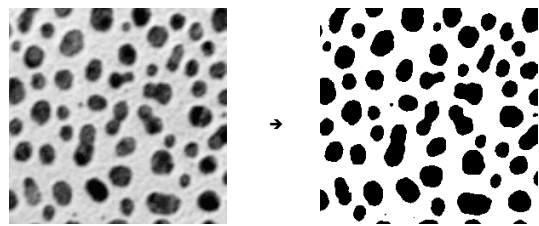


VC3 – Image Processing Binarisierung

Prof. Dr. Klaus Jung



Binarisierung



2 © Klaus Jung

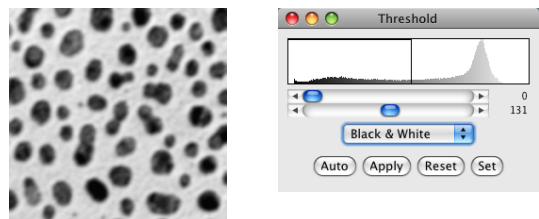
Quelle: Kai Uwe Barthel

Binarisierung

- Erfassung von Objekten oder Text z. B. für
 - Fax-Übertragung
 - OCR (Optical Character Recognition)
 - Segmentierung
 - Objekte zählen
- Problematik:
 - Unterschiedliche Kontrast- und Helligkeitsverhältnisse in Bildern und Dokumenten
- Schwellwertverfahren
 - fester Schwellwert
 - lokal adaptiver Schwellwert
 - Angepasst an die Eigenschaften der Umgebung des betrachteten Punktes

3 © Klaus Jung

Schwellwert (Treshold)



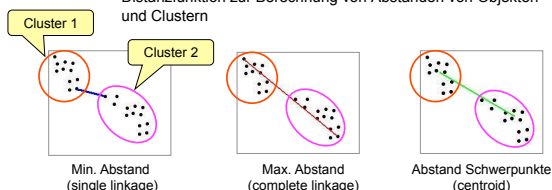
- Manuelle Einstellung eines Schwellwerts
 - Optimal: Histogramm hat zwei deutlich getrennte Peaks

4 © Klaus Jung

Quelle: Kai Uwe Barthel

Bestimmung des Schwellwerts

- Clustering Algorithmen (allgemein)
 - Objekte als Zufallsvariablen
 - Darstellung als Vektoren im Vektorraum der Eigenschaftsausprägungen (Features)
 - Finden von Clustern (Punktwolken)
 - Distanzfunktion zur Berechnung von Abständen von Objekten und Clustern



5 © Klaus Jung

Clustering Algorithmus

- k-means-Algorithmus
 1. Initialisierung: (zufällige) Auswahl von k Clusterzentren
 2. Zuordnung: Jedes Objekt wird dem ihm am nächsten liegenden Clusterzentrum zugeordnet
 3. Neuberechnung: Es werden für jeden Cluster die Clusterzentren neu berechnet als Schwerpunkt der Objekte
 4. Wiederholung: Falls sich nun die Zuordnung der Objekte ändert, weiter mit Schritt 2, ansonsten fertig
- Spezialfall: Iso-Data-Algorithmus

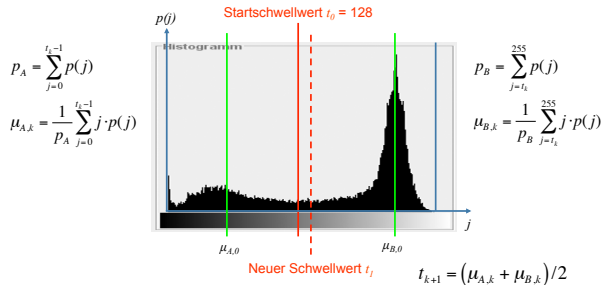
6 © Klaus Jung

Quelle: Wikipedia

1. Histogramm erstellen.
2. Startet bei irgendeinem Startwert (z.B. 128) = t_0
3. Histogramm wird geteilt unter (a) und über (b) dem Threshold
4. Schwerpunkt von a finden -> Mittelwert bilden im Bereich a & b
5. Rechne die Mitte zwischen diesen beiden Mittelwert-Punkten aus & wird t_1
Wiederhole bis t_{neu} gleich t_{alt} -> Aufpassen wegen Rundungen ->
Abbruchbedingung mit Abweichungsprüfung z.B. machen

Iso-Data-Algorithmus

Gebe t_0 vor und berechne t_1, t_2, t_3 usw. bis keine Änderung



7 © Klaus Jung

Weitere Verfahren

- Otsu-Thresholding
 - Ähnlich dem Iso-Data-Algorithmus, aber mit Varianz
- Max Entropy
 - Ähnlich dem Iso-Data-Algorithmus, aber mit Entropy
- Suche nach Modalwerten
 - Relative Maxima im Histogramm (uni-, bi-, multimodal)
- Mixture Modeling
 - Histogramm durch andere Verteilungen approximieren

8 © Klaus Jung

Otsu-Thresholding

$$p_A = \sum_{j=0}^{t-1} p(j) \quad p_B = \sum_{j=t}^{255} p(j)$$

$$\mu_A = \frac{1}{p_A} \sum_{j=0}^{t-1} j \cdot p(j) \quad \mu_B = \frac{1}{p_B} \sum_{j=t}^{255} j \cdot p(j)$$

$$\sigma_A^2 = \frac{1}{p_A} \sum_{j=0}^{t-1} (j - \mu_A)^2 \cdot p(j) \quad \sigma_B^2 = \frac{1}{p_B} \sum_{j=t}^{255} (j - \mu_B)^2 \cdot p(j)$$

- Varianz innerhalb von A und B

$$\sigma_t^2 = p_A \sigma_A^2 + p_B \sigma_B^2 \quad \text{Minimieren}$$

- Varianz zwischen A und B

$$\sigma_z^2 = p_A (\mu_A - \mu)^2 + p_B (\mu_B - \mu)^2 \quad \text{Maximieren}$$

$$\mu = \sum_{j=0}^{255} j \cdot p(j)$$

$$\sigma^2 = \sigma_t^2 + \sigma_z^2 = \text{const}$$

9 © Klaus Jung

Max Entropy Verfahren

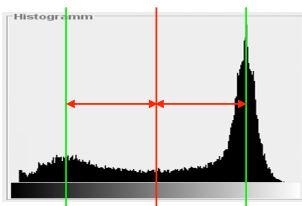
- Entropie in den Klassen A und B maximieren
 - Idee: Entropie H als Maß für den Informationsgehalt

$$H_A = - \sum_{j=0}^{t-1} p(j) \cdot \log_2 p(j) \quad H_B = - \sum_{j=t}^{255} p(j) \cdot \log_2 p(j)$$

10 © Klaus Jung

Suche nach Modalwerten

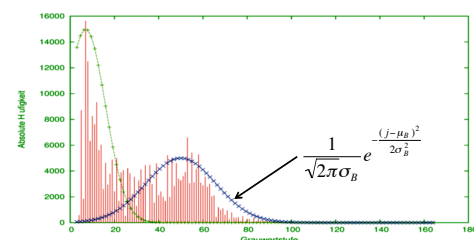
- Beispiel: Histogramm hat 2 relative Maxima
 - → bimodale Verteilung



11 © Klaus Jung

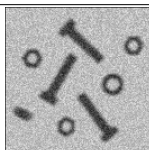
Mixture Modeling

- Verteilungsfunktion durch andere approximieren
 - Hier: Zwei Gauß'sche Normalverteilungen

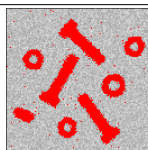


12 © Klaus Jung

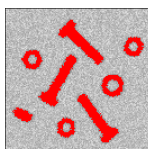
Beispiele



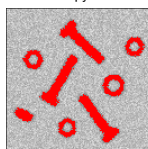
Iso Data



Max Entropy



Otsu



Mixture Modelling