

Laboratorium 6: Detekcja punktów zainteresowania i cechy lokalne

Operator SIFT

1. Wstęp

Bardzo często w przetwarzaniu obrazów mamy do czynienia z problemem rozpoznawania obiektów na podstawie cech lokalnych, czy też punktów/obszarów zainteresowania. Punkty te mogą zostać wykorzystane np. do wyszukiwania obrazów (porównywania obrazów) lub też do identyfikacji struktur podobnych na obrazach – np. identyfikacji podobnych struktur komórkowych na obrazach pozyskanych w badaniach mikrobiologicznych. Punkty zainteresowania powinny być odporne na transformacje geometryczne oraz potencjalną zmianę rozkładu jasności na obrazie, aby mogły być użyte do celów rozpoznawania (zob. rys. 1).



Rys. 1. Kandydat na 'punkt zainteresowania' (punkt narożny) wraz z jego otoczeniem, do potencjalnej identyfikacji obrazu 'Lena'. Odpowiednio liczny zbiór punktów kluczowych na obrazie zwiększy prawdopodobieństwo identyfikacji tego obrazu.

Poniżej główne cechy, które powinien spełniać właściwy punkt zainteresowania:

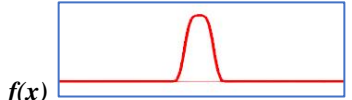

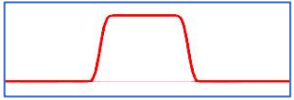
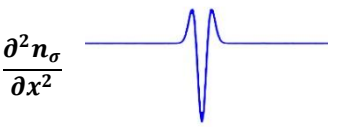


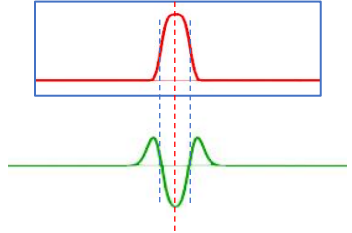
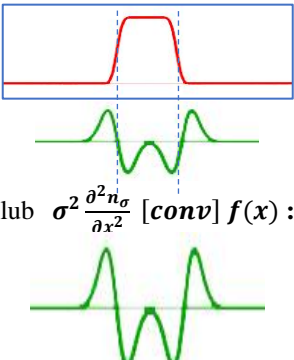
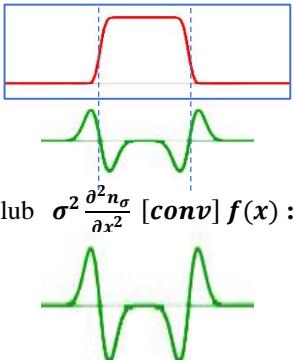
- Posiada odpowiednie otoczenie – tj. jego otoczenie powinno posiadać 'bogaty' w cechy kontekst, np. wysoką wariancję jasności lub kolorów, wysoką wartość entropii / miar rozmytości, wprowadzając pewnego rodzaju unikatowość tego obszaru,
- Odpowiednio określoną reprezentację (sygnaturę) – tj. w celu określania podobieństwa punktów, należy móc je opisać w odpowiedni sposób,
- Posiada dobrze określoną pozycję na obrazie,
- Powinien być odporny na podstawowe transformacje geometryczne,
- Powinien być odporny na zmiany jasności.

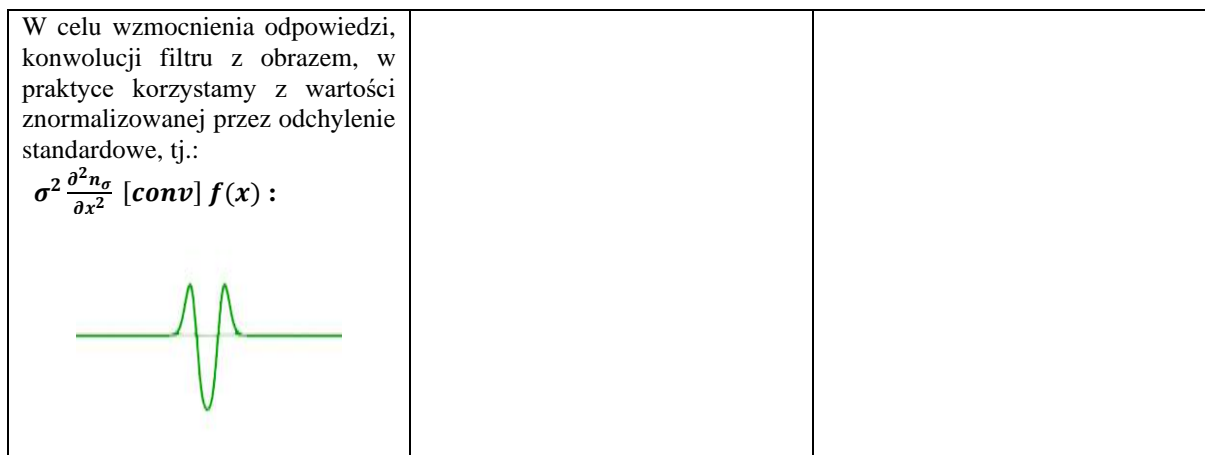
Z powyższych powodów, bardzo ciekawym rozwiązaniem w kierunku szukania cech identyfikujących obrazy na poziomie lokalnym są tzw. BLOB-y (ang. *image blobs*). Skrót BLOB oznacza *binary-large-object* i dotyczy opisu odpowiednio dużej porcji informacji jako całość, zazwyczaj multimedialnej. W przypadku przetwarzania obrazów, 'BLOB obrazu' związany jest z podobszarami obrazu o podobnej semantyce w sensie różnorodnej informacji obrazowej.

2. Laplasjan funkcji Gaussa (ang. Laplacian of Gaussian)

Bardzo ciekawym podejściem do detekcji krawędzi, a następnie również identyfikacji blobów obrazu jest laplasjan funkcji gaussowskiej, czyli wykorzystanie własności drugiej pochodnej funkcji gaussowskiej.

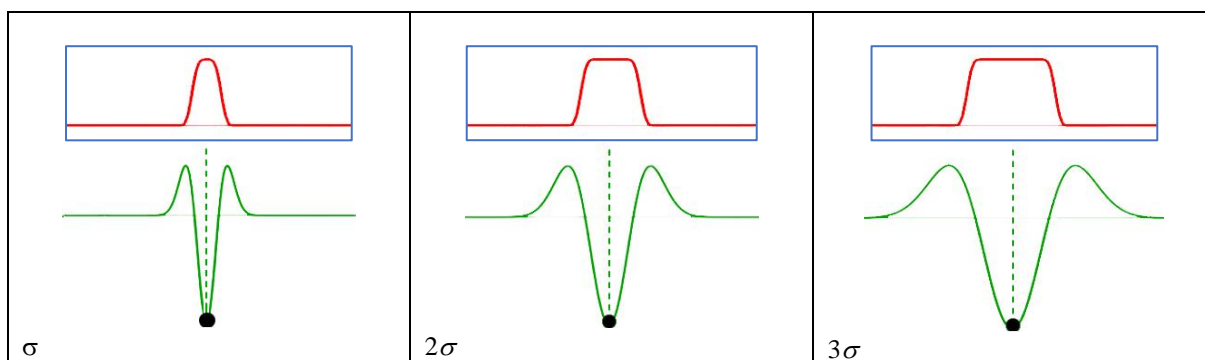
Już wcześniej poznaliśmy możliwości detekcji krawędzi za pomocą operatorów pierwszej i drugiej pochodnej obrazu. Aby zrozumieć wykorzystanie drugiej pochodnej funkcji Gaussa dla obrazu, rozważmy rysunek poniżej (rys. 2). Dla uproszczenia rozważamy przypadek 1D (wybrany wiersz obrazu).

 <p style="text-align: center;">Blob 1</p> <p>Założmy, że mamy blob o sygnaturze krawędzi tak jak powyżej (tj. obraz prezentujący przejście przez krawędź).</p>	 <p style="text-align: center;">Blob 2</p>	 <p style="text-align: center;">Blob 3</p>
 <p style="text-align: center;">$\frac{\partial^2 n_\sigma}{\partial x^2}$</p> <p>Postać drugiej pochodnej funkcji Gaussa (oznaczonej jako: n_σ).</p>		
 <p style="text-align: center;">$\frac{\partial^2 n_\sigma}{\partial x^2} [conv] f(x)$</p> <p>Konwolucja filtru z obrazem. Można zauważyć spełnienie własności przejścia przez zero, umożliwiające identyfikację krawędzi. Dodatkowo, blob może zostać zidentyfikowany (wartość x) poprzez znalezienie lokalnego ekstremum.</p>	 <p>lub $\sigma^2 \frac{\partial^2 n_\sigma}{\partial x^2} [conv] f(x) :$</p> <p>Wynik konwolucji jest zgodny z oczekiwaniem, niemniej jednak w przypadku 2 i 3, trudniej o jednoznaczną identyfikację blobu poprzez wartość ekstremum.</p>	 <p>lub $\sigma^2 \frac{\partial^2 n_\sigma}{\partial x^2} [conv] f(x) :$</p> <p>Wynik konwolucji jest zgodny z oczekiwaniem, niemniej jednak w przypadku 2 i 3, trudniej o jednoznaczną identyfikację blobu poprzez wartość ekstremum.</p>



Rys. 2. Ilustracja zastosowania operatora drugiej pochodnej funkcji Gaussa do identyfikacji krawędzi oraz ewentualnej lokalizacji blobu obrazu na podstawie wartości lokalnego ekstremum wyniku konwolucji obrazu z filtrem.

Ciekawym rozwiązaniem powyższego problemu, czyli możliwości identyfikacji pozycji blobu 2 i 3, stanowi możliwość zwiększenia wartości parametru odchylenia standardowego. Okazuje się, że można zidentyfikować odpowiednie wartości σ , dla których osiągniemy żądany efekt, co więcej wartość ta jest proporcjonalna do wielkości blobu. Tzn. jeśli przyjmiemy, że blob 2 stanowi podwojenie grubości krawędzi blobu 1, a blob 3 potrojenie, to uzyskamy żądany efekt dla wartości blobu 2 i 3, odpowiednio 2σ i 3σ (zob. rys. 3).



Rys. 3. Możliwość identyfikacji bloba, poprzez zwiększenie wartości parametru σ .

Podsumowując powyższe własności, możemy wprowadzić algorytm identyfikacji blobów na obrazach, za pomocą własności drugiej pochodnej funkcji Gaussowskiej:

Zakładając przypadek 1D :

Wejście: $f(x)$

1. Oblicz $\sigma^2 \frac{\partial^2 n_\sigma}{\partial x^2} [\text{conv}] f(x)$, dla różnych wartości parametru σ : $\{\sigma_0, \sigma_1, \dots, \sigma_k\}$ (w praktyce: $\sigma_k = \sigma_0 \cdot s^k$, gdzie s jest stałą oraz $k=0,1,2,\dots$),
2. Znajdź: $(x^*, \sigma^*) = \arg \max_{(x, \sigma)} \left| \sigma^2 \frac{\partial^2 n_\sigma}{\partial x^2} [\text{conv}] f(x) \right|$.

Wyjście: x^* - lokalizacja blobu; σ^* - wielkość blobu.

Powyższe rozumowanie można rozszerzyć do funkcji dwóch zmiennych w bardzo łatwy sposób, mianowicie poprzez zastosowanie laplasjanu funkcji Gaussa (Laplacian of Gaussian, LoG) lub znormalizowany (NLoG).

Laplasjan funkcji dwóch zmiennych: $\nabla^2 f(x, y) = \frac{\partial^2}{\partial x^2} f(x, y) + \frac{\partial^2}{\partial y^2} f(x, y)$,

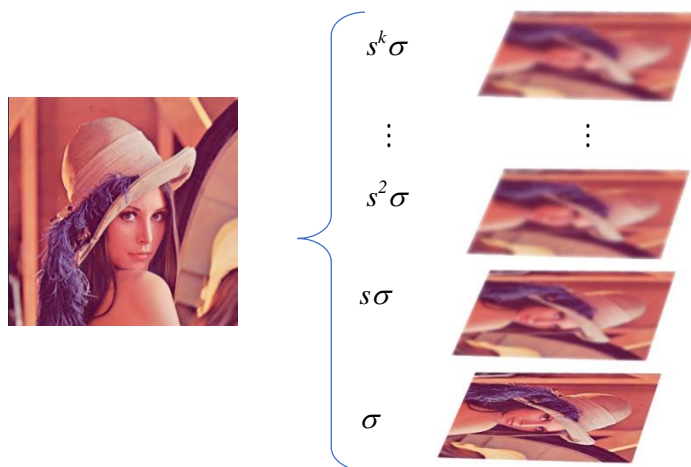
LoG: $\nabla^2 n_\sigma$,

NLoG: $\sigma^2 \nabla^2 n_\sigma$.

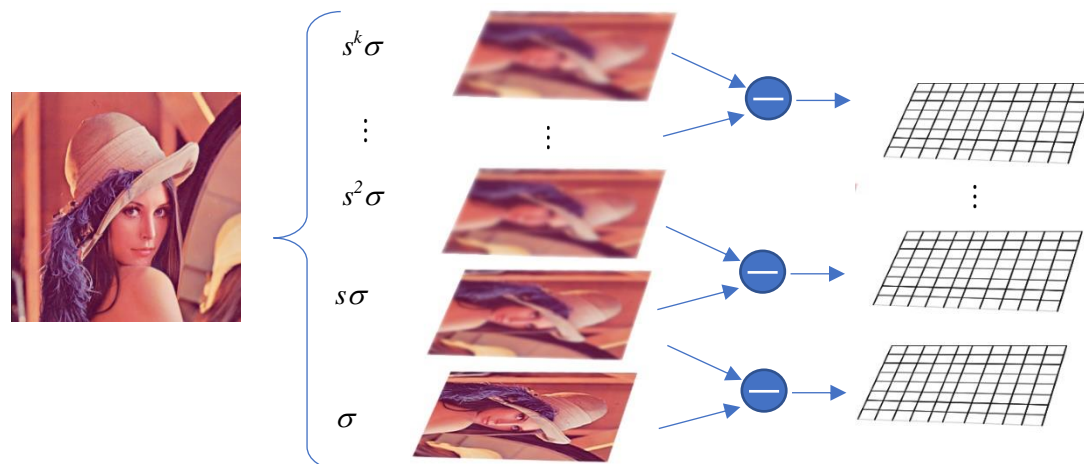
3. Operator SIFT (Scale-invariant feature transform)

Rozpoznanie blobów obrazu za pomocą operatora SIFT, wprowadzonego przez Davida G. Lowe'a [1] bezpośrednio wykorzystuje powyżej opisane własności. Poniżej przedstawiono proces identyfikacji obszarów zainteresowania obrazu za pomocą operatora SIFT:

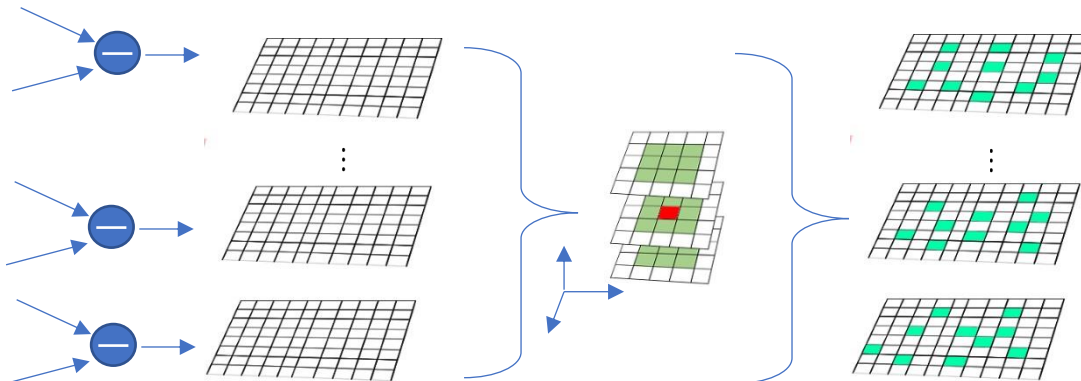
1. Reprezentacja obrazu dla różnych wartości parametru σ (stosując konwolucję z funkcją Gaussa dla różnych wartości σ):



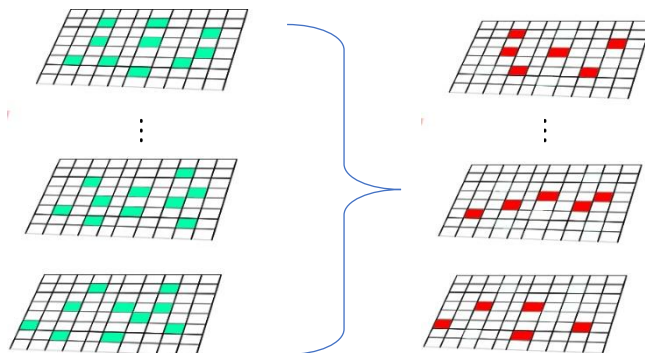
2. Zastosowanie różnicy obrazów (wykonanie operacji DoG - *difference of Gaussians*) aproksymującej NLoG, tj. $\text{DoG} = (s n_\sigma - n_\sigma) \approx (s - 1) | \text{NLoG} |$:



3. Znalezienie lokalnych wartości maksymalnych poruszając się maską $3 \times 3 \times 3$, tj. nie tylko dla każdej nowo wygenerowanej warstwy osobno, ale również uwzględniając skalę (kolejne warstwy dla przeskalowanej wartości σ):

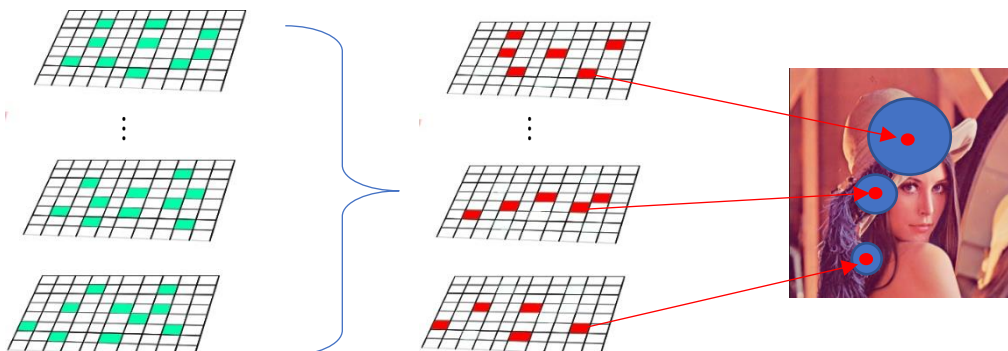


4. Progowanie powyższych wartości w celu uzyskania tylko ‘mocnych ekstremów lokalnych’.



5. Identyfikacja bloków.

Posiadając zbiór punktów reprezentujące bloby na obrazie, jesteśmy w stanie je zaprezentować oraz wykorzystać do dalszej analizy. Mamy nie tylko współrzędne bloba na obrazie, ale również jego wielkość odpowiadającą wykorzystanej skali parametru σ :



Poniżej, kilka przykładów prezentujących działanie operatora SIFT na obrazach (dla tego samego obrazu detekcja może się różnić, np. w zależności od przyjętej koncepcji wyboru lokalnych ekstremów, liczby warstw lub innych założeń) :



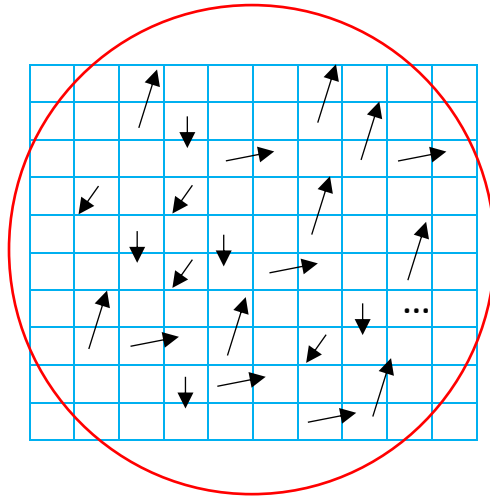
4. Operator SIFT - deskryptor

Rozpoznając błozy na obrazie za pomocą operatora SIFT, możemy następnie rozpocząć ich opis np. w celu parowania obrazów za pomocą punktów kluczowych. Należy zaznaczyć, że punkty kluczowe identyfikujące błozy za pomocą operatora SIFT spełniają kryteria punktów zainteresowania opisanych we wstępie. Zidentyfikowane błozy są odporne na zmianę skali – efekt zmiany skali można skorygować wykorzystując informacje dotyczące wielkości błoza. Efekt transformacji obrotu można skorygować rozpoznając kierunek dominującego wektora gradientu błoza. Analizując również tylko informacje o kierunku gradientów w błoze, uniezależniamy go od ewentualnych zmian rozkładu jasności na obrazie.

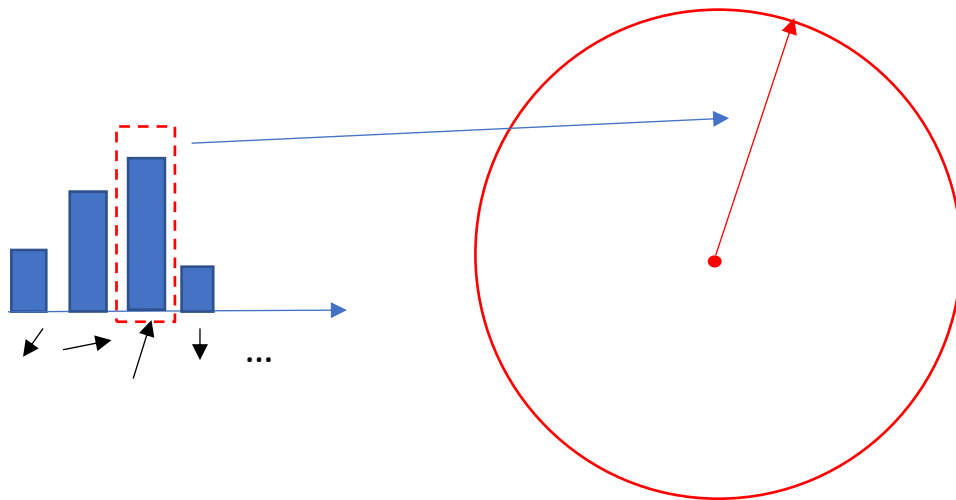
Posiadając tego typu bloby, można przygotować ich opis do dalszej analizy. Standardowo, opis dotyczy rozpoznania kierunków (orientacji) gradientów w blobie, co można wykorzystać do:

- Rozpoznania ‘kierunku dominującego’,
- Wykorzystania odpowiednich ‘stref’ bloba do analizy orientacji gradientów w strefie, oraz parowania punktów na podstawie odległości histogramów gradientów (uwzględniając tylko orientację gradientu).

Rozpoznania kierunku ‘dominującego’:



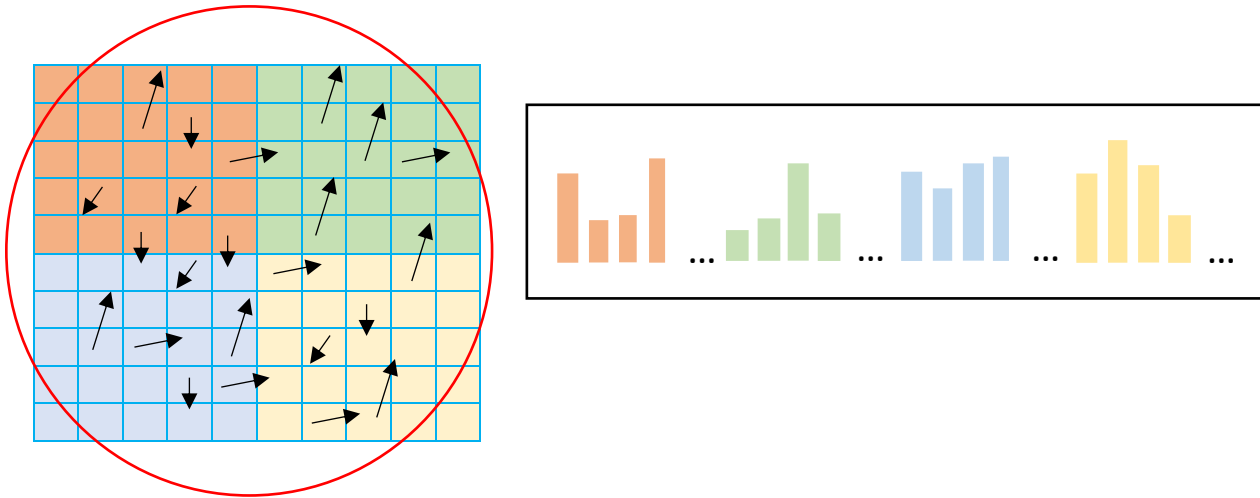
W każdym punkcie bloba możemy obliczyć kierunek wektora gradientu (patrz lab. ‘*detekcja krawędzi*’). Następnie, generując histogram ‘kierunków’ możemy wyznaczyć dominujący kierunek, który przyjmujemy za kierunek wyznaczony przez bloba (rys. 4):



Rys. 4. Dominujący kierunek rozpoznanego bloba.

Opis punktu kluczowego (sygnatura, deskryptor):

Do tego problemu można podejść w dość prosty sposób – podział na strefy analizowanego bloba w celu zwiększenia unikatowości bloba, wyznaczenie histogramu kierunków gradientów, a następnie konkatencja histogramów (*zob. rys. 5*). W taki sposób wyznaczymy deskryptor, który można użyć do parowania przy wykorzystaniu odpowiednich metryk dla histogramów.



Literatura:

[1] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* , 60, 2 (2004), pp. 91-110.