

GMT Tutorium - Probeklausur

Lea Hering

Universität Tübingen

11. Februar 2022

Aufgabe 1: Bildanalyse

1. Nennen Sie vier Anwendungsgebiete der Bildanalyse.

- Zeichenerkennung
- Qualitätsprüfung in der industriellen Produktion
- Medizinische Bildanalyse
- Luftbilddauswertung
- Fahrzeugsteuerung
- Gesichtserkennung
- Robotik
- Inhaltsbasierte Bildsuche im Internet

Aufgabe 1: Bildanalyse

2. Was ist ein Histogramm und welche Bildeigenschaften sind aus ihm ersichtlich? Erklären Sie kurz.

Ein Histogramm ist eine graphische Darstellung der Häufigkeit einzelner Werte eines Merkmals, wie z.B. die Grau- oder Farbwerte eines Bildes. Daraus ist der Kontrastumfang und die Helligkeit des Bildes ersichtlich. Zudem können Aussagen über die Dynamik des Bildes getroffen werden.

3. Vergrößert sich der lokale Kontrast bei einer Anhebung des globalen Kontrastes? Erklären Sie kurz.

Ja, der lokale Kontrast vergrößert sich immer bei einer Anhebung des globalen Kontrastes, wobei der lokale Kontrast nicht analog zum globalen vergrößert wird.

Aufgabe 1: Bildanalyse

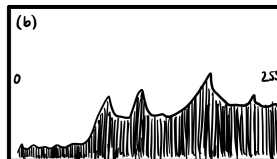
4. Sind Punktoperatoren geeignet, um den lokalen Kontrast zu erhöhen? Erklären Sie kurz.

Ja, beispielsweise die Logarithmus-Transformation

5. Skizzieren Sie die Histogramme zu den folgenden Bildbeschreibungen:

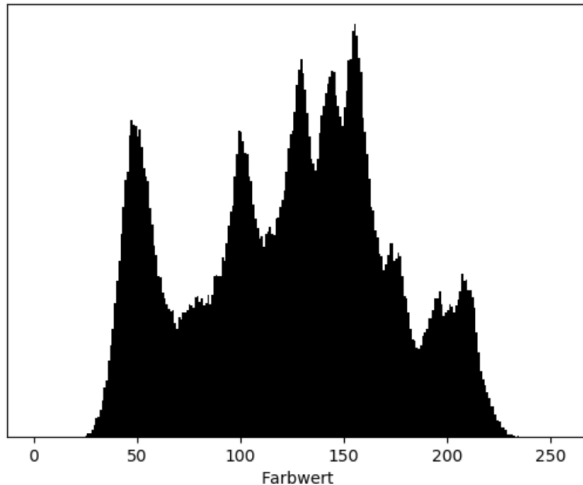
(a) Stark überbelichtetes Bild

(b) Bild mit hohem Kontrast und hoher Dynamik



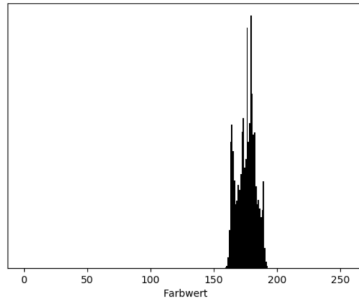
Aufgabe 1: Bildanalyse

6. Betrachten Sie das folgende Histogramm eines Graustufenbildes.



Aufgabe 1: Bildanalyse

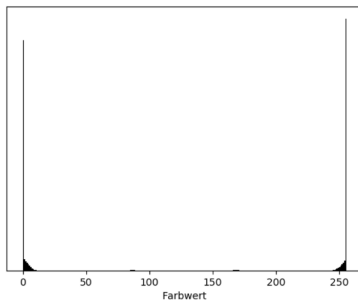
6.(a) Welche Veränderungen zum Originalbild?



Helligkeit erhöht, Kontrast verringert

Aufgabe 1: Bildanalyse

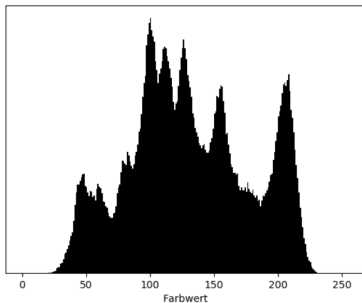
6.(b) Welche Veränderungen zum Originalbild?



Erhöhung des Kontrasts, dadurch Verringerung der Dynamik

Aufgabe 1: Bildanalyse

6.(c) Welche Veränderungen zum Originalbild?



Invertierung des Bildes

Aufgabe 2: Filter

1. Welcher der folgenden Operatoren ist ein Kirsch-Operator?
Markieren Sie Ihre Auswahl deutlich.

$$h_1 = \begin{pmatrix} 5 & 3 & 3 \\ 5 & 0 & 3 \\ 5 & 3 & 3 \end{pmatrix}$$

$$h_3 = \begin{pmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{pmatrix}$$

$$h_2 = \begin{pmatrix} -3 & 0 & 4 \\ -3 & 3 & 4 \\ -3 & 0 & 4 \end{pmatrix}$$

$$h_4 = \begin{pmatrix} -5 & -3 & 3 \\ -5 & 0 & 3 \\ -5 & 3 & 3 \end{pmatrix}$$

h_3 ist ein Kirsch-Operator

Aufgabe 2: Filter

2. Gegeben seien die Grauwerte eines 3×3 -Grauwertbildes. Berechnen Sie das Ergebnis für das mittlere Pixel nach Anwendung des Laplace-Operators.

Geben Sie den Rechnungsweg für an. Schreiben Sie das Ergebnis als eine Ganzzahl.

$$\text{LaPlace} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 4 & 6 & 4 \\ 7 & \mathbf{0} & 23 \\ 15 & 30 & 35 \end{pmatrix}$$

$$0 \cdot 4 + 1 \cdot 6 + 0 \cdot 4 + 1 \cdot 7 + (-4) \cdot 0 + 1 \cdot 23 + 0 \cdot 15 + 1 \cdot 30 + 0 \cdot 35 = 66$$

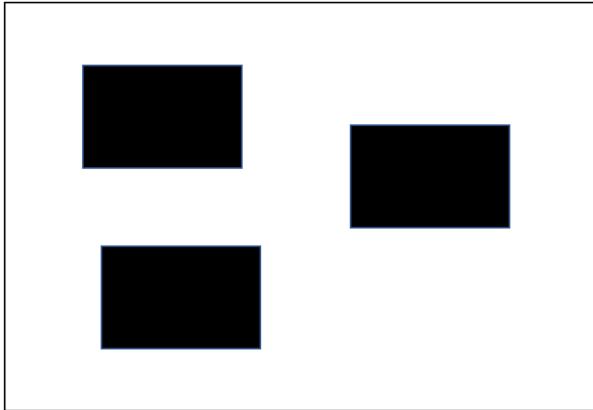
Aufgabe 2: Filter

3. Erklären Sie den Unterschied in der Wirkungsweise des Gauß- und des Laplace-Filters bei lokalen Unterschieden eines Bildes.

Lokale Unterschiede sind die Wertunterschiede der Eigenschaften eines Pixels zu denen seiner Nachbar-Pixeln. Während beim Gaußfilter lokale Unterschiede geglättet werden, werden bei einer Differenzbildung des Laplace-Filters lokale Unterschiede verstärkt.

Aufgabe 2: Filter

4. Skizzieren Sie das folgende Bild nach Anwendung des SobelX Filters.



Aufgabe 2: Filter

4. Skizzieren Sie das folgende Bild nach Anwendung des SobelX Filters.



Aufgabe 3: Nyquist-Shannon-Abtasttheorem

Gegeben sei eine sinusförmige Schwingung mit der Frequenz $f = 1000$ Hz.

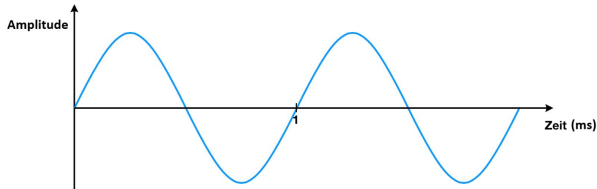
1. Wie hoch muss nach dem Nyquist-Shannon-Abtasttheorem die Abtastfrequenz f_{abtast} mindestens sein?

Die Abtastfrequenz f_{abtast} muss nach dem Nyquist-Shannon-Abtasttheorem mindestens $2 \times f = 2 \cdot 1000 \text{ Hz} = 2000 \text{ Hz}$ betragen.

Aufgabe 3: Nyquist-Shannon-Abtasttheorem

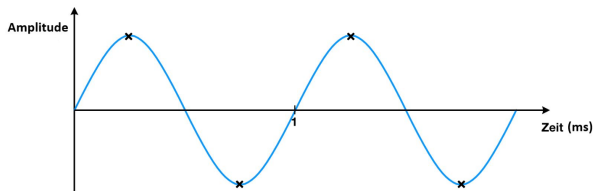
2. Im folgenden ist eine Sinus-Schwingung mit der Frequenz $f = 1000$ Hz abgebildet.

Gehen Sie von einer minimalen Abtastfrequenz nach Nyquist-Shannon-Abtasttheorem aus. Zeichnen Sie die Abtastpunkte direkt in der abgebildeten Grafik ein. Skizziere die resultierende Schwingung nach Abtastung.



Aufgabe 3: Nyquist-Shannon-Abtasttheorem

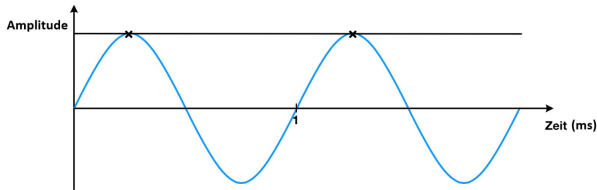
- 2.(a) Die resultierende Schwingung ist entweder identisch zur gegebenen Schwingung oder ein Rechtecksignal mit der gleichen Frequenz. Die Abtastpunkte können alle gleichermaßen auf der Sinus-Kurve verschoben sein.



Aufgabe 3: Nyquist-Shannon-Abtasttheorem

2.(b) Nun sei die Abtastfrequenz $f_{\text{abtast}} = f = 1000$ Hz.

Die resultierende Schwingung entspricht einer Geraden durch die Abtastpunkte. Die Abtastpunkte können alle gleichermaßen auf der Sinus-Kurve verschoben sein.



Aufgabe 4: Bildkompression

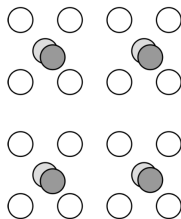
1. Auf welcher Eigenschaft des menschlichen Sehens basiert das Verfahren der Unterabtastung? Erklären Sie kurz.

Das menschliche Auge kann Farben geringer auflösen als Helligkeitsstufen. D.h., dass Helligkeit besser wahrgenommen wird als Farbunterschiede.

Mit Farbunterabtastung wird daher die Chrominanz mit einer geringeren Abtastrate gespeichert, ohne dass der subjektive optische Qualitätsverlust stark darunter leidet. Der Vorteil ist hierbei die Reduktion der benötigten Datenmenge eines Bildes.

Aufgabe 4: Bildkompression

2. Welches Unterabtastungsverfahren ist im unteren Diagramm dargestellt und in welchem Bildformat wird es verwendet?
Das Diagramm symbolisiert 4×4 Pixel, wobei jeder einzelne Pixel eine Helligkeitsinformation besitzt, während sich eine Gruppe aus vier Pixeln die Farbinformationen teilt.



- (a) Unterabtastung 4:2:0
(b) Bildformat: JPEG

Aufgabe 4: Bildkompression

3. Welche Vorteile bietet die Diskrete Kosinustransformation (DCT) gegenüber der Diskreten Fourier-Transformation (DFT) bei der Bildkompression? Erklären Sie kurz.

Die diskrete Fouriertransformation beschreibt ein beliebiges endliches, aber periodisches Signal als eine Summe von gewichteten periodischen Funktionen (Basisfunktionen) mit unterschiedlichen Frequenzen (Frequenzraumrepräsentation). Diese Transformation bewirkt, dass möglichst viel Information durch eine kleine Anzahl von Elementen ausgedrückt wird (Informationsverdichtung).

Aufgabe 4: Bildkompression

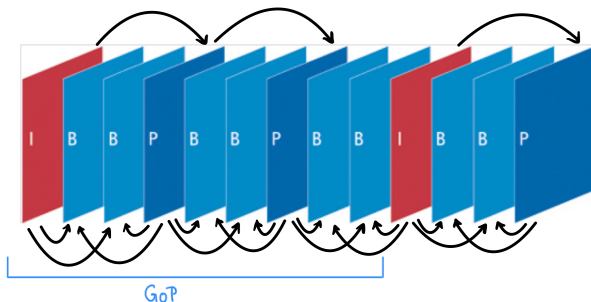
3. Die diskrete Kosinustransformation transformiert ebenfalls ein endliches Signal zu einer Summe aus Basisfunktionen. Während in der DFT die Basisfunktionen aus einer Menge an komplexen Sinus-artigen Exponentialfunktionen bestehen, sind sie bei der DCT ausschließlich Kosinusfunktionen. Ein zu transformierendes Signal kann damit in einer Menge aus Kosinusfunktionen ausgedrückt werden, welche sich (im Gegensatz zur DFT) nur durch eine kleine Anzahl an Koeffizienten unterscheiden. Dies führt zu einer nochmals gesteigerten Informationsverdichtung.
Weiterhin führt die DCT zu einer Energieverdichtung, da die Amplitude schnell mit steigender Frequenz sinkt. Ein weiterer Vorteil ist, dass sich der Nullpunkt (Grundton) in einer Ecke befindet und damit eine einfache und für die RLE-Kompression günstige Anordnung (Zick-Zack) vorgenommen werden kann.

Aufgabe 5: Video-Kompression

1. Nennen Sie die drei wichtigsten Frame-Typen im MPEG-1-Format und beschreiben Sie kurz deren Funktion.
 - ▶ I-Bild (Intra-Coded Picture): Ein I-Bild entspricht einem konventionellen Standbild. Er dient als Anker für den wahlfreien Zugriff. Dieser Bildtyp wird nur per JPEG-Verfahren komprimiert, benötigt aber keine Information von Nachbarsbildern um dekodiert zu werden.
 - ▶ P-Bild (Predicted Picture): P-Bilder beinhalten lediglich die Änderungen zum vorherigen Bild. Sie benötigen daher Informationen von vorausgegangenen I-Bildern oder P-Bildern, um dekodiert zu werden. Sind sind dadurch aber stärker komprimierbar als I-Bilder.
 - ▶ B-Bild (Bi-Predictive Picture): B-Bilder sind abhängig von den vorausgegangenen und folgenden I- oder P-Bildern, lassen sich aber dadurch nochmals stärker komprimieren und erreichen die höchste Kompression.

Aufgabe 5: Video-Kompression

2. (a) Zeichnen Sie die verschiedenen Abhängigkeiten der Frames untereinander mit Pfeilen in die obenstehende Grafik ein.



Aufgabe 5: Video-Kompression

2. (b) Was ist eine Group of Pictures? Erklären Sie kurz und markieren Sie gegebenenfalls alle Group of Pictures im obenstehenden Diagramm.

Eine Group of Pictures (Bildgruppe) ist eine Gruppe von in Abhängigkeit untereinander kodierten, aufeinander folgenden Bildern/Frames im Bilderstrom eines Videos. Basierend auf den oben beschriebenen Bildtypen, setzt sich eine Bildgruppe aus mehreren dieser zusammen, sodass ein Kompromiss zwischen Kompression und Datenstromrobustheit erzielt wird. Die B-Bilder sind sowohl von Vorgängern als auch von Nachfolgern abhängig. Der Empfänger dieser Videosequenz muss daher einen genügend großen Puffer zur Verfügung haben und einige Bilder im Voraus laden, ehe ein B-Bild dekodiert und angezeigt werden kann. P-Bilder hängen nur vom Vorgänger ab. Falls ein Übertragungsfehler innerhalb der Bildgruppe auftaucht, können (erst) ab der folgenden Bildgruppe wieder alle Bilder dekodiert werden. Weiterhin können I-Bilder ohne größeren Dekodieraufwand direkt angesprochen und angezeigt werden.

Aufgabe 6: Gestenerkennung

1. In welchen zwei Kategorien kann die Sensorik zur Gestenerkennung eingeteilt werden? Nennen Sie zu jeder Kategorie jeweils Vor- und Nachteile.
 - ▶ Am Körper befindlicher Sensorik (z.B. Datenhandschuh)
 - + Direkter Zugang zu Merkmalspunkten aus den Sensordaten. Damit sind Bewegungsvektoren direkt ableitbar.
 - + Hohe Präzision
 - Sensoren/Geräte müssen am Körper getragen werden
 - Keine Erkennung ohne das Zutun der Person möglich (Person muss z.B. Datenhandschuh zunächst anziehen)

Aufgabe 6: Gestenerkennung

1. In welchen zwei Kategorien kann die Sensorik zur Gestenerkennung eingeteilt werden? Nennen Sie zu jeder Kategorie jeweils Vor- und Nachteile.
 - Mittels externer (optischer) Sensorik (z.B. Time-of-Flight-Kamera, Kinect, etc.)
 - + Kein Zutun der Person notwendig
 - + Vollständige Bewegungsfreiheit
 - Merkmalspunkte müssen zunächst algorithmisch erkannt werden
 - Größere Fehlerrate als beim direkten Auslesen aus Körpersensoren

Aufgabe 6: Gestenerkennung

2. Dynamic Time Warping (DTW) ist ein wichtiges Verfahren bei der Gestenerkennung.

(a) Erklären Sie das generelle Prinzip hinter DTW kurz.

Dynamic-Time-Warping (DTW) ist ein Algorithmus, welcher zwei Wertfolgen (auch mit unterschiedlichen Längen) aufeinander abbildet.

Zwischen diesen Wertfolgen kann mit DTW ein Differenzmaß aufgestellt werden. Daraufhin sucht der DTW-Algorithmus den kostengünstigsten Weg vom Anfang zum Ende beider Wertfolgen.

Aufgabe 6: Gestenerkennung

2. Dynamic Time Warping (DTW) ist ein wichtiges Verfahren bei der Gestenerkennung.
 - (b) Welchen Vorteil bietet DTW bei der Gestenerkennung? Erklären Sie kurz.

Bei der optischen Gestenerkennung aus einer Bildfolge können aus mehreren Gründen, wie z.B. unterschiedliche Samplingrate, veränderte Perspektive, sprunghafte Bewegungen, die Gelenkpunkte nicht immer 1:1 auf den Punkten der nachfolgenden Bildern gemappt werden.

Hier kann DTW genutzt werden, um bei solchen ungünstigen Bedingungen ein Mapping zwischen den Gelenkpunkten zweier Frames zu erhalten. Der Algorithmus liefert hierbei das rechnerisch bestmögliche Mapping.