Real-Time Image Processing

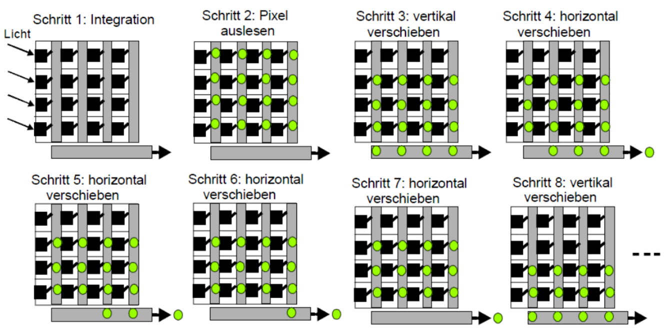
Zusammenfassung

Joel von Rotz / Quelldateien

Einführung

CCD Sensoren wird der photoelektrische Effekt genutzt. Dabei wird bei Wellenlänge 1100nm (IR) Elektronen-Lochpaare erzeugt. Die Anzahl erzeugten Elektronen entspricht der Intensität des einfallenden Lichtes und werden beim Lesen aus dem Sensor herausgeschoben $\rightarrow$ *Eimerkettenprinzip*

CMOS Sensoren Identisch zu CCD, ausser das jedes Pixel eine eigene Verstärker-Stufe besitzt. Dies erlaubt einen grösseren Dynamikbereich durch die Anpassung der individuellen Verstärker.



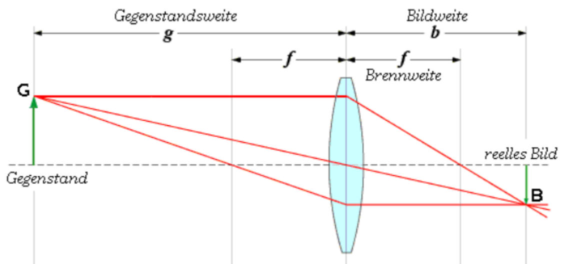
Quantisierung

Bild-Auflösung abhängig von Pixelanzahl des Sensors. Dynamikbereich/Grau**tiefe**-Auflösung abhängig von Grenzen der A/D Wandler.

Reduzieren der Bild-Auflösung reduziert den Dynamik-Bereich nicht, sondern schneidet die obere und untere Tiefengrenze ab.

Rasterung

$$\frac{1}{g} + \frac{1}{b} = \frac{1}{f} \qquad \frac{B}{G} = \frac{b}{g}$$



Globale Charakterisierung von Bildern

Mittelwerk **Varianz**

$$\mu_I = \frac{1}{M \cdot N} \sum_{m,n} I_{m,n} \qquad \sigma_I^2 = \frac{1}{M \cdot N} \sum_{m,n} (I_{m,n} - \mu_I)^2$$

Histogramm Das Histogramm gibt die absolute oder relative Häufigkeit $p_i(g)$ aller Grauwerte $g \in [0, 255]$ eines Bildes an:

$$0 \le p_I(g) \le 1 \quad \forall g \quad \rightarrow \sum_g p_I(g) = 1$$

Die kumulierte oder Summenhäufigkeit $h_I(g)$:

$$h_I(g) = \sum_{g' \le g} p_I(g')$$

Punkt Operationen & Bildverknüpfungen

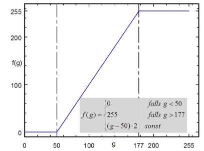
Look Up Tables LUTs werden verwendet, um Transformationen von diskreten Werten (Ganzzahlen) zu machen.

Lineare Grauwerttransformationen

Lassen sich anhand Notation schreiben $f : [0, 255] \rightarrow [0, 255] \subset \mathrm{IR}$

Spreizung Um ein Grauwert-Intervall (z.B. [50,166]) auf das gesamte Intervall [0,255] zu verteilen

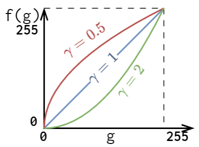
$$f(g) = \begin{cases} 0 & \text{falls } g < g_1 \\ 255 \cdot \frac{g - g_1}{g_2 - g_1} & \text{falls } g \in [g_1, g_2] \\ 255 & \text{falls } g > g_2 \end{cases}$$



Nichtlineare Operationen

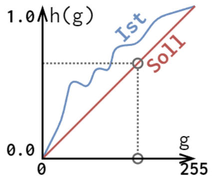
Gammakorrektur um menschliche Helligkeitswahrnehmung zu korrigieren.

$$f(g) = 255 \cdot \left[\frac{g}{255}\right]^\gamma$$



Histogrammausgleich Teilt die Wahrscheinlichkeit der Grauwerte gleichmässig auf

$$f(g) = g_{max} \cdot \sum_{g'=0}^{g} p_I(g')$$



Arithmetische & Logische Bildverknüpfungen

Differenzbildung

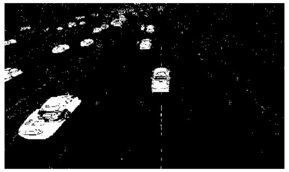
Für Bewegungsdetektion!

$$\begin{aligned} \Delta I_{k+n} &= 1/2 \cdot (255 + \overbrace{I_{k+n}}^{\text{neuer}} - \overbrace{I_k}^{\text{älter}}) \\ &= 1/2 \cdot (255 + I((k+n) \cdot \Delta t) - I(k \cdot \Delta t)) \end{aligned}$$

Schwellwert

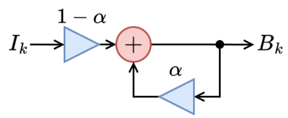
$$f(g) = \begin{cases} 255 & \text{falls } g < g_1 \vee g_2 < g \\ 0 & \text{falls } g \in [g_1, g_2] \end{cases}$$

$$\text{mit} \quad g_1 = 128 - \text{thres}, \quad g_2 = 128 + \text{thres}$$



Hintergrundschätzung

$$B_k = \alpha \cdot B_{k-1} + (1 - \alpha) \cdot I_k$$
$$\alpha \in (0, 1)$$

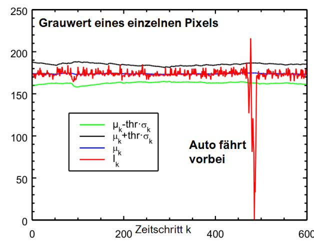


$$\rightarrow \Delta I_k = 1/2 \cdot (255 + B_k - I_k)$$

Statistische Analyse

Mittelwert μ_k & Varianz σ_k des Bildes

$$\begin{aligned} \mu_k &= \alpha \cdot \mu_{k-1} + (1 - \alpha) \cdot I_k \\ \sigma_k &= \alpha \cdot \sigma_{k-1} + (1 - \alpha) \cdots \\ & (\mu_k - I_k)^T \cdot (\mu_k - I_k) \end{aligned}$$



Filteroperatoren im Ortsraum

Lineare Filter

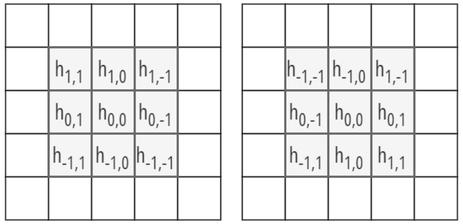
Faltung (convolution)

$$I \otimes h \quad : \quad I_{m,n} \to \sum_{p=-u}^{u} \sum_{q=-v}^{v} I_{m-p,n-q} \cdot h_{p,q}$$

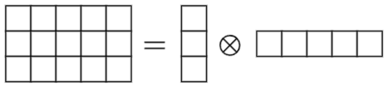
$$\begin{aligned} \text{Kommutativität}: &\quad I \otimes J = J \otimes I \\ \text{Assoziativität}: &\quad (I \otimes J) \otimes K = I \otimes (J \otimes K) \\ \text{Distributivität}: &\quad I \otimes (J + K) = I \otimes J + I \otimes K \\ \text{Asso. mit Faktor}: &\quad a \cdot (I \otimes J) = \begin{cases} (a \cdot I) \otimes J \\ I \otimes (a \cdot J) \end{cases} \end{aligned}$$

Damit die Indizierung verständlicher ist, wird anstatt Faltung die Korrelation (correlation) verwendet:

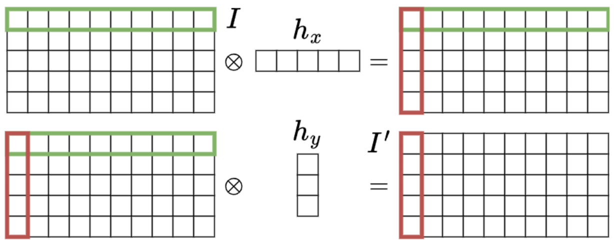
$$I \,\tilde{\otimes}\, h \quad : \quad I_{m,n} \to \sum_{p=-u}^{u} \sum_{q=-v}^{v} I_{m+p,n+q} \cdot h_{p,q}$$



Separierbare Masken können mit der Faltung zu einer gänzlichen Maske zusammengeführt werden.



Anhand der Assoziativität können zwei sukzessive Faltungen mit eindimensionalen Masken durchgeführt werden.



Tiefpass / Glätten

Mittelung eines Pixels mit den Nachbarpixeln ist die einfachste Rauschunterdrückung.

$$TP = \frac{1}{\sum \text{Mask}} \cdot \text{Mask}$$

$$\text{Rechteck} = \frac{1}{9} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \text{Gauss} = \frac{1}{16} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Hochpass / Kantenhervorhebung

Via Ableitungen $\partial I / \partial x$ & $\partial I / \partial y$ können Kanten hervorgehoben werden. Folgend ist eine rauschanfällige Version...

$$h_x = 1/2 \cdot \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \qquad h_y = 1/2 \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

... welche folgend mit Prewitt und Sobel Filter verbessert wurde.

Prewitt-Filter gleichmässige Pixelgewichtung

$$h_x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \otimes [-1\ 0\ 1] = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

$$h_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \otimes [1\ 1\ 1] = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Sobel-Filter gewichtet die Pixel, die sich näher an der Mitte der Maske befinden, mehr.

$$h_x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \otimes [-1\ 0\ 1] = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$h_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \otimes [1\ 2\ 1] = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Bildschärfung

β Schärfungsparameter

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

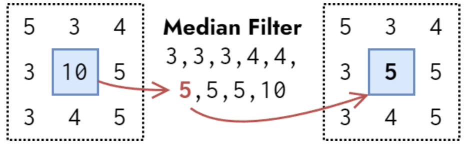
$$I = \beta \cdot L$$

Nichtlineare Filter

Rangordnungsoperatoren

Pixel im Frame werden der Grösse nach geordnet und dann wird der Pixelwert mit einer Strategie bestimmt.

Median Filter



Min/Max Filter Mit dem gleichen Prinzip, einfach werden entsprechend der maximale oder minimale Wert genommen. Zusätzlich kann das **lokale** Maxima und Minima ermittelt werden.

$$\text{LocMin} = (I == \min_{R \times R}(I))$$
$$\text{LocMax} = (I == \max_{R \times R}(I))$$

Morphologische Operationen

Dilation $\oplus$ Pixel setzen, wenn mindestens ein Pixel gesetzt ist. Ist Kommutativ $I \oplus J = J \oplus I$ und Assoziativ $(I \oplus J) \oplus K = I \oplus (J \oplus K)$.

$$I \oplus h = \{(m,n) \mid (\hat{h})_{m,n} \cap I \neq \{\}\}$$

> **!** **(Un)symmetrische Strukturelemente**
>
> $\hat{h}$ bedeutet (Punkt-)Spiegelung. Bei symmetrischen Strukturelementen ist dies belanglos. Bei Unsymmetrischen muss für die Dilation (& Schliessung) immer mit dem gespiegelten Strukturelement gearbeitet werden.

Erosion $\ominus$ Pixel setzen, wenn alle Pixel im Frame gesetzt sind. Ist weder kommutativ, noch assoziativ.

$$I \ominus h = \{(m,n) \mid (h)_{m,n} \subset I\}$$

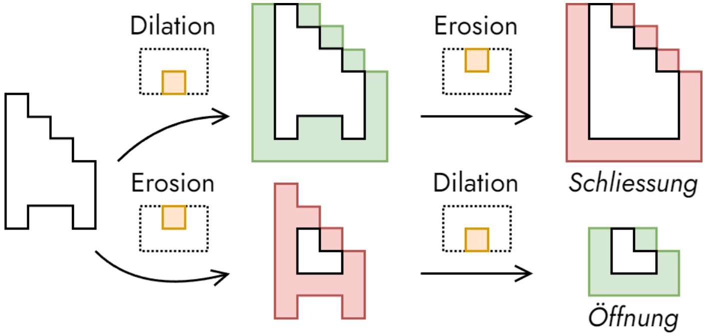
Kettenregel: $I \ominus (h \oplus g) = (I \ominus h) \ominus g$

Schliessung $\bullet$ Mehr vom Vordergrund zeigen. Ist idempotent $(I \bullet g) \bullet g = I \bullet g$

$$I \bullet h = (I \oplus h) \ominus h$$

Öffnung $\circ$ Mehr vom Hintergrund zeigen. Ist idempotent $(I \circ g) \circ g = I \circ g$

$$I \circ h = (I \ominus h) \oplus h$$



$$I - g \subseteq I \circ g \subseteq I \subseteq I \bullet g \subseteq I \oplus g$$

Hit- & Miss Operation

Mit Hit- & Miss Operationen wird geprüft, wo die Nachbarschaftsumgebung eines Pixels genau der Struktur der Einsen und Nullen im Struk-Element entspricht. Das Strukturelement besteht gleichzeitig aus der Hit-Maske (mit Einsen ; h_1) und Miss-Maske (mit Nullen ; h_0).

$$I \pm h = (I - h_1^i) \cap (I^C - h_0)$$

I^C=Komplement aller Pixel ungleich Null im Bild I. Strukturelemente werden solange angewendet, bis keine Operation gemacht werden kann und dies für alle i Strukturelemente!

$$\underset{h^1 =}{\begin{bmatrix} 0 & 0 & \\ 0 & \color{red}{1} & 1 \\ & 1 & \end{bmatrix}} \quad \underset{h^2 =}{\begin{bmatrix} 0 & 0 & 0 \\ & \color{red}{1} & \\ 1 & 1 & 1 \end{bmatrix}} \quad \underset{h^3 =}{\begin{bmatrix} & 1 & \\ 0 & \color{red}{1} & 1 \\ 0 & 0 & \end{bmatrix}} \quad \underset{h^4 =}{\begin{bmatrix} 0 & & 1 \\ 0 & \color{red}{1} & 1 \\ 0 & & 1 \end{bmatrix}} \dots$$

Fouriertransformation

Die 2D Fourier Transformation ist eine Faltung von zwei 1D Fourier Transformationen

! Crash Kurs

$\mathrm{rect}\left(\frac{t}{T}\right)$	$\lvert T\rvert \cdot \mathrm{si}(\pi T f)$	$\lvert T\rvert \cdot \mathrm{si}\left(\frac{T}{2}\,\omega\right)$
$\mathrm{si}\left(\pi \frac{t}{T}\right)$	$\lvert T\rvert \cdot \mathrm{rect}(T f)$	$\lvert T\rvert \cdot \mathrm{rect}\left(\frac{T}{2\pi}\,\omega\right)$
$\Lambda\left(\frac{t}{T}\right)$	$\lvert T\rvert \cdot \mathrm{si}^2(\pi T f)$	$\lvert T\rvert \cdot \mathrm{si}^2\left(\frac{T}{2}\,\omega\right)$
$\mathrm{si}^2\left(\pi \frac{t}{T}\right)$	$\lvert T\rvert \cdot \Lambda(T f)$	$\lvert T\rvert \cdot \Lambda\left(\frac{T}{2\pi}\,\omega\right)$

2D DFT $t \rightarrow f \Leftrightarrow f \rightarrow t$

$$\hat{h}_{lk} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} h_{mn} \cdot e^{-i\cdot 2\pi \cdot \left[\frac{l\cdot m}{M} + \frac{k \cdot n}{N}\right]}$$

$$\Longleftrightarrow$$

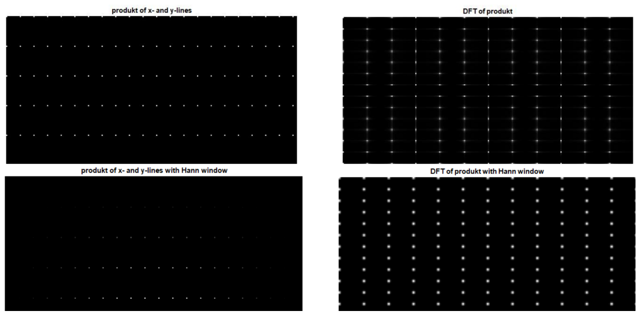
$$h_{mn} = \frac{1}{M\cdot N}\sum_{l=0}^{M-1}\sum_{k=0}^{N-1} \hat{h}_{lk} \cdot e^{i\cdot 2\pi \cdot \left[\frac{l\cdot m}{M} + \frac{k \cdot n}{N}\right]}$$

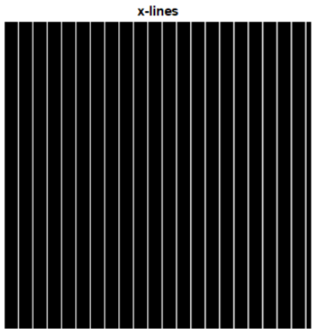
Parseval Theorem Wie in der 1D gilt auch das Parseval Theorem in 2D

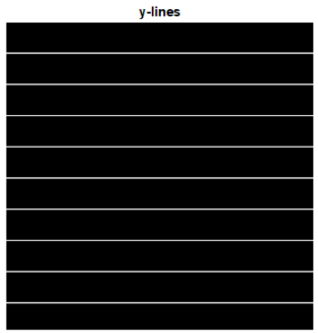
$$M\cdot N \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} h_{mn}^2 = \sum_{j=0}^{M-1}\sum_{l=0}^{N-1} \lvert \hat{h}_{lk}\rvert^2$$

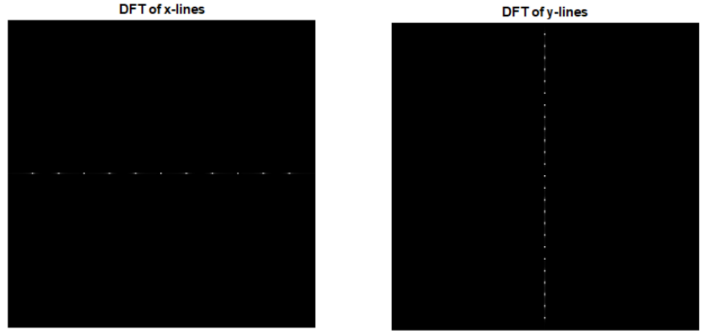
Da die DFT für periodische Signale ausgelegt ist, können Artefakte entstehen. Mit einem **Hann**-Filter alle Pixel um die Bildmitte schwächen (im Bild Pixelwert $\rightarrow 0$)

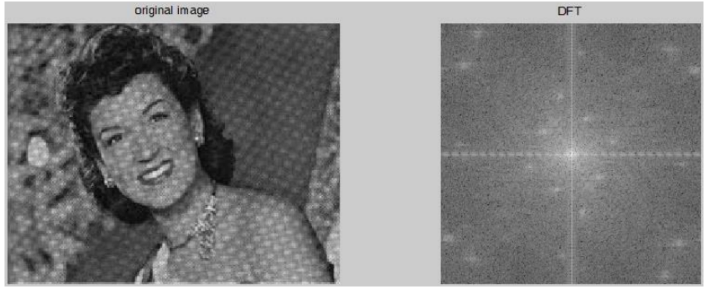
$$w_n = \frac{1}{2} \cdot \left[1 - \cos\left(\frac{2\pi \cdot n}{N-1}\right)\right]$$

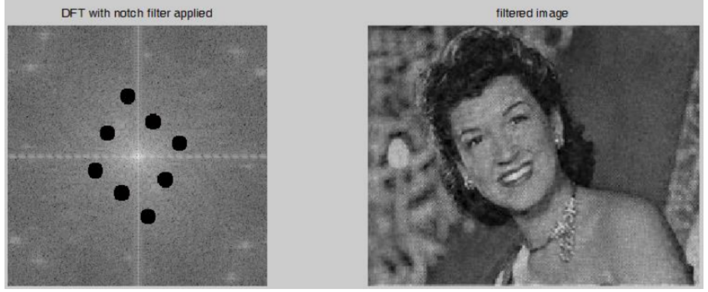












Segmentierung und Merkmalsextraktion

Otsu / Automatische Schwellwertbestimmung

Wahrscheinlichkeit ω_n, Mittelwert μ_n, Varianz σ_n^2

Klasse C_0

$$\omega_0 = \sum_{g=0}^{K} p_I(g) \qquad \mu_0 = \frac{1}{\omega_0}\sum_{g=0}^{K} p_I(g)\cdot g$$

$$\sigma_0^2 = \frac{1}{\omega_0}\sum_{g=0}^{K} p_I(g)\cdot (g-\mu_0)^2$$

Klasse C_1

$$\omega_1 = \sum_{g=K+1}^{255} p_I(g) \qquad \mu_1 = \frac{1}{\omega_1} \sum_{g=K+1}^{255} p_I(g) \cdot g$$

$$\sigma_1^2 = \frac{1}{\omega_1} \sum_{g=K+1}^{255} p_I(g) \cdot (g - \mu_1)^2$$

Bestimmung

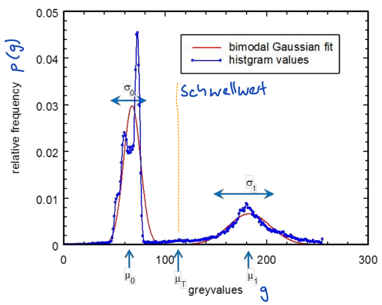
Intra-Klassen Varianz σ_W^2, *Inter-Klassen* Varianz σ_B^2, globaler Mittelwert μ_T

$$\sigma_W^2 = \omega_o \cdot \sigma_0^2 + \omega_1 \cdot \sigma_1^2 \quad K_{opt} = \underset{K \in [0,255]}{\arg\min} \{\sigma_W\}$$

$$\sigma_B^2 = \omega_o \cdot (\mu_0 - \mu_T)^2 + \omega_1 \cdot (\mu_1 - \mu_T)^2$$

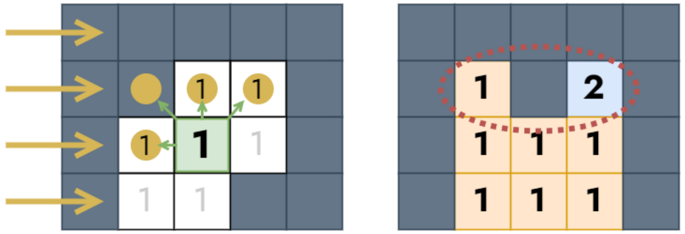
$$\mu_T = \sum_{g=0}^{255} p_I(g) \cdot g$$

$$\boxed{\begin{array}{l} K_{opt} = \underset{K \in [0,255]}{\arg\max} \{\sigma_B\} \\ \sigma_B^2 = \omega_0 \cdot \omega_1 \cdot (\mu_0 - \mu_1)^2 \end{array}}$$



Region Labeling

Bitmap Bild wird von einer Seite durchgelabelt $\rightarrow$ Wird ein Bit erkannt, werden die Nachbarspixel geprüft. Ist bereits einer der Nachbaren in einer Regions of Interest (ROI), wird dieses Label übernommen, ansonsten wird eine neue Region kreiert.



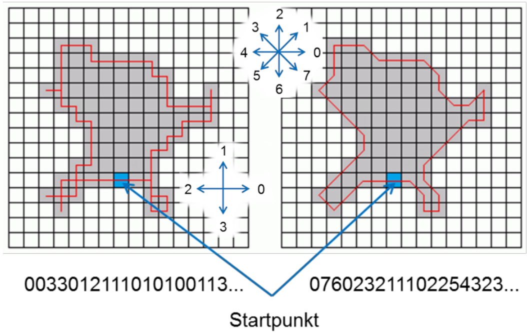
ROIs müssen dem Pixel-Wertebereich des Bildes entsprechen $\rightarrow$ 1-Byte Grauwertbild erlaubt 255 ROIs / 2-Byte erlaubt $2^{16} - 1$ ROIs.

Region Labeling hat grossen Speicherbedarf $\rightarrow$ Kettencodes

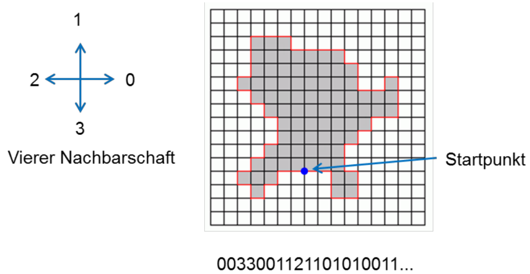
Kontour - Kettencodes

Es wird der **Startpunkt** gemerkt und die Richtungen (N-Bits $\rightarrow$ 2^N Richtungen)

Freeman Code Mittelpunkt vom Pixel



Eine "korrekte" Beschreibeng des Rands mit Crack Codes als Trennlinie zwischen Vorder- und Hintergrund.



Merkmalextrakiton

Fläche A

$$A = \sum_{I_{mn} \in \text{ROI}} 1 = \underbrace{-\oint_{\text{Rand}} y(x)\,dx}_{\text{mit Crack Codes}}$$

$$= \underbrace{\sum_{I_{mn} \in \{2-\text{Seg.}\}} m - \sum_{I_{mn} \in \{0-\text{Seg.}\}} m}_{\text{Vertikale Richtungen haben kein Einfluss}}$$

Massenmittelpunkt (x_s, y_s)

$$x_s = \frac{1}{A} \sum_{I_{mn} \in \text{ROI}} n \qquad y_s = \frac{1}{A} \sum_{I_{mn} \in \text{ROI}} m$$

Orientierung M

$$M = \begin{bmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{bmatrix}$$

$$M_{yy} = \frac{1}{A} \sum_{I_{mn} \in \text{ROI}} (m^2 - y_s^2) \qquad M_{xx} = \frac{1}{A} \sum_{I_{mn} \in \text{ROI}} (n^2 - x_s^2)$$

$$M_{xy} = M_{yx} = \frac{1}{A} \sum_{I_{mn} \in \text{ROI}} (m \cdot n - x_s \cdot y_s)$$

Bounding Box

ROI wird mit einem Rechteck umfasst.

Linien Segmentierung und Merkmalsextraktion

Kantendetektion

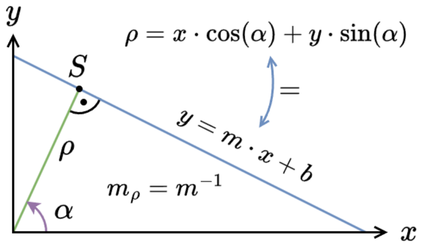
Problem: Kanten deutlich grösser als ein Pixel, globaler Schwellwert ergibt nicht in allen Bereichen des Bildes zufriedenstellende Kanten, zusammengehörige Kanten unterbrochen bei zu hohem Schwellwert.

Canny Kantendetektionsalgorithmus

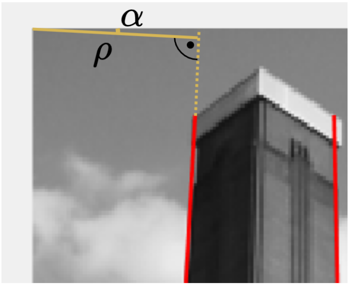
1 **Glättung** mittels Gaussfilter 2 **Kantenfilter** mittels Sobel oder Prewitt 3 **Bestimmung der lokalen Maxima** entlang der Richtung der Gradienten (Pixel selektiert, wenn |Gradient| grösser als Nachbarpixel) 4 **Kantenextraktion** anhand "oberen" Schwellwert (lok. Maxima über diesem wird es als Kante betrachtet) und sobald Pixel gefunden ist entlang der Kante mittels "unterem" Schwellwert zur Kante zugeordnet, durch Hysterese kann dynamisch auf Kontrast-schwankungen im Bild reagiert werden.

Linienextraktion mittels Hough-Transformation

Die Gradienten-Informationen (Betrag und Richtung) in neuen Raum transferieren und dort kann aufgrund von Ansammlungen auf die Formen im Originalbild geschlossen werden.

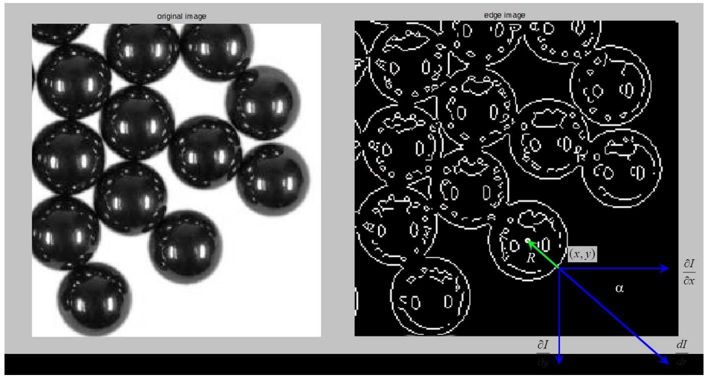


$$\begin{aligned} \alpha &\in [-90^\circ, 90^\circ] \\ \rho &\in [-D, D] \\ D &= \sqrt{\Delta x^2 + \Delta y^2} \end{aligned}$$



Detektion von Kreisen mittels Hough

$$x_c = x - R \cdot \frac{\partial I}{\partial x} \cdot \left[\frac{dI}{dr}\right]^{-1} \qquad y_c = y - R \cdot \frac{\partial I}{\partial y} \cdot \left[\frac{dI}{dr}\right]^{-1}$$



Farbe