

Lib-Indexer

EDBV WS 2019/2020: AG_C_3

Anand Eichner (11808244)

Laurenz Edmund Fiala (11807869)

Anna Nieto-Berezhinskaya (01223066)

Aleksandar Vucenovic (01635282)

5. Januar 2020

1 Gewählte Problemstellung

(1-1,5 Seiten)

entspricht dem (aktualisierten) Konzept

1.1 Ziel

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

1.2 Eingabe

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

1.3 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 Label in einem Viereck umschließen

1.4 Voraussetzungen und Bedingungen

- Die Bücher müssen gerade (+/- 5°) stehen.
- Das Bild darf nicht mehr als 30° 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.
- Ein Weißabgleich muss durchgeführt worden sein.

Neue Anforderungen:

- Der Hintergrund der Labels muss weiß sein (Intensität muss über dem globalen Otsu-Threshold liegen)

1.5 Methodik

Methodik- Pipeline

1. Hough-Transformation
Finden der Regalfächer zum Korrigieren der Perspektive
2. Perspektivenkorrektur
tuwe 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

1.6 Evaluierungsfragen

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.

1.7 Zeitplan

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 Arbeitsteilung

Name	Tätigkeiten
Anna Nieto-Berezhinskaya	?
Anand Eichner	houghTransform (+ locals), PerspectiveCorrection (+ locals),
Laurenz Edmund Fiala	label_detection (+ locals),
Aleksandar Vučenović	ocr (+ locals), ncc (+ locals), ssd_naive (+ locals), preprocessing (+ locals)

3 Methodik

(2-3 Seiten)

Hier wird die verwendete Methodik in der Theorie vorgestellt:

Welche Methodik wurde verwendet? Warum eignet sich diese Methodik für die gewählte Problemstellung? Habt ihr Methoden verändert (Einschränkungen, Abwandlungen, Parameter), wenn ja wie? etc.

Die erwähnten Methoden werden zum größten Teil auf Beschreibungen in Büchern oder wissenschaftlichen Artikeln beruhen. Daher ist hier auch der richtige Platz für Zitate. Die hier zitierten Publikationen sollten mittels Abkürzung bzw. Nummer referenziert sein und sich in der Referenzliste am Ende des Berichts über diese Bezeichnung finden lassen.

Ein Beispielsatz (inkl. entsprechender Literaturangabe am Ende des Berichts): Interest

Points wurden mittels Scale Invariant Feature Transform [?] detektiert.

Bei der Verwendung von Latex gestaltet sich das Zitieren besonders einfach - siehe Beispielssatz im Source der Latex-Vorlage.

Wichtig in diesem Abschnitt ist, dass sich der Leser Eures Berichts mit den verwendeten Methodiken auskennt und weiß, weshalb ihr diese Methodiken verwendet habt und keine anderen. Es soll dem Leser helfen den nächsten Abschnitt des Berichts besser zu verstehen.

4 Implementierung

(1-X Seiten)

Hier gebt ihr einen Überblick über Eure Implementierung:

Wie habt ihr die im vorhergehenden Abschnitt vorgestellte Methodik praktisch umgesetzt? Wie werden die einzelnen Methoden kombiniert (zB. Implementierungspipeline)?

Hier ist Platz für Implementierungsdetails wie zB. gewählte Parameter.

Wie startet der User das Programm? Welche Parameter hat der User zu setzen?

Auch in diesem Abschnitt können Referenzen und Zitate notwendig sein.

Wichtig in diesem Abschnitt ist, dass der Leser Eures Berichts versteht wie ihr Euer Projekt in MATLAB umgesetzt habt um sich auch im Quelltext leichter zurechtfinden zu können.

5 Evaluierung

(2-X Seiten)

Hier stellt ihr Euren Datensatz vor und beantwortet Evaluierungsfragen:

z.B. Fakten zum Datensatz: Anzahl der Bilder, Größe der Bilder, Quelle des Datensatzes (falls selbst aufgenommen: Aufnahmegerät, Einstellungen,... / falls nicht selbst erstellt: Datenbank vorstellen... → Referenzen!)

Diskussion der Evaluierungsfragen: Beantwortung der Fragen, Diskussion anhand von Beispielen, Diskussion von Grenzfällen: für welche Bilder funktioniert die Implementierung, für welche nicht? Worin unterscheiden sich diese Bilder? Warum funktionieren sie nicht? etc.

Evaluiert wird der ganze Datensatz, nicht nur einzelne Bilder. Einzelne Bilder können zum Aufzeigen von Fehlern/Problemen/besonders guten Ergebnissen... genutzt werden.

Zur Evaluierung gehört auch das Testen der einzelnen Methodiken (separat), mit Erwähnung eventueller Einschränkungen.

6 Schlusswort

(max. 1 Seite)

Hier fasst ihr Ergebnisse Eures Projekt zusammen:

Welche Schlussfolgerung lässt sich ziehen? Gibt es offene Probleme? Wie lässt sich Eure Lösung noch verbessern? etc.

Webseiten werden als Fußzeilen (an jener Stelle wo sie verwendet werden) eingebunden, nicht als Literature!