

Wizja maszynowa

Segmentacja obrazu

wykład 14

Adam Szmigielski

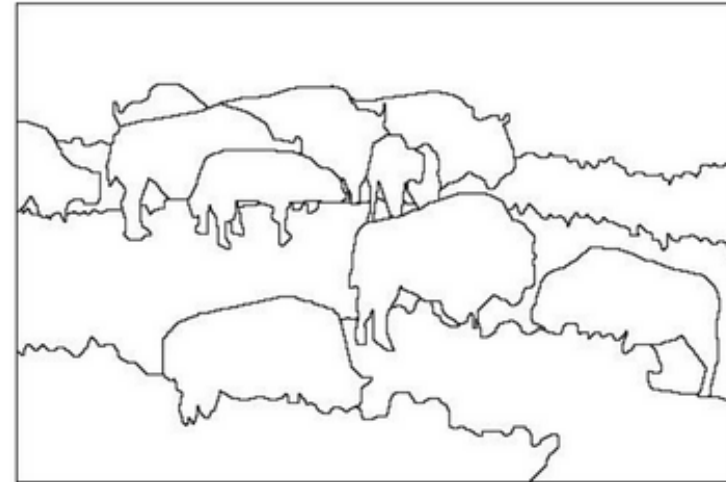
aszmigie@pjawst.edu.pl

materiały: *ftp(public) : //aszmigie/WM*

Segmentacja obrazu

- Segmentacja obrazu to zadanie znalezienia grup pikseli, które “pasują do siebie”.
- W statystykach ten problem jest znany jako *analiza skupień* i jest szeroko badanym obszarem z setkami różnych algorytmów

Segmentacja obrazu

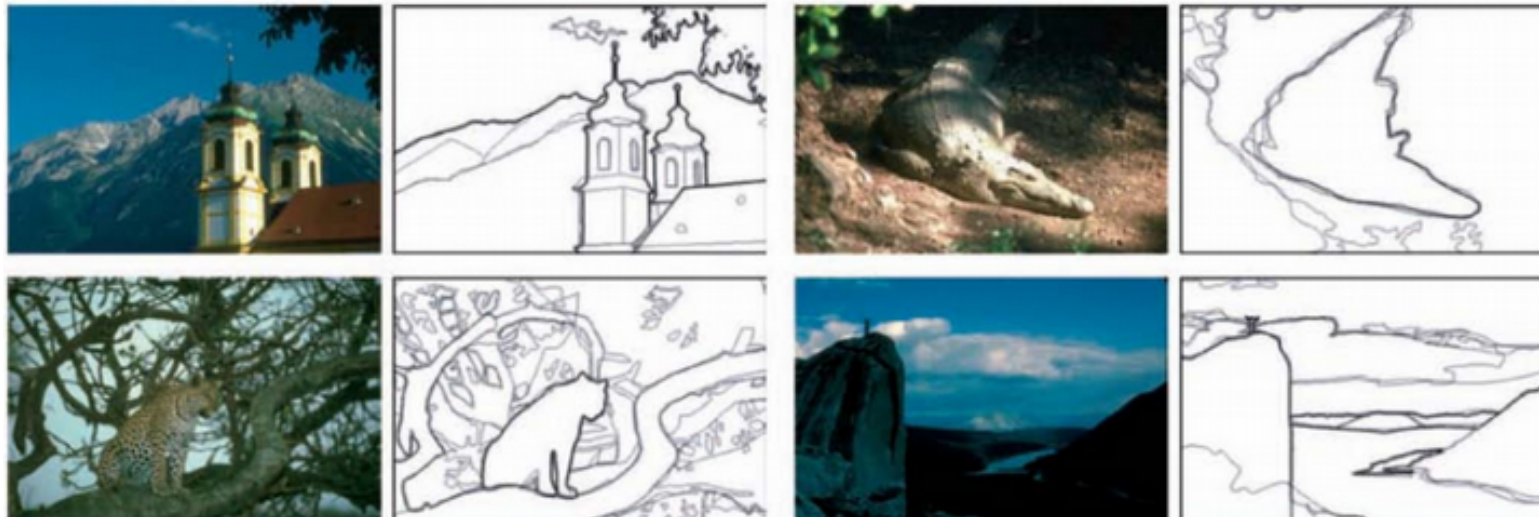


- Metody segmentacji obrazu będą szukać obiektów, które albo mają w sobie pewną miarę jednorodności, albo mają pewną miarę kontrastu z obiektami na ich granicy,
- Miary jednorodności i kontrastu mogą obejmować takie cechy, jak poziom szarości, kolor i tekstura.

Wykrywanie krawędzi: pierwszy krok do segmentacji obrazu

- Celem segmentacji obrazu jest znalezienie regionów, które reprezentują obiekty lub znaczące części obiektów
- Podział obrazu na obszary odpowiadające interesującym obiektom jest niezbędny do interpretacji i zrozumienia sceny
- Identyfikacja rzeczywistych obiektów, pseudo-obiektów, cieni lub faktyczne znalezienie czegokolwiek interesującego na obrazie wymaga jakiejś formy segmentacji

Edge detection



We can define an edge as a location of *rapid intensity variation*.

- A mathematical way to define the slope and direction of a surface is through its gradient,

$$J(x) = \nabla I(x) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)(x).$$

- The local gradient vector J points in the direction of steepest ascent in the intensity function,

- Its magnitude is an indication of the slope or strength of the variation.

Edges calculation

The gradient of the smoothed image can therefore be written as

$$J_{\sigma}(x) = \nabla[G_{\sigma}(x) \star I(x)] = [\nabla G_{\sigma}](x) \star I(x)$$

- We can convolve the image with the horizontal and vertical derivatives of the Gaussian kernel function,

$$\nabla G_{\sigma}(x) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)(x) = [-x - y] \frac{1}{\sigma^3} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

The parameter σ indicates the width of the Gaussian

- The desired directional derivative is equivalent to second gradient $J_{\sigma}(x)$:

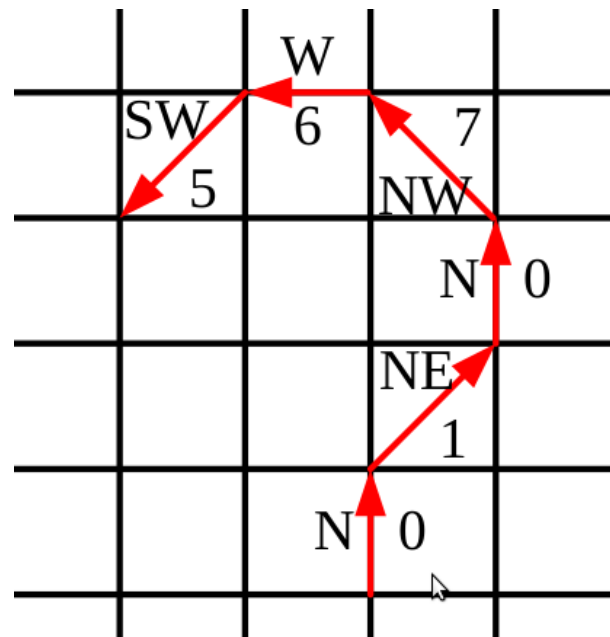
$$S\sigma(x) = J_{\sigma}(x) = [\nabla^2 G_{\sigma}](x) \star I(x)$$

The gradient operator dot product with the gradient is called the *Laplacian of Gaussian* (LoG)

- *The Laplacian of Gaussian* can be replaced with *Difference of*

Gaussian (DoG).

Edge linking



- Linking the edgels into chains involves picking up an unlinked edgel and following its neighbors in both directions.
- More compactly - A chain code encodes a list of connected points lying on an $N8$ grid using a three-bit code corresponding to the eight cardinal directions (N, NE, E, SE, S, SW, W, NW)
- Once the edgels have been linked into chains, we can apply an optional thresholding with hysteresis to remove low-strength contour segments (Canny).

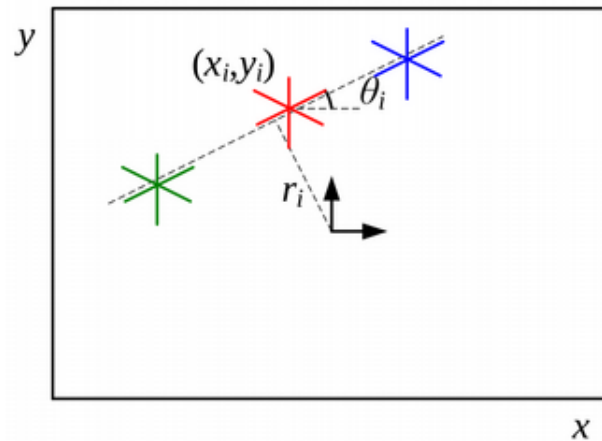
Edge smoothing

- A more useful representation is the *arc length parameterization* of a contour, $x(s)$, where s denotes the arc length along a curve,
- Arc-length parameterization can also be used to smooth curves in order to remove digitization noise,
- An alternative approach, based on selectively modifying different frequencies in a wavelet decomposition,
- Curve smoothing with a Gaussian kernel.

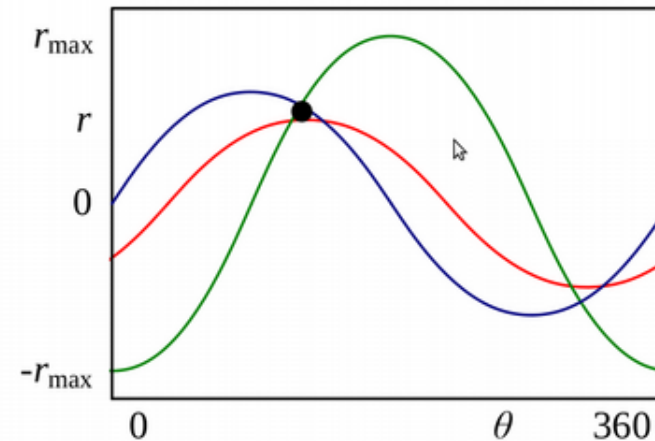
Lines

- Detecting and matching these lines can be useful in a variety of applications (pose estimation, in indoor environments, road lines, buildings etc.).
- In many applications it is preferable to approximate curve with a simpler representation (e.g., as a piecewise-linear polyline)
- If a smoother representation or visualization is desired, either approximating or interpolating splines or curves can be used.

Hough transforms



(a)



(b)

- (a) Each point votes for a complete family of potential lines
 $r_i(\theta) = x_i \cos \theta + y_i \sin \theta$
- (b) Each pencil of lines sweeps out a sinusoid in (r, θ) ; their intersection provides the desired line equation
- Each edge point votes for all possible lines passing through it, and lines corresponding to high accumulator are examined for potential line fits.

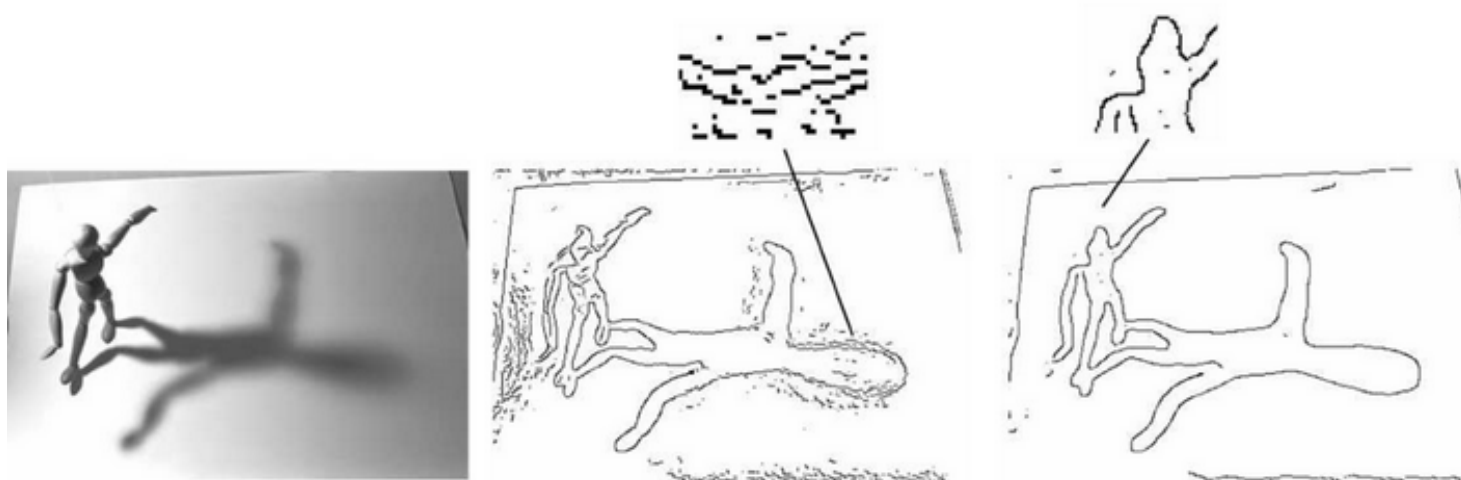
Wykrywanie krawędzi

- **Cel:** Zidentyfikuj nagłe zmiany (nieciągłości) na obrazie
 - Intuicyjnie większość informacji semantycznych i informacji o kształcie z obrazu można zakodować na krawędziach,
 - Bardziej kompaktowy niż piksele
- **Idealny:** rysowanie linii przez artystę (artysta korzysta również z wiedzy na poziomie przedmiotu)

Metody wykrywania krawędzi

- Operatory gradientowe
 - Roberts
 - Prewitt
 - Sobel
- Gradient Gaussowski (Canny)
- Laplasjan z Gaussa (Marr-Hildreth)
- Detektor krawędzi oparty na modelu (Haralick)

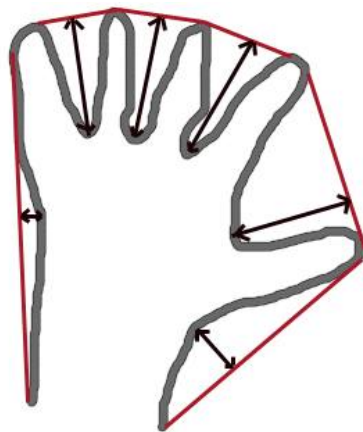
Co to są kontury?



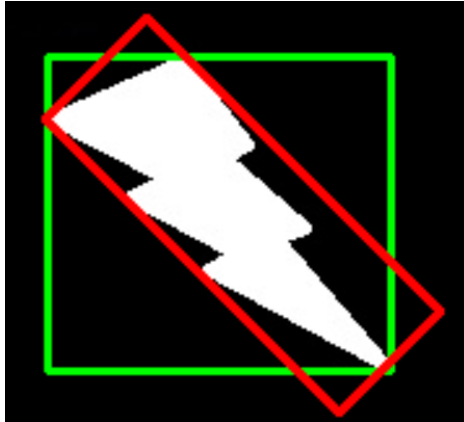
- Kontury można po prostu wyjaśnić jako krzywą łączącą wszystkie ciągłe punkty (wzdłuż granicy), mającą ten sam kolor lub intensywność.
- Kontury są użytecznym narzędziem do analizy kształtów oraz wykrywania i rozpoznawania obiektów.

Cechy konturu - opencv

- Momenty
- powierzchnia konturu
- Obwód konturu - zwany także długością łuku.
- Aproksymacja konturu - aproksymuje kształt konturu do innego, prostszego kształtu
- Wypukła pokrywa - wygląda podobnie do przybliżenia konturu,



- **Prostokąt ograniczający** - *prosty prostokąt ograniczający i prostokąt obrócony*



- **Minimalny krąg otaczający**



- Dopasowywanie elipsy



- Dopasowanie linii



Właściwości konturu

- **Aspect Ratio** - stosunek szerokości do wysokości prostokąta ograniczającego

$$AspectRatio = \frac{Width}{Height}$$

- **Extent** stosunek powierzchni konturu do ograniczającego obszaru prostokąta.

$$Extent = \frac{ObjectArea}{BoundingRectangleArea}$$

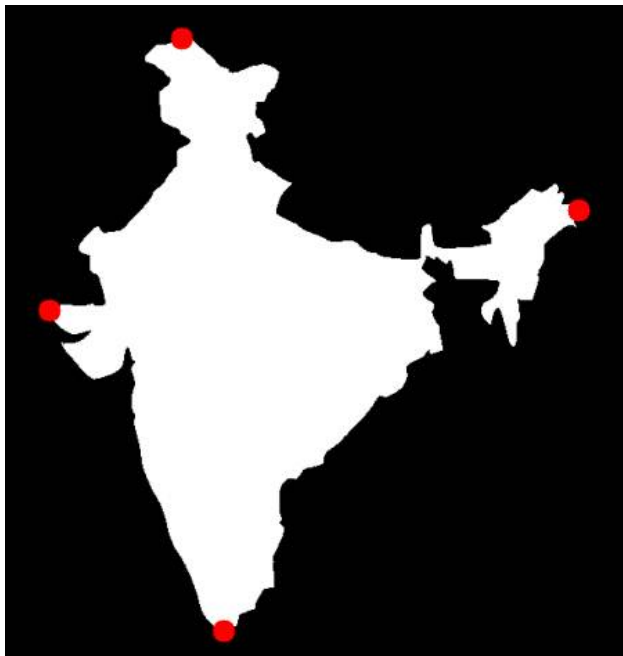
- **Solidity** stosunek powierzchni konturu do wypukłej powierzchni pokrywy.

$$Solidity = \frac{ContourArea}{ConvexHullArea}$$

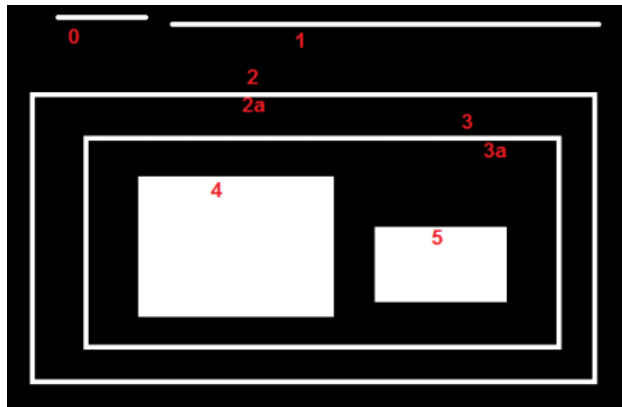
- **Equivalent Diameter** - średnica okręgu o powierzchni równej powierzchni konturu:

$$EquivalentDiameter = \sqrt{\frac{4 \times ContourArea}{\pi}}$$

- **Orientation** to kąt, pod jakim skierowany jest obiekt,
- Wartość maksymalna, wartość minimalna i ich lokalizacje,
- Średni kolor lub Średnia intensywność,
- **Extreme Points** oznacza najwyższy, najniższy, najbardziej prawy i lewy punkt obiektu.



Hierarchia konturów - opencv



- Kiedy niektóre kształty znajdują się wewnątrz innych kształtów (zewnątrzny jako rodzic i wewnętrzny jako dziecko), wówczas kontury obrazu mają ze sobą pewien związek - ta relacja nazywana jest **hierarchią**.
- Kontury 0, 1, 2 są zewnętrzne lub najbardziej zewnętrzne. Contour-2a można uznać za dziecko contour-2. Contour-3 jest dzieckiem konturu-2 i znajduje się w kolejnej hierarchii. Kontury 4,5 są elementami potomnymi konturu-3a i znajdują się na ostatnim poziomie hierarchii.

Najpopularniejsze techniki segmentacji

- *active contours*,
- *level sets*,
- *region splitting and merging*,
- *mean shift* (mode finding),
- *normalized cuts* - podział na podstawie wskaźników podobieństwa pikseli,
- *binary Markov random fields* - rozwiązano za pomocą cięć graficznych.

Aktywne kontury

- **Snakes** to minimalizująca energię, dwuwymiarowa krzywa splajnu, która ewoluuje (przesuwa się) w kierunku cech obrazu, takich jak mocne krawędzie,
- **Intelligent scissors** pozwalają użytkownikowi szkicować w czasie rzeczywistym krzywą, do której przylega granice obiektów,
- **Level set techniques** ewoluują krzywą jako zerowy zbiór funkcji charakterystycznej, co pozwala im łatwo zmieniać topologię i uwzględniać statystyki oparte na regionach.

Węże (Snakes)

- Węże to dwuwymiarowe uogólnienie splajnów 1D minimalizujących energię:

$$E_{int} = \int \alpha(s) \|f_s(s)\|^2 + \beta(s) \|f_{ss}(s)\|^2 ds,$$

gdzie s jest długością łuku wzdłuż krzywej $f(s) = (x(s), y(s))$ i $\alpha(s)$ i $\beta(s)$ to funkcje ważenia pierwszego i drugiego rzędu,

- Możemy zdyskretyzować tę energię, próbkując równomiernie położenie początkowej krzywej wzdłuż jej długości, aby uzyskać,

$$E_{int} = \sum_i \alpha(i) \|f(i+1) - f(i)\|^2 / h^2 + \beta(i) \|f(i+1) - 2f(i) + f(i-1)\|^2 / h^4,$$

gdzie h jest wielkością kroku ponownego próbkowania krzywej wzdłuż iteracji długości łuku.

Energia wewnętrzna splajnu

- Snake jednocześnie minimalizuje zewnętrzny potencjał oparty na obrazie i ograniczeniach.
- Potencjały oparte na obrazie są sumą kilku terminów:

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term},$$

gdzie *line* termin przyciąga węza do ciemnych grzbietów, *edge* termin przyciąga go do silnych gradientów (krawędzi) i *term* przyciąga go do zakończeń linii.

Węże (Snakes)

- Większość systemów używa tylko terminu krawędzi, który może być bezpośrednio proporcjonalny do gradientów obrazu,

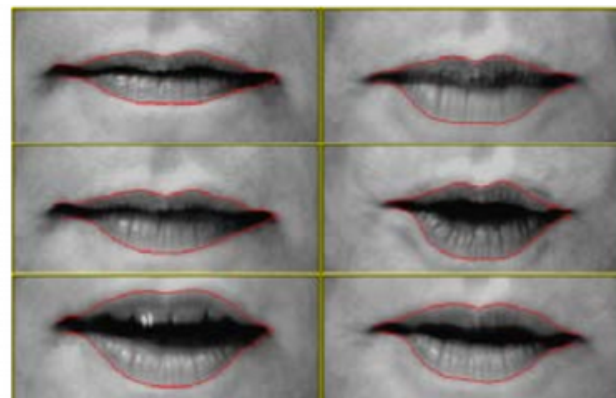
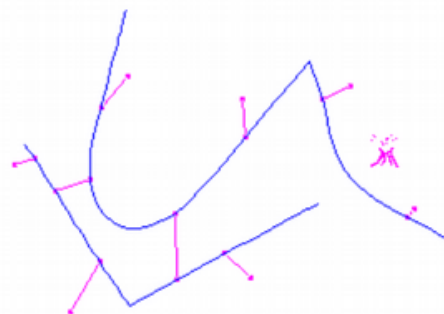
$$E_{edge} = \sum_i -\|\nabla I(f(i))\|^2,$$

lub do wygładzonej wersji obrazu Laplasjanu,

$$Edge = \sum_i |(G_\sigma \star \nabla^2 I)(f(i))|^2.$$

Czasami wyodrębnia się krawędzie, a następnie używa się mapy odległości do krawędzi jako alternatywy dla tych dwóch pierwotnie proponowanych potencjałów.

Węże - przykład

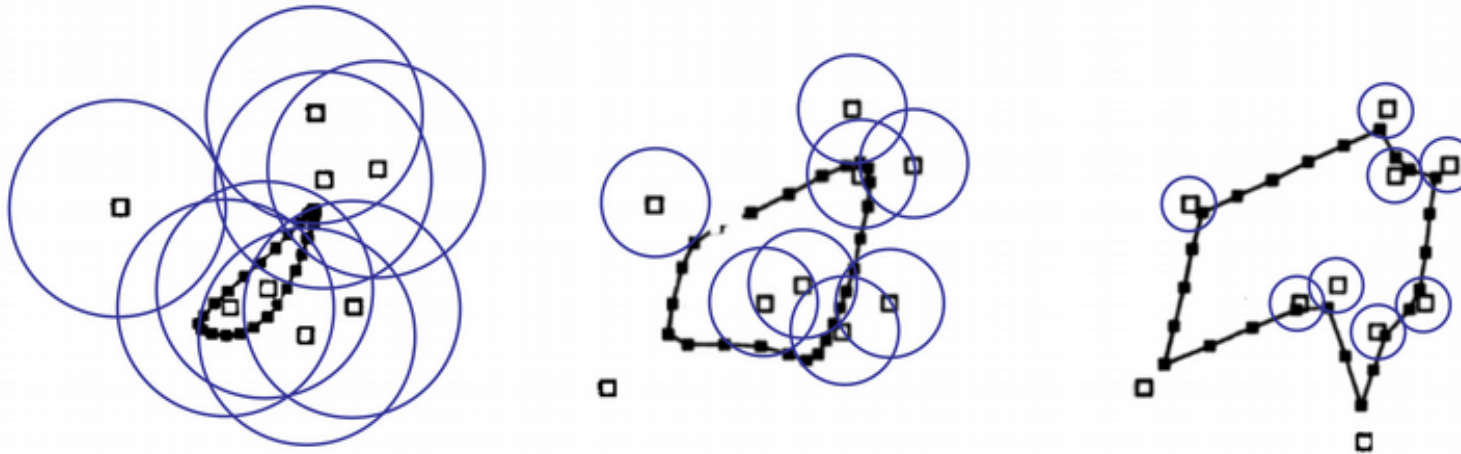


- W aplikacjach interaktywnych można również dodawać różne wiązania umieszczone przez użytkownika, np. Siły przyciągające (sprężyste) skierowane do punktów zakotwiczenia $d(i)$,

$$E_{spring} = k_i ||f(i) - d(i)||^2,$$

Gdy węże ewoluują, minimalizując swoją energię, często “poruszają się” i “pełzają”, co stanowi ich popularną nazwę.

Elastyczna siatka ang. Elastic net



- Ramy minimalizacji zużycia energii są oparte na problemie komiwojażera,
- Wąż, który jest zmuszony przejść przez każde miasto, mógłby rozwiązać ten problem (bez żadnych gwarancji optymalności)
- Zamknięte kwadraty połączone prostymi odcinkami to punkty trasy.
- Niebieskie kółka wskazują przybliżony zakres siły przyciągania każdego miasta.

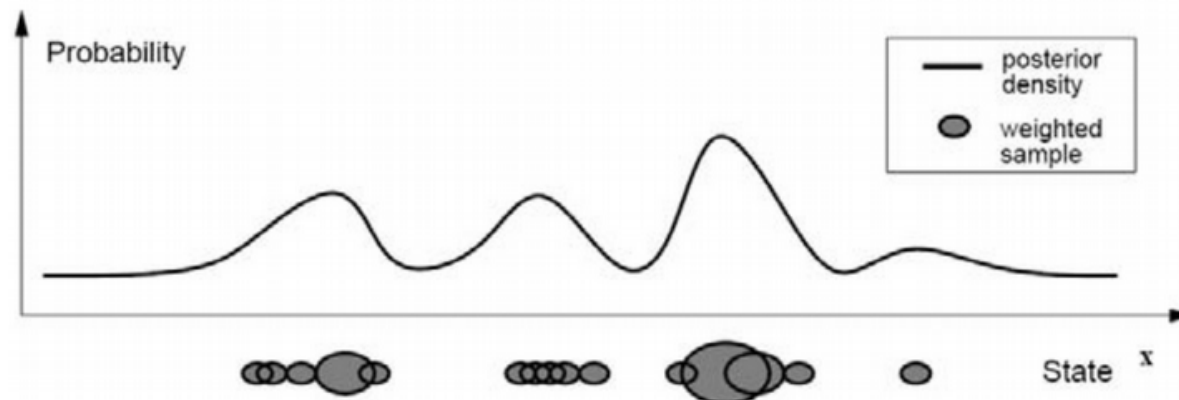
Dynamiczne węże

- Obiekt będący przedmiotem zainteresowania jest śledzony od klatki do klatki, gdy odkształca się i ewoluuje.
- Sensowne jest użycie oszacowań z poprzedniej klatki do przewidywania i ograniczania nowych oszacowań.
- Jednym ze sposobów jest użycie filtrowania Kalmana, co daje formułę o nazwie **węże Kalmana**.
- Filtr Kalmana oparty jest na liniowym dynamicznym modelu ewolucji parametrów kształtu,

$$x_t = Ax_{t-1} + w_t,$$

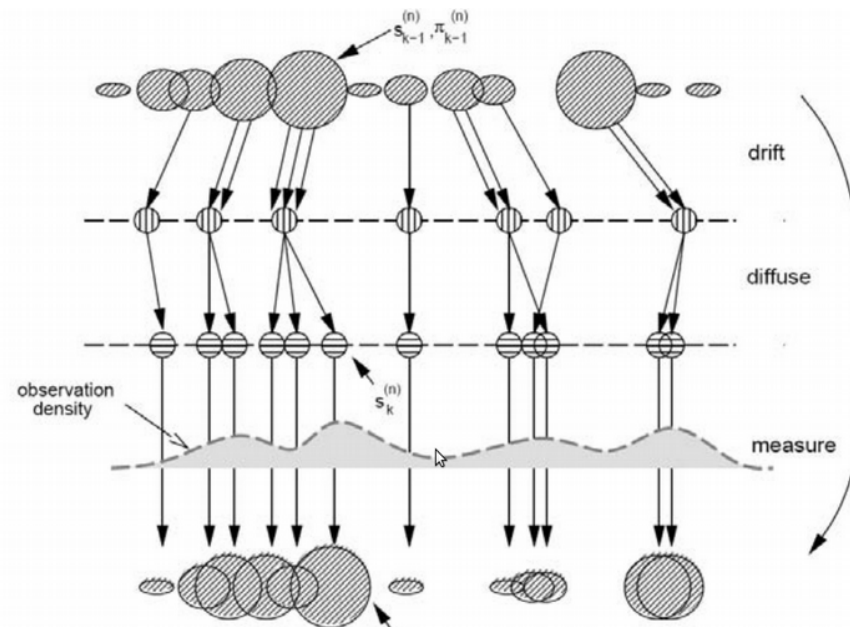
gdzie x_t i x_{t-1} są bieżącymi i poprzednimi zmiennymi stanu, A to liniowa macierz przejścia, a w to szum (często modelowany jako Gaussian).

Filtry cząstek



- Techniki filtrowania cząstek przedstawiają rozkład prawdopodobieństwa za pomocą zbioru ważonych próbek punktowych,
- Aby zaktualizować lokalizacje próbek zgodnie z dynamiką liniową (dryft deterministyczny),
- Środki próbek są aktualizowane i generowanych jest wiele próbek dla każdego punktu.

Kondensacja dzięki zastosowaniu filtra cząstek



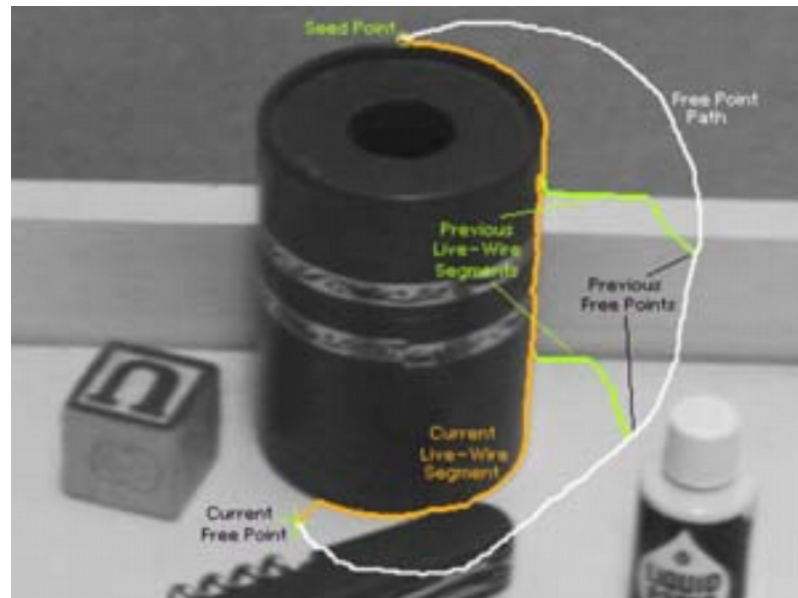
- Lokalizacje są przenoszone losowymi wektorami pobranymi z dystrybucji,
- Na koniec wagi tych próbek mnoży się przez gęstość prawdopodobieństwa pomiaru,
- Bierzemy każdą próbkę i mierzymy jej prawdopodobieństwo przy aktualnych (nowych) pomiarach.

Nożyce

Aktywne kontury pozwalają użytkownikowi z grubsza określić granicę zainteresowania i mieć system ewoluować kontur w kierunku dokładniejszej lokalizacji.

- Gdy użytkownik rysuje zgrubny kontur, system oblicza i rysuje lepszą krzywą, która przylega do krawędzi o wysokim kontraście,
- Obraz jest najpierw wstępnie przetwarzany w celu skojarzenia niskich kosztów z krawędziami (połączenie przejścia przez zero, wielkości gradientu i orientacji gradientu w celu obliczenia tych kosztów).
- Następnie, gdy użytkownik śledzi zgrubną krzywą, system w sposób ciągły przelicza ścieżkę o najniższych kosztach.

Nożyczki - przykład



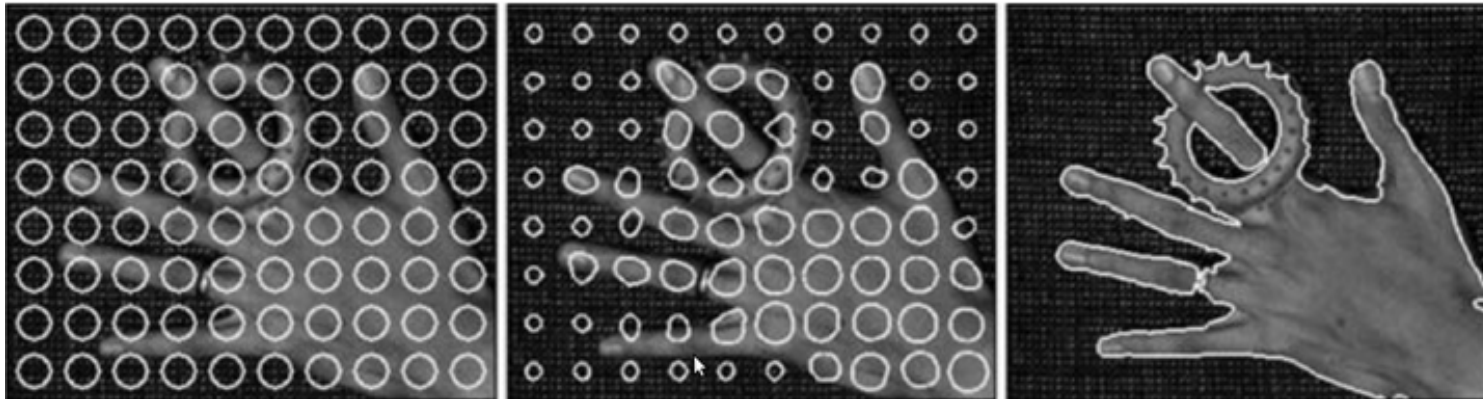
- Gdy mysz śledzi białą ścieżkę, nożyczki podążają pomarańczową ścieżką wzdłuż granicy obiektu,
- Zielone krzywe pokazują pozycje pośrednie.

Zestawy poziomów

Alternatywną reprezentacją takich zamkniętych konturów jest użycie zestawu poziomów, w którym przejście przez zero funkcji charakterystycznej definiuje krzywą.

- Zestawy poziomów ewoluują w celu dopasowania i śledzenia interesujących obiektów poprzez modyfikację podstawowej **funkcji osadzania** (funkcji 2D) $\phi(x, y)$ zamiast krzywej $f(s)$,
- Alternatywnym podejściem jest ponowne przedstawienie problemu w ramach segmentacji, gdzie energia mierzy spójność statystyk obrazu wewnątrz i na zewnątrz segmentowanych regionów.

Zestawy poziomów - przykład



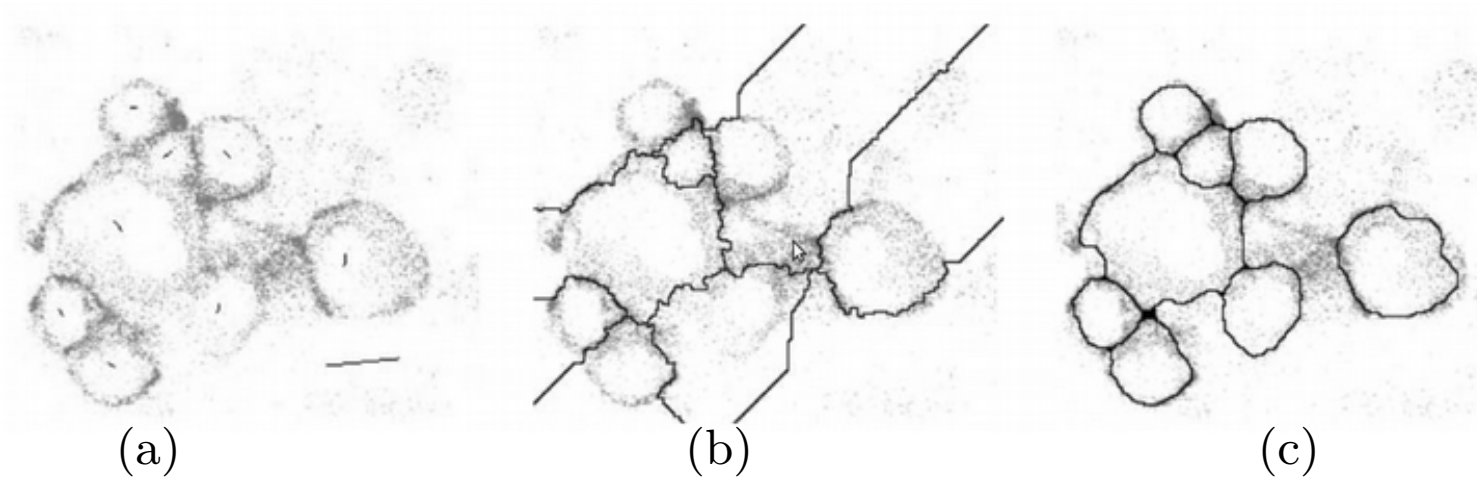
- Gaussa są używane do modelowania rozkładu pikseli pierwszego planu i tła.
- Początkowe okręgi ewoluują w kierunku dokładnej segmentacji pierwszego planu i tła, dostosowując ich topologię w miarę ich ewolucji.

Dział wodny

Technika związana z progowaniem, ponieważ działa na obrazach w skali szarości do obliczenia zlewiska.

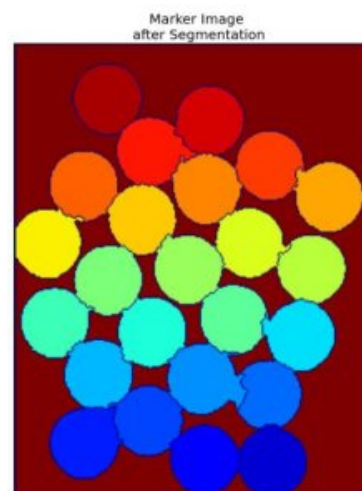
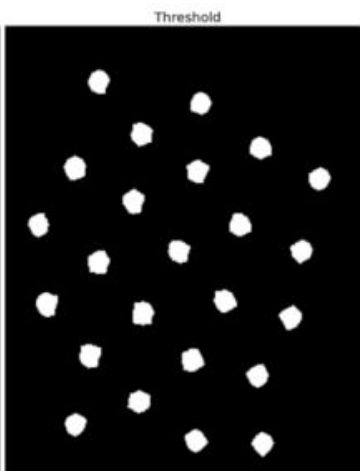
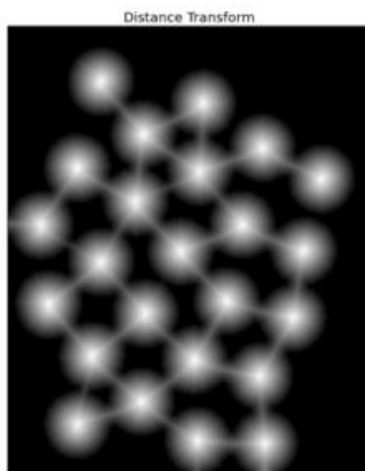
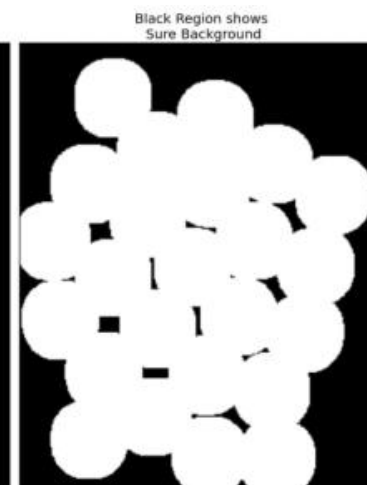
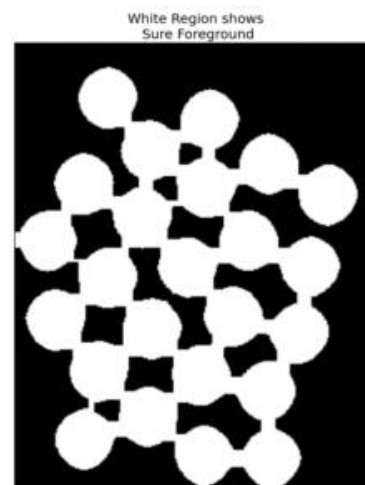
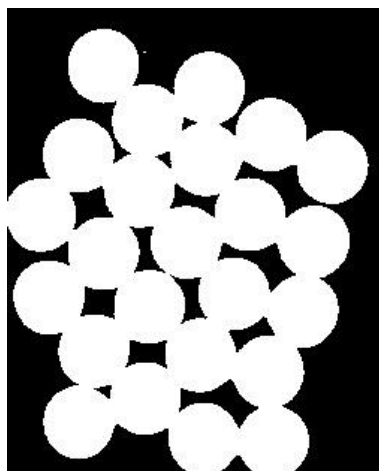
- Najprostszą możliwą techniką segmentacji obrazu w skali szarości jest wybranie progu, a następnie obliczenie połączonych komponentów.
- Ta technika dzieli obraz na kilka zlewni, które są regionami obrazu (interpretowanymi jako pole wysokości lub krajobraz),
- Ponieważ obrazy rzadko mają ciemne obszary oddzielone jaśniejszymi krawędziami, segmentacja zlewni jest zwykle stosowana do wygładzonej wersji obrazu gradientu wielkości.

Zlewisko - przykład



- (a) Oryginalny obraz mikroskopii konfokalnej z zaznaczonymi nasionami (segmenty linii);
- (b) Standardowa segmentacja działu wodnego;
- (c) Lokalnie ograniczona segmentacja zlewiska.

Segmentacja obrazu za pomocą algorytmu zlewiska - opencv



Segmentacja oparta na wykresach

Algorytm ten wykorzystuje względne różnice między regionami, aby określić, które z nich należy scalić.

- Rozpoczyna się od miary niepodobieństwa między pikselami $w(e)$, która mierzy na przykład różnice intensywności między sąsiadami $N8$.
- Dla dowolnego regionu R , jego wewnętrzna różnica jest definiowana jako największa waga krawędzi w minimalnym drzewie rozpinającym regionu,

$$Int(R) = \min_{e \in R} w(e).$$

- Dla dowolnych dwóch sąsiednich regionów z co najmniej jedną krawędzią łączącą ich wierzchołki, różnica między tymi regionami jest definiowana jako minimalna waga krawędzi łączącej te dwa regiony,

$$Dif(R_1, R_2) = \min_{e=(v_1 \in R_1, v_2 \in R_2)} w(e).$$

- Ich algorytm łączy dowolne dwa sąsiednie regiony, których różnica jest

mniej niż minimalna wewnętrzna różnica tych dwóch regionów,

$$MInt(R_1, R_2) = \min(Int(R_1) + \tau(R_1), Int(R_2) + \tau(R_2)),$$

gdzie $\tau(R)$ jest heurystyczną karą regionu

Techniki wyszukiwania średniej i modów

Techniki znajdowania zmiany średniej i trybu, takie jak k-średnie i mieszanin Gaussa.

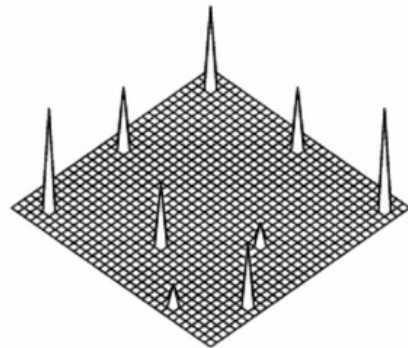
- K-średnie i mieszaniny technik Gaussa wykorzystują parametryczny model funkcji gęstości,
- Gęstość to superpozycja niewielkiej liczby prostszych rozkładów (np. Gaussa), których położenie (środkie) i kształt (kowariancja) można oszacować.
- Z drugiej strony przesunięcie średnie wygładza rozkład i znajduje wartości szczytowe,
- Ponieważ modelowana jest pełna gęstość, podejście to nazywa się *nieparametryczne*.

Globalna optymalizacja

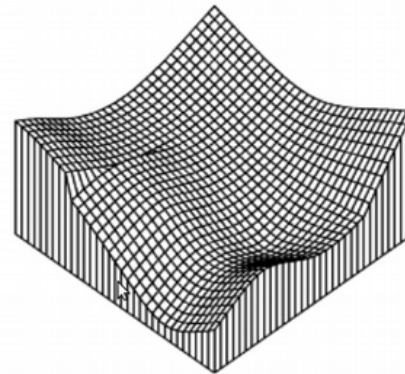
Możemy najpierw sformułować cele pożądaney transformacji za pomocą jakiegoś kryterium optymalizacji, a następnie znaleźć rozwiązanie, które najlepiej spełnia to kryterium.

- *Regularyzacja* lub *wariacyjne metody*, konstruuje ciągłą globalną funkcję energetyczną, która opisuje pożądanę cechy, a następnie znajduje rozwiązanie minimalnej energii,
- Formuluje problem przy użyciu statystyki bayesowskiej, modelując zarówno zaszumiony proces pomiaru, który wytworzył obrazy wejściowe, jak i *wcześniejsze założenia* dotyczące przestrzeni rozwiązań (często kodowane przy użyciu *Markov Random Field*)

Regularyzacja - źle uwarunkowane problemy



(a)



(b)

- Jeśli użyjemy interpolacji wielomianowej, tego rodzaju problemy (a) są źle uwarunkowane,
- Ponieważ próbujemy odzyskać nieznaną funkcję $f(x, y)$, z której pobrano próbkę punktu danych $d(x_i, y_i)$, takie problemy są również często nazywane *odwrotnymi problemami*,
- Ponieważ próbujemy odzyskać pełny opis z ograniczonego zestawu próbek.

Energetyczne miary funkcji

- Dla funkcji jednowymiarowych $f(x)$ możemy scałkować kwadratową pierwszą pochodną funkcji,

$$\varepsilon_1 = \int f_x^2(x) dx$$

albo może całkować drugą pochodną do kwadratu,

$$\varepsilon_2 = \int f_{xx}^2(x) dx$$

używamy indeksów do oznaczenia zróżnicowania.

- Takie miary energii to przykłady *funkcjonałów*, czyli operatorów odwzorowujących funkcje na wartości skalarne. Nazywa się je również często *metodami wariacyjnymi*.

Kara za dane

- W ujęciu płynności regularyzacja wymaga również pewnego rodzaju *kary za dane*,
- Dla interpolacji danych rozproszonych, termin danych mierzy odległość między funkcją $f(x, y)$ a zbiorem punktów danych $d_i = d(x_i, y_i)$,

$$\varepsilon_d = \sum_i [f(x_i, y_i) - d_i]^2.$$

- Aby uzyskać globalną energię, którą można zminimalizować, dwa terminy energetyczne są zwykle dodawane,

$$\varepsilon = \varepsilon_d + \lambda \cdot \varepsilon_s,$$

gdzie ε_s to *kara za gładkość*, a λ to *parametr regularyzacji*.

Energia danych dwuwymiarowych dyskretnych

- Dwuwymiarowa dyskretna energia danych jest zapisywana jako

$$E_d = \sum_{i,j} w(i,j)[f(i,j) - d(i,j)]^2,$$

gdzie lokalne wagi $w(i,j)$ kontrolują siłę wymuszania ograniczenia danych.

- Te wartości są ustawione na zero, gdy nie ma danych i można je ustawić na odwrotną wariancję pomiarów danych, gdy są dane
- Całkowita energia dyskretyzowanego problemu może być teraz zapisana w postaci kwadratowej

$$E = E_d + \lambda E_s = x^T A x - 2x^T b + c,$$

gdzie $x = [f(0,0) \dots f(m-1, n-1)]$ nazywany jest *wektorem stanu*.

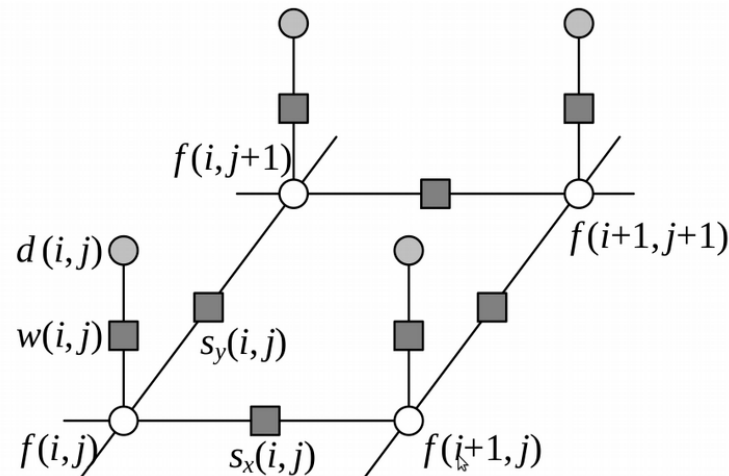
Minimalizowanie kwadratowej formy energii

Celem jest zminimalizowanie energii $E_s = x^T A x - 2x^T b + c$,

- Rzadka symetryczna macierz dodatnio-określona A nazywana jest *Hessjanem*, ponieważ koduje drugą pochodną funkcji energii,
- Nazywamy b ważonym wektorem danych. Zminimalizowanie powyższej formy kwadratowej jest równoważne rozwiązaniu rzadkiego układu liniowego

$$Ax = b,$$

Regularyzacja pierwszego rzędu



Funkcje energii dyskretnej gładkości stają się $E_1 =$

$$= \sum_{i,j} s_x(i, j) [f(i+1, j) - f(i, j) - g_x(i, j)]^2 + s_y(i, j) [f(i, j+1) - f(i, j) - g_y(i, j)]^2$$

- Białe kółka to niewiadome $f(i, j)$, podczas gdy ciemne kółka to dane wejściowe $d(i, j)$.
- W interpretacji sieci rezystancyjnej wartości d i f kodują napięcia wejściowe i wyjściowe, a czarne kwadraty oznaczają rezystory, których przewodnictwo jest ustawione na $s_x(i, j)$, $s_y(i, j)$ i $w(i, j)$.

Pola losowe Markowa ang. Markov random fields

- Model bayesowski może oddzielnie modelować zaszumiony proces tworzenia (pomiaru) obrazu i wcześniejszy model statystyczny w przestrzeni rozwiązań.
- Losowe modele pól Markowa można definiować na zmiennych dyskretnych, takich jak etykiety obrazów (gdzie zmienne nie mają odpowiedniego uporządkowania), dla których regularyzacja nie ma zastosowania.

Reguła Bayesa

- Rozkład posteriori dla danego zestawu pomiarów y , $p(y|x)$, w połączeniu z poprzednikiem $p(x)$ nad niewiadomymi x , jest podane przez

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

gdzie $p(y) = \sum_x p(y|x)p(x)$ jest stałą normalizującą używaną do tworzenia rozkładu $p(x|y)$ (unormowane do 1).

- Biorąc ujemny logarytm z obu stron, otrzymujemy

$$-\log p(x|y) = -\log p(y|x) - \log p(x) + C$$

który jest *ujemnym logarytmem wiarygodności posteriori*

Maksymalizacja posteriori - MAP

- Aby znaleźć najbardziej prawdopodobne (maksymalne a posteriori) rozwiązanie x przy pewnych pomiarach y , po prostu minimalizujemy to ujemne prawdopodobieństwo log, które można również traktować jako energię,

$$E(x, y) = E_d(x, y) + E_p(x).$$

- Pierwszy wyraz $E_d(x, y)$ to **energia danych** lub kara za dane - mierzy ujemne log prawdopodobieństwo, że dane zostały zaobserwowane przy nieznanym stanie x .
- Drugi człon $E_p(x)$ to **energia priori** - odgrywa rolę analogiczną do energii gładkości w regularyzacji.

Przetwarzanie obrazu - pole losowe Markowa

- Nieznane x to zestaw pikseli wyjściowych

$$x = [f(0, 0) \dots f(m - 1, n - 1)],$$

a dane są (w najprostszym przypadku) pikselami wejściowymi

$$y = [d(0, 0) \dots d(m - 1, n - 1)]$$

- Prawdopodobieństwo $p(x)$ to rozkład Gibbsa lub Boltzmanna, którego prawdopodobieństwo logarytmiczne ujemne można zapisać jako sumę potencjałów interakcji parami,

$$E_p(x) = \sum_{\{(i,j),(k,l)\} \in N} V_{i,j,k,l}(f(i,j), f(k,l)),$$

gdzie $N(i, j)$ oznacza sąsiadów piksela (i, j) .

- Być może trzeba będzie oszacować energię w większym zbiorze klik, które zależą od kolejności pola losowego Markowa

Pole losowe Markowa - przykład binarny

- Przykładami takich pól są 1-bitowe (czarno-białe) dokumenty skanowane
- Aby odszumiać zeskanowany obraz, ustawiliśmy karę za dane, aby odzwierciedlić zgodność między zeskanowanym a ostatecznym obrazem,

$$E_d(i, j) = w\delta((i, j), d(i, j))$$

- Kara za gładkość, aby odzwierciedlić zgodność między sąsiednimi pikselami

$$E_p(i, j) = E_x(i, j) + E_y(i, j) = s\delta(f(i, j), f(i+1, j)) + s\delta(f(i, j), f(i, j+1)).$$

- Najprostszym podejściem jest wykonanie zejścia gradientowego. To podejście jest znane jako klasyfikacja kontekstowa.

Segmentacja obrazu - MRF



- Użytkownik rysuje kilka czerwonych kresek na pierwszym planie i kilka niebieskich w tle.
- System oblicza rozkłady kolorów dla pierwszego planu i tła oraz rozwiązuje binarny MRF.
- Wagi gładkości są modulowane przez gradienty intensywności (krawędzie), co sprawia, że jest to warunkowe pole losowe (CRF).

Cięcia wykresów i metody oparte na energii

Algorytmy segmentacji to chęć grupowania pikseli o podobnych statystykach i wyznaczania granic między pikselami w różnych regionach.

Cięcia wykresów i metody oparte na energii

energię odpowiadającą problemowi segmentacji można zapisać jako:

$$E(f) = \sum_{i,j} E_r(i,j) + E_b(i,j),$$

gdzie termin region

$$E_r(i,j) = E_S(I(i,j); R(f(i,j)))$$

jest ujemnym logarytmicznym prawdopodobieństwem, że intensywność pikseli (lub kolor) $I(i,j)$ jest zgodna ze statystykami regionu $R(f(i,j))$ i terminem granicznym:

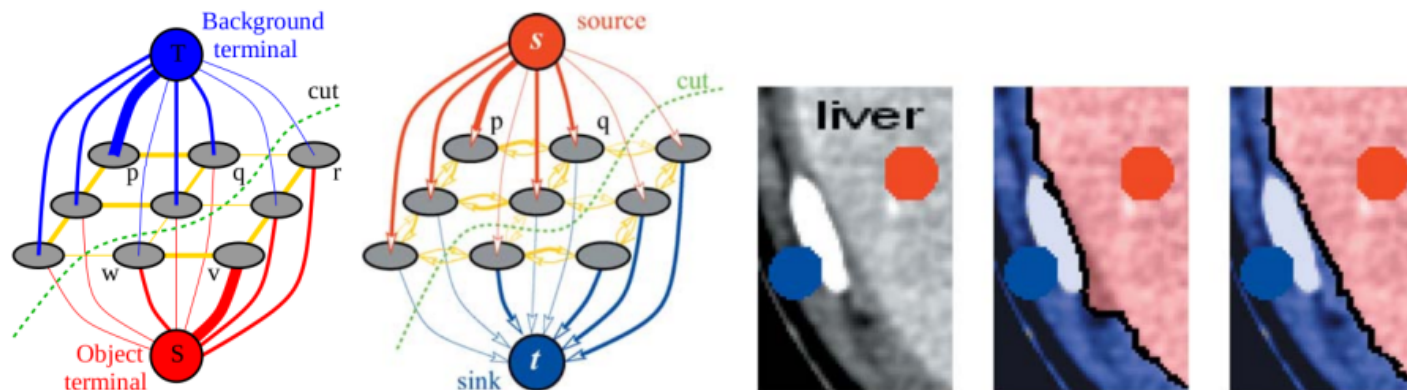
$$E_b(i,j) = s_x(i,j)\delta(f(i,j) - f(i+1,j)) + s_y(i,j)\delta(f(i,j) - f(i,j+1))$$

mierzy niespójność między $N4$ sąsiadami modulowanymi przez lokalne warunki gładkości poziomej i pionowej $s_x(i,j)$ i $s_y(i,j)$. Statystyki regionu

mogą wyglądać następująco:

$$E_S(I; \mu_k) = ||I - \mu_k||^2.$$

Cięcia wykresów i metody oparte na energii - przykład



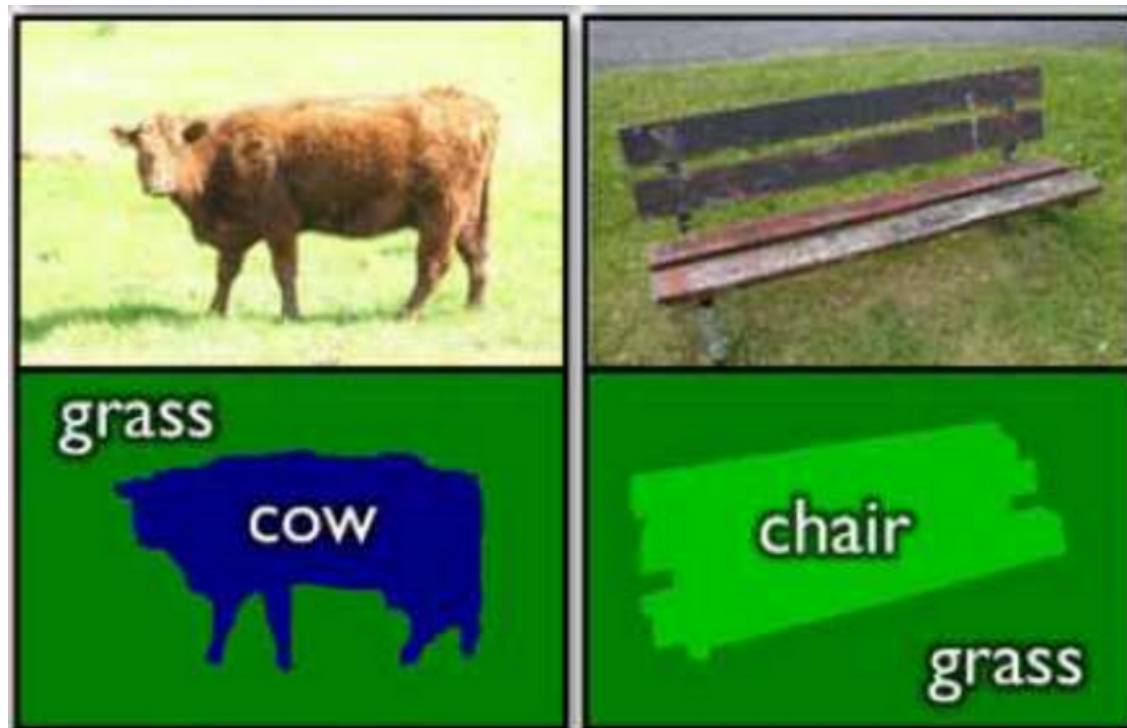
- Użytkownik najpierw zaznacza piksele w tle i obszary obiektów,
- Te piksele stają się następnie nasionami, które wiążą węzły na wykresie $S - T$ ze źródłem i etykietami ujścia S i T ,
- Krawędzie wykresu pochodzą z regionu i granicznych warunków energetycznych.

Interaktywne wyodrębnianie pierwszego planu za pomocą algorytmu GrabCut - opencv



- Pierwszy gracz i piłka nożna są ujęci w niebieskim prostokącie.
- Następnie wykonuje się ostatnie poprawki białymi kreskami (oznaczające pierwszy plan) i czarnymi kreskami (oznaczającymi tło).

Segmentacja semantyczna



- Trudną wersją ogólnego rozpoznawania obiektów i zrozumienia sceny jest jednocześnie wykonywanie rozpoznawania i dokładnej segmentacji granic,
- Ciąg dalszy nastąpi ...

Zadania na laboratoria

Zaimplementuj algorytm segmentacji *Watershed* w dwóch wersjach

- Gdzie pola ostrości są wybierane automatycznie (można użyć zdjęć monet),
- Gdzie obszary segmentacji są zaznaczane ręcznie (inny obraz niż w przypadku 1)
- Spróbuj zastosować tę technikę do liczenia monet na tacy - poprzednio zadanie.