

SMobEye: Mobile Objekt- und Texterkennung

Fachgebiet: Human Centered Engineering

Betreuer: Prof. Roger Cattin

Experte: Dr. Martin Schmid (Kantonsspital Luzern)

Diese Arbeit siedelt sich im Gebiet der Bilderkennung, genauer bei der Objekt- und Texterkennung an. Ziel war es, die Grundlagen für eine mobile Software zu entwickeln, welche blinden und sehbehinderten Menschen in ihrem Alltag hilft bei der Umgebungserkennung und -Interaktion. Anhand dieses Projekts soll evaluiert werden, ob eine solche Software mit den limitierten Ressourcen von Tablets oder Smartphones möglich ist. Als Beispielapplikation wurde die Erkennung von Türklingeln und die Suche nach einem bestimmten Namen implementiert.

Die Zielstellung dieser Arbeit wurde zusammen einer blinden Lehrperson der Blindenschule in Zollikofen erarbeitet. In einem Interview wurde diskutiert, welche Probleme im Alltag einer blinden oder sehbehinderten Person mit Hilfe von Smartphones oder Tablets gelöst oder zumindest vereinfacht werden können. Aus der Liste der möglichen Problemen wurde schlussendlich das Auffinden einer bestimmten Türklingel gesucht.

Diese Aufgabe wird in mehrere Schritte unterteilt. Zuerst muss das aktuelle Bild von der Rückenkamera erfasst werden. Für die Erkennung von Kanten und von Text ist wichtig, dass das Bild möglichst wenig Unschärfe aufweist. Da nicht immer ein Bildstabilisator vorhanden ist, werden die in den meisten Geräten verbauten Beschleunigungssensoren verwendet um festzustellen, ob sich das Gerät in Bewegung befindet und deswegen mit der Aufnahme gewartet werden soll.

Anschliessend werden im Bild Türklingeln gesucht. Für die Knöpfe wurde dazu ein formbasierter Ansatz mit einer Kantenerkennung und der Hough Transformation implementiert. Auch bei der Erkennung der Namensschilder wird auf einen auf der Form basierenden Ansatz gesetzt. Zuerst wird das Rauschen im Bild reduziert. Anschliessend werden in dem Graubild Rechtecke gesucht.

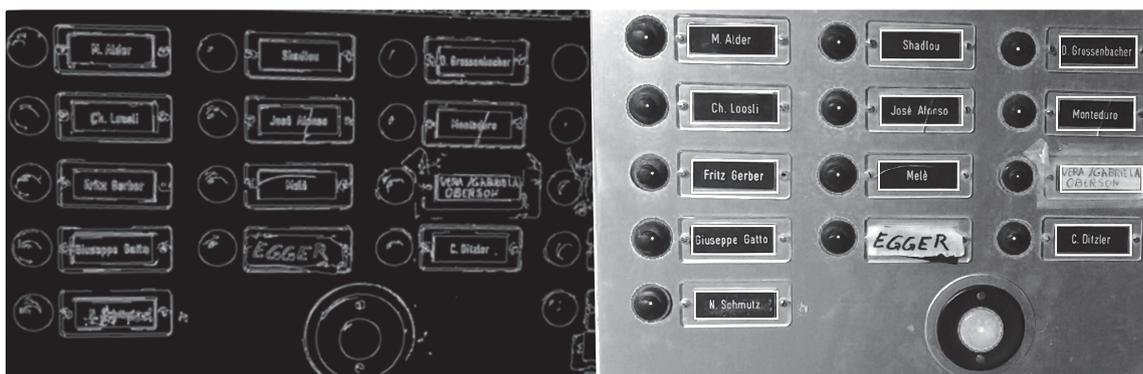
Dazu werden aus dem Bild mehrere Binärbilder mit unterschiedlichem Schwellwert erstellt. In diesen wird dann nach Konturen gesucht, welche den Kriterien für ein Rechteck entsprechen. Das Resultat dieser Erkennung wird in dem Bild dargestellt. Nachdem die Position der Türklingel und des Namensschildes gefunden worden ist, muss der Text auf dem Schild ausgelesen werden. Dazu wird die freie OCR Software Tesseract verwendet. Die Namensschilder werden nach einer Vorverarbeitung, welche dunklen Text auf hellem Grund erstellt, ausgelesen.

Nun kann der Anwender nach einem bestimmten Text oder Teiltex suchen. Die Suche arbeitet nicht exakt, sondern unscharf. Die gefundenen Texte werden mit dem vom Anwender eingegebenen Text verglichen. Die Ähnlichkeit wird mit den Algorithmen von Monge-Elkan und der Levensthein Distanz bewertet. Der Treffer mit der grössten Ähnlichkeit ist das vom Anwender gesuchte Namensschild.

Die vorgestellten Ansätze und Teile des Codes können auch für ähnliche Probleme im Alltag verwendet werden. Beispiele sind die Suche nach Informationen auf einem Fahrplan, die Erkennung und das Auslesen von Strassenschildern und Hausnummern oder das Finden von Preisschildern in einem Ladengeschäft.



Christian Loosti



Resultat einer Türklingelerkennung