Superpixel

17. Januar 2017

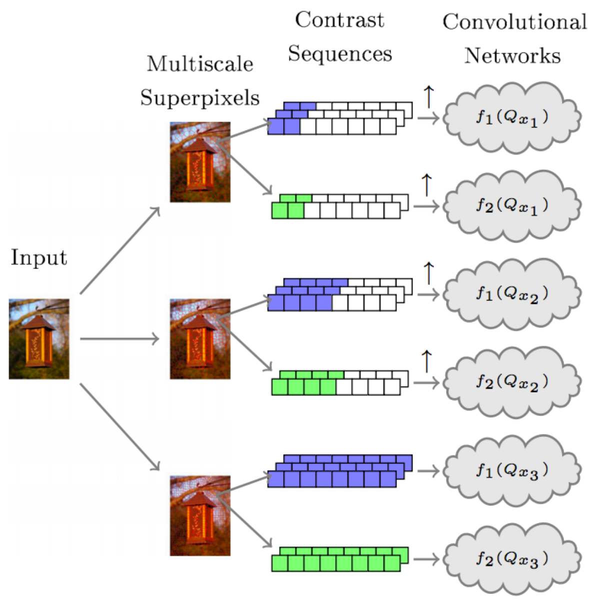
1 Einleitung

- *Superpxiel*: Menge von n Pixeln $S_i = \{t_1, \ldots, t_n\}$, wobei $t_i \in \{1, \ldots, N\}$ jeweils einen Pixel beschreibt und die Menge von S_i räumlich verbunden ist
- Menge von Superpixeln: $S = \{S_1, \ldots, S_m\}$, sodass $S_i \cap S_j = \emptyset$ für alle i, j und $\cup_i S_i = \cup_j t_j$
- Nachbarschaft: $(S_i, S_j) \in \mathcal{N}$, wenn S_i und S_j räumlich verbunden sind
- Vorteile:
  - Superpixel bieten eine Möglichkeit, die Größe des Problems zu minimieren
  - CNNs auf Bildern sind rauschend
  - große Netze auf Bildern mit Megapixeln rechnen langsam
- Nachteil: Superpixel haben einen bestimmten Fehlergrad
- $\Rightarrow$ finde den besten Ausgleich zwischen Größe und Fehlergrad

2 Lernen von Superpixeln

- **SuperCNN:** anstatt eines Bilders wird eine Sequenz von Superpixeln in das CNN gefüttert
- Problem: kontextbezogene Informationen gehen verloren (Methoden wie Superpixel Lattices adressieren dieses Problem, opfern aber Genauigkeit)
- $\Rightarrow$ zwei Kernel sollen Information wiederherstellen:
  1. *Spatial Kernel:* beschreibt Einzigartigkeit der Farben
  2. *Range Kernel:* beschreibt Farbverteilung
- zusätzlich: Multiscale Strukur des Netzes mit *Shared Weights*

- SuperCNN berechnet für individuelles Bild in etwa genauso lange wie klassische CNNs auf Bildern (0.45s)



- Vorteile:
  - benötigt weniger Trainingsdaten
  - Trainingsdaten werden generalisierter genutzt $\Rightarrow$ Netz fällt es leichter, für unbekannte Bilder Gemeinsamkeiten zu erkennen
  - gleiche oder bessere Performance

3 Features

3.1 Momente

Momente sind in der Bildverarbeitung bestimmte gewichtete Milttelwerte aus den Helligkeitswerten der einzelnen Pixel eines Bildes. Sie werden gewoehnlich so gewaehlt, dass sie gewuenschte Eigenschaften des Bildes widerspiegeln oder gewisse geometrische Interpretationen besitzen. Momente sind hilfreich, um einzelne Objekkte in einem **segmentierten** Bild zu beschreiben.

3.1.1 Nicht-zentrierte Momente

$$M_{ij} = \sum_x \sum_y x^i y^j g(x, y) \tag{1}$$

- **Flaeche:** M_{00}
- **Schwerpunkt:** $\frac{M_{10}}{M_{00}}$ und $\frac{M_{01}}{M_{00}}$

3.1.2 Zentrale Momente

Zentrale Momente sind invariant bezueglich Translationen.

$$\mu_{ij} = \sum_x \sum_y (x - \overline{x})^i (y - \overline{y})^j g(x,y) \tag{2}$$

wobei $\overline{x}$, $\overline{y}$ Schwerpunkt.

- $\mu_{00} = M_{00}$

Informationen ueber die Ausrichtung des Bildes koennen gewonnen werden, indem man zuerst die drei zentralen Momente zweiten Grades verwendet, um eine Kovarianzmatrix zu berechnen

$$\text{cov}[I(x,y)] = \begin{pmatrix} \frac{\mu_{20}}{\mu_{00}} & \frac{\mu_{11}}{\mu_{00}} \\ \frac{\mu_{11}}{\mu_{00}} & \frac{\mu_{02}}{\mu_{00}} \end{pmatrix} = \begin{pmatrix} \mu'_{20} & \mu'_{11} \\ \mu'_{11} & \mu'_{02} \end{pmatrix} \tag{3}$$

Die Eigenvektoren dieser Matrix entsprechen der grossen und kleinen Halbachse der Helligkeitswerte. Die Eigenwerte der Kovarianzmatrix sind

$$\lambda_i = \frac{\mu'_{20} + \mu'_{02}}{2} \pm \frac{\sqrt{4\mu'^2_{11} + (\mu'_{20} - \mu'_{02})^2}}{2} \tag{4}$$

Die *Exzentrizitaet* (engl. *Eccentricity*) des Bildes ist $\sqrt{1 - \frac{\lambda_2}{\lambda_1}}$. Sie ist ein Mass fuer das Groessenverhaeltnis der beiden Hauptachsen. Bei runden Objekten ist sie nahe an 0, bei laenglichen Objekten nahe an 1.

3.1.3 Skalierungsinvariante Momente

Es koennen Momente η_{ij} mit $i+j \geq 2$ konstruiert werden, die invariant bezueglich Skalierung und Translation sind, indem man das entsprechende zentrale Moment durch das entsprechend skalierte Moment vom Grad 0 teilt.

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{1+\frac{i+j}{2}}} \tag{5}$$

3.1.4 Rotationsinvariante Momente

Es ist weiterhin moeglich, Momente zu konstruieren, die zusaetzlich invariant bezueglich einer Bildrotation sind. Haeufig benutzt wird die *Hu*-Menge invarianter Momente.

- $I_1 = \eta_{20} + \eta_{02}$: Traegheitsmoment um den Schwerpunkt des Bildes, wenn die Helligkeitswerte der Pixel als physikalische Dichte interpretiert werden

3.2 Polygon-Features

- **Area:** Anzahl an Pixeln im Segment
- **Bounding Box Area:** Flaeche der vertikalen/horizontalen gelegenen Bounding Box um das Segment
- **Convex Area:** Anzahl an Pixeln der konvexen Huelle des Segments

- **Local Centroid:** Zentrum/Schwerpunkt des Segments relativ zur Bounding Box
- **Eccentricity:** Groessenverhaeltnis der beiden Hauptachsen zueinander $\in [0, 1]$
- **Major Axis Length:** $4\sqrt{\lambda_1}$
- **Minor Axis Length:** $4\sqrt{\lambda_2}$

Daraus koennen weitere Features ermittelt werden:

- **Extent:** $\frac{\text{Area}}{\text{Bounding Box Area}}$
- **Solidity:** $\frac{\text{Area}}{\text{Convex Area}}$