

Passanten mithilfe von Webcams erfassen und zählen

Jeroen Staab

Vorhaben

In meiner Masterarbeit im Sommer 2017 möchte ich Personen mithilfe von Webcams erfassen und zählen. Eine Vielzahl Studien aus den unterschiedlichsten Forschungsfelder beschäftigt sich mit digitaler Bildverarbeitung. Der Lehrstuhl für Fernerkundung des Instituts für Geographie und Geologie, Würzburg, ist ein moderner und erfolgreicher Forschungszweig des deutschen Zentrums für Luft und Raumfahrt (DLR). Zusammen mit dem dortigen Mitarbeiter Dr. Michael Thiel habe ich schon oft über mögliche Anwendungsfelder vor unserer Haustüre nachgedacht. Ist es nicht schade, dass obwohl wir einen solch starken Lehrstuhl haben die Stadt selbst, bzw. die Humangeographie so wenig involviert ist? Als Master-Student für Humangeographie interessiert mich die Synthese.

Ziel

Ziel meiner Masterarbeit für angewandte Humangeographie ist, eine anwendungsnahe Brücke zwischen Humangeographie und digitaler Naherkundung zu schlagen. Dabei möchte ich so viel wie möglich bestehende Technik nutzen, um kostengünstig Passanten zu erfassen und zu zählen. Objektive, kontinuierliche Zeitreihen über lange Zeiträume über Besucherströme erlauben es, aussagekräftige Studien über den wirtschaftlichen Nutzen von Umweltschutzgebieten zu veröffentlichen.

Ziel der Umsetzung selbst ist es, interdisziplinäres unter einen Hut zu bekommen. Zusammen mit Dr. Michael Thiel vom Lehrstuhl für Fernerkundung möchte ich ein Computer-Programm programmieren, welches universal einsetzbar ist. Der Lehrstuhl für Geographie und Regionalforschung von Prof. Dr. Job soll meine Arbeit betreuen und erhält dadurch im Gegenzug eine fertig konfigurierte Software Lösung, welche in aktuellen Forschungsprojekte eingebunden ist.

Aktueller Stand

Zahlreiche Studien aus Automobilindustrie, Sicherheitssektoren und Computer-Wissenschaften beschäftigen sich mit der Detektion von Passanten. Diese verwenden unterschiedlichste Sensoren (Kamera, Video, Laser,...) und Algorithmen (Image Differenzierung, Objekt Erkennung, maschinelles Lernen). Bisher habe ich auf der einfachsten Stufe minütlich Fotos voneinander abgezogen, um Veränderungen aufzuzeigen. Aufgrund der hohen temporalen Auflösung fielen schwankende Wetterbedingungen (Helligkeit) nur sehr gering ins Gewicht. Diese also stabile und sehr universell einsetzbare Methode zählt allerdings nur die Anzahl signifikant veränderter Pixel, die absolute Anzahl Menschen im Bild lässt sich jedoch davon ableiten.

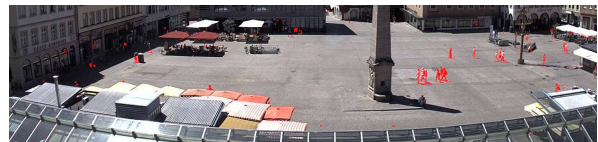


Abbildung 1. Detektionen überlagern das Originalbild in Rot

Entwickelt wird die Methode, wie in der Fernerkundung üblich, zunächst vor der eigenen Haustüre. In diesem Fall im Stadtzentrum der Stadt Würzburg. Seit knapp einem Jahr erfassen wir minütlich den unteren Marktplatz. Erste Betrachtungen der Zeitreihe machen einen plausiblen Eindruck.

Bislang wurde das Projekt in *R* und *Linux* programmiert. Das rührt daher, dass ich viel Erfahrung aus dem Vegetations-Monitoring mitbringe und u.a. an Phenopix mitgearbeitet habe. Deshalb sind Daten- und Zeitserien-Struktur identisch, sowie die nachfolgenden statistischen Analysen kompatibel.

Probleme & Lösungsansätze

Im Detail offenbart die obige Abbildung Ungenauigkeiten. Sowenig wie eine andere Zählverfahren wird auch eine Klassifikation nie zu 100% korrekt sein, doch lässt sich der Fehler mit einer ausführlichen Genauigkeitsanalyse quantifizieren.

Schon jetzt sehe ich drei Fehlerquellen, wobei es in der Fernerkundung für alle drei Lösungen gibt, welche sich übertragen lassen. Die Erste ist der vollautomatisiert Zahlalgorithmus an sich, welcher mit größter Sorgfalt implementiert werden muss. Zwei unterschiedliche Methoden stehen zur Disposition: (a) Bislange verwende ich ein einfaches ImageDifferenzierung, es ist schnell zu prozessieren und liefert LIVE Auskünfte über die aktuelle Situation und ist relativ stabil gegenüber Umwelteinflüssen. Selbst wenn jemand die Kamera bewegen würde, wäre die Erfassung in der darauf folgenden Minute wieder korrekt. Wie jedoch erwähnt werden hier nicht die Anzahl Passanten sondern nur die Anzahl signifikant veränderter Pixel gezählt. Stehende, bzw. sitzende Menschen werden nicht erfasst. (b) Deutlich cooler wäre ein Objekt orientiertes Klassifikationsverfahren wie ich es bereits in meiner Bachelorarbeit bei Dr. Michael Thiel angewendet habe. Hier werden Passanten explizit als solche identifiziert und gezählt. Auch weitere Klassen wie Fahrradfahrer, Rollstuhl/Kinderwagen, Auto, etc. wären möglich.

Zweitens steht und fällt jede Zeit-Serien Analyse mit der Datendichte, sondern wirkt sich letztere auch positiv auf die Klassifikationsgenauigkeit aus. Videos erfassen meist 15-30 Bilder pro Sekunde und werden z.B. in London eingesetzt um Menschen gar zu tracken. In unserem Fall jedoch, wo die Kamera an abgelegenen Standorten steht, Bilder entweder hochladen oder direkt auswerten müssen, müssen über Strom- und Bandbreiten Ressourcen nachgedacht werden. Wir sind schließlich nicht das englische Ministerium für Innere Sicherheit sondern Geographen. Bislange verwende ich ein Foto pro Minute, im Jahr macht das 120GB. Wie mit Johannes Schamel diskutiert könnten je nach erfassten Blickwinkel auch eine schnellere Repititionsrate notwendig sein, um die Grundgesamtheit der auf schmalen Trails wandernden Touristen zu erfassen. Um die Bildfrequenz konfigurieren zu können brauchen wir eine eigene Kamera. Eine Teil-Fragestellung meiner Arbeit soll sein, den Einfluss der Aufnahmefrequenz

auf das Ergebnis zu quantifizieren.

Abschließend sehe ich noch eine Vielzahl weiterer Schwierigkeiten auf das Projekt zukommen. Gerade in heterogenem Gelände, mit einer Kamera unterhalb von Baumwipfeln sehe ich herausfordernde Illuminationsschwankungen auf den Klassifikator zukommen. Diese lassen sich jedoch mit Filter-Verfahren aus der Fernerkundung glätten. Auch die Anbindung des Messgeräts an Strom und Internet muss sichergestellt sein. Hier können wir die Erfahrungen von Prof. Dr. Job seinem Team mit vergrabenen Trittzählern einbinden.

Aufbau der Masterarbeit

Wie aufgezeigt soll meine Masterarbeit soll verschiedene Disziplinen unter einen Hut bringen - Humangeographen verwenden aktuelle Methoden aus den Computer-Wissenschaften um ihre Fragestellungen objektiv quantitativ, kostengünstig und effizient zu lösen. Leider ist das Feld aber noch sehr neu und dementsprechend wichtig ist es mir die Brücke auch auf beiden Seiten gut verankert zu wissen. Um den unterschiedlichen Arbeitsmethoden aus Fernerkundung und Regionalforschung gerecht zu werden, möchte ich meine Arbeit aus zwei Teilen zusammensetzen:

1. Digitaler Werkzeugkasten

Der methodische Teil besteht zum einen aus einem mehrere Tausend Zeilen Quellcode umspannenden Software-Paket. Ich möchte versuchen es als OpenSource Projekt zu führen, darüber zu bloggen und vor allem es anschließend unter einer Creative-Commons Lizenz frei zugänglich machen.

Zum anderen besteht er aus der Software-Dokumentation. Diese soll in englischer Sprache erfolgen und von der Länge etwa der Hälfte der gesamten Arbeit entsprechen.

Programmieren möchte ich in einer kreativen Umgebung, besonders die angestrebte Objekt-Erkennung zu implementieren ist ein Meilenstein. Dazu möchte ich mich in die offene Python Bibliotheken namens OpenCV einarbeiten. In der Fernerkundung haben Hooman Latifi, Dimi Dimov und Steven Hill Erfahrung mit der Software, ihnen möchte ich mich anschließen.

2. Anwendungsbeispiel Nationalpark

Die zweite Hälfte meiner Masterarbeit präsentiert die generierten Ergebnisse. Ich verfasse den Text wahlweise auf Deutsch oder auch auf Englisch, je nachdem wie Prof. Dr. Job das präferiert. Identifizierte Fehlerquellen und resultierende Fehlerbalken werden ausführlich diskutiert. Des weiteren sollen die Webcam-Ergebnisse mit bestehenden Zahlergebnissen verglichen werden und mithilfe einer multiplen Regression genaue Besucherzahlen modelliert werden.

Gerne bin ich bereit den Aufbau-, Betrieb- und Wartung der Kamera im Gelände mit zu übernehmen und die entsprechende Verantwortung zu tragen.

Material

Realisiert werden kann die Datenerfassung im Gelände mit verschiedenen Kamerasystemen, wovon im Folgenden ein günstiges und ein teureres Setup vorgeschlagen wird. Wichtig ist bei beiden Systemen auf Witterungsbeständigkeit, infrastrukturelle Anbindung (Strom + UMTS) und Diebstahlschutz zu achten. Als dritter Materialpunkt wird hier noch auf Server-Infrastruktur hingewiesen.

1. Setup

Im Vegetations-Monitoring haben sich Kameras der Marke *Stardot* bewährt. 1500€ reichen für einen Zählstandort:

- Die *NetCam XL* ist eine äußerst robuste Webcam, sie hält Temperaturschwankungen von -40°C bis 50°C stand. Direkt vom Hersteller kostet sie in der 5 Megapixel Variante 1099\$. Es gibt sie auch ein Modell für 1299\$, welches auch Nachts Daten erheben kann, doch halte ich diese Funktion für unser Projekt überflüssig.
- Jede Kamera braucht eine Wetterfeste Hülle, bspw. die *StarDot Weatherproof Enclosure* für 159\$.
- Für weiteres Zubehör (Internetanschluss, Gestänge/Baummontage, evtl. Tarnung) sollten noch einmal 350€ beantragt werden.

2. Setup

Einen absoluten Low-Cost Ansatz (609,60€) verfolgt die zweite Setup-Variation. Der RaspberryPi 3 ist höchst universell einsetzbar und kann flexibel für die jeweilige Erhebung programmiert werden.

- Ein vernünftiges Starterset, inklusive 5 Megapixel Webcam, Netzteil und primärer Hülle ist bei Conrad für 119,95€ unter der Bestell Nr. 1435680 - 62 zu kaufen.
- Auch hier wird eine Wetterfeste Hülle gebraucht. Aufgrund der kompakten Größe lässt sich der RaspberryPi 3 inklusive Kamera bspw. auch obige Stardot Hülle von umgerechnet 139,65€ einbauen.
- Für weiteres Zubehör (Internetanschluss, Gestänge/Baummontage, Tarnung und ggf. Heizung) sollten noch einmal 350€ beantragt werden.

Server Infrastruktur

Die Kameras selbst verfügen nur über einen kleinen internen Speicher und übertragen die Daten regelmäßig zu einem Server. Diese kann man sich mieten, etwa bei der Serverprofis GmbH für 4,99€ im Monat das Paket "Private XL 5.1". Doch macht es meiner Meinung nach, aus Sicherheitsgründen, deutlich mehr Sinn den Server in das Hochschulnetzwerk zu integrieren. Ein eigener Server hat darüber hinaus noch viele weitere Vorteile wie bspw. einen verschlüsselten SSL Zugang. Der Lehrstuhl für Fernerkundung würde diesem Projekt einen ideal konfigurierten Server zur Verfügung stellen. Zur Datensicherung bedarf es einer internen 3,5 Zoll Festplatte mit einem Terabyte für 50€.