

Systemy Wizyjne

Opracowanie do zaliczenia cz. 2

Zakres : Lab 6 – Lab 11



Laboratorium 6

Zagadnienia :

- Segmentacja,
- Progowanie,
- Czynniki wpływające na jakość binaryzacji,
- Binaryzacja a progowanie,
- Typy binaryzacji,
- Binaryzacja globalna,
- Binaryzacja lokalna,
- Binaryzacja dwuprogowa

Pytania :

- Co to jest segmentacja ? Po co się ją stosuje ? Jakie są metody segmentacji ? (3 z lab.) Podaj trzy przykładowe kryteria podziału dla segmentacji.
- Progowanie - jakie są zalety metody, na podstawie czego się ją przeprowadza, jak działa ?
- Co ma największy wpływ na jakość binaryzacji i tym samym co może sprawić, że klasyczne metody mogą zawieść ? (2 czynniki)
- Czy binaryzacja i progowanie to jest to samo ? Jeśli nie, to czym się różnią ?
- Jakie znasz typy binaryzacji ? (2) Opisz je / ich działanie.
- Jakie znasz trzy metody wyznaczania progu w binaryzacji globalnej ? Opisz krótko te metody.
- Binaryzacja lokalna – Czym się różni od globalnej ? Jak działa w klasycznej postaci ? W jakich sytuacjach się stosuje ? Jak działa metoda Sauvola – czyli co dodatkowo wykorzystuje w stosunku do klasycznej metody ?
- Binaryzacja dwuprogowa – jak działa ?

Segmentacja

Jednym z najważniejszych etapów, podczas analizy obrazów, jest segmentacja - podział obrazu na rejony według pewnego kryterium (jasność, kolor, tekstura) . Obraz na etapie wstępnego przetwarzania to zbiór nieuporządkowanych pikseli. Obraz po segmentacji to zbiór obiektów.

Metody segmentacji

Metody segmentacji:

- **tresholding (progowanie)**
- **region growing (rozrost obszaru)**
- **region splitting (podział obszaru)**
- **„Split & Merge”**
- **edge linking**
- **edge following**
- **pyramid edge detection**

Na laborkach były : Przez rozrost, binaryzacja przez prog. , podział i scalanie

Progowanie i binaryzacja

Najprostszą (i też najczęściej wykorzystywaną) metodą segmentacji jest binaryzacja. Do jej głównych zalet zalicza się: intuicyjność, prostotę, łatwość implementacji i szybkość wykonywania. Jest ona etapem wielu algorytmów analizy obrazów. Pozwala na znaczną redukcję informacji w obrazie (np. z zakresu 0-255 do 0-1).

Binaryzacja najczęściej realizowana jest poprzez progowanie. Na przykład: dla obrazu w odcieniach szarości ustala się próg na 'k'. Wszystkie piksele o wartości większej od 'k' zostają uznane za obiekty istotne, a pozostałe za tło. Możliwe jest też ustalenie dwóch progów (binaryzacja dwuprogowa, o tym kilka linijek niżej).

Binaryzację najczęściej wykonuje się na podstawie histogramu. Podstawą określania progu binaryzacji jest zazwyczaj analiza histogramu ("ręczna" lub automatyczna).

Na łatwość przeprowadzenia binaryzacji największy wpływ mają dwa czynniki:

- szum
- niejednorodne oświetlenie

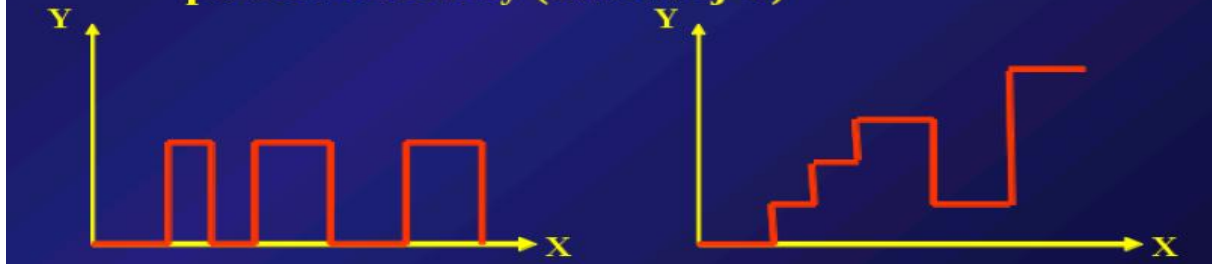
Należy pamiętać, że progowanie i binaryzacja to nie to samo !

Progowanie

- zakłada użycie pewnej liczby progów i dotyczy dziedzin (zmiennych X)

Binaryzacja

- określa liczbę poziomów na wyjściu i dotyczy przeciwdziedzin (zmiennej Y)



Czyli binaryzacją określa się sam proces zamiany skali obrazka np. z 0-255 na 0-1. Nie jest powiedziane, że trzeba to zrobić progowaniem, ale progowanie jest jedną z metod (najczęściej) wykorzystywanych do binaryzacji.

Typy binaryzacji :

Binaryzacja

globalna = stały próg dla całego obrazu

lokalna = próg ustalany niezależnie dla fragmentów obrazu

W automatycznym systemie rozpoznawania obrazów potrzebne są metody, które same potrafią wyznaczać próg binaryzacji. Taką metodą jest

Metoda Otsu

Dokonuje ona podziału histogramu na dwie klasy, tak aby wariancja pomiędzy tymi klasami była minimalna.

Metoda Kittler

Minimalizacja funkcji błędu zbudowanej z dystrybuanty prawdopodobieństwa i odchylenia standardowego, koncepcyjnie podobna do Otsu.

Metoda Yen

Maksymalizacja entropii obiektów i tła.

Istnieje kilkadziesiąt metod wyznaczania tego progu. Powyższe metody wykorzystują binaryzację globalną.

Globalna – dla wszystkich pikseli jeden próg

Binaryzacja lokalna – dla każdego piksela osobny próg

Analiza wyników binaryzacji dla obrazów "rice.png" i "katalog.bmp" pokazuje, że globalna binaryzacja nie najlepiej działa dla obrazów o niejednorodnym oświetleniu. Dla obu obrazów trudno również wyznaczyć odpowiedni próg "ręcznie".

Metodą, która pozwala poprawić wyniki binaryzacji, jest binaryzacja lokalna (niekiedy zwana adaptacyjną). Polega ona na wyznaczeniu progu osobno dla każdego piksela na podstawie jego otoczenia.

Przyjmujemy okno o ustalonym rozmiarze iluś pikseli. Najprostsza wersja binaryzacji lokalnej zakłada, że próg w danym oknie to średnia z pikseli w danym oknie.

Jakość działania binaryzacji lokalnej można poprawić wyznaczając próg za pomocą metody Sauvola. Wykorzystuje ona, oprócz średniej, informację o odchyleniu standardowym w danym oknie.

$$T = \text{średnia} \cdot [1 \pm k(\text{odch.st}/R - 1)]$$

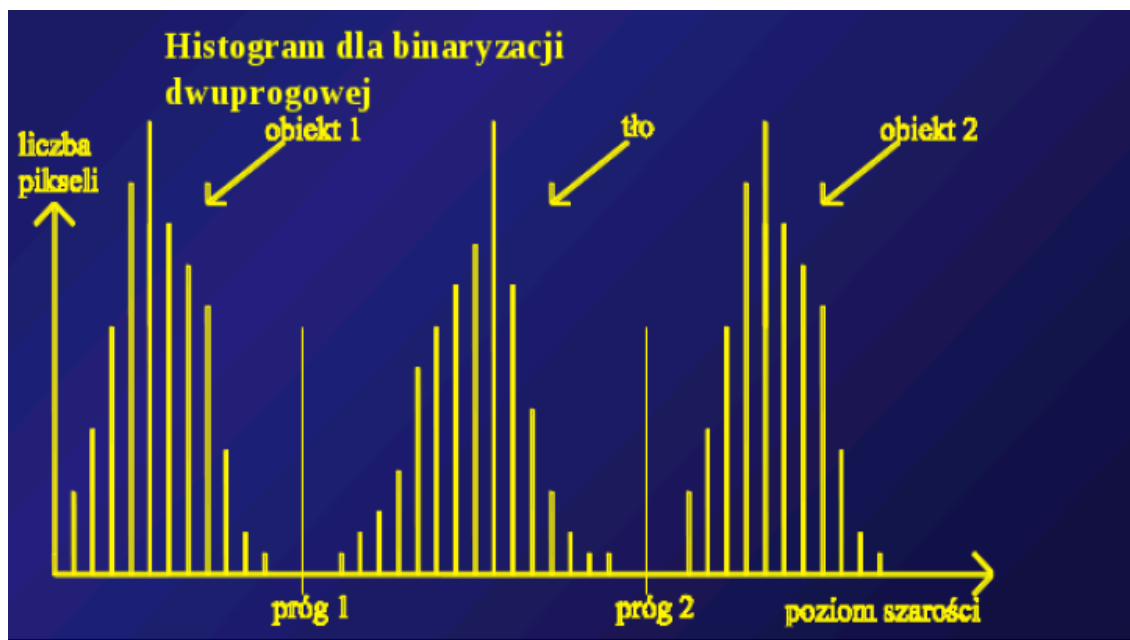
gdzie k i R to parametry (R zwykle 128, a k na początek przyjmij 0.15).

Binaryzacja dwuprogowa

Binaryzację można przeprowadzić wykorzystując dwa progi - wybieramy w ten sposób przedział jasności (piksele w nim zawarte klasyfikujemy jako obiekty).

W matlabie istnieje opcja podglądu wartości piksela, dzięki temu można wyznaczyć odpowiednie progi (zakres).

Przykład



Zadanie dodatkowe - binaryzacja z wykorzystaniem średniej ruchomej. (można uzupełnić)

Laboratorium 7

Zagadnienia :

- Segmentacja przez rozrost

Pytania :

- Powstanie czego w obrazie powoduje segmentacja ?
- Jaki jest algorytm segmentacji przez rozrost i jakie ma dwa miejsca które trzeba sprecyzować ?
- Jak można przyjąć kryteria podobieństwa w algorytmie segmentacji przez rozrost ? W jaki sposób można porównywać piksele ?

Proces segmentacji możemy określić jako operację, która powoduje, że:

- każdy piksel obrazu zostaje przypisany do jednego ze zbiorów wynikowych
- powstają rejony stykających się ze sobą (sąsiadujących wg sąsiedztwa 4 lub 8-punktowego) pikseli
- każde dwa regiony są rozłączne (nie mają wspólnych pikseli)
- piksele w ramach jednego regionu są do siebie podobne względem jakiejś własności (jasność, kolor, tekstura)
- piksele sąsiadujących regionów różnią się względem własności z poprzedniego punktu (jasność, kolor, tekstura)

Segmentacja przez rozrost

Algorytm segmentacji przez rozrost jest następujący:

1. wybieramy piksel startowy, oznaczamy go jako należący do nowej klasy
2. Sprawdzamy, czy sąsiedzi, nie przypisani do żadnej klasy, są podobni do piksela startowego (wg wybranej własności / kryterium : jasność / kolor / tekstura)

3. Sąsiedzi podobni są oznaczani jako należący do tej samej klasy co piksel startowy
4. Sąsiedzi podobni stają się nowymi pikselami startowymi
5. Jeżeli w obrazie pozostały piksele nieprzypisane do żadnej klasy to zaczynamy od punktu 1.

Powyższy algorytm ma dwa niejasno sprecyzowane miejsca. Po pierwsze - jak wybrać punkt startowy? Po drugie co to znaczy podobne?

Teoretyczna odpowiedź na pierwsze pytanie - dowolny piksel może stać się startowym (można go np. wylosować). Oznacza to, że metoda przez rozrost dla tego samego obrazu może dać różne wyniki w zależności od wyboru / wylosowania pikseli startowych. I tak już musi zostać :)

Natomiast odpowiedź na drugie pytanie jest bardziej złożona. Należy przyjąć jakieś kryterium jednolitości. Można wybrać pewną stałą wartość dla danej klasy (np. wziętą z piksela startowego) i dołączać piksele różniące się od niej o z góry zadaną wartość.

Można też liczyć nie względem piksela startowego, ale względem aktualnie rozpatrywanego. Wtedy wartość dla klasy nie jest stała - zmienia się z kolejno przyłączanymi pikselami. Można też jako wartość odniesienia przyjąć średnią z otoczenia piksela (podobnie jak robiliśmy przy binaryzacji lokalnej).

Laboratorium 8

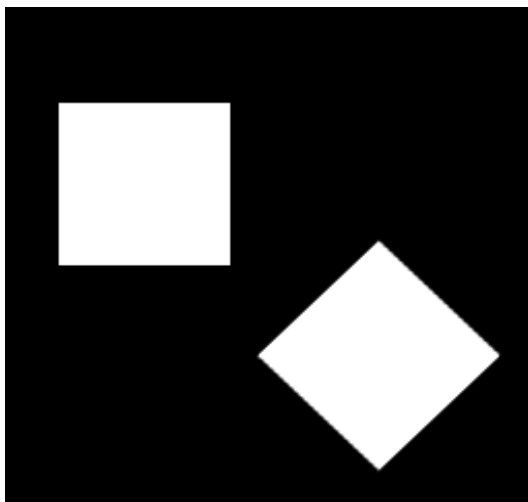
Zagadnienia :

- Indeksacja z tablicą sklejeń,
- Tablica sklejeń,
- Współczynniki kształtu

Pytania :

- Indeksacja z tablicą sklejeń – Podaj algorytm indeksacji – 2 kroki + 3 możliwości przy nadawaniu etykiety. Po co się ją stosuje ? Następnym krokiem, który jest wykonywany przy robieniu indeksacji jest zapis etykiet do Ostatnim krokiem algorytmu jest drugie przejście po obrazie i scalenie etykiet o tej samej
- Czym jest tablica sklejeń, jaki ma rozmiar i w jaki sposób jest wypełniana?
- Jakie są parametry obiektów i jak je wyznaczać ? O czym należy przy wyznaczaniu pamiętać ?
- Czym są współczynniki kształtu ? Jak je wyznaczyć ? Od czego zależą ?
Jakie znasz współczynniki kształtu (4 z lab. + 1 z wykł.) ? Jak na ich podstawie wyznaczyć dany kształt z obrazu ?

Większość dotychczas poznanych i wykorzystywanych algorytmów przetwarzania obrazu wykonywała operacje na całym obrazie (lub co najwyżej dokonywany był podział obiekt/obiekty - tło).



Patrząc na powyższy obraz człowiek widzi dwa kwadraty. Dla automatycznego systemu rozpoznawania obrazu przejście od etapu obiekty/tło do wyróżnienia dwóch kwadratów nie jest takie oczywiste i wymaga zastosowania jakiejś formy indeksacji - czyli przypisania każdemu z pikseli (tu białych) jakiejś etykiety.

W tej części ćwiczenia zaprezentowany zostanie algorytm indeksacji z tablicą sklejeń.

Opis algorytmu:

- 1) Na wejściu mamy obraz po binaryzacji - zakładamy, że piksele białe ('1') należą do obiektów, tło jest czarne (0).
- 2) W pierwszej iteracji obraz analizujemy linia po linii, aż do napotkania pierwszego piksela o wartości '1'. W tej sytuacji dokonujemy analizy otoczenia (otoczenia można przyjąć 4-elementowe, czyli boki bez rogów lub 8-elementowe, czyli boki z rogami) piksela:

A	B	C
D	X	

X - rozważany piksel

A,B,C,D - otoczenie piksela

L - etykieta

Możliwe są następujące przypadki dla otoczenia A,B,C,D :

- a) wszystkie należą do tła. Wtedy znaleziony piksel X należy do nowego obiektu - nadajemy mu zatem etykietę $L+1$ ($X = L+1$) - przez L rozumiemy poprzednią etykietę. (przed uruchomieniem algorytmu L należy zainicjować wartością 1).
- b) jeden lub więcej pikseli ma przypisaną aktualną etykietę L. Wtedy rozważanemu pikselowi przypisujemy etykietę L ($X = L$)
- c) w otoczeniu występują piksele o różnych etykietach np. L1 i L2. Wtedy przyjmuje się zasadę, że rozważanemu pikselowi X przypisuje się mniejszą z wartości L1 i L2 ($X = \min(L1, L2)$)

W wyniku indeksacji uzyskujemy obraz, na którym wyróżnionych jest więcej obiektów niż w rzeczywistości się na nim znajduje. Dlatego konieczny jest drugi etap indeksacji wykorzystujący tzw. tablicę sklejeń.

Tablica sklejeń

Tablica sklejeń powinna mieć rozmiar równy liczbie etykiet, które mogą wystąpić na obrazie (w naszym wypadku 255 - zera nie są etykietowane).

Na początku tablicę inicjujemy wartościami 0. U góry adresy – możliwa liczba etykiet. (Nie jestem pewny, ale liczba tych etykiet, to chyba liczba białych pikseli na obrazie, w skrajnym przypadku każdy ma inną etykietę ?).

1	2	3	4	...	255
0	0	0	0	...	0

W tablicy zapisujemy następującą informację:

- w sytuacji gdy dodajemy nową etykietę (przypadek a)) w tablicy pod "adresem" etykiety zapisujemy "wartość" etykiety np.

1	2	3	4	...	255
1	0	0	0	...	0
1	2	3	4	...	255
1	2	3	4	...	0

w sytuacji wystąpienia konfliktu (przypadek c)) w tablicy sklejeń zapisujemy pod "adresem" wyższej z etykiet wartość etykiety niższej:

1	2	3	4	...	255
1	1	3	4	...	0

Zapis tej informacji wykonujemy podczas pierwszej iteracji algorytmu indeksacji. Następnie wykonujemy drugą iterację po obrazie (wstępnie poetykietowanym) i zamieniamy etykiety zgodnie z tablicą sklejeń. W wyniku tej operacji powinniśmy otrzymać obraz poprawnie poindeksowany.

W matlabie istnieje funkcja, która pozwala wykonać indeksację, jest to *bwlabel*.

Współczynniki kształtu

Parametry lokalne - dotyczące obiektów:

- powierzchnia obiektu
- średnica obiektu
- długość obwodu

Parametry globalne - dotyczące obrazu:

- liczba obiektów na jednostkę powierzchni
- udział procentowy wybranych elementów obrazu

Liczba obiektów:

- wynik indeksacji/etykietowania
- uwzględnienie „ramki bezpieczeństwa”

Powierzchnia obiektu:

- liczba punktów
- pole zawarte pomiędzy środkami pikseli

Obwód:

- liczba punktów konturu zewnętrznego
- uwzględnienie narożników
- średnia z długości mierzonych wewnątrz i na zewnątrz
- aproksymacja wielokątem lub krzywymi

Jeśli obwód wyznaczamy na podstawie liczby pikseli na obwodzie, to zawsze wyjdzie nam mniejszy niż w rzeczywistości więc należy pamiętać o narożnikach.

Współczynniki kształtu są pewnymi parametrami liczbowymi opisującymi kształt obiektu. Pozwala to na użycie ich do automatycznego rozpoznawania obiektów. Teoretycznie współczynniki kształtu (dobre) powinny być niezależne od zmiany położenia obiektu, jego orientacji i wielkości.

W praktyce okazuje się jednak, że wartości współczynników zmieniają się w pewnym zakresie dla tych samych obiektów. Wynika z tego konieczność uwzględnienia pewnej tolerancji wartości współczynników kształtu wykorzystywanych w praktycznych zastosowaniach analizy i rozpoznawania obrazu.

W ćwiczeniu wykorzystane zostaną następujące współczynniki, każdy współczynnik nadaje się do czegoś innego najlepiej, np. jeden najlepiej wykrywa koła inny trójkąty itd. :

Compactness - stosunek pola obiektu do pola najmniejszego prostokąta w jakim się obiekt mieści

R_{min}/R_{max} - pierwiastek stosunku minimalnej odległości konturu od środka ciężkości do maksymalnej odległości konturu od środka ciężkości

Blair - Bliss - $S / \sqrt{2 * \pi * \sum r_i^2}$

gdzie S - pole powierzchni obiektu, r_i - odległości pikseli obrazka od środka ciężkości, i - numer piksela obiektu.




Haralick - pierwiastek z $\frac{\sum d_i^2}{[(n * \sum d_i^2) - 1]}$ gdzie d_i – odległości pikseli konturu od środka ciężkości, n - liczba pikseli konturu.

Inne popularne :

Współczynnik Malinowskiej:

$$R_M = \frac{L}{2 \cdot \sqrt{\pi \cdot S}} - 1$$

L - obwód obiektu
S - pole powierzchni obiektu

		
W3=0	W3=0.116	W3=0.389

Po wyznaczeniu współczynników aby wysegmentować określony obiekt należy określić przedziały dla danego współczynnika.

Laboratorium 9

Zagadnienia :

- Segmentacja przez podział i scalanie

Pytania :

- Jaka jest fundamentalna różnica pomiędzy algorytmem segmentacji przez rozrost, a przez podział i scalanie ? (Chodzi o początek algo.)
- Podaj algorytm segmentacji przez podział i scalanie.
- Kiedy 2 obszary uznamy za jednorodne ? Jakie jest kryterium ?

Segmentacja przez podział i scalanie

Algorytm

Poprzedni algorytm segmentacji (przez rozrost) 'startował' z pojedynczego piksela dochodząc do całego obrazu. Obecnie implementowany algorytm zaczyna od obrazu kończąc na małym obszarze (teoretycznie można 'dojść' aż do pojedynczego piksela, ale nie jest to ani konieczne ani efektywne)

Algorytm segmentacji przez podział i scalanie jest następujący:

1. Dzielimy obszar na 4 podobszary (w pierwszym kroku obszarem jest cały obraz) prowadząc prostą pionową i poziomą przez środek obszaru.
2. Sprawdzamy, czy podobszary są jednorodne wg przyjętego kryterium.
3. Jeśli obszar P_i nie jest jednorodny, a jest większy od przyjętego minimum to przejdź do punkt 1. Jeśli obszar jest jednorodny, to idź do punktu 4.
4. Sprawdź czy obszary sąsiednie są 'jednorodne', jeśli tak to wtedy należy je połączyć (faza łączenia)

Kiedy 2 obszary są jednorodne ?

Można policzyć wybraną własność dla pikseli obu obszarów (czyli dla sumy obszarów) i jeżeli kryterium dla pikseli obu obszarów jest spełnione to obszary uznajemy za jednorodne (łączymy je).

Można także wybrać 'słabszy' wariant (w sensie siły sprawdzania) - jeżeli własności każdego z obszarów z osobna niewiele się różnią od siebie to obszary uznajemy za jednorodne. W naszej implementacji użyjemy 'słabszego' wariantu. Za kryterium jednorodności w implementacji można uznać np. odchylenie standardowe od średniego odcienia koloru.

Laboratorium 10

Zagadnienia :

- Przekształcenie Fouriera dla obrazów cyfrowych
- F obraz (amplitudy i przesunięcia fazowego)

Pytania :

- Gdzie na obrazie fazy i amplitudy są wysokie, a gdzie niskie częstotliwości ?
- Jak z wykresu amplitudy odczytywać fale z których składa się obraz oryginalny ?
- Do czego może być przydatna t. Fouriera ? (2)
- Jak jest realizowana filtracja w dziedzinie częstotliwości ?
- W jakich przypadkach ta filtracja może być lepsza od zwykłej ?
- Jakiego typu filtry można zastosować w dziedzinie częstotliwości ?

(Najpierw przeczytać poniższą teorię, w trakcie czytania będą odnośniki do linków poniżej, które mogą być pomocne w zrozumieniu, ale nie muszą)

Link 1. Obraz jako złożenie fal :

<https://www.youtube.com/watch?v=mEN7DTdHbAU>

Jak odczytywać wykresy + parametry fali :

<https://www.youtube.com/watch?v=xhO8iz2qCOE>

W transformacie Fouriera zakładamy, że obraz jest sumą fal różniących się od siebie parametrami (objaśnione dobrze w link nr 1). Parametry tych fal to :

- Amplituda,
- Faza,
- Częstotliwość,
- Orientacja,

(Wszystkie powyższe parametry są dobrze objaśnione w link 2)

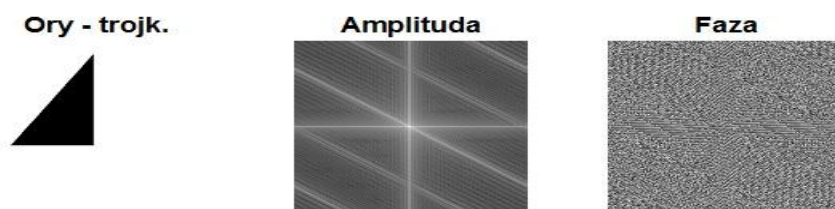
Po wykonaniu dwuwymiarowej transformaty Fouriera, która dana jest wzorem:

Dla obrazu cyfrowego o rozmiarach $M \cdot N$ pikseli jego transformatę Fouriera określa się następująco:

$$F(i, k) = \beta_L \cdot \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} L(m, n) \cdot \exp\left(\frac{-j \cdot 2 \cdot \pi \cdot m \cdot i}{M}\right) \cdot \exp\left(\frac{-j \cdot 2 \cdot \pi \cdot n \cdot k}{N}\right)$$

$$\text{dla } i=0,1,\dots,M-1; k=0,1,\dots,N-1$$

Otrzymujemy graficzną reprezentację fali z których składa się obraz.
Przykładowo :



Co przedstawiają te wykresy ? W środku każdego wykresu mamy (punkt 0,0) określający zerową częstotliwość i zerowe przesunięcie. Jeśli ze środka wykresu poprowadzimy wektor do dowolnego punktu (piksela) obrazu to kąt jaki ten wektor tworzy z osią x będzie określał orientację fali, natomiast długość tego wektora będzie określał wartość częstotliwości, a wartość piksela amplitudę lub przesunięcie fazowe w równaniu fali.

Można więc powiedzieć, że każdy piksel to osobna fala (objaśnione dobrze w link nr 2).

Oczywiście mniejsze częstotliwości znajdują się blisko środka wykresu, a większe przy krajach.

Wyróżniające się na wykresie amplitudy punkty odpowiadają falom tworzącym obraz. (Czarne punkty to amplituda = 0, jeśli amplituda fali równa się 0 to nie jest to tak naprawdę fala)

Z wykresu amplitudy można wnioskować o położeniu krawędzi na obrazie. Np. na powyższym wykresie amplitudy widać dwie linie poziomą i pionową, które odpowiadają krawędziom poziomej i pionowej trójkąta (orientacja tych fali ma kąt 0 stopni lub 90 stopni). Natomiast krawędź pod kątem 45 stopni jest reprezentowana przez linie ukośne.

Możliwe jest również wykonanie odwrotnej transformaty Fouriera

Operacja dwuwymiarowej odwrotnej transformacji Fouriera zdefiniowana jest wzorem :

$$L(m,n) = \beta_F \cdot \sum_{i=0}^{M-1} \sum_{k=0}^{N-1} F(i,k) \cdot \exp\left(\frac{j \cdot 2 \cdot \pi \cdot m \cdot i}{M}\right) \cdot \exp\left(\frac{j \cdot 2 \cdot \pi \cdot n \cdot k}{N}\right)$$

$$\text{dla } m=0,1,\dots,M-1; n=0,1,\dots,N-1$$

, która pozwala przejść z postaci po przekształceniu transformatą na obraz oryginalny.

Wszystko fajnie, ale po co ta transformata ? Pozwala ona na filtrację dolno i górno przepustową z wykorzystaniem masek konwolucji. Odbywa się to zgodnie ze wzorkiem :

$$g(x,y) * f(x,y) = G(u,v) F(u,v)$$

Po lewej maska i obraz oryginalny, po prawej odpowiednio transformaty Fouriera.

(* oznacza konwolucję), po prawej mnożenie w dziedzinie częstotliwości

Kilka uwag:

- Operacja filtracji w dziedzinie częstotliwości może okazać się bardziej efektywna w niektórych przypadkach - a dokładniej jeżeli operacje fft2 i ifft2 zajmą mniej czasu niż klasyczna konwolucja (zazwyczaj dla dużego obrazu, z dużą maską).
- Sama filtracja w dziedzinie częstotliwości to mnożenie całego obrazu przez jedną maskę

- W przypadku filtracji w dziedzinie częstotliwości zakłada się, że obraz "zawija się" na brzegach - co powoduje pewne artefakty
- W dziedzinie częstotliwości "działają" tylko filtry określane jako liniowe - filtry medianowe, maksymalne, minimalne itp. nie mają swoich odpowiedników.

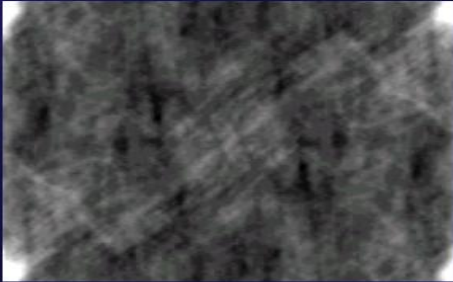
Interpretując filtrację górno i dolno przepustową na F obrazie, filtracja dolnoprzepustowa będzie się wiązać z pozostawieniem punktów bliskich środka obrazu, a górnoprzepustowa z pozostawieniem punktów bliskich krawędziom.


Sposób obliczania amplitudy i fazy :

TRANSFORMATA FOURIERA

PRZYKŁADY

Amplituda a faza

$$A(i, k) = |F(i, k)|$$


$$\phi(i, k) = \frac{F(i, k)}{A(i, k)}$$


Laboratorium 11

Zagadnienia :

- Transformata Hougha

Pytania :

- Co to jest transformata Hougha ?
- Po co jest stosowana ?
- Czym jest przestrzeń Hougha ?
- Jak reprezentowana jest prosta w przestrzeni Hougha a jak pęk prostych ?
- Gdzie na obrazie przestrzeni Hougha szukać krawędzi ?
- Opisz ogólnie jaki jest algorytm tej metody.
- Czy sposób reprezentacji prostej w tej transformacji ma znaczenie ? Jeśli tak, to która reprezentacja jest lepsza i dlaczego jedna z nich nie działa poprawnie ?

Obrazy do opracowania pochodzą z poniższego linku w którym jest bardzo fajnie wyjaśniona idea Transformaty Hougha. W opracowaniu przepisywałem trochę z instrukcji z lab. , a trochę to co usłyszałem w filmiku. Link poniżej.

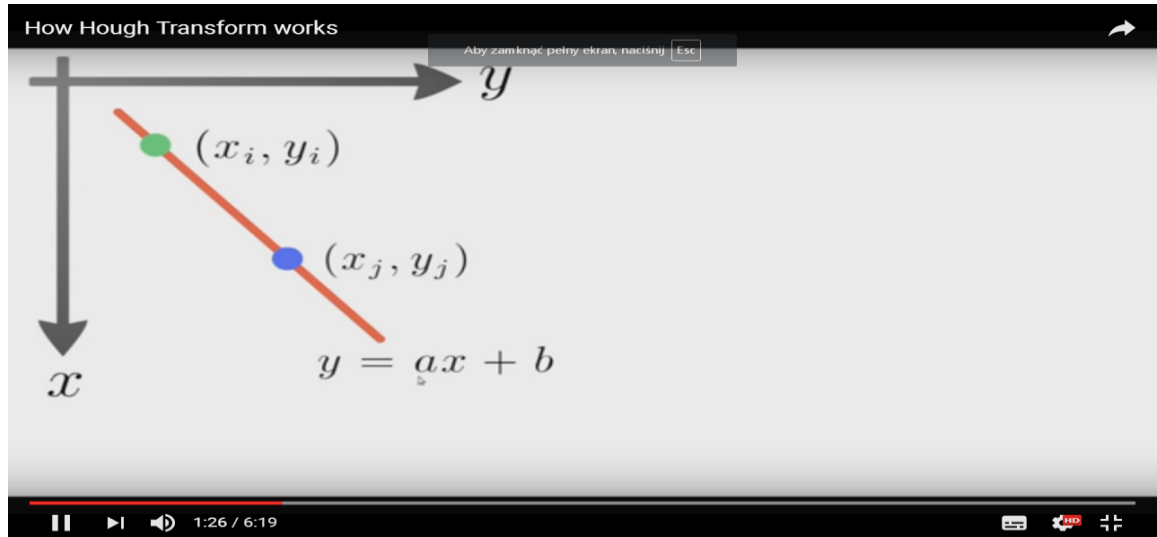
Link 1 :

<https://www.youtube.com/watch?v=4zHbl-fFII>

Transformata Hougha (Hafa) jest to metoda pozwalająca na wykrycie krawędzi w obrazie. Aby móc zastosować transformatę Hougha najpierw należy doprowadzić obraz do postaci $0 - 1$.

Idąc od początku...

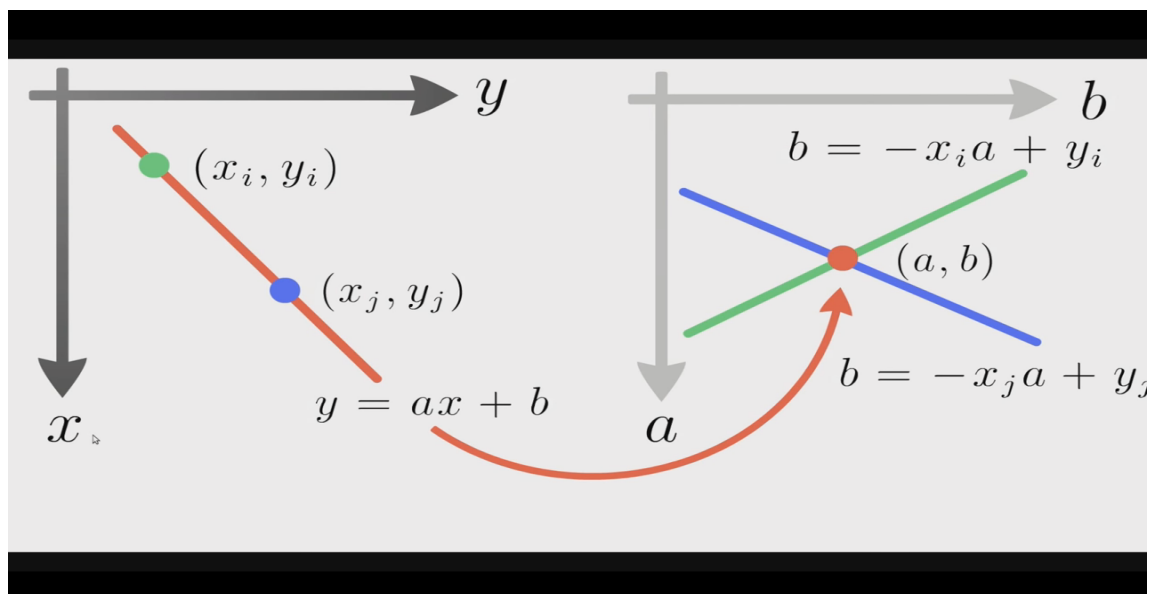
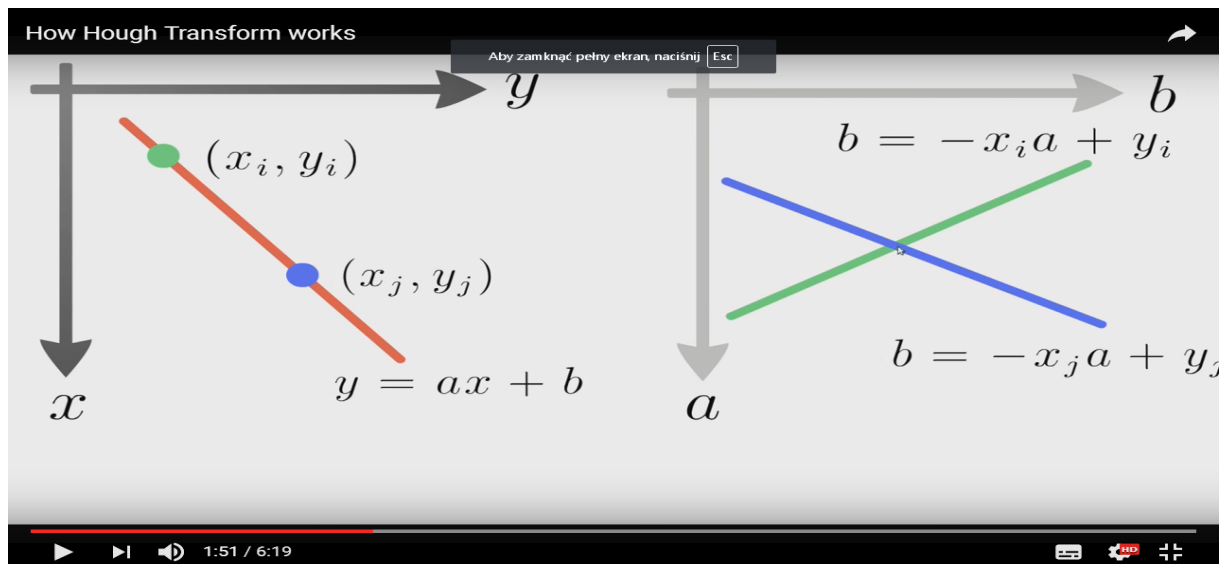
Prosta na obrazie może być zapisana klasycznym równaniem jak poniżej, czyli $y = ax + b$ (układ na rysunkach jest przekręcony o -90 stopni) :



Rozpatrzmy teraz układ współrzędnych z osiami opisującymi parametry a i b odpowiadającymi współczynnikom równania prostej $y = ax + b$. Jak na takim układzie współrzędnych można przedstawić naszą prostą ?

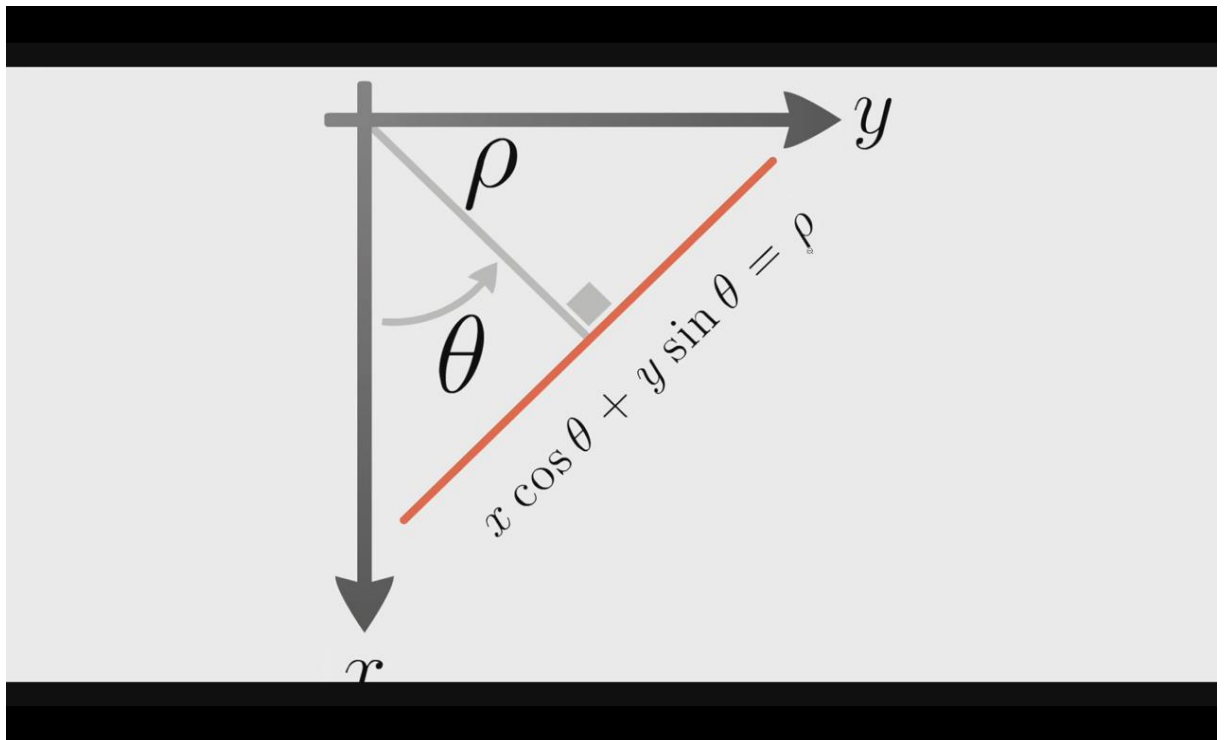
A no w prosty sposób, bo widać, że prosta taka jak na powyższym rysunku ma ściśle określoną parę parametrów (a, b) , czyli istnieje tylko jedna taka para parametrów a i b dla której można narysować taką prostą. Jak zatem przedstawić taką prostą w układzie współrzędnych z osiami a i b ? A no wystarczy zaznaczyć punkt określający jednoznacznie parametr a i parametr b .

Dobrze, a teraz pomyślmy jak w takim układzie współrzędnych z osiami a i b przedstawić punkt ? Np. patrząc na poniższy rysunek zakładamy, że chcemy przedstawić w układzie a, b punkt zielony (x_i, y_i) . Jak to zrobić ? A no zaznaczamy wszystkie możliwe parametry wszystkich możliwych prostych, które przechodzą przez ten punkt. Takie zaznaczenie tworzy prostą w układzie a, b . Analogicznie zaznaczamy punkt niebieski (x_j, y_j) w układzie a, b . Punkt przecięcia tych prostych określa parametry a, b które odpowiadają prostej przechodzącej przez punkt zielony i niebieski.



Jeśli mielibyśmy mnóstwo punktów na prostej na której znajduje się punkt zielony i niebieski, to nadal taka krawędź byłaby jednym tym samym punktem w układzie a, b .

Na zajęciach zaczynamy od innej postaci prostej w której 'ro' oznacza odległość prostej od początku układu współrzędnych , a teta oznacza kąt pomiędzy prostą a osią OX :



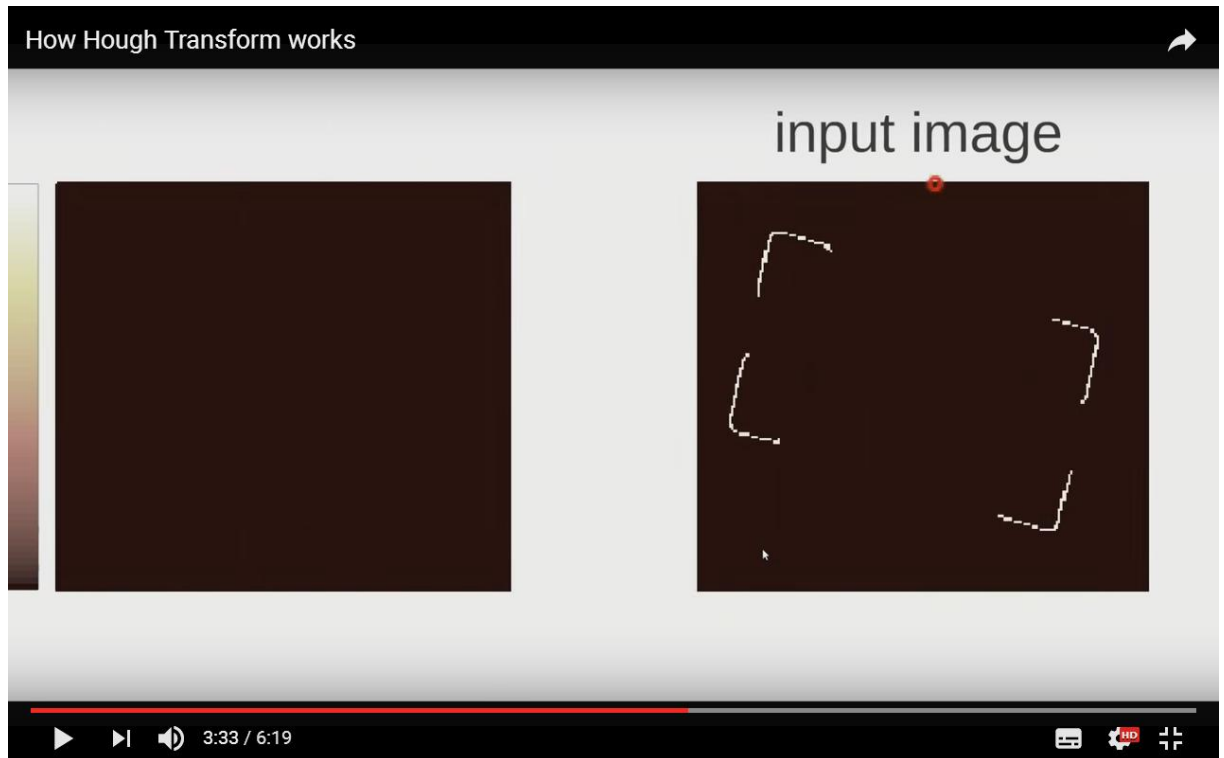
Po co taka reprezentacja ? Przecież to tylko komplikowanie sobie życia. Zauważyć należy, że taka reprezentacja jest w stanie reprezentować prostą równoległą do osi OY czego nie jesteśmy w stanie zapisać w reprezentacji $y = ax + b$. Ma to duże znaczenie przy algorytmie transformacji Hougha.

Wiemy więc już, że prostej w przestrzeni kartezjańskiej odpowiada punkt w przestrzeni Hougha (przestrzeni parametrów teta i ro) . Punktowi z przestrzeni kartezjańskiej odpowiada prosta w przestrzeni Hougha.

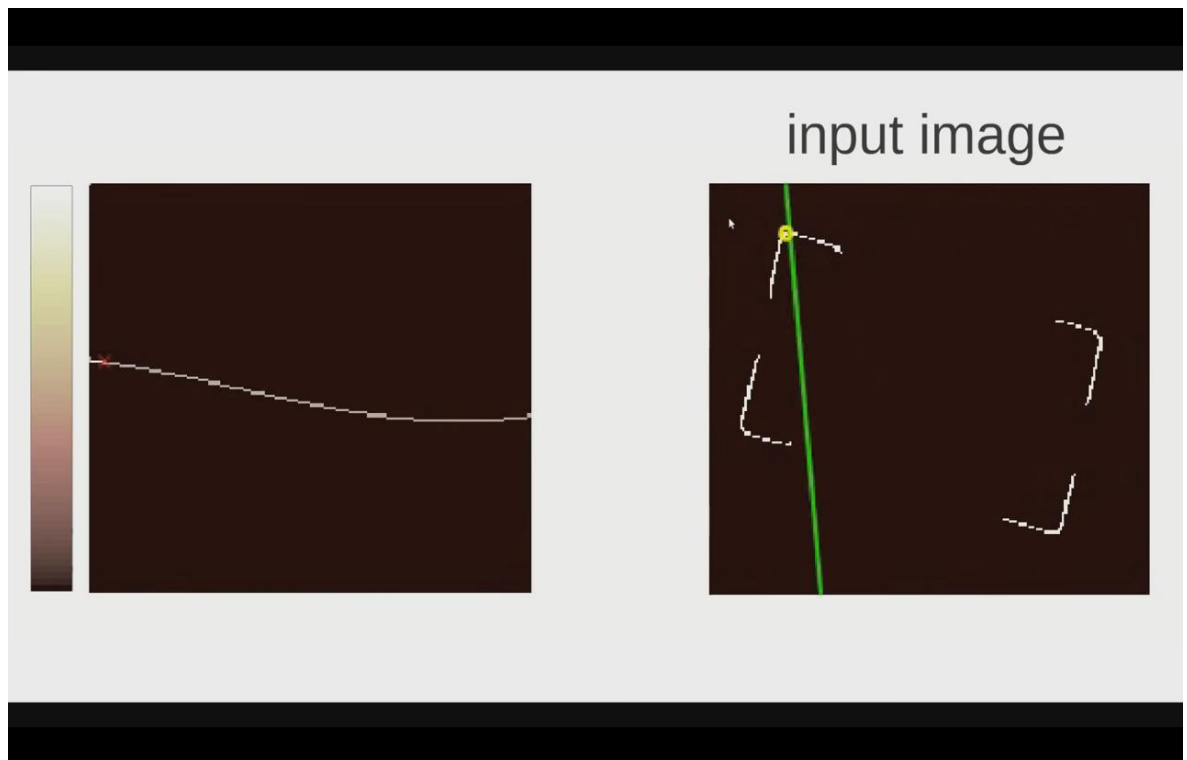
Natomiast pękowi prostych przechodzących przez punkt w przestrzeni 'ro' i 'teta' odpowiada krzywa sinusoidalna w przestrzeni Hougha. (Trzeba to sobie trochę wyobrazić, ciężko jest to wyjaśnić, trzeba by było spróbować narysować sobie samemu taki wykres punkt po punkcie...)

A teraz jak działa algorytm transformaty Haugha ?

Zakładamy, że mamy wejściowy obraz binarny taki jak poniżej po prawej stronie :

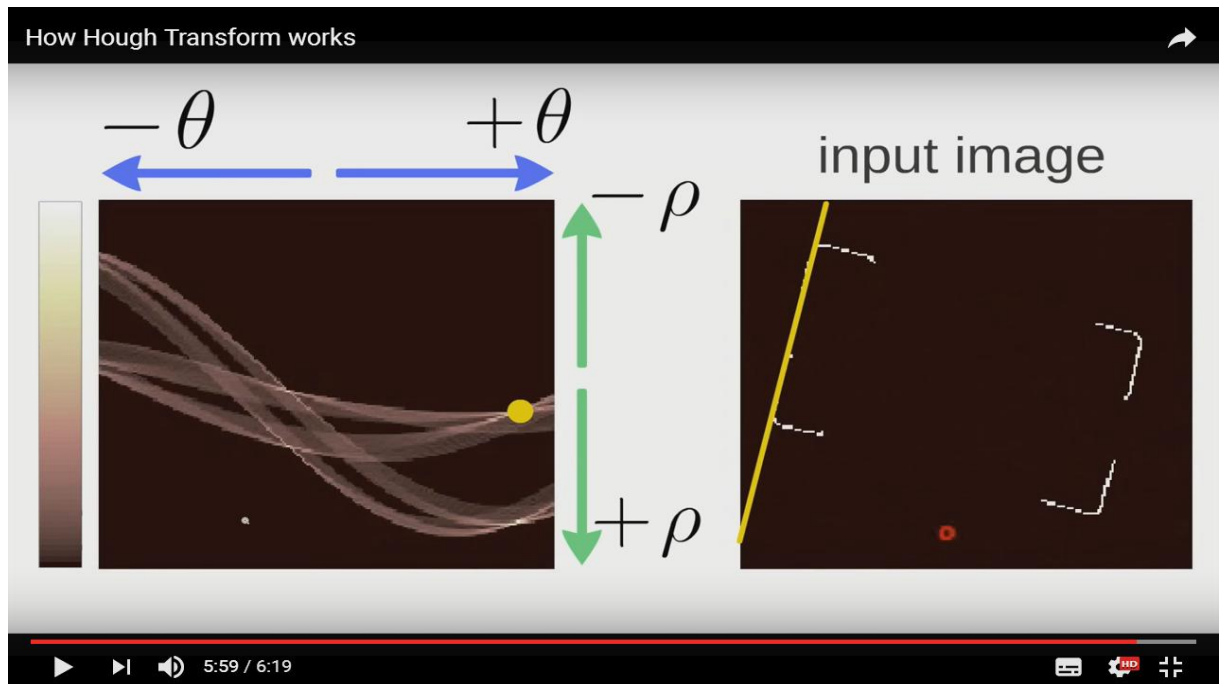


Przez każdy badany (różny od zera, czyli biały) punkt obrazu prowadzony jest pęk prostych, przechodzących przez ten punkt. Każda z tych prostych transformowana jest do przestrzeni Hougha i tworzy tam punkt o współrzędnych ρ i θ jak na obrazku poniżej. Jak wiemy pęk prostych tworzy w przestrzeni Hougha prostą sinusoidalną. (Ta zielona prosta to przykładowa przechodząca przez piksel, rozważa się wszystkie przez niego przechodzące – stąd po lewej krzywa sinusoidalna) .



Przestrzeń Hougha jest przestrzenią akumulacyjną tzn. punkty sinusoidalnych krzywych, wygenerowanych dla punktów obrazu pierwotnego dodają się w miejscach, w których krzywe te przecinają się. Powstałe w ten sposób (w przestrzeni Hougha) maksima odpowiadają zbiorom punktów, należących do jednej prostej. Współrzędne ρ, θ tego maksimum jednoznacznie określają położenie prostej na obrazie pierwotnym.

Przykładowo na poniższym obrazku jest zaznaczone wykrycie jednej z krawędzi w przestrzeni Hougha i na obrazie rzeczywistym.



Na zajęciach można było wybrać parametry takie jak : rho resolution oraz theta resolution. Fragment instrukcji poniżej :

- wykorzystaj funkcję `hough`
- ustal parametry: `'RhoResolution', 0.1`, - rozdzielczość promienia wodzącego
`'ThetaResolution', 0.5` - rozdzielczość kąta (w stopniach)
- funkcja zwraca macierz **H** (przestrzeń Hougha) oraz dwa wektory **theta** i **rho**

Dlaczego proste $y = ax + b$ nie nadają się do T. Hougha ? Teraz to już oczywiste. Przyjmując postać prostej $y = ax + b$ nie można zapisać prostych pionowych, równoległych do osi OY, a co za tym idzie nie można ich przetransformować do przestrzeni Hougha, a więc nie można nimi wykryć krawędzi pionowych.