

# **Konzept: Lib-Indexer**

**EDBV WS 2019/2020: AG\_C\_3**

Anand Eichner (11808244)

Laurenz Edmund Fiala (11807869)

Anna Nieto-Berezhinskaya (01223066)

Aleksandar Vucenovic (01635282)

Jansen Wu (01226578)

28. Oktober 2019

## **1 Ziel**

Das Projekt soll Bücher in einem Bücherregal erkennen, in Bücher-Koordinaten umwandeln und nach ihrem Label abspeichern.

## **2 Eingabe**

JPG-Bild eines Bücherregals mit Büchern, auf denen eindeutige Labels (schwarz-auf-weiß) kleben.

## **3 Ausgabe**

visuell:



Abbildung 1: Input-Bild

textuell:

Strukturierte Klartext-Datei mit Inhalt:

- Standort der Bücher (in Büchern zum Ursprung - links-oben)
- Vier Pixel-Vektoren, die ein Buch in einem Viereck umschließen

## 4 Voraussetzungen und Bedingungen

- Die Bücher müssen gerade ( $+/- 5^\circ$ ) stehen.
- Das Bild darf nicht mehr als  $30^\circ$  von der Waagrechte abweichen.
- Das Bild muss eine für die Texterkennung der Labels ausreichende Auflösung aufweisen (Abhängig von der Entfernung).
- Das Bild muss farbig sein.
- Das Bild muss ausreichend hell sein.
- Wenn möglich ist ein Weisausgleich durchzuführen.

## 5 Methodik

Methodik- Pipeline

1. Hough-Transformation  
*Finden der Regalfächer zum Korrigieren der Perspektive*
2. Perspektivenkorrektur  
*Mittels Transformationsmatrix aus HT berechnet*
3. Eckenerkennung  
*Finden der Ecken von Labels*

4. Integral imaging  
*Finden von Labels innerhalb eines Akzeptanzbereichs, es werden nur Bereiche zwischen verschiedenen, zuvor erkannten, Ecken überprüft*
5. Eigene Heuristik  
*Einordnen von Labels in Buch-Koordinaten*
6. Optical Character Recognition  
*Erkennen von Text auf den Labels in den zuvor erkannten Bereichen*

## 6 Evaluierung

### Interaktion zwischen den Komponenten:

- *Werden die Regalfächer korrekt erkannt?*  
Voraussetzungen: Seite der Regalfächer, die zur Kamera zeigt, ist schwarz.  
Ergebnis: An jedem Fach liegt eine Gerade an.
- *Wird die Perspektive korrekt angepasst?*  
Voraussetzungen: Korrekte Geraden der Regalfächer.  
Ergebnis: Bücher und Labels sind im Bild weitestgehend rechteckig.
- *Werden alle Labels erkannt?*  
Voraussetzungen: Perspektivenkorrigiertes Bild  
Ergebnis: Bounding-Boxes der gefundenen Labels
- *Sind die Bounding Boxes der gefundenen Labels korrekt? Ist der gesamte Text darin enthalten?*  
Voraussetzungen: Korrekt erkanntes Label oder ein false-positive.  
Ergebnis: Vier Vektoren, die den gesamten Text umschließen (und nicht mehr). Bei false-positives ist das Ergebnis nicht relevant, jedoch sollte es nicht zu groß sein (z.B. das gesamte Bild überdecken).
- *Werden die Labels korrekt in Bücher-Koordinaten umgewandelt?*  
Voraussetzungen: Bounding Boxes der Labels sind korrekt.  
Ergebnis: Bücher-Koordinatensystem als 2D-Array mit Ursprung links-oben.
- *Werden die Labels der TU-Bibliothek korrekt gelesen und in Text umgewandelt?*  
Voraussetzungen: Label Bounding-Boxes wurden korrekt berechnet (enthalten keinen unnötigen Text).  
Ergebnis: String-Repräsentation des Labels. Erwartete Korrektheit: > 90% für typische Datensätze.
- *Wird die Wahrscheinlichkeit der Label-Korrektheit angemessen berechnet?*  
Voraussetzungen: Korrekt in Text umgewandeltes Label.  
Ergebnis: Floating-point Wert im Intervall [0, 1]. Alle Labels mit Wahrscheinlichkeit 0 wurden entfernt.

### **Qualitative Evaluierungsfragen:**

- Welche Muster müssen wir in unseren Daten initial erkennen?
- Wie muss das Inputbild umgewandelt werden, um eine möglichst effektive Eckenerkennung zu ermöglichen?
- Wie versichern wir, dass unser Bild perspektivisch korrekt transformiert wurde?
- Welches Indexing ist sinnvoll, und wie signalisieren wir dem User die Bedeutung unserer Indexe?(Buch-Koordinaten)
- Welches Pre-Processing ist notwendig, um eine Texterkennung zu ermöglichen?

### **Quantitative Evaluierungsfragen:**

- Wieviele Bücher erkennen wir durchschnittlich?
- Welche Auflösung ist notwendig, um akzeptable Ergebnisse zu erzielen?
- Wie wirkt sich der Winkel, in dem das Foto aufgenommen wurde auf die Verwertbarkeit der Labels aus (Erkennung & OCR)?
- Wie wirkt sich der Grad der Befülltheit des Regals auf die Label-Erkennung aus?
- Wie wirken sich die Farben und Kontraste der Bücher auf unseren OCR aus?

## **7 Datenbeispiel**



Abbildung 2: Input-Bild

## **8 Zeitplan**

### **Literatur**

- [1] Mike Stephens Chris Harris. A combined corner and edge detector. *Alvey Vision Conference*, 1988. Zugriff 28.10.2019; Verantwortliche: Anna Nieto-Berezhinskaya, Laurenz E. Fiala.

Meilenstein	abgeschlossen am	Arbeitsaufwand in h
Prototyp	10.11.	30
Hough-Transformation	17.11.	50
Perspektivenkorrektur	17.11.	6
Labelerkennung	25.11.	80
Labels in Buch-Koordinaten	1.12.	4
Optical Character Recognition	10.12.	110
Labels filtern	15.12.	6
Daten in Output-Format umwandeln	15.12.	4
Tests	18.12.	5
Evaluierung	20.12.	5

- [2] Kuo-Liang Chung, Teh-Chuan Chen, and Wen-Ming Yan. New memory- and computation-efficient hough transform for detecting lines. *Pattern Recognition*, 37(5):953 – 963, 2004. Zugriff 28.10.2019; Verantwortliche: Anand Eichner, Jansen Wu.
- [3] John Illingworth and Josef Kittler. A survey of efficient hough transform methods. In *Alvey Vision Conference*, pages 1–8, 1987. Zugriff 28.10.2019; Verantwortliche: Anand Eichner, Jansen Wu.
- [4] Robert Sablatnig. Evc-cv-11. morphologische operationen und bildsegmentierung slides. *Computer Vision Lab, Technische Universität Wien*, 2019. Zugriff 28.10.2019; Verantwortliche: Anna Nieto-Berezhinskaya, Laurenz E. Fiala.
- [5] Robert Sablatnig. Evc-cv-8. bildmerkmale - interest points slides. *Computer Vision Lab, Technische Universität Wien*, pages 17–36, 2019. Zugriff 28.10.2019; Verantwortliche: Anna Nieto-Berezhinskaya, Laurenz E. Fiala.