

# READ

## Recognition and Enrichment of Archival Documents

### D5.14. ScanREAD

Günter Mühlberger (UIBK), Markus Diem (CVL), Stefan Fiel (CVL), Florian Kleber (CVL)

Distribution:

<http://read.transkribus.eu/>

---

**READ**  
**H2020 Project 674943**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 674943



<b>Project ref no.</b>	H2020 674943
<b>Project acronym</b>	<b>READ</b>
<b>Project full title</b>	<b>Recognition and Enrichment of Archival Documents</b>
<b>Instrument</b>	H2020-EINFRA-2015-1
<b>Thematic Priority</b>	EINFRA-9-2015 - e-Infrastructures for virtual research environments (VRE)
<b>Start date / duration</b>	01 January 2016 / 42 Months

<b>Distribution</b>	Public
<b>Contractual date of delivery</b>	M12
<b>Actual date of delivery</b>	
<b>Date of last update</b>	30.11.2016
<b>Deliverable number</b>	D5.14
<b>Deliverable title</b>	ScanREAD
<b>Type</b>	Software
<b>Status &amp; version</b>	1.0
<b>Contributing WP(s)</b>	-
<b>Responsible beneficiary</b>	ULCC
<b>Other contributors</b>	UIBK, CVL
<b>Internal reviewers</b>	UCL, StAZh
<b>Author(s)</b>	Günter Mühlberger, Markus Diem, Stefan Fiel, Florian Kleber
<b>EC project officer</b>	Martin MAJEK
<b>Keywords</b>	Crowd-Scanning, Android App, DocScan, Transkribus

## Table of Contents

Executive Summary .....	4
1. Transkribus DocScan .....	4
1.1. Development Environment.....	4
1.2. Functionality .....	4
1.3. Resources .....	5
2. Methodology .....	6
2.1 Page Detection .....	6
2.2 Sharpness .....	6
2.3 Illumination .....	6
3. Transkribus ScanTent .....	6
4. Outlook.....	7
5. References.....	8

## Executive Summary

An Open Source Android app for the scanning of documents is developed. *Transkribus DocScan* detects automatically the page area and gives feedback of defined quality criteria (page area, perspective distortion, rotation, sharpness). An auto-shoot mode captures a picture if all criteria are fulfilled. The app will integrate a batch processing to allow the scanning of entire documents and stores metadata of the document (origin, title, etc.). *Transkribus DocScan* will integrate an interface to *Transkribus* to allow the user to log in and upload the images to the *Transkribus Cloud*. In addition to the app a *Transkribus ScanTent* is developed which is a universal scan equipment for mobile phones to provide homogenous scan conditions (e.g. uniform illumination).

The development of the Crowd-Scanning app was originally foreseen as a sub-contract for the Russian company ABBYY. This plan was changed due to the fact that CVL was able to reuse recently developed technology. Thus, the app can be provided as open source on github. To improve the functionality of the app and planned *ScanTent* the task will be continued in Task 8.2.

### 1. Transkribus DocScan

The main target group of *Transkribus DocScan* are archivists, scholars and researchers analysing historical documents. The developed app allows the use of a common mobile phone to scan documents and to ensure a certain quality of the pictures. Thus, no special equipment is needed and the app makes it easy for people to digitise documents (no specific knowledge is needed). The app will allow to organize the taken scans in collections and the users can choose to upload selected collections to *Transkribus* where they can be automatically transcribed or used as training data for HTR, etc.

In the next sections the development environment as well as the methodology for the page detection and quality criteria are described. Furthermore, the *Transkribus ScanTent* prototype is presented and an outlook is given.

#### 1.1. Development Environment

The app is developed in Android Studio (Version 2.1). To allow an efficient image processing OpenCV library (Version 3.1.0) is used. The OpenCV functions are called via Java Native Interfaces (JNI). A (Gradle) Buildscript is used to compile the native C++ source files and OpenCV functions with Android SDK. The camera is addressed using the `android.hardware.camera` class which is backward compatible with Android 4.0.

#### 1.2. Functionality

*Transkribus DocScan* has a live view and detects in real time the document page area. Furthermore, it evaluates following quality criteria:

- Perspective Distortion
- Size

- Sharpness
- Illumination

These criteria give the user feedback, to assure a certain quality of the picture which will be uploaded to Transkribus. The quality control minimizes the preprocessing steps (e.g. skew estimation, contrast enhancement) for the document processing in Transkribus. Transkribus DocScan has 2 modes: a manual mode allows users to shoot a picture at any time. The auto shoot mode only captures an image if all criteria are fulfilled.

Figure 1 shows one image of Transkribus DocScan with the menu to select the desired shooting mode, and shows also a detected page (yellow frame) and the focus measure regions (green = in focus, red = out of focus). If all quality criteria are fulfilled, the app displays *OK* (lower left corner) as feedback. If a quality criterion is not obtained, user commands are displayed. Figure 1 (right side) shows that the user must move closer with the camera.

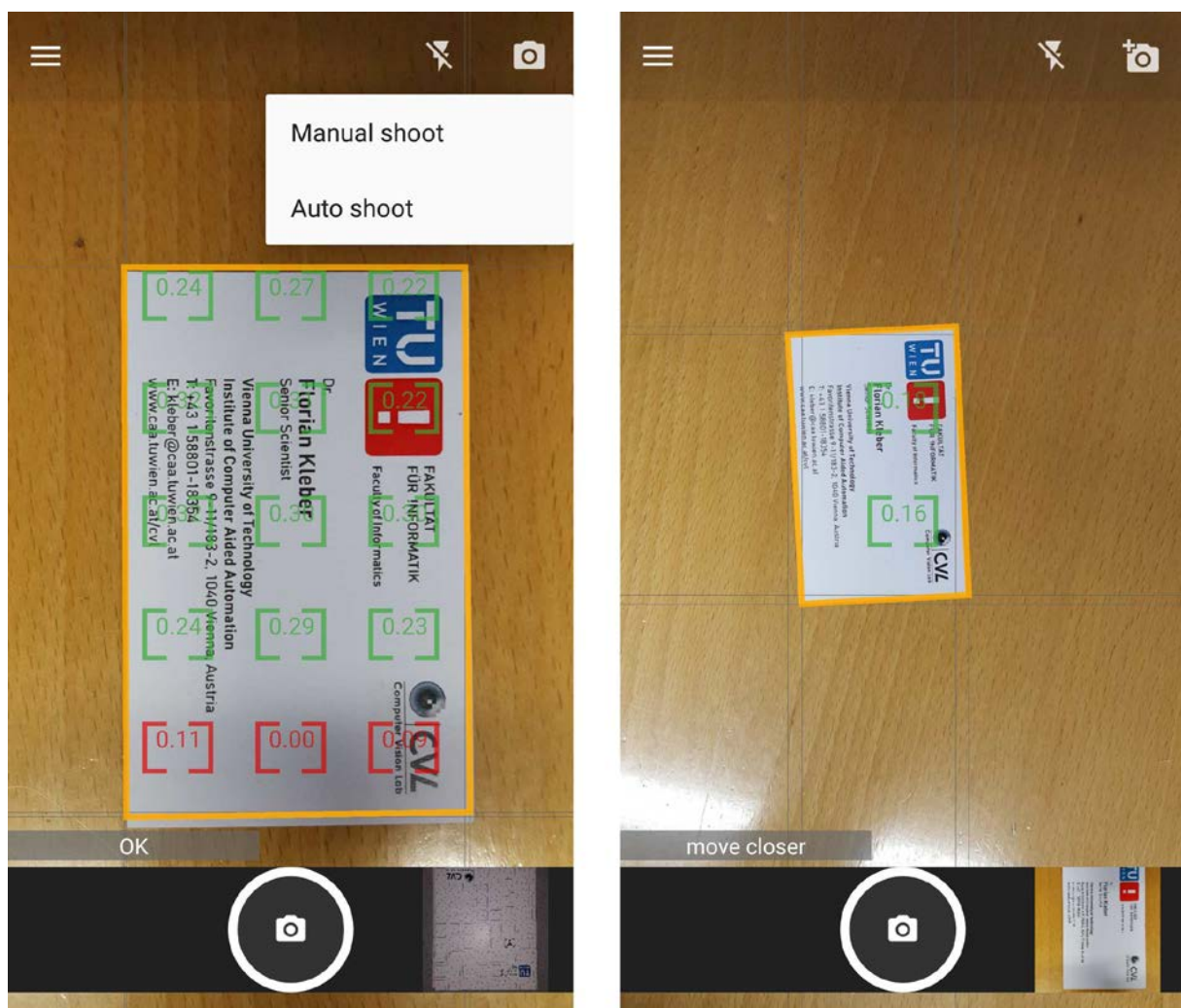


Figure 1: Screenshot of Transkribus DocScan with detected page area and focus measures. The right side shows the user feedback of the app to ensure certain quality criteria.

### 1.3 Resources

The Transkribus DocScan App is OpenSource and available at the Transkribus Github repository:

<https://github.com/TUWien/DocScan>

## 2. Methodology

The following sections describe the methodology of the Transkribus DocScan app quality criteria control. An overview of Smartphone Document Capture methodologies is presented in [1].

### 2.1 Page Detection

To localize the page area within an image, a blob based detection of rectangles is implemented. The image is binarized with N different thresholds, which increases the robustness against different illuminations and allows also a stable rectangle detection in low contrast images. After the thresholding all foreground regions are analysed. To minimize the memory allocation and the computational complexity all regions are represented as contours which are approximated by a polygon. Rectangle candidates are polygons which contain 4 edges and are convex. Additionally, all angles must have appr. 90°, thus applying the following threshold:

$$\max_{i=1,\dots,n} \cos(\theta_i) < 0.3$$

for all points  $p_i$  of a polygon, where  $\theta_i$  is the angle of the i-th corner. After detecting all rectangle candidates overlapping candidates are filtered. The page detection can be computed in real time on a smart phone and has the advantage of detecting generic rectangles in an image without the necessity of computationally complex machine learning approaches. Based on the detected rectangle also the quality criteria perspective distortion and page size is calculated.

### 2.2 Sharpness

To determine the sharpness of the page region, the image is subdivided into smaller regions and based on a focus measure operator the sharpness is calculated. Like Peruz et al. [2] we chose a gradient based approach. To get normalized values, the page region is binarized using Otsu and the gradient values of the binarized image are taken as reference (normalization factor).

### 2.3 Illumination

In order to detect uneven illumination, scanlines are sampled every N pixels within the page rectangle detected. For each scanline, the sum of all gradients is computed. Hence, a flat illumination will result in a low value that represents sensor noise. Note that edges of characters cancel themselves. The scanlines are computed horizontally and vertically which allows for detecting all gradients possible.

## 3. Transkribus ScanTent

In order to support the user of the Transkribus DocScan the idea of a “scan-tent” came up. This is a piece of equipment which holds the smartphone above the document to improve the quality of images taken by the user. The Transkribus ScanTent has two main advantages:

- 1) The user needs not to hold the smartphone with his hands, but has both hands free to keep the pages of a (bound) document still and also to turn the pages.

- 2) The quality of the image is more likely to be of higher quality since the camera is not moved and also the distance to the document is stable.
- 3) Shadows and sub-optimal lightning conditions are avoided, the images are always taken in a homogenous environment.

A prototype of a portable ScanTent has been developed. Figure 2 shows the tent which can be placed on a desk. On the bottom the scan area is marked and a ruler is added to allow the calculation of the scale. A LED illumination provides a uniform illumination. On the right side of Figure 2 the smartphone on the top of the tent is shown. The detected page is marked with the yellow rectangle.

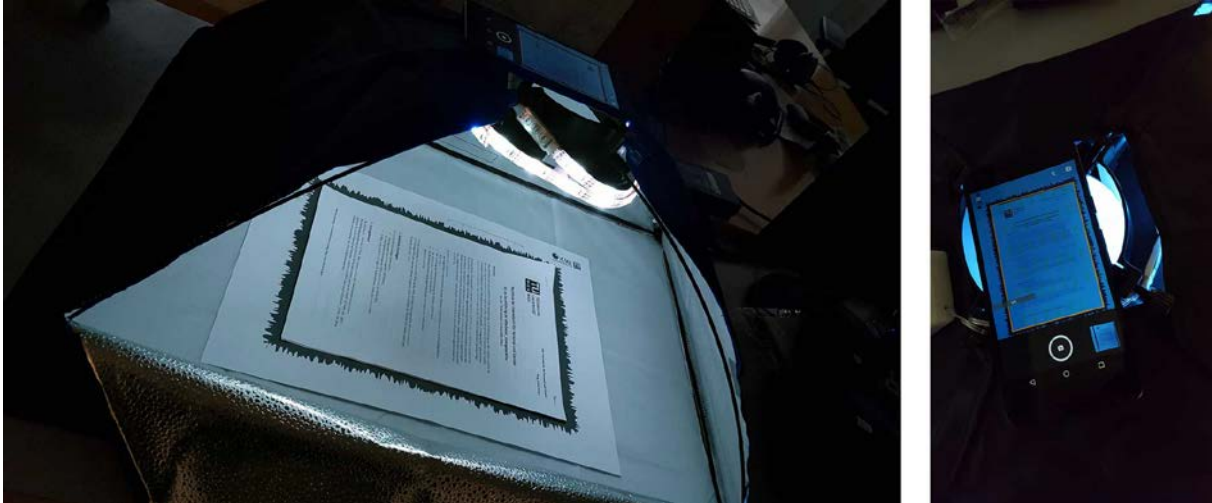


Figure 2: Transkribus ScanTent

The design allows an efficient acquisition of entire books using the planned batch mode of the Transkribus DocScan app.

Note: The realisation of this idea follows the considerations laid down for task 8.1. Open Innovation Forum.

## 4. Outlook

To improve the functionality of the app and planned ScanTent the task will be continued in Task 8.2. Future developments include a batch mode for the Transkribus DocScan app. Additionally, a user interface to store metadata of the documents captured will be implemented, as well as an interface to Transkribus to allow a user to log in and upload documents and document series.

For the Transkribus ScanTent a universal mount for the smartphone will be designed and 3D plotted. Furthermore, the illumination system will be improved.

## 5. References

- [1] J. Burie, J. Chazalon, M. Coustaty, S. Eskenazi, M. Luqman, M. Mehri, N. Nayef, J. Ogier, S. Prum and M. Rusinol, "ICDAR2015 Competition on Smartphone Document Capture and OCR (SmartDoc)," in *13th International Conference on Document Analysis and Recognition (ICDAR)*, 2015.
- [2] S. Pertuz, D. Puig and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, pp. 1415-1432, 2013.