedem, K., El-Sana, J., Dinstein, I.: WebGT: an interactive web-based system for historical document ground truth generation. In: 2013 12th International Conference on Document Analysis and Recognition (ICDAR), pp. 305–308. IEEE (2013)
32. Valsecchi, F., Abrate, M., Bacciu, C., Piccini, S., Marchetti, A.: Text encoder and annotator: an all-in-one editor for transcribing and annotating manuscripts with RDF. In: Sack, H., Rizzo, G., Steinmetz, N., Mladeníć, D., Auer, S., Lange, C. (eds.) *ESWC 2016. LNCS*, vol. 9989, pp. 399–407. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-47602-5\\_52](https://doi.org/10.1007/978-3-319-47602-5_52)

33. Antonacopoulos, A., Karatzas, D., Bridson, D.: Ground truth for layout analysis performance evaluation. In: Bunke, H., Spitz, A.L. (eds.) DAS 2006. LNCS, vol. 3872, pp. 302–311. Springer, Heidelberg (2006). [https://doi.org/10.1007/11669487\\_27](https://doi.org/10.1007/11669487_27)
34. Vats, E., Hast, A.: On-the-fly historical handwritten text annotation. In: Proceedings of the 2017 Workshop on Human-Document Interaction (2017, in press)
35. Wei, H., Seuret, M., Liwicki, M., Ingold, R.: The use of Gabor features for semi-automatically generated polygon-based ground truth of historical document images. *Digit. Scholarsh. Humanit.* **32**(1), i134–i149 (2017)
36. Romero, V., Bosch, V., Hernández, C., Vidal, E., Sánchez, J.A.: A historical document handwriting transcription end-to-end system. In: Alexandre, L.A., Salvador Sánchez, J., Rodrigues, J.M.F. (eds.) IbPRIA 2017. LNCS, vol. 10255, pp. 149–157. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-58838-4\\_17](https://doi.org/10.1007/978-3-319-58838-4_17)
37. Terrades, O.R., Toselli, A.H., Serrano, N., Romero, V., Vidal, E., Juan, A.: Interactive layout analysis and transcription systems for historic handwritten documents. In: 10th ACM Symposium on Document Engineering, pp. 219–222 (2010)
38. Serrano, N., Pérez, D., Sanchis, A., Juan, A.: Adaptation from partially supervised handwritten text transcriptions. In: Proceedings of the 2009 International Conference on Multimodal Interfaces, ICMI-MLMI 2009, pp. 289–292. ACM, New York (2009)
39. Serrano, N., Giménez, A., Sanchis, A., Juan, A.: Active learning strategies for handwritten text transcription. In: International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction, ICMI-MLMI 2010, pp. 48:1–48:4. ACM, New York (2010)
40. Romero, V., Toselli, A.H., Vidal, E.: Multimodal Interactive Handwritten Text Transcription, vol. 80. World Scientific, Singapore (2012)
41. <https://www.prhlt.upv.es/wp/project/2016/idoc>
42. Bosch, V., Toselli, A.H., Vidal, E.: Semiautomatic text baseline detection in large historical handwritten documents. In: 2014 14th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 690–695. IEEE (2014)
43. Giotis, A.P., Sfikas, G., Gatos, B., Nikou, C.: A survey of document image word spotting techniques. *Pattern Recogn.* **68**, 310–332 (2017)
44. Vats, E., Hast, A., Singh, P.: Automatic document image binarization using Bayesian optimization. In: Proceedings of the 2017 Workshop on Historical Document Imaging and Processing. ACM (2017, in press)
45. Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. In: Proceedings of the 2013 conference on Computer Supported Cooperative Work, pp. 1301–1318. ACM (2013)
46. Romero, V., Fornés, A., Serrano, N., Sánchez, J.A., Toselli, A.H., Frinken, V., Vidal, E., Lladós, J.: The ESPOSALLES database: an ancient marriage license corpus for off-line handwriting recognition. *Pattern Recogn.* **46**(6), 1658–1669 (2013)
47. Fernández-Mota, D., Almazán, J., Cirera, N., Fornés, A., Lladós, J.: BH2M: the Barcelona historical, handwritten marriages database. In: 2014 22nd International Conference on Pattern Recognition (ICPR), pp. 256–261. IEEE (2014)