Analiza i przetwarzanie obrazów i wideo, wykład 7 ~~(8)~~

Detekcja obszarów zainteresowania. Ekstrakcja cech lokalnych

J. Sas

AI-Tech, 2022/2023

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"

Agenda:

- Klasyfikacja cech wykorzystywanych w APOW
- Detektory Moraveca i Harrisa
- Detektory SIFT i SURF
- Cechy tekstury obrazu
- Deskryptory obrazów wykorzystujące rozkłady cech lokalnych

Ekstrakcja cech z obrazu

Ekstrakcja cech stanowi syntezę danych wizualnych, mającą na celu wydobycie informacji istotnej dla użytkownika.

- Występuje w praktycznie każdej metodzie analizy obrazu.
- Stanowi kluczowy element metod analizy obrazu.
- Dostarcza podstawowych informacji do dalszego przetwarzania.
- Powinna być **dostosowana** do rozwiązywanego problemu.
- Jest wykonywana dla określonego **regionu zainteresowania**.
- Wielka różnorodność cech i metod ekstrakcji opisanych w literaturze.

Podział cech ze względu obszar ekstrakcji

- **Globalne**:
  - jeden obszar zainteresowania obejmujący cały obraz.
  - charakteryzują cały obraz, pomijając szczegóły.
- **Semi–globalne**, **semi–lokalne**:
  - reprezentują duże i istotne części obrazu.
  - zazwyczaj jest kilka do kilkunastu regionów.
  - mogą odzwierciedlać obiekty znajdujące się na obrazie.
  - mogą dzielić obraz w sztywny, ustalony sposób.
- ...

Podział cech ze względu obszar ekstrakcji

- **Globalne**: ...
- **Semi–globalne**, **semi–lokalne**: ...
- **Lokalne**:
  - reprezentują własności niewielkich fragmentów obrazu.
  - cecha opisuje własności **obszaru zainteresowania** - (ang. *Region of Interest - RoI)* wokół **punktu zainteresowania/kluczowego** (ang. *Keypoint* lub *Point of Interest - PoI*).
  - można zastosować do detekcji obiektów w obrazie - znalezione punkty kluczowe składają się na jeden obiekt (jak je łączyć?).
  - ROI mogą się wzajemnie nakładać (nie są rozłączne).
  - dla uproszczenia analizy RoI mogą być aproksymowane określoną figurą geometryczną (np. elipsą).
  - duża liczba PoI/RoI w obrazie - rzędu kilkaset do kilku tysięcy.
  - wyznaczanie PoI i otaczających je RoI jest formą segmentacji obrazu - segment: otoczenie punktu kluczowego, jak np. w MSER.

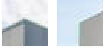
Cechy – niezmienniczość

Niezmienniczość

Niezmienniczość – stałość cechy przy określonych przekształceniach obrazu.

- Niezmienniczość cech znacznie ułatwia dalszą analizę,
- W różnych zadaniach oczekujemy niezmienności względem różnych czynników,
- Wybór cech należy oprzeć o ich niezmienniczość.
- Przykłady niezmienniczości względem typowych czynników:
  - od zmiany jasności,
  - od zmiany natężenia kolorów,
  - od zmiany warunków zewnętrznego oświetlenia,
  - od obrotu,
  - od przesunięcia,
  - od skalowania,
  - od przekształcenia afinicznego (obroty, przesunięcia i skalowania).

Cechy – niezmienniczość

A)  B)  C) 

Detektor narożników Moraveca

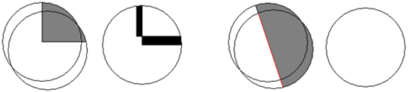
Dlaczego narożniki są ważne jako punty kluczowe?

- narożniki i T-złączenia (T-junctions) często są charakterystycznymi punktami wyznaczającymi kształty
- lepiej identyfikują kształty niż krawędzie
- jest ich znacznie mniej niż punktów na krawędziach
- pozostają w zbliżonych relacjach geometrycznych po transformacjach geometrycznych obrazu:
    - np. przy zmianie warunków obserwacji (rzutowania) sceny na powierzchnię obrazu

Detektor narożników Moraveca

Zasada:

- Rozważamy niewielki fragment (np. kołowy) obrazu
- Wyliczamy w nim średnią jasność (lub wartość innego atrybutu)
- W przypadku obszaru (prawie) jednorodnego niewielka zmiana położenia fragmentu nie prowadzi do istotnych zmian wartości średniej, niezależnie od kierunku zmiany położenia,
- W przypadku krawędzi przesunięcie wzdłuż krawędzi nie powoduje zmiany, przesunięcie prostopadle do krawędzi powoduje największą zmianę,
- W przypadku narożnika - nie ma takiego kierunku przesunięcia, który nie powodowałby znacznych zmian wartości średniej, tzn. przesunięcie w dowolnym kierunku daje duże zmiany średniej.



Detektor narożników Moraveca

Zmiana przy przesunięciu fragmentu:

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y') \in W(x,y)} w(x', y')(I(x' + \Delta_x, y' + \Delta_y) - I(x', y'))^2 \quad (1)$$

- $w(x', y')$ - wektor wagi dla pozycji $(x', y') \in W$
- (Δ_x, Δ_y) - wektor "próbnego" przesunięcia okna.
- kierunki próbne: $(\Delta_x, \Delta_y) = (1, 0), (0, 1), (1, 1), (1, -1)$
- $Q(x, y) = \min\{E(x, y, 1, 0), E(x, y, 0, 1), E(x, y, 1, 1), E(x, y, 1, -1)\}$
  - im $Q(x, y)$ większe tym bardziej prawdopodobne wystąpienie narożnika (min bo dla narożnika zmiany w każdym z kierunków duże)
- jako punkt narożnikowy (kluczowy) przyjmujemy wystąpienie lokalnego maximum $Q(x, y)$ o wartości powyżej ustalonego progu

Detektor narożników Harrisa

Wady detektora Moraveca:

- badamy tylko cztery kierunki
- podatność na szumy
- lokalne maksima E mogą być nieprecyzyjne (fałszywe narożniki przy maksimach nieznacznie różniących się od otoczenia)

Detektor narożników Harrisa

Wady detektora Moraveca:

- badamy tylko cztery kierunki
- podatność na szumy
- lokalne maksima E mogą być nieprecyzyjne (fałszywe narożniki przy maksimach nieznacznie różniących się od otoczenia)

Rozwinięcie i udoskonalenie detektora Moraveca:

- zmniejszenie podatności na szumy przez wstępne filtrowanie wygładzające (Gauss)
- brak ograniczenia rozważanych kierunków przesunięcia okna
- inna zasada ostatecznej decyzji o wykryciu narożnika

Detektor narożników Harrisa

- Różnica średniej jasności przy przesunięciu o wektor(Δ_x, Δ_y), W - obszar wokoło (x,y); $w(x', y')$ - waga punktu w średniej ważonej:

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(x', y')(I(x' + \Delta_x, y' + \Delta_y) - I(x', y'))^2$$

- Wyrażamy za pomocą gradientu (I_x, I_y):

$$I(x' + \Delta_x, y' + \Delta_y) - I(x', y') = [I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T$$

- Sumę dla dowolnego kierunku przesunięcia testowego (Δ_x, Δ_y) wyrażamy za pomocą macierzy sumy iloczynów pochodnych cząstkowych $Q(x, y)$

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(u, v)([I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T)^2 = [\Delta_x, \Delta_y]Q(x, y)[\Delta_x, \Delta_y]^T$$

Detektor narożników Harrisa

- Różnica średniej jasności przy przesunięciu o wektor(Δ_x, Δ_y), W - obszar wokoło (x,y); $w(x', y')$ - waga punktu w średniej ważonej:

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(x', y')(I(x' + \Delta_x, y' + \Delta_y) - I(x', y'))^2$$

- Wyrażamy za pomocą gradientu (I_x, I_y):

$$I(x' + \Delta_x, y' + \Delta_y) - I(x', y') = [I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T$$

- Sumę dla dowolnego kierunku przesunięcia testowego (Δ_x, Δ_y) wyrażamy za pomocą macierzy sumy iloczynów pochodnych cząstkowych $Q(x, y)$

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(u, v)([I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T)^2 = [\Delta_x, \Delta_y]Q(x, y)[\Delta_x, \Delta_y]^T$$

Detektor narożników Harrisa

- Różnica średniej jasności przy przesunięciu o wektor(Δ_x, Δ_y), W - obszar wokoło (x,y); $w(x', y')$ - waga punktu w średniej ważonej:

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(x', y')(I(x' + \Delta_x, y' + \Delta_y) - I(x', y'))^2$$

- Wyrażamy za pomocą gradientu (I_x, I_y):

$$I(x' + \Delta_x, y' + \Delta_y) - I(x', y') = [I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T$$

- Sumę dla dowolnego kierunku przesunięcia testowego (Δ_x, Δ_y) wyrażamy za pomocą macierzy sumy iloczynów pochodnych cząstkowych $Q(x, y)$

$$E(x, y, \Delta_x, \Delta_y) = \sum_{(x',y')\in W(x,y)} w(u, v)([I_x(x', y'), I_y(x', y')][\Delta_x, \Delta_y]^T)^2 = [\Delta_x, \Delta_y]Q(x, y)[\Delta_x, \Delta_y]^T$$

Detektor narożników Harrisa

Macierz Q (dla uproszczenia pominięto wagi $w(x', y')$):

$$Q(x,y) = \begin{matrix} a & b \\ b & c \end{matrix} = \begin{bmatrix} \sum_W I_x(x,y)^2 & \sum_W I_x(x,y) I_y(x,y) \\ \sum_W I_x(x,y) I_y(x,y) & \sum_W I_y(x,y)^2 \end{bmatrix}$$

- $E(x,y,\delta) = \delta Q(x,y) \delta^T; \delta = (\Delta_x, \Delta_y)$ dla ustalonego (x,y) wyraża jak zmienia się średnia jasność w oknie W przy przesunięciu o δ.
- Szukamy takich (x,y) aby zmiany przy niewielkim przesunięciu w dowolnym kierunku były duże.
- Równanie $E(\delta) = \delta Q(x,y) \delta^T = 1; \delta = (\Delta_x, \Delta_y)$ na płaszczyźnie kierunków δ opisuje elipsę $a\Delta_x^2 + 2b\Delta_x\Delta_y + c\Delta_y^2 = 1$ gdzie wektory własne macierzy Q są kierunkami osi a wartości własne λ_i są odwrotnościami długości osi:
$$d_i = 1/\sqrt{\lambda_i}; i = 1,2$$

Detektor narożników Harrisa

Macierz Q (dla uproszczenia pominięto wagi $w(x', y')$):

$$Q(x,y) = \begin{matrix} a & b \\ b & c \end{matrix} = \begin{bmatrix} \sum_W I_x(x,y)^2 & \sum_W I_x(x,y) I_y(x,y) \\ \sum_W I_x(x,y) I_y(x,y) & \sum_W I_y(x,y)^2 \end{bmatrix}$$

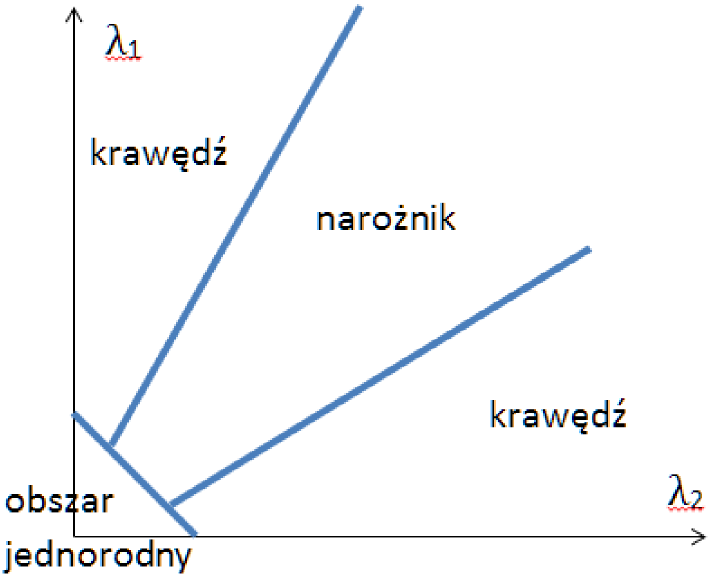
- Równanie $E(\delta) = \delta Q(x,y)\delta^T = 1; \delta = (\Delta_x, \Delta_y)$ na płaszczyźnie kierunków δ opisuje elipsę $a\Delta_x^2 + 2b\Delta_x\Delta_y + c\Delta_y^2 = 1$ gdzie wektory własne macierzy Q są kierunkami osi a wartości własne są odwrotnościami długości osi.
- Jeśli δ będzie wektorem jednostkowym to mała wartość $E(\delta)$ oznacza, że aby dojść do krawędzi elipsy należy przesunąć się znacznie (brak zmian w tym kierunku) - tak będzie jeśli idziemy wzdłuż długiej osi elipsy (mała wartość własna)
- Jeśli δ będzie wektorem jednostkowym to duża wartość $E(\delta)$ oznacza, że aby dojść do krawędzi elipsy należy przesunąć się nieznacznie (szybkie zmiany w tym kierunku) - tak będzie jeśli idziemy wzdłuż krótkiej osi elipsy (duża wartość własna)

Detektor narożników Harrisa

Macierz Q:

$$Q(x,y) = \begin{matrix} a & b \\ b & c \end{matrix} = \begin{bmatrix} \sum_W I_x(x_y)^2 & \sum_W I_x(x_y) I_y(x,y) \\ \sum_W I_x(x_y) I_y(x,y) & \sum_W I_y(x,y)^2 \end{bmatrix}$$

- Mamy punkt narożnikowy jeśli wartości różnicy $E(\delta)$ przy dowolnym kierunku $\delta = (\Delta_x, \Delta_y)$ są duże
- Jeśli jedna z osi elipsy jest długa (mała wartość własna) to duże zmiany położenia w tym kierunku dają małe zmiany wartości c
- Jeśli obie osie elipsy są krótkie (duże wartości własne) to nawet małe zmiany położenia w tym kierunku dają duże zmiany c - PUNKT NAROŻNIKOWY
- Jeśli jedna z osi elipsy jest długa (mała wartość własna) a druga krótka (duża wartość własna) to mamy do czynienia z krawędzią
- Jeśli obie osie elipsy są długie (małe wartości własne) to mamy do czynienie z obszarem jednorodnym lub gładko cieniowanym


λ_1
krawędź
narożnik
krawędź
obszar
jednorodny
λ_2

Macierz Q:

$$Q(x,y) = \begin{matrix} a & b \\ b & c \end{matrix} = \begin{bmatrix} \sum_W I_x(x_y)^2 & \sum_W I_x(x_y) I_y(x,y) \\ \sum_W I_x(x_y) I_y(x,y) & \sum_W I_y(x,y)^2 \end{bmatrix}$$

- Można wykazać, że obie wartości własne są duże, gdy następująca funkcja osiąga dużą wartość:

$$R(x,y) = det(Q(x,y)) - \alpha trace(Q(x,y))^2 \tag{2}$$

$$R(x,y) = \lambda_1 \lambda_2 - \alpha(\lambda_1 + \lambda_2)^2 \tag{3}$$

gdzie α empirycznie ustalono na około (0.4-0.8)

- Ostatecznie uznajemy że występuje punkt narożnikowy tam gdzie R(x,y) osiąga lokalne maksimum

Zastosowane odmiany detektorów:

- Harris-Affine - podobnie jak w detektorze Harrisa, ale:
  - wyznaczane są maksima w przestrzeni skal oraz
  - odpowiedź detektora maksymalizuje się względem obrotu i nierównomiernego skalowania wzdłuż osi.

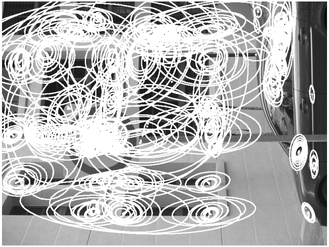
  Szczegóły w: K. Mikolajczyk, C. Schmid: Scale & Affine Invariant Interest Point Detectors. 2004.

- Hessian-Affine - podobnie jak w detektorze Harris-Affine, ale ...
  - zamiast macierzy pierwszych pochodnych wykorzystuje się macierz drugich pochodnych (hesjan) - I_{xx}, I_{xy}, I_{yx}, I_{yy} oraz jego wyznacznik

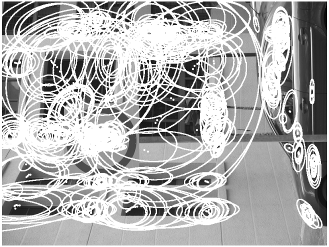
  Szczegóły w: K. Mikolajczyk: Detection of local features invariant to affine transformations, PhD thesis, 2002 (http://perception.inrialpes.fr/Publications/2002/Mik02/Mikolajczyk_these2002.pdf)

Punkty kluczowe – przykład

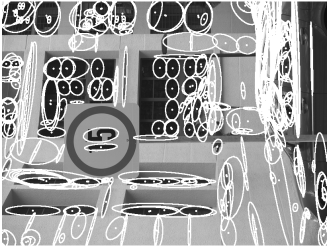
(a) obraz



(b) Harris-Affine



(c) Hessian-Affine



(d) MSER

Punkty kluczowe – przykład

Punkty kluczowe – po co to jest?

Do czego mogą posłużyć punkty kluczowe:

- identyfikowanie rejonów zainteresowania
- rozpoznawanie obiektów
- znajdowanie obszarów podobnych
- śledzenie obiektów/modelowanie ruchu ($\rightarrow$ np. SLAM)
- rekonstrukcja 3D (np. Structure from motion)
- ...

Co jest potrzebne a czego punkty kluczowe same w sobie nie mają?

Punkty kluczowe – po co to jest?

Do czego mogą posłużyć punkty kluczowe:

- identyfikowanie rejonów zainteresowania
- rozpoznawanie obiektów
- znajdowanie obszarów podobnych
- śledzenie obiektów/modelowanie ruchu ($\rightarrow$ np. SLAM)
- rekonstrukcja 3D (np. Structure from motion)
- ...

Co jest potrzebne a czego punkty kluczowe same w sobie nie mają?

SIFT - Scale Invariant Feature Transform

Cechy:

- Metoda znajdowania punktów kluczowych i wyznaczania ich **deskryptorów** (lokalnych wektorów cech)
- Punkty zainteresowania mają cechy narożników
- Wymagana niezmienniczość względem wielkości obrazu (skali) i obrotów
- Niezmienniczość względem przesunięcia uzyskiwana z zasady (obliczenia dla otoczenia znalezionego punktu kluczowego)

SIFT - Scale Invariant Feature Transform

Etapy:

- Konstrukcja przestrzeni skal (uniezależnienie od wielkości detalu w obrazie)
- Wyznaczanie różnic w przestrzeni skal (DoG - *Difference of Gaussians*)
- Znajdowanie punktów kluczowych (ekstrema w trójwymiarowej przestrzeni DoG) - z aproksymacją dla dokładniejszej lokalizacji w obrazie o pierwotnej rozdzielczości przez rozwinięcie w szereg Taylora
- Odrzucanie punktów o niskim kontraście (pozostają tylko te o wysokich wartościach DoG oraz wysokich obydwu wartościach własnych macierzy pochodnych - jak w detektorze Harrisa)
- Uniezależnienie od obrotu - wyznaczenie kierunku dominującego gradientu poprzez budowę histogramu kierunków gradientów w otoczeniu punktu kluczowego,
  - ewentualne zwielokrotnienie punktu jeśli istnieje więcej niż jedno ewidentne maksimum - z jednego punktu kluczowego generujemy kilka deskryptorów

SIFT - Scale Invariant Feature Transform

Etapy:

- Konstrukcja przestrzeni skal (uniezależnienie od wielkości detalu w obrazie)
- Wyznaczanie różnic w przestrzeni skal (DoG - *Difference of Gaussians*)
- Znajdowanie punktów kluczowych (ekstrema w trójwymiarowej przestrzeni DoG) - z aproksymacją dla dokładniejszej lokalizacji w obrazie o pierwotnej rozdzielczości przez rozwinięcie w szereg Taylora
- Odrzucanie punktów o niskim kontraście (pozostają tylko te o wysokich wartościach DoG oraz wysokich obydwu wartościach własnych macierzy pochodnych - jak w detektorze Harrisa)
- Uniezależnienie od obrotu - wyznaczenie kierunku dominującego gradientu poprzez budowę histogramu kierunków gradientów w otoczeniu punktu kluczowego,
    - ewentualne zwielokrotnienie punktu jeśli istnieje więcej niż jedno ewidentne maksimum - z jednego punktu kluczowego generujemy kilka deskryptorów

SIFT - Scale Invariant Feature Transform

Etapy:

- Konstrukcja przestrzeni skal (uniezależnienie od wielkości detalu w obrazie)
- Wyznaczanie różnic w przestrzeni skal (DoG - *Difference of Gaussians*)
- Znajdowanie punktów kluczowych (ekstrema w trójwymiarowej przestrzeni DoG) - z aproksymacją dla dokładniejszej lokalizacji w obrazie o pierwotnej rozdzielczości przez rozwinięcie w szereg Taylora
- Odrzucanie punktów o niskim kontraście (pozostają tylko te o wysokich wartościach DoG oraz wysokich obydwu wartościach własnych macierzy pochodnych - jak w detektorze Harrisa)
- Uniezależnienie od obrotu - wyznaczenie kierunku dominującego gradientu poprzez budowę histogramu kierunków gradientów w otoczeniu punktu kluczowego,
  - ewentualne zwielokrotnienie punktu jeśli istnieje więcej niż jedno ewidentne maksimum - z jednego punktu kluczowego generujemy kilka deskryptorów

SIFT - Scale Invariant Feature Transform

Etapy:

- Konstrukcja przestrzeni skal (uniezależnienie od wielkości detalu w obrazie)
- Wyznaczanie różnic w przestrzeni skal (DoG - *Difference of Gaussians*)
- Znajdowanie punktów kluczowych (ekstrema w trójwymiarowej przestrzeni DoG) - z aproksymacją dla dokładniejszej lokalizacji w obrazie o pierwotnej rozdzielczości przez rozwinięcie w szereg Taylora
- Odrzucanie punktów o niskim kontraście (pozostają tylko te o wysokich wartościach DoG oraz wysokich obydwu wartościach własnych macierzy pochodnych - jak w detektorze Harrisa)
- Uniezależnienie od obrotu - wyznaczenie kierunku dominującego gradientu poprzez budowę histogramu kierunków gradientów w otoczeniu punktu kluczowego,
  - ewentualne zwielokrotnienie punktu jeśli istnieje więcej niż jedno ewidentne maksimum - z jednego punktu kluczowego generujemy kilka deskryptorów

SIFT - Scale Invariant Feature Transform

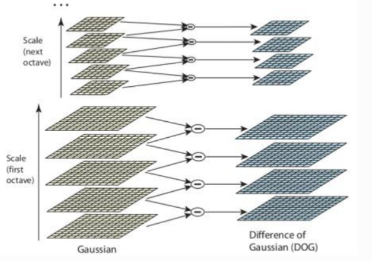
Etapy:

- Konstrukcja przestrzeni skal (uniezależnienie od wielkości detalu w obrazie)
- Wyznaczanie różnic w przestrzeni skal (DoG - *Difference of Gaussians*)
- Znajdowanie punktów kluczowych (ekstrema w trójwymiarowej przestrzeni DoG) - z aproksymacją dla dokładniejszej lokalizacji w obrazie o pierwotnej rozdzielczości przez rozwinięcie w szereg Taylora
- Odrzucanie punktów o niskim kontraście (pozostają tylko te o wysokich wartościach DoG oraz wysokich obydwu wartościach własnych macierzy pochodnych - jak w detektorze Harrisa)
- Uniezależnienie od obrotu - wyznaczenie kierunku dominującego gradientu poprzez budowę histogramu kierunków gradientów w otoczeniu punktu kluczowego,
  - ewentualne zwielokrotnienie punktu jeśli istnieje więcej niż jedno ewidentne maksimum - z jednego punktu kluczowego generujemy kilka deskryptorów

SIFT - Budowa przestrzeni skal i DoG

- Dążymy do wykrycia krawędzi filtrem Laplace'a ale z uniknięciem efektu szumu - stosujemy DoG
- Potrzebujemy wariantów obrazów o różnym stopniu rozmycia - skala rozmycia w bardzo dużym zakresie
- Stosowanie filtrów konwolucyjnych o dużych rozmiarach (duże σ) kosztowne obliczeniowo
- Zamiast filtrowania filtrami o dużej masce - zastosować zmniejszanie obrazu przez uśrednianie pikseli (operacja mało kosztowna)

Rysunek 1: *Wyznaczanie DoG w przestrzeni skal*

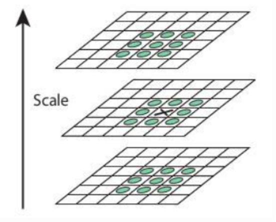


Źródło: Dokumentacja OpenCV

SIFT - Wyznaczanie kandydujących punktów kluczowych

- Odejmujemy od siebie sąsiednie obrazy w ciągu skal - otrzymujemy DoG dla różnych skal
- W/w operacja odpowiada detekcji krawędzi
- Znajdujemy maksima w 3D
- Dla dokładniejszej lokalizacji punktu (ze względu na zastosowaną *piramidę obrazów*) "uciąglamy" funkcję jasności przez aproksymację z szeregu Tylora

Rysunek 2: *Wyznaczanie 3D maksimów*

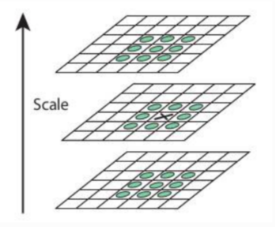


Źródło: Dokumentacja OpenCV

SIFT - Wyznaczanie kandydujących punktów kluczowych

- ...
- W praktyce: tak otrzymanych punktów kandydujących wciąż za dużo - stosujemy do otrzymanych kandydatów detektor narożników Harrisa
- Punkty, które spełniają kryterium narożnika detektora Harrisa uznajemy za poszukiwane punkty kluczowe

Rysunek 3: *Wyznaczanie 3D maksimów*

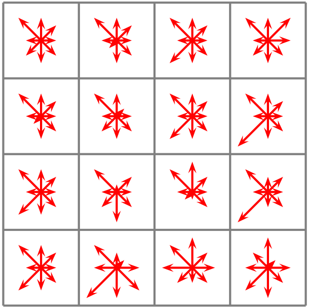


Źródło: Dokumentacja OpenCV

SIFT - niezmienniczość względem obrotu

- W otoczeniu punktu kluczowego budujemy 8-binowy histogram gradientów
- Znajdujemy kierunek dominujący w histogramie gradientów (maksimum)
- Sektor odpowiadający maksimum uznajemy za kierunek dominujący
- Jeśli w zakresie 80% globalnego maksimum znajdują się inne lokalne maksima to punkt traktujemy jako wielokrotny - z wieloma kierunkami dominującymi
- Przy dalszym wyznaczaniu deskryptora obracamy lokalnie obraz o kąt dominujący

SIFT - wyliczanie deskryptora



- Region zainteresowania - kwadrat 16x16 otaczający punkt kluczowy
- Region zainteresowania podzielony na fragmenty (tutaj: 16x4x4)
- W każdym subregionie 4x4 budujemy 8-binowy histogram gradientów
- Siła krawędzi (w histogramie) dodatkowo ważona odległością od środka
- Wektor cech – złączenie otrzymanych znormalizowanych histogramów (normalizacja L1 do 1)
- Razem $8 \times 16 = 128$ cech
- Dodatkowo współrzędne regionu zainteresowania.

Ciąg dalszy nastąpi!