

Kolokwium K2T2

Zaawansowana Analiza Obrazu, Wideo i Ruchu

Zadanie 1

Odczytaj wartości z referencyjnej mapy głębi wygenerowanej przy użyciu symulatora CARLA (depth.png). Plik został zapisany z użyciem 24-bitowego formatu głębi stosowanym w oprogramowaniu CARLA ([link do dokumentacji](#)).

Przyjmij następujące parametry:

- Maksymalna odległość mapy głębi: 1000 m,
- Minimalna odległość mapy głębi: 0,
- Kodowanie 24 bity (uint24).

Mapa głębi zapisana została przy użyciu 24-bitowego kodowania (uint24) w postaci obrazu RGB w formacie PNG.

R	G	B	uint24	
00000000	00000000	00000000	0	min (near)
11111111	11111111	11111111	16777215	max (far)

Aby odczytać wartość w zadanej jednostce dla konkretnego piksela możesz posłużyć się poniższym wzorem, gdzie za maksymalna odległość powinna zostać wyrażona w odpowiedniej jednostce np. 1000 m.

$$\text{odleglosc} = R + G*256 + B*256*256 / (256*256*256 - 1) * \text{maksymalna_odleglosc}$$

R, **G**, **B** to poszczególne składowe koloru.

Uwaga: domyślna kolejność składowych dla OpenCV to BGR.

Podaj odczytane wartości głębi oraz przyjęte współrzędne dla wskazanych na zdjęciu (info.jpg) punktów (P1, P2, P3), pamiętaj o właściwej jednostce. Zapisz przekonwertowaną głębię do pliku PNG (8bit), gdzie wartość jasności 250 odpowiada odległości 50 metrów. Pamiętaj by przed zapisanie do pliku PNG odpowiednio przyciąć zakres głębi do przedziału (0 – 50 metrów). Odległości powyżej

50 metrów w pliku zapisz jako wartość 255 (maksymalna jasność dla 8bit). Stosując powyższą metodę uzyskasz rozdzielczość głębi 0,2 [m].

Jako wynik dołącz następujące pliki:

Plik z1.TXT – plik tekstowy

(P1) X, Y = [?, ?]
depth P1 = ? [m]

(P2) X, Y = [?, ?]
depth P2 = ? [m]

(P3) X, Y = [?, ?]
depth P3 = ? [m]

z1depth.png – głębia 8-bit

Zadanie 2

Wyznacz mapę rozbieżności (disparity map) na podstawie referencyjnej mapy głębi (depth map). Przyjmij poniższe założenia.

- Kanoniczny układ kamer stereo
- Poziome pole widzenia kamery (HFOV): 90 stopni
- Odległość pomiędzy kamerami (baseline): 0,1 [m]
- Maksymalna odległość dla mapy głębi: 1000 [m]
- Minimalna odległość mapy głębi: 0
- 24 bitowe kodowanie mapy głębi (uint24)

Wynikową mapę rozbieżności zapisz w postaci 8 bitowego obrazu PNG, gdzie poziom jasności odpowiada przesunięciu w pikselach w zakresie od 0 do 255.

Jako wynik dołącz następujące pliki:

Plik z2.TXT – plik tekstowy

```
(P1) X, Y = [?, ?]  
disparity P1 = ? [px]  
(P2) X, Y = [?, ?]  
disparity P2 = ? [px]  
(P3) X, Y = [?, ?]  
disparity P3 = ? [px]
```

z2disparity.png – mapa rozbieżności

Zadanie 3

Przekonwertuj referencyjną mapę głębi **depth.png** z zadania 1 na kolorową chmurę punktów w formacie PLY. W wynikowej chmurze zamieść tylko te punkty, dla których odległość od kamery nie przekracza 50 metrów. Poszczególne punkty chmury powinny odpowiadać kolorom pikseli z pliku **left.png**. Skorzystaj z parametrów zdefiniowanych w zadaniu 2. Otwórz plik przy użyciu oprogramowania CloudCompare, MeshLab lub skryptu Python bazującego na bibliotece Open3D. Wykonaj zrzuty ekranu widoku chmury punktów na wprost (z3a.png), z góry (z3b.png) oraz z dowolnej perspektywy (z3c.png).

Jako wynik dołącz następujące pliki:

z3.ply – plik z chmurą punktów

z3a.png – zrzut ekranu z wizualizacją chmury punktów (widok na wprost)

z3b.png – zrzut ekranu z wizualizacją chmury punktów (widok z góry)

z3c.png – zrzut ekranu z wizualizacją chmury punktów (dowolna perspektywa)

Zadanie 4

Korzystając z metody **StereoSGBM** udostępnionej przez bibliotekę OpenCV wyznacz mapę rozbieżności oraz mapę głębi. Dobierz odpowiednie parametry metody **StereoSGBM**. Porównaj otrzymane wyniki z referencyjnymi wartościami rozbieżności zapisanymi w formacie PFM (disp0.pfm).

<https://vision.middlebury.edu/stereo/data/scenes2014/datasets/Playtable-perfect/>

im0.png – obraz z kamery lewej

im1.png – obraz z kamery prawej

disp0.pfm – referencyjna mapa rozbieżności

calib.txt – plik z parametrami kamer

Plik **calib.txt** zawiera parametry kamer. Odległość bazowa pomiędzy kamerami (baseline) zdefiniowana została w milimetrach.

```
cam0=[3979.911 0 1244.772; 0 3979.911 1019.507; 0 0 1]
cam1=[3979.911 0 1369.115; 0 3979.911 1019.507; 0 0 1]
doffs=124.343
baseline=193.001
width=2964
height=2000
ndisp=270
isint=0
vmin=23
vmax=245
dyavg=0
dymax=0
```

Poniżej dla przykładu przedstawiona została macierz parametrów wewnętrznych kamery lewej (cam0) według wartości z pliku **calib.txt**.

$$\text{camera matrix}_{\text{cam0}} = \begin{bmatrix} 3979.911 & 0 & 1244.772 \\ 0 & 3979.911 & 1019.507 \\ 0 & 0 & 1 \end{bmatrix}$$

Pliki w formacie PFM możesz wczytać przy użyciu biblioteki OpenCV stosując flagę `cv2.IMREAD_UNCHANGED`, a następnie dla ułatwienia przekonwertować przy użyciu biblioteki **numpy** do tablicy przy użyciu **np.asarray**.

```
groundtruth = cv2.imread('data/stereo/disp0.pfm', cv2.IMREAD_UNCHANGED)
groundtruth = np.asarray(groundtruth)
groundtruth = groundtruth / 256
```

Podaj wartość referencyjnej (`disparity_gt`) i wyznaczonej rozbieżności (`disparity`) dla punktu o współrzędnych `X=1550`, `Y=900` lub punktu w okolicy o niezerowej wartości referencyjnej. Dla wskazanego punktu oblicz (`depth_gt`) na podstawie danych referencyjnych oraz odległość (`depth`) na podstawie wyznaczonej rozbieżności dla punktu. Wyniki zapisz w metrach. Zapisz wyjściową mapę rozbieżności oraz głębię do plików PNG 8bit.

Jako wynik dołącz następujące pliki:

Plik `z4.TXT` – plik tekstowy

```
X, Y = [?, ?]  
depth_gt = ? [m]  
depth = ? [m]  
disparity_gt = ? [px]  
disparity = ? [px]
```

`z4disparity.png` – mapa rozbieżności

`z4depth.png` – głębia

Zadanie 5

Korzystając z własnej implementacji metody bazującej na dopasowaniu bloków omówionej podczas wykładów wyznacz mapę rozbieżności oraz mapę głębi dla pary obrazów kanonicznego układu stereo. Porównaj rozbieżność dla wybranego punktu z wartością referencyjną odczytaną z pliku PFM. Jeśli Twoja metoda estymacji rozbieżności zajmuje zbyt dużo czasu, spróbuj przyciąć zdjęcie np. do ¼ pierwotnego rozmiaru lub przeskaluj (2 lub 4 krotnie) – jeśli przycinasz zdjęcie, wybrany obszar powinien uwzględniać wskazany punkt, jeśli skalujesz zdjęcie przeskaluj również wskazane współrzędne tak by odpowiadały pierwotnej lokalizacji i zapisz współczynnik skali (*scale*). Podaj wartości głębi i rozbieżności dla punktu o współrzędnych *X=1550*, *Y=900* lub najbliższego punktu, dla którego Twój algorytm prawidłowo wyliczył rozbieżność. Zapisz mapę rozbieżności i głębię do plików PNG 8bit.

Jako wynik dołącz następujące pliki:

Plik z5.TXT – plik tekstowy

```
X, Y = ?  
scale = ?  
depth_gt = ? [m]  
depth = ? [m]  
disparity_gt = ? [px]  
disparity = ? [px]
```

z5disparity.png – mapa rozbieżności

z5depth.png – głębia