

# Przetwarzanie obrazów:

Główne założenia przedmiotu:

- Temat ma mniejszy próg wejściowy niż OpenGL (jest łatwiejszy):
- Na Labach, to co jest potrzebne to aplikacja która pozwala wczytać obraz, przetworzyć go, oraz zapisać obraz wynikowy (ma to być aplikacja desktopowa)  
Dostaniemy 2 wersje szkieletu takowej aplikacji, które wczytują pliki obrazów w jakimś formacie oraz są w stanie je zapisać (dostaniemy gotowca opartego o bibliotekę okienkową Qt! C++) / lub można napisać własnoręcznie w dowolnym języku.
- Celem przedmiotu jest algorytmika przetwarzania obrazów a nie tworzenie aplikacji desktopowych!
- Można pracować w zespołach 2-osobowych.

## ---== WYKŁAD 1 ==---

### *Wprowadzenie do teorii koloru. Postrzeganie barwy*

Obraz cyfrowy - źródło informacji, zapisane w formie tablicy pixeli.

Podczerwień i ultrafiolet są niewidoczne dla oka.

Do przetwarzania obrazów przede wszystkim przyda nam się pochodna funkcji (2-ch zmiennych) - na obraz będziemy patrzeć jak na funkcję w ogólności.

Przyda się również pojęcie entropii. I dekompozycja macierzowa (Algebra liniowa)

---

Zakładamy percepcję wzrokową - obserwator ludzki pod względem odbioru świata jest pod wieloma względami kiepskim obserwatorem. Widzimy gorzej niż niektóre stworzenia na świecie. Dlatego też musimy się do tej sytuacji dostawać

Odbiorcą obrazów jest człowiek - dlatego w tworzeniu urządzeń i algorytmów dla obrazów - uwzględniamy biologię człowieka, budowę i działanie oka oraz mózgu.

Do zadań widzenia komputerowego można zaliczyć między innymi:

- wyróżnianie na obrazach istotnych fragmentów (segmentacja obrazu)
- Śledzie położenia obiektów oraz jego wewnętrzną strukturę czy np.: śledzenie gestów / orientacji obiektów (wymaga głębokości)

Efekty filmowe aktualnie robi się na poziomie jednej klatki filmowej, na poziomie 2D.

Źródłem obrazu jest światło. Z punktu widzenia fizyki jest to bardzo prosta rzecz. Mamy przestrzeń jednowymiarową - jedna długość fali / jedna częstotliwość - oznacza jeden kolor. (kolor -> to co nazywamy światłem widzialnym) Zakres fal widzialnych:

- Jedna częstotliwość = jeden kolor.
- Czopki reprezentują widzenie dzienne.
- Pręciki reprezentują widzenie nocne.

---

Modele widzenia są 3 elementowe (RGB) - jest kilkanaście klasycznych modeli i też ogólny (np.: „podkowa”) Te modele mówią nam, jaki kawałek przestrzeni kolorów jest w stanie nam wygenerować albo odebrać (wykryć) jakieś urządzenie.

Na podkowie, oko ludzkie identyfikuje (odbiera percepcyjnie), pewne obszary/grupy współrzędnych - jako całkowicie jednolite (elipsy MacAdama) (nie dostrzegamy różnicy pomiędzy kolorami)

## ----== WYKŁAD 2 ==----

### Operacje punktowe. Histogram

#### Podkowa (Slajd 7):

Rzut przestrzeni barw - ukazuje nam barwy bazowe ze spektrum fal widzialnych. (w nim mamy wszystkie możliwe barwy do wygenerowania na komputerze) Na brzegach zaznaczone są długości fali, oraz wartości maksymalne) Narożniki to RGB i kolor całej podkowy rozkłada się gradientowo.

Barwy dopełniające - wybieramy punkt barwy nas interesujący, następnie prowadzimy odcinek przez punkt bieli i otrzymujemy dany kolor (negatyw) Zsumowanie wybranej barwy i barwy dopełniającej da nam kolor bieli.

funkcje które używamy do przekształceń są ciągłe, a kolory w świecie cyfrowym są ograniczone przez skończoną liczbę 3 bajtów na których są zapisane. Możliwe do wygenerowania jest jedynie  $256^3$  kolorów.

Desaturacja - zmienienie obrazu na obraz w odcieniach szarości

## Operacje punktowe: (prostsze koncepcyjnie i implementacyjnie)

Pobierają one jeden pixel obrazu oryginalnego i zmieniają za pomocą algorytmiki jego wartość na inną, 1 na 1 (Wartość jednego piksela zostaje zamieniona na inną). W analogi do funkcji, operacje te pobierają tylko jedną informację lokalną [punktową (X,Y)] i generują nową wartość koloru na podstawie funkcji.

Dlatego - nie sprowadzamy tego do myślenia o pixelach, ponieważ pod spodem mamy modelowanie przez funkcje. Jest to funkcja jednej zmiennej.

Percepcja oka i innych receptorów ma naturę całkową - my całkowicie niezależnie wykonujemy operacje na 3 składowych (w przypadku RGB) - i następnie wkładamy efekt do jednego worka, nazywanego siatkówką oka. 3 receptory to odbierają a następnie jest to dopiero całkowane na poziomie mózgu.

Z punktu widzenia algorytmicznego oznacza to - że możemy myśleć zawsze, budując algorytmy, o pojedynczych składowych - weźmiemy sobie osobno R, osobno G i osobno B. Sprowadzając to do operacji punktowej - to bierzemy sobie liczbę typu czerwień i przetwarzam ją na inną liczbę. Następnie tym samym algorytmem przetwarzamy kolejno G i B

(jeśli chcemy pixel przetworzyć aby miał inną wartość) Model jedno bajtowy jest więc z zakresu od 0 do 255 po s próbkowaniu i zakres możliwych wartości również jest od 0 do 255. Dlatego obraz takiej funkcji rysujemy w kwadracie 255 x 255 jak na slajdzie (5)

Slajd 5 -> Do określonej wartości koloru (od 0 do 80) funkcja odwzorowuje wszystkie wartości w 0 (w czerń). Ta funkcja mówi: „zabijam pierwszą 80 do 0 (w kontekście szarości -> do czerni), a od 175 do 255 jest skalowane do maksymalnej energii - (w kontekście szarości -> do bieli!). W klasycznym myśleniu jest to operacja kontrastu.

2-ga funkcja jest to obraz binarny - czyli biel i czerń.

(slajd 6) L = maksymalna składowa (255)

Przykładowe typowe funkcje dla operacji punktowych:

- 1)  $Y = X$  (nic nie zmienia w obrazie / odwzorowanie identycznościowe) - identity
- 2)  $Y = -X + 255$  (Negatyw obrazu - zmieniamy wartość piksela na kolor przeciwny)
- 3) Inverslog - w ogólności ta funkcja bardziej zabija/obniża wartość pierwszej części zakresu (czyli składowych ciemniejszych - stają się one dużo bardziej ciemne niż były) a składowe wyższe dużo słabiej się osłabiają (nieznacznie ściemniamy to co było jasne)
- 4) Log - pierwszy zakres znacznie zwiększa jasność a w drugim zakresie jasność jest delikatnie wzmacniana.

Odpowiednikiem operacji punktowych w GIMPIE, jest zakładka - „kolory”

## Operacje „sploty”:

Żeby wygenerować nową wartość dla jednego pixela - będzie wymagane pobranie informacji nie tylko z tego konkretnego pixela ale również z pewnej jego otoczki (z grupy pixeli, otaczających dany pixel - lokalnego pola)

Cel  $g(x,y)$  - jest źródłem informacji do wygenerowania nowego koloru pixela -> będzie to grupa otaczających go pixeli - najczęściej jego otoczka. (Nazywamy to splotem.

## Histogram:

Jest to liczbowa charakterystyka obrazu (możliwe od wyliczenia na podstawie obrazu 2D - można przedstawić jako wykres) (Myślenie pokroju - weź obraz i na jego podstawie opisz interesującą nas cechę za pomocą liczb)

Pixel reprezentuje nam 3 składowe koloru (RGB) - dziedzina jest przestrzenna.

Histogram przesuwą nas do osi jednowymiarowej - ile koloru/danej częstotliwości znajduje się na obrazie. (dziedzina częstotliwościowa - dzięki temu mamy wizję - ile jest w obrazie % danej typu energii. (np.: ilość czerwieni, czy ilość zielonego lub niebieskiego)

**Ciekawostka?:** [ Wiąże się to z wykonywaniem transformatów na obrazach. Pomaga to w bardzo dobrym i dokładnym odsumianiu obrazów. (filtr dolnopuszczający - przepuszcza niskie częstotliwości i kasują/filtrują wysokie częstotliwości - stosuje się to w klasycznych kompresjach np. JPEG) ]

Histogram - jest to pewna statystyka - ale bardzo istotna - jest źródłem wstępnej informacji do wielu innych algorytmów. Informacja ta mówi nam ile statystycznie jest pixeli danej jasności w obrazie.

Histogram mówi nam również - czy zdjęcie jest dobrze zrobione. Czy zdjęcie wykorzystuje pełne spektrum informacji które w oryginalnym świecie są obecne. (nie powinno być braków - wartości powinny być dla pełnej skali)

Wykonując różnego rodzaju operacje na obrazie, np. operacje punktowe - jego histogram może się zmieniać i mogą pojawiać się w nim luki!

# ----== WYKŁAD 3 ==----

## Filtry splotowe

### Filtry splotowe:

Ogólne założenie jest takie, że generując nową wartość koloru dla konkretnego pixela (X,Y) i dla tego wygenerowania bierzemy informacje z samego pixela oraz z pewnej jego otoczki. Ta otoczka zwykle jest symetryczna, więc w języku kwadratu ma długość nieparzystą. Np.: zwykle otoczka przyjmuje rozmiar 3 x 3 lub 5 x 5. W rozmyciu Gaussowskim nawet bierze się 11 x 11. Im więcej bierzemy tym więcej obliczeń wykonujemy! (Czas i ilość obliczeń!)

Slajd 3 - SPLOT FUNKCJI -> tłumaczone w języku 2D -> chcemy wygenerować nową wartość w konkretnym pixelu, i bierzemy sobie w jego otoczkę i mówimy: wezmę sobie informacje z otoczki 3 x 3. Mając to 3x3 budujemy tzw. Maskę filtra, czyli nową tablicę w której wprowadzamy pewne wartości. Nakładamy sobie centrum tej maski na pixel który chcemy wyliczyć, następnie mnożymy wartości (kolor w danym pixelu przed wartość odpowiadającą temu pixelowi na masce) (9 iloczynów, a następnie je ze sobą sumujemy!) to jest dokładnie działanie splotu, jest to tak naprawdę liniowe mieszanie. Wyliczenie średniej z otoczki + centrum - > wizualnie oznacza to ROZMYCIE!

Im otoczka większa - tym większe rozmycie! Wygładzenie obrazu. (Jednostajne wygładzenie.)

W GIMPIE - kategoria - filtry - wszystkie oprócz: „artystycznych” opierają się o splot.