

Aufgabe 1

Computer Vision

Hochschule Darmstadt, Sommersemester 2020

Roman Kessler

30. April 2020

1 Aufgabe 1.1

Im folgenden wird der Vorgang skizziert, wie ich aus einem Bild der Fingerkuppe mit mittelmäßiger Auflösung und Beleuchtungsbedingungen den Fingerabdruck sichtbar gemacht habe. Untersuchungsobjekt ist mein linker Zeigefinger. Für erste Übungen mit *OpenCV* und Bilderarbeitung in *Python 3* und zur besseren Reproduzierbarkeit der Ergebnisse wurde die Aufgabe in Python durchgeführt.

Im ersten Schritt wurde das Schwarz-Weiß Bild auf den zu interessierenden Bereich zurecht geschnitten ("cropped"), und das Histogramm dargestellt (siehe Abbildung 1).

An dem Histogramm erkennt man schon eine bimodale Verteilung der Grauwerte in den Bereichen 100 und 200 (Abbildung 1, rechts). Um den Grauwertbereich besser auszunutzen, entscheide ich mich, das Histogramm zu strecken. Dadurch wird die Verteilungsfunktion zu einer (etwa) linearen Funktion gemacht. Die Dichte der Grauwerte wird von 0 zu 255 gestreckt (Abbildung 2).

Durch die Streckung des Histogramms werden die Muster der Fingerkuppe für das menschliche Auge schon besser ersichtlich. Als nächstes wurde eine adaptive Histogramm-Angleichung versucht. Dazu wird das Bild in kleinere Kacheln (8x8 Pixel) geteilt, und eine Histogrammangleichung für jede Kachel separat durchgeführt. Wir sehen noch einmal eine leichte Verbesserung in der Erkennung der Strukturen (Abbildung 3, links) und eine noch gleichmäßigere und stufenlosere Verteilung der Grauwerte (Abbildung 2, rechts, auch wenn diese für das menschliche Auge vermutlich kaum zu erkennen ist).

Weiterführend könnte man noch eine Grauwertangleichung durchführen, die wie ein Gradient über das Bild läuft, von links oben (dunkle Bereiche) nach rechts unten (helle Bereiche), um das Bild homogener zu machen.

2 Aufgabe 1.2

Um die Rillen des Innengewindes einer Schraubenmutter zu überprüfen, würde ich mich für die Fourier-Transformation entscheiden.

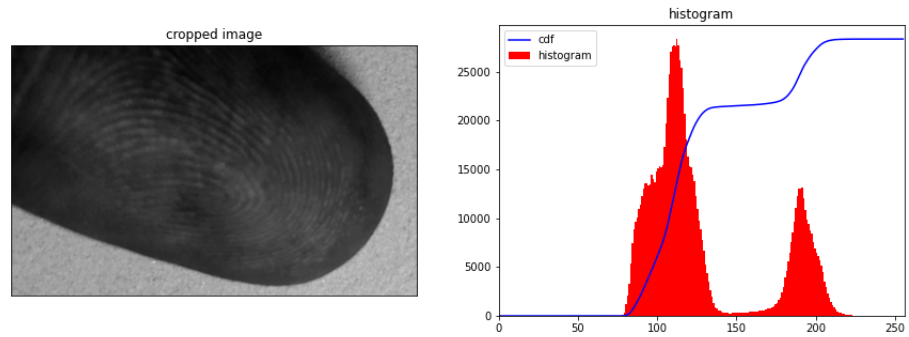


Abbildung 1: Ursprungsbild

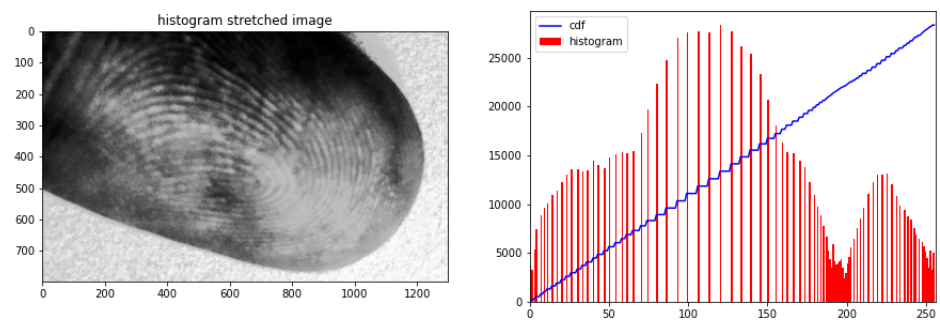


Abbildung 2: Histogramm-gestrecktes Bild

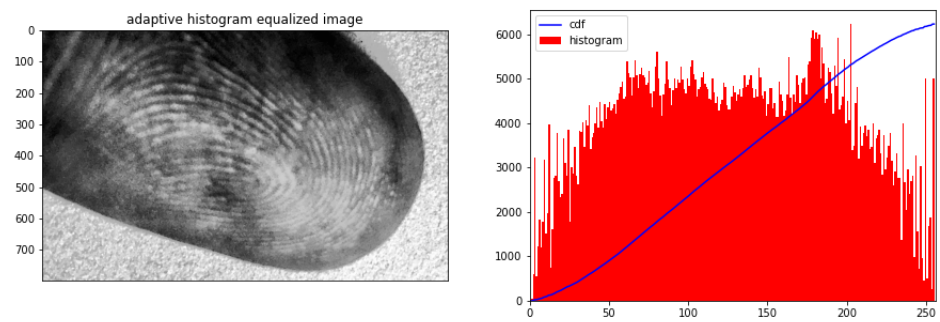


Abbildung 3: Adaptive Histogrammangleichung

Bei einer sauberen Fräsung, sollte ich in der Gewindenfrequenz, und in den Oberwellen dazu große Amplituden erhalten. Weiterhin sollte ich weniger starke Amplituden in den Frequenzen dazwischen erhalten. Je besser das Signal-zu-Rausch Verhältnis ist (also hohe Amplitude in den "guten" Frequenzen, niedrige in den umliegenden Frequenzen dazwischen), desto sauberer der Schliff. Ich sollte dabei jedoch auf die Richtung achten (orthogonal zur Bohrung). Um vielleicht einen Phasenverschiebung bei der Gewindefräsung auszuschließen, sollte ich ebenfalls die Phase der Fourier Transformation betrachten und mit meiner zu erwartenden Phase abgleichen.

3 Aufgabe 1.3

Es sollte sowohl mit weißen Gesichtern auf schwarzem Hintergrund, als auch mit schwarzen Gesichtern auf weißem Hintergrund trainiert werden. Andere Farbkombinationen sollten auch möglich sein, aber das Training könnte vielleicht weniger effizient erfolgen. Ebenfalls sollten auch Bilder im Trainingsdatensatz sein, welche mehr als 1 Gesicht gleichzeitig zeigen, damit der resultierende Algorithmus sich nicht gezwungen sieht, sich für eines der Gesichter zu entscheiden.

4 Aufgabe 1.4

Zunächst spielten für meine Einordnung des Bildes markante Merkmale des Stiers, wie Hörner, Schweif, Körper, Haltung der Beine eine Rolle. Weiterhin die gesamte Haltung des Körpers, die auf Aggression schließen lässt. Den Menschen erkennen wir auch sofort an typischen Merkmalen. Da er eine selbstbewusste und exponierte Körperhaltung hat, leiten wir ab, dass die Situation gewollt ist, und die Szenerie keine Fluchtsituation darstellt.

Ich denke ein CNN könnte hier anhand der Körper-Merkmale der beiden Figuren die einzelnen Figuren erkennen. Das Gesamtbild einordnen zu können, käme auf die Art des Netzwerkes und die Daten an, mit denen es trainiert wurde, und, ob es auf Szenen oder auf einzelne Objekte trainiert wurde.

Systematisch, um einzelne Objekte zu erkennen, würde man zunächst alle Kanten im Bild segmentieren. Weiterhin Anhand der Helligkeitswerte der Objekte (hier: Stier dunkel, Hintergrund hell) die verschiedenen Teilbereiche des Bildes (z.B. Pixel) zu zusammengehörigen Objekten clustern (=segmentieren).

5 Aufgabe 1.5

Wir betrachten die Abbildung 4. Wie können wir daraus ein zusammenhängendes Bild generieren - Weichzeichnen mit einer Faltung mit z.B. einem Gaussfilter oder Mittelwert-Filter. Es könnte aber auch ein "Maximum-Value" Filter verwendet werden, da die Dunklen Bereiche anschließend durchgängig gemacht werden, und die hellen Bereiche durchgängig hell bleiben.

Ein CNN, welches das Ziel hat, das Objekt zu Klassifizieren (z.B. binär: Gesicht / kein Gesicht), würde implizit wahrscheinlich ähnlich vorgehen. Denn ein CNN hat üblicherweise mehrere Faltungsschichten (Convolutional Layers). Da Bilder üblicherweise verrauscht sind, würden einige der Convolution Filters sicher auch während dem Training zu einer Verwaschung des Bildes führen.

6 Aufgabe 1.6

Interpretationsmöglichkeiten

1. Viereck
2. Schlangenlinie
3. Zusammengesetztes Objekt aus beidem

Wir würden aufgrund des Gestaltungsgesetzes "**Gesetz der Kontinuität / Gesetz der guten Fortsetzung**" wahrscheinlich die Schlangenlinie als vordergründiges / interessierendes Objekt wahrnehmen. Ein Computer Vision Algorithmus arbeitet hier i.d.R. anders, und würde eher das Viereck als Objekt von Interesse kategorisieren.

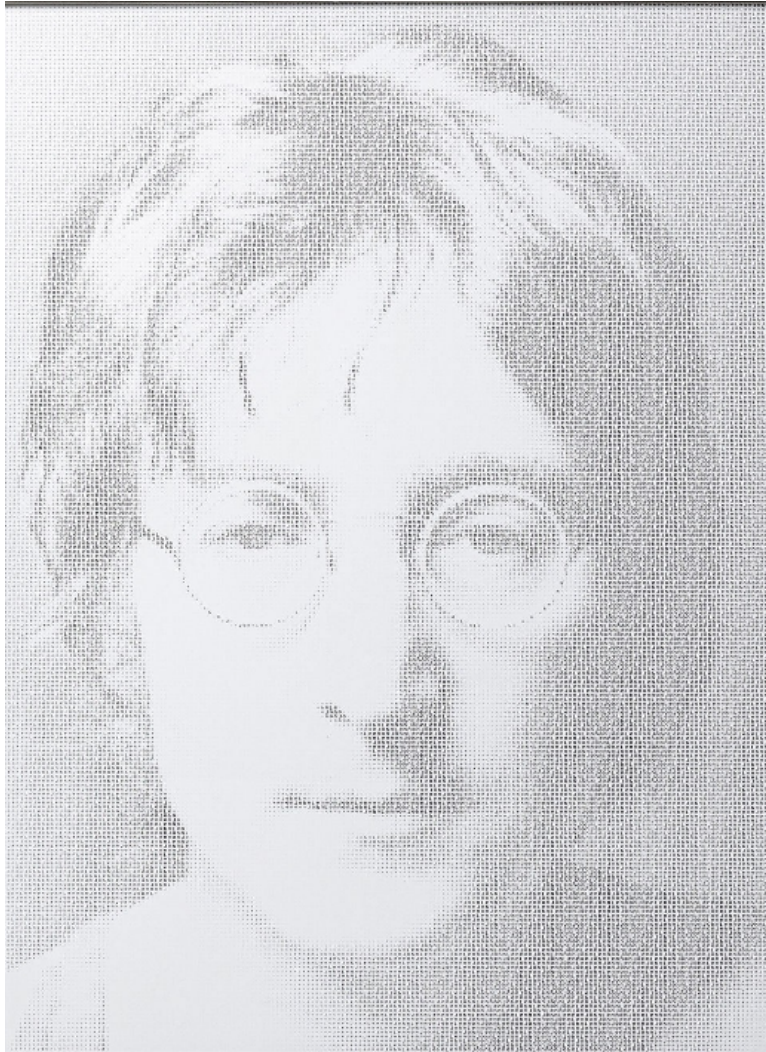


Abbildung 4: John Lennon