

Segmentácia

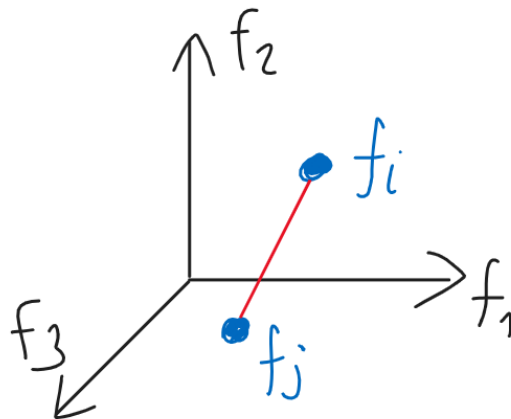
Contents

- GrabCut
- Zdroje

GrabCut

Metóda segmentácie **GrabCut** je založená na grafoch. Cieľom je rozdeliť pixely na dve disjunktné množiny. Vo väčšine prípadov to môžeme interpretovať ako oddelenie objektu od pozadia.

Každý pixel má priradený svoj **feature vector** (vektor príznakov). Tento vektor je n-rozmerný zoznam čísel. Napr. pixel i môže mať priradený 5-rozmerný vektor pozostávajúci z troch farebných zložiek a súradnice pixela (R, G, B, x, y) . Potom konkrétny pixel môže mať priradený vektor $\mathbf{f}_i = (255, 0, 120, 30, 40)$, pričom farba pixela je $(255, 0, 120)$ a jeho súradnica je $[30, 40]$. Vektor príznakov môže byť tvorený ľubovoľnými číselnými charakteristikami, napr. gradient v danom pixeli, vzdialenosť k najbližšej hrane a pod.



Vektor príznakov.

Rozdielnosť pixelov (pixel dissimilarity) je definovaná ako euklidovská vzdialenosť vektorov, ktoré k nim prislúchajú:

$$S(\mathbf{f}_i, \mathbf{f}_j) = \sqrt{\sum_k (f_{ik} - f_{jk})^2}$$

Afinita pixelov i, j (ich príbuznosť) je následne vyjadrená ako:

$$w(i, j) = A(\mathbf{f}_i, \mathbf{f}_j) = e^{\frac{-1}{2\sigma^2} S(\mathbf{f}_i, \mathbf{f}_j)}$$

Parameter σ nastavuje afinitu pixelov. Čím je menšia hodnota S , tým je afinita pixelov A väčšia.

Obrázok ako graf

Uvažujme ohodnotený graf $G = (V, E)$ s ohodnotením hrán $w : E \rightarrow \mathbb{R}$, kde:

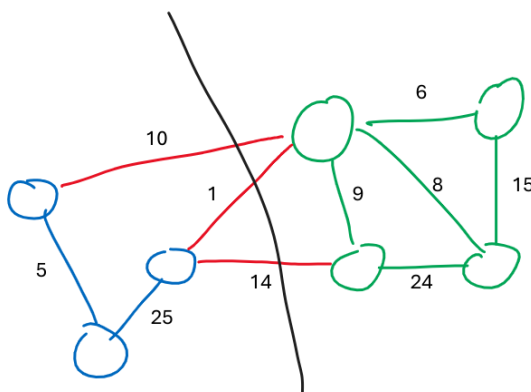
- množina vrcholov V : každý pixel v obrázku je vrcholom v grafe.
- množina hrán E : hrana existuje medzi ľubovoľnou dvojicou pixelov (vrcholov). V praxi je vhodné obmedziť počet hrán, a teda každý vrchol je spojený hranou iba s vrcholmi vo svojom okolí.
- funkcia ohodnotenia hrany w : váha hrany $w(i, j) = A(\mathbf{f}_i, \mathbf{f}_j)$ je definovaná ako afinita pixelov (vrcholov).

Rez grafu

Rez grafu $C = (V_A, V_B)$ je rozdelenie množiny vrcholov grafu V na dve disjunktné množiny V_A a V_B .

Cena rezu grafu je označená ako suma ohodnotení hrán, ktoré budú *prerezané*:

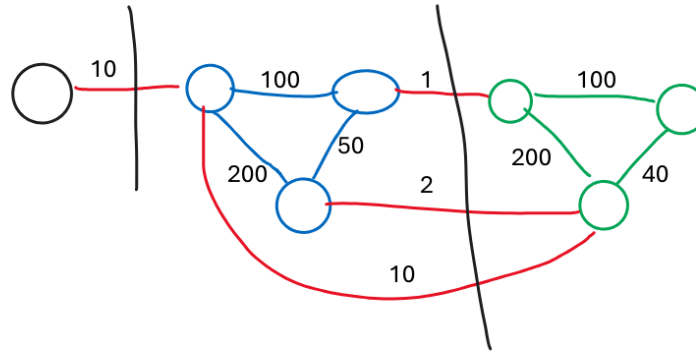
$$cut(V_A, V_B) = \sum_{u \in V_A, v \in V_B} w(u, v)$$



Rez grafu znázornený čiernou farbou. Cena rezu je súčet ohodnotení červených hrán $10+1+14=25$.

Min-cut: Úlohou je minimalizovať cenu rezu grafu. V ideálnom prípade budú jednotlivé podmnožiny vrcholov tvorené pixelmi, ktoré sú navzájom podobné (majú veľkú hodnotu afinity) a graf bude prerezaný na hranách s nízkou hodnotou afinity.

Avšak tento prístup má tendenciu preferovať odseparovanie malých segmentov.



Vpravo je znázornený žiadaný rez grafu, ktorý rozdelí množinu vrcholov na dve podobne veľké podmnožiny. Cena tohto rezu je 13. Avšak preferovaný bude rez, ktorý oddelí jeden vrchol (rez vľavo s cenou 10).

Min-Cut Max-Flow

Existuje veta, ktorá tvrdí, že cena minimálneho rezu je rovná maximálnemu toku v sieti (Min-Cut Max-Flow theorem). Na nájdenie maximálneho toku v sieti slúži napr. Ford-Fulkerson metóda, ktorá je implementovaná Edmonds–Karp algoritmom.

Sieť N je definovaná ako usporiadaná štvorica $N = (D, c, s, t)$, kde D je orientovaný graf, $c : E(D) \rightarrow \mathbb{N}$ je kapacitná funkcia, s je počiatkový a t cieľový vrchol.

Tok v sieti N je funkcia $f : E(D) \rightarrow \mathbb{N}$, pre ktorú platí:

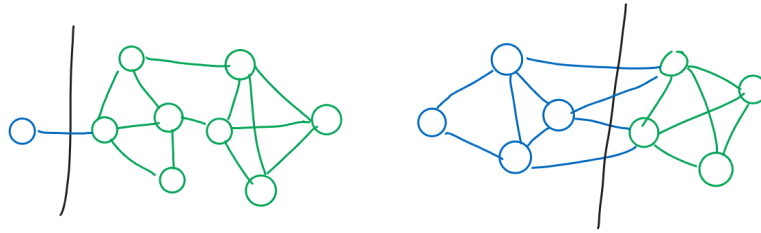
1. kapacitné obmedzenie: $\forall e \in E(D) : 0 \leq f(e) \leq c(e)$
2. zachovanie toku: $\forall v \in V(D) \setminus \{s, t\} : \sum_{y \in N^+(x)} f(x, y) = \sum_{y \in N^-(x)} f(y, x)$

Normalizovaný rez grafu

Normalizovaný rez je aplikovaný, aby sa preferovali podgrafy s väčším počtom vrcholov.

Hodnota $assoc(V_A, V)$ vyjadruje ako silno sú vrcholy z podmnožiny V_A prepojené s vrcholmi V .

$$assoc(V_A, V) = \sum_{u \in V_A, v \in V} w(u, v)$$



Modrou farbou sú zobrazené hrany, ktoré budú zarátané do hodnoty *assoc*. Rez vľavo má túto hodnotu nízku v porovnaní s rezom vpravo (konkrétna hodnota závisí od ohodnotenia hrán).

Cena normalizovaného rezu grafu je vyjadrená nasledovne:

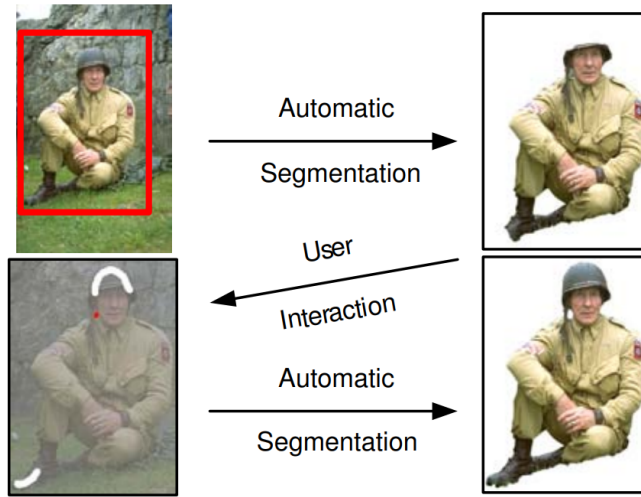
$$NCut(V_A, V_B) = \frac{cut(V_A, V_B)}{assoc(V_A, V)} + \frac{cut(V_A, V_B)}{assoc(V_B, V)}$$

Minimalizácia *NCut* je NP úplný problém. Avšak existuje rýchla aproximácia založená na vlastných vektoroch [Shi, 2000].

GrabCut

GrabCut (implementovaná v OpenCV) [Rother, 2004] je interaktívna metóda založená na iteratívnom rezaní grafu.

- **Inicializácia ROI** - obdĺžnikom sa manuálne označí oblasť, kde sa nachádza objekt. Každý pixel môže byť v jednom z troch stavov (pozadie, objekt alebo neznámy). Všetky pixely mimo obdĺžníka sú pozadie, vo vnútri obdĺžníka sú neznáme.
- Metóda urobí automatickú segmentáciu na základe rezu grafu. V OpenCV je výsledná maska pozostávajúca zo 4 hodnôt (0 - pozadie, 1 - objekt, 2 - pravdepodobne pozadie, 3 - pravdepodobne objekt).
- **Interakcia používateľa** - na príslušných miestach, ktoré neboli korektné segmentované, používateľ označí časť pixelov, ktoré patria pozadiu alebo objektu (nastaví hodnoty 0 alebo 1 niektorým pixelom).
- Metóda vykoná rez grafu na základe aktualizovanej masky.



Metóda GrabCut.

Zdroje

- [First Principles of Computer Vision](#), Perception | Image Segmentation
- [GrabCut OpenCV](#)
- [Shi, 2000] Shi, J. and Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8), pp.888-905.
- [Rother, 2004] Rother, C., Kolmogorov, V. and Blake, A., 2004. "GrabCut" interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, 23(3), pp.309-314.