

Automatisierte Aufbereitung archäologischer Grabungsfotos mittels Computer Vision

Simon Metzger

Masterarbeit

zur Erlangung des akademischen Grades Master of Arts im
Studiengang Digitale Methodik der Geistes- und
Kulturwissenschaften

Johannes-Gutenberg-Universität Mainz und Hochschule Mainz

Zusammenfassung

Im Format abstract

Inhaltsverzeichnis

1 Einleitung	2
1.1 Grabung Kapitol	2
1.2 Datensatz vorstellen	2
1.3 Material vorstellen	2
1.3.1 Tafeln	2
1.3.2 Tafelvergleiche	4
1.3.3 Schrift	4
1.4 Pipeline	5
2 Objektdetektion	6
2.1 Detektionsverfahren	6
2.1.1 Feature Detection	6
2.1.2 CNN	6
2.1.3 Contours	6
2.2 Cropverfahren	7
2.2.1 simple crop	7
2.2.2 Hough	7
2.3 Zusammenfassung	7
3 Texterkennung	8
3.1 Theorie Texterkennung	8
3.2 Software: Tesseract	8
3.3 Pre-Processing	8
3.4 OCR	8
3.5 Evaluation	8
4 Ergebnisse	10
5 Fazit	10

1 Einleitung

Einleitung und Fragestellung
[Hough, 1962]

1.1 Grabung Kapitol

Grabungsverlauf bis 2014 (recherchieren)
Übernahme durch DAI (recherchieren)

1.2 Datensatz vorstellen

Herkunft
Umfang
Fragestellungen des Projektes

1.3 Material vorstellen

1.3.1 Tafeln

Die Verwendung von Tafeln zur Dokumentation von Fund- und Grabungsarealen ist in allen, im weitesten Sinne grabenden, Wissenschaften weit verbreitet. So setzt auch die Archäologie diese Methode ein. Dabei werden neben den zu dokumentierenden Gebieten verschiedenste Formen von Tafeln oder Schildern platziert, auf denen Zeit und Ort der Aufnahme sowie weitere bild- und motivbezogene Informationen festgehalten werden können. Der Vielfalt von Form und Material der Tafeln ist dabei keine Grenze gesetzt.

Bei den Tafeln, die Gegenstand dieses Projektes sind, handelt es sich um Schiefertafeln mit einem Holzrahmen, die mit Kreide beschriftet wurden (Vgl. Abb 1). Für die Detektion der Tafeln ergeben sich daraus folgende Faktoren:

1. Die Tafeln haben grundsätzlich eine rechteckige Form.
2. Durch die Breite des Rahmens können bis zu zwei Rechtecke erkannt werden, ein Inneres und ein Äußeres.
3. Durch die große Differenz zwischen dem hellen Holzrahmen und der dunklen Schieferplatte sollte der innere Rand in der Regel gut detektierbar sein.



Abbildung 1: Beispiel eines Fotos der verwendeten Tafel. GOT bezeichnet die Kampagne, darunter folgt das Datum. US ist die Abkürzung für *unità stratigrafica*, die stratigrafische Einheit.

Die im Beispielbild gezeigte Tafel stellt ein Idealbild dar: Die Tafel nimmt einen relativ großen Teil des Originalbildes ein. Sie ist frontal vor der Kamera positioniert. Die Beleuchtung ist gut und indirekt. Keines der weiteren Bildelemente verdeckt die Tafel. Diese Beschreibung impliziert schon die Problemfelder, die bei der Detektion beachtet werden müssen:

1. Die Tafel ist unter Umständen stark rotiert (Vgl. Abb 2).
2. Die Distanz der Tafel zur Kamera und damit ihre Größe im Bild kann stark variieren.
3. Der Rahmen der Tafel kann teilweise verdeckt oder anderweitig durch Gegenstände überlagert sein (Vgl. Abb 2).
4. Die Farbe des Tafelrahmens kann dazu führen, dass sie sich nicht klar vom Hintergrund abhebt, was die Detektion des (äußeren) Randes erschweren kann.
5. Unregelmäßigkeiten im Rahmen, die auf grobe Verarbeitung oder Abnutzung zurückzuführen sind, können die Detektion erschweren.
6. Die Beleuchtung kann zu Problemen führen. Grundsätzlich sind alle Fotos hell und gut ausgeleuchtet, direktes Licht kann sich aber negativ auf die Kontraste auswirken.
7. Weitere Gegenstände, die den Spezifika der Tafeln entsprechen, können im Bild vorhanden sein.



Abbildung 2: Schwierigere Detektion: Rotation und teilweise verdeckter Rahmen.

Teilweise werden die hier genannten Probleme auch bei der Texterkennung wieder relevant. Auf diese und auf weitere wird an geeigneter Stelle zurückgegriffen.

1.3.2 Tafelvergleiche

Im Rahmen der Arbeit wurden weitere Tafeln exemplarisch dem Algorithmus unterzogen. Dabei handelte es sich um Aufnahmen der späteren Grabungen des Deutschen Archäologischen Instituts am Kapitol in Rom sowie um vergleichbare Fotos von Bodenuntersuchungen der Gruppe Terrestrische Ökohydrologie der Friedrich-Schiller-Universität Jena. Der ursprüngliche Gedanke dahinter war eine möglichst universale Detektion von Tafeln aller Art anzustreben. Die unterschiedlichen Daten konnten dabei vor allem Stärken und Schwächen der letztlich gewählten Technik aufzeigen.

Die Tafeln beider Projekte sollen im Folgenden kurz vorgestellt werden, um das Spektrum der Komplexität evtl. Vergleiche zu Tafeln aus späterer Grabung als Positivbeispiel:

besser gearbeitete Tafeln

besser lesbare Schrift

evtl. Vergleiche zu Tafeln der Bodenkunde als Negativbeispiel:

Tafel schwierig durch Form und Farbe

Klarsichthülle: Reflektion und Formveränderung

oft verdeckt

Bilder zur Veranschaulichung einfügen

1.3.3 Schrift

Kreide auf Schiefer Probleme wie Handschrift, Verwischung, Karomuster

1.4 Pipeline

Struktur der Arbeit wie Pipeline: Bildakquise Objekterkennung Crop-Verfahren
Pre-Processing OCR Evaluation Ergebnis

2 Objektdetektion

Das folgende Kapitel befasst sich mit den ersten beiden Schritten in der automatisierten Analyse der Grabungsfotos: der Erkennung der Schiefertafeln und ihrer Extraktion aus dem Gesamtbild.

Zunächst werden verschiedene Möglichkeiten der Erkennung präsentiert und diskutiert. Der Schwerpunkt liegt hier auf dem schlussendlich umgesetzten Verfahren. Abschließend wird das mit der Tafeldetektion verbundene Ausschneiden der gefundenen Tafeln aus dem Gesamtbild vorgestellt.

2.1 Detektionsverfahren

Aus den oben genannten Punkten und den Anforderungen der Aufgabenstellung lässt sich als erster Arbeitsschritt die Detektion der Tafeln ableiten. Die wichtigste Prämisse ist dabei, dass Falsch-Negative, also nicht erkannte Tafeln, vermieden werden. Grund dafür ist, dass diese für die weitere Bearbeitung komplett verloren sind. Falsch-Positive sollten ebenfalls weitestgehend ausgeschlossen werden. Diese können jedoch in späteren Bearbeitungsschritten, bspw. der Texterkennung, erkannt und aussortiert werden. Daher ist dieses Kriterium von niedrigerer Priorität. Dieses Kapitel wird sich daher verschiedenen Verfahren widmen, mit denen rechteckige, beschriftete Objekte erkannt werden können. Diese Verfahren werden kurz vorgestellt und ihre Ergebnisse in einer ersten, explorativen Umsetzung präsentiert. Anschließend wird begründet, warum diese Verfahren im Rahmen des Projektes nicht zum Einsatz kommen. Final wird dann das Kontur-basierte Erkennungsverfahren erläutert, mit dem die besten Ergebnisse erzielt wurden.

2.1.1 Feature Detection

Was ist Feature detection? Wie funktioniert sie? Was war die Idee hinter dem Ansatz? Wie sehen die Ergebnisse aus?

2.1.2 CNN

Hier werden CNNs vorgestellt. Was sind CNNs? Was können sie, wie funktionieren sie? Warum habe ich sie ausprobiert, was war die Idee dahinter? Warum habe ich nicht selbst trainiert? Wie sehen die Ergebnisse aus?

2.1.3 Contours

Was sind Contours? Worin besteht die Grundidee? Wie wurde diese Idee umgesetzt? -> Flowchart Zweispurigkeit der Ansätze: iterativ und adaptiv. Erklären warum.

`rect_detect` als Finale, in dem die beiden Ansätze wieder zusammengeführt werden

2.2 Cropverfahren

Was ist die Aufgabe beim Crop? Worin liegen hier die Schwierigkeiten? Auch hier wieder Zweispurigkeit der Ansätze erklären

2.2.1 simple crop

Was ist die Idee? Wie wurde sie umgesetzt? Wo liegen die Probleme?

2.2.2 Hough

Was ist die Idee? Wie wurde sie umgesetzt? Wo liegen die Probleme?

2.3 Zusammenfassung

evtl zusammenfassen wie vorgegangen wurde, warum dieser Weg gut ist und was das Wichtigste Ergebnis ist

3 Texterkennung

3.1 Theorie Texterkennung

Ursprünge der Texterkennung

Aktueller Forschungsstand

Wechsel auf machine learning

Schwerpunkte

Handschrift vs. Druckschrift

3.2 Software: Tesseract

evtl ALternativen

Tesseract: was kann es, wie funktioniert es

Wechsel von Charactererkennung via CV zu Zeilenerkennung via machine learning

wie funktioniert die Box-detection?

3.3 Pre-Processing

Preprocessing: besondere Herausforderungen, vorgehen, beide Varianten vorstellen

3.4 OCR

normales Modell

eigenes Modell

Vergleich: Tafeln aus späterer Grabung (gesetzte Lettern)

evtl. Vergleich Tafeln Bodenkunde

3.5 Evaluation

(Kapitel evtl. vorziehen wegen Bedeutung für den ganzen Prozess -*i* auch als Maß für Objektdetektion)

Evaluation: Vorgehen, Überlegungen

Verwendete Kennzahlen

Vergleichbarkeit der Ergebnisse

4 Ergebnisse

Ergebnisse aus dem kompletten Datensatz präsentieren

5 Fazit

Auswertung des geschriebenen Codes Materialkritik Ausblick und weitere Ideen

Literatur

[Hough, 1962] Hough, P. V. C. (1962). Method and Means for Recognizing Complex Patterns. *U.S. Patent*, (3,069,654).