

Computer Vision: Interessante Punkte

D. Schlesinger – TUD/INF/KI/IS

- Harris Detektor
- Maximally Stable Extremal Regions

Ein Detektor soll:

- eine relativ kleine Anzahl besonderer Punkte erzeugen
 - möglichst invariant sein zu
 - affinen Transformationen (Rotation, Skalierung etc.)
 - Diskretisierung
 - Änderung der Farbcharakteristika (Beleuchtung, Kontrast etc.)
-

Wie ähnlich ist das Bild $I(x, y)$ in einer Position (x, y) zu sich selbst, wenn es um $(\Delta x, \Delta y)$ verschoben wird?

Dies wird durch **Autokorrelation Funktion** beschrieben:

$$c(x, y, \Delta x, \Delta y) = \sum_{(u,v) \in W(x,y)} w(u, v) \left(I(u, v) - I(u + \Delta x, v + \Delta y) \right)^2$$

- $W(x, y)$ ist ein Fenster um (x, y)
- $w(u, v)$ ist der Faltungskern

z.B. eine Konstante oder Gaussian: $\exp\left[-\frac{(u-x)^2+(v-y)^2}{2\sigma^2}\right]$

Die Bildfunktion $I(u+\Delta x, v+\Delta y)$ wird durch lineare Terme der Taylor-Reihe approximiert:

$$\begin{aligned} I(u + \Delta x, v + \Delta y) &\approx I(u, v) + \frac{\partial I(u, v)}{\partial x} \Delta x + \frac{\partial I(u, v)}{\partial y} \Delta y \\ &= I(u, v) + [I_x(u, v), I_y(u, v)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \end{aligned}$$

mit partiellen Richtungsableitungen $I_x(u, v)$ und $I_y(u, v)$ an der Stelle (u, v) .
Die Autokorrelationsfunktion wird zu:

$$\begin{aligned} c(x, y, \Delta x, \Delta y) &= \sum_{(u,v) \in W(x,y)} \left(I(u, v) - I(u + \Delta x, v + \Delta y) \right)^2 \\ &\approx \sum_{(u,v) \in W(x,y)} \left([I_x(u, v), I_y(u, v)] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\ &= [\Delta x, \Delta y] Q(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \end{aligned}$$

mit

$$Q(x, y) = \begin{bmatrix} \sum_W I_x(x, y)^2 & \sum_W I_x(x, y) I_y(x, y) \\ \sum_W I_x(x, y) I_y(x, y) & \sum_W I_y(x, y)^2 \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

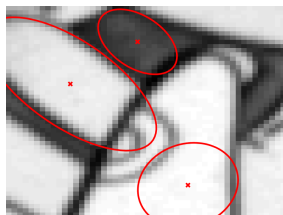
Die Autokorrelationsfunktion

$$c(x, y, \Delta x, \Delta y) = [\Delta x, \Delta y] Q(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

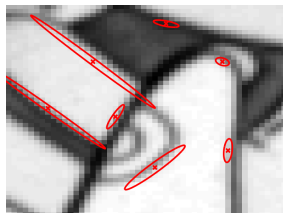
ist eine quadratische Funktion in Δx und Δy

- Isolinien sind Ellipsen, da symmetrisch positiv definit
- Eigenwerte $\lambda_1, \lambda_2 > 0$ definieren die Ausdehnung
- Eigenvektoren definieren die Rotation
(werden nicht gebraucht, denn rotationsinvariant)

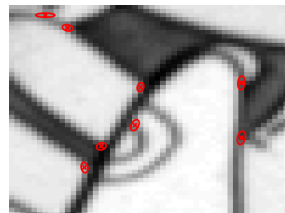
Ellipsen mit $c(x, y, \Delta x, \Delta y) = 1$:



Homogene Regionen
Beide λ -s sind klein



Kanten
Ein λ ist groß, ein ist klein

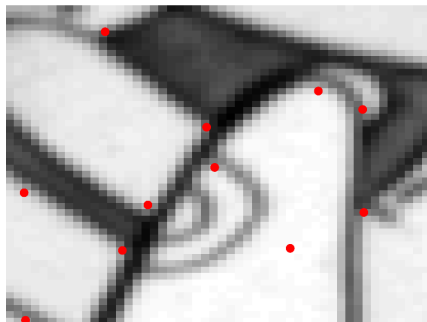


Ecken
Beide λ -s sind groß

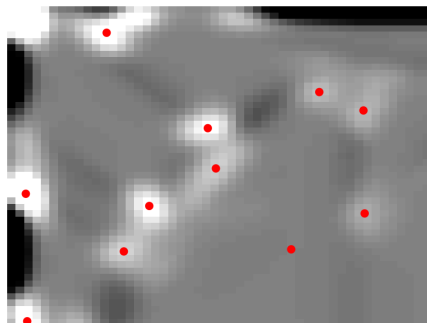
„Eckigkeit“ ist eine Charakteristik von $Q(x, y)$

$$\lambda_1 \lambda_2 = \det Q(x, y) = AC - B^2, \quad \lambda_1 + \lambda_2 = \text{trace} Q(x, y) = A + C$$

$$\text{Vorschlag von Harris: } H = \lambda_1 \lambda_2 - 0.04(\lambda_1 + \lambda_2)^2$$



Das Bild $I(x, y)$



Eckigkeit $H(x, y)$

Die besondere Punkte sind lokale Maxima der Eckigkeit



Suche nach den lokalen Maxima:

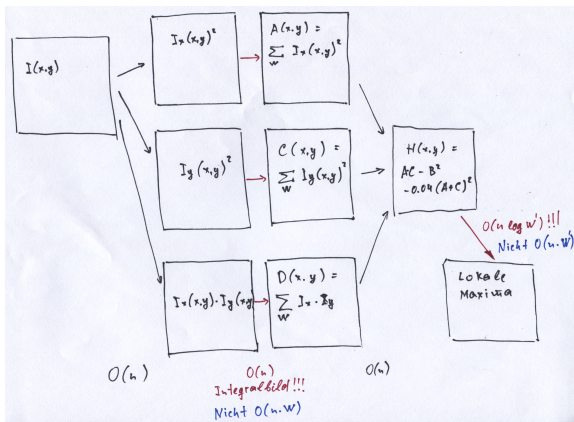
```
for ( alle Pixel  $p$  )
    current = computeH( $p$ );
    flag = true;
    for ( alle  $p' \in W'(p)$  )
        if ( computeH( $p'$ ) > current ) flag = false;
    if ( flag ) „ $p$  ist besonders“;
```

Berechnung der Eckigkeit:

```
computeH( $p$ )
    for ( alle  $p' \in W'(p)$  )
         $A+ = I_x(p)^2$ ;
         $C+ = I_y(p)^2$ ;
         $B+ = I_x(p)I_y(p)$ ;
    return  $AC - B^2 - 0.04(A + C)$ 
```

Zeitkomplexität: $O(n \cdot W \cdot W')$ – zu langsam :-(
(so darf man nicht programmieren!!!)

Harris Detektor, besserer Algorithmus



- Nichts mehrmals berechnen
- „Integralbild“-Verfahren für Summationen
- Spezielle Datenstrukturen für das Finden lokaler Maxima
⇒ $O(n \log W')$

Maximally Stable Extremal Regions (MSER)

Das Bild I ist eine Abbildung $I : D \subset \mathbb{Z}^2 \rightarrow S$.

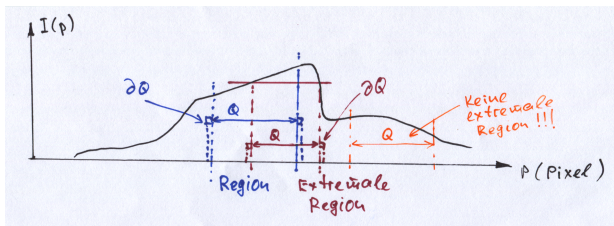
Die extreme Regionen sind definierbar, wenn:

1. S ist vollständig geordnet

(zum Beispiel die Menge der Grauwerte $\{0 \dots 255\}$ oder $S = \mathbb{R}$),

2. Eine Nachbarschaftsrelation $A \subset D \times D$ ist definiert

(zum Beispiel die 4-Nachbarschaft, d.h. $pAq \Leftrightarrow |p_x - q_x| + |p_y - q_y| \leq 1$)



Ein **Region** Q ist eine zusammenhängende Teilmenge aus D , d.h.

für jedes Paar $p, q \in Q$ gibt es eine Folge $p, a_1, a_2 \dots a_n, q$ so dass $pAa_1, a_iAa_{i+1}, a_nAq$.

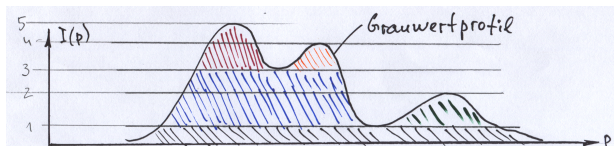
Die (äußere) **Grenze** ∂Q ist eine Teilmenge aus $D \setminus Q$, so dass

für alle Pixel $q \in \partial Q$ gibt es mindestens ein Pixel $p \in Q$ mit pAq .

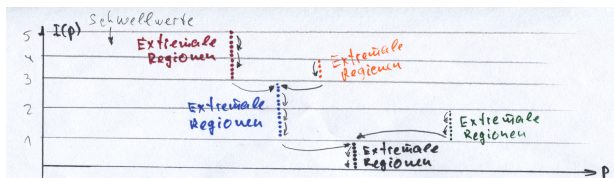
Ein Region $Q \subset D$ ist **extremal** wenn für alle Paare $p \in Q, q \in \partial Q$ $I(p) > I(q)$ gilt.

Maximally Stable Extremal Regions

Eine extremale Region ist eine Zusammenhangskomponente im binarisierten Bild.



Die Menge aller extremalen Regionen bildet eine Baumstruktur:

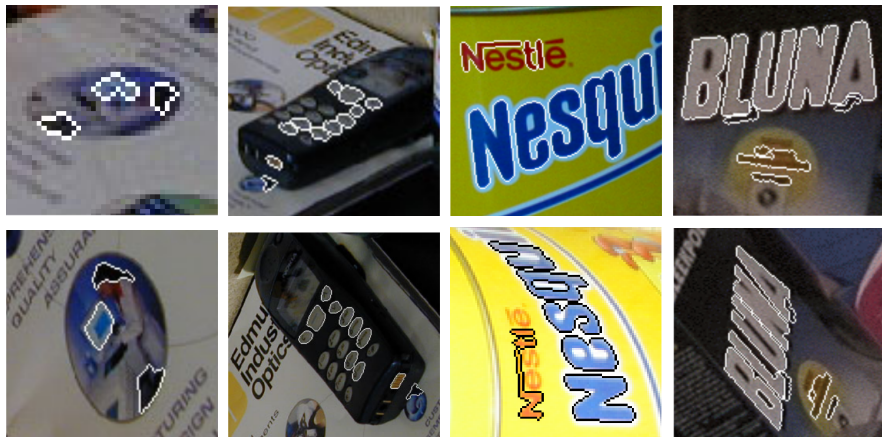


Sei $Q_1 \subset Q_2 \subset \dots \subset Q_i \dots$ eine „Folge“ der ineinander geschachtelten extremalen Regionen
Extremale Region Q_{i^*} ist **maximal stabil extremal** wenn

$$q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$$

einen lokalen Minimum in i^* hat ($|\cdot|$ ist die Kardinalität, $\Delta \in S$ ist ein Parameter).

Maximally Stable Extremal Regions



Eigenschaften:

- Invariant zu affinen Transformation der Grauwerte (Skalierung und Verschiebung).
- Covariant zu elastischen Transformationen des Definitionsbereiches.
- Sowohl kleine als auch Große Strukturen werden detektiert.

Naiver Algorithmus:

for (alle Schwellwerte)

 Binarisiere das Bild

 Erzeuge alle Zusammenhangskomponente

 Bilde den Baum

 Verfolge alle Pfade vom Wurzel zu Blätter, finde MSER-s

Zeitkomplexität – $O(|S|n)$.

Besserer Algorithmus:

1. Sortiere Pixel nach Grauwerten

2. Platziere die Pixel der Reihenfolge nach ins Bild,

 aktualisiere Zusammenhangskomponente und den Baum

 durch Vereinigung der Teilmengen.

1. – $O(n)$ mit BINSORT,

2. – $O(n \log n)$ mit union-find Algorithmus.

0.14 Sekunden auf Linux PC mit Athlon XP 1600+ für ein 530×350 Bild.

Andere Verfahren:

Kanten Detektoren (Sobel, Laplace usw.), Differenz von Gaussianen, Blob-Detektoren ...

Literatur:

- Chris Harris & Mike Stephens: A Combined Corner and Edge Detector (1988)
- J. Matas, O. Chum, M. Urban, T. Pajdla: Robust Wide Baseline Stereo from Maximally Stable Extremal Regions (BMVC 2002)
- K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir: A Comparison of Affine Region Detectors (IJCV 2006)