

PRACA PRZEJŚCIOWA 2

Specjalność: **TECHNIKI MULTIMEDIALNE**

Instytut prowadzący specjalność: **Instytut Metrologii i Inżynierii Biomedycznej**

Instytut prowadzący pracę: **Instytut Automatyki i Robotyki**

Temat pracy: Program do automatycznego pozycjonowania głębokości napisów w sekwencji obrazów stereoskopowych

Zakres pracy:

1. Zapoznanie się z istniejącymi rozwiązaniami umożliwiającymi obliczania głębokości z obrazów stereoskopowych
2. Stworzenie założeń programu i sposobu podawania wyników
3. Napisanie oprogramowania obliczającego mapę głębi scen w sekwencjach obrazów stereoskopowych
4. Sprawdzenie i ocena prawidłowości działania gotowego oprogramowania, na wzorcowych sekwencjach wideo

Podstawowe wymagania:

1. Program ma za zadanie umieszczać podane na wejściu napisy w odpowiednim momencie w sekwencji obrazów oraz w wyliczonej głębokości sceny.
2. Program ma za zadanie na wyjściu zapisywać plik (np. w formacie XML) ze skalowalną współrzędną głębokości dla poszczególnych zakresów czasowych sekwencji, podanych na wejściu razem z napisami
3. Materiał stereoskopowy, na którym mają być prowadzone obliczenia, ma być nagrany na kamerach po kalibracji i po procesie retyfikacji

Literatura:

- dokumentacja bibliotek przetwarzania obrazu OpenCV
- „Image processing and recognition” - Robert Sitnik
- *Cyfrowe przetwarzanie obrazu* – M.Kujawińska
- dokumentacja środowiska programistycznego C++
- „OpenCV learning” - O'Reily

Nazwisko dyplomanta:

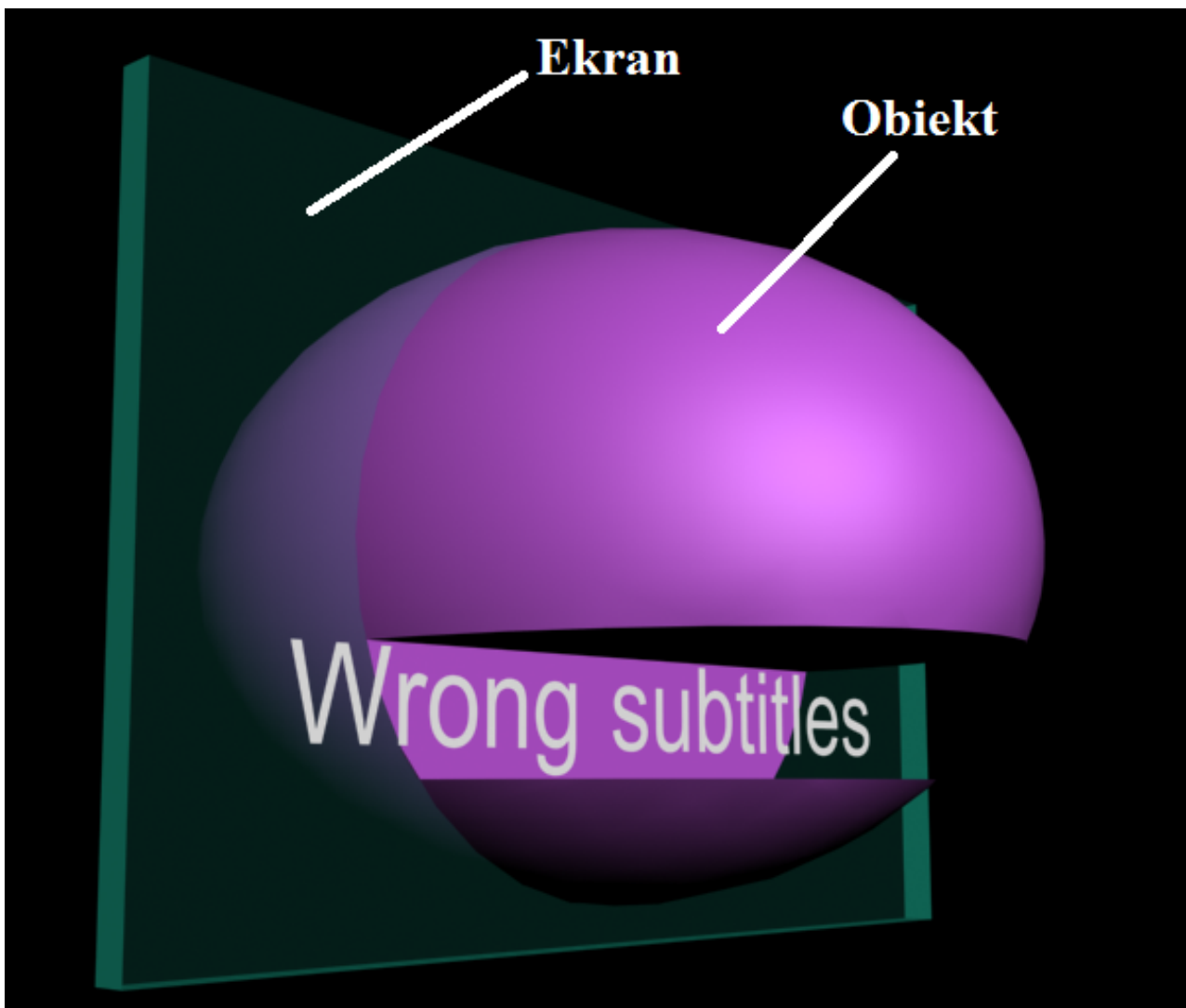
Michał Wielgosz

Nazwisko promotora:

prof. nzw. dr hab. inż. Barbara Putz

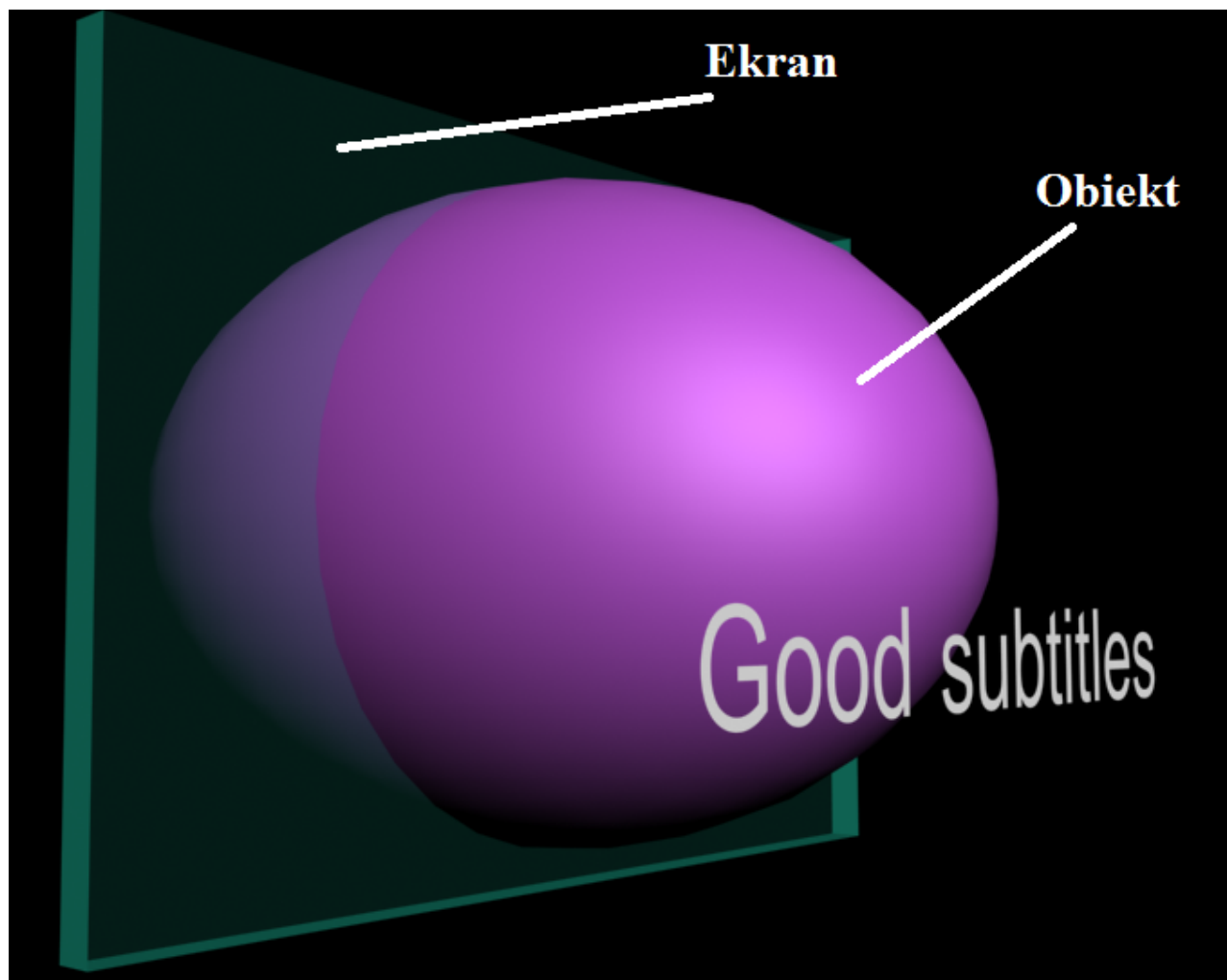
Problem:

Coraz częściej produkowane są filmy, animacje stereoskopowe. Podczas ich oglądania w kinach prawie w każdym wypadku można zaobserwować pewien problem. Poniżej przedstawione są dwa obrazki. Pokazują one ekran wyświetlający obraz stereoskopowy, prosty obiekt występujący w scenie stereoskopowej i napisy. Na pierwszym rysunku można zauważyć że napisy wstawione są w scenę w zerowej głębokości, czyli na płaszczyźnie ekranu. Jest to najprostszy sposób umieszczania napisów, ponieważ wprowadza się je w jednakowym miejscu zarówno na klatce dla lewego i prawego oka. Tekst wyświetlany w taki zwykły sposób, przecinający obiekt będący w scenie, co widać na poniższym rysunku.



Takie przenikanie, wcinanie się napisów w obiekty wychodzące przed ekran w kierunku widza, jest niepożądanym efektem. Podczas oglądania filmu z takimi napisami, czuje się dyskomfort. Dodatkowo w pewien sposób niszczy się strukturę głębi poszczególnych obiektów w kadrze. Przez takie napisy, przerywana jest ciągłość głębi obiektów w scenie.

Dużo lepszym rozwiązaniem jest ustawianie napisów przed najbliższym obiektem występującym w kadrze ujęcia. Dobrze pokazuje to drugi rysunek:



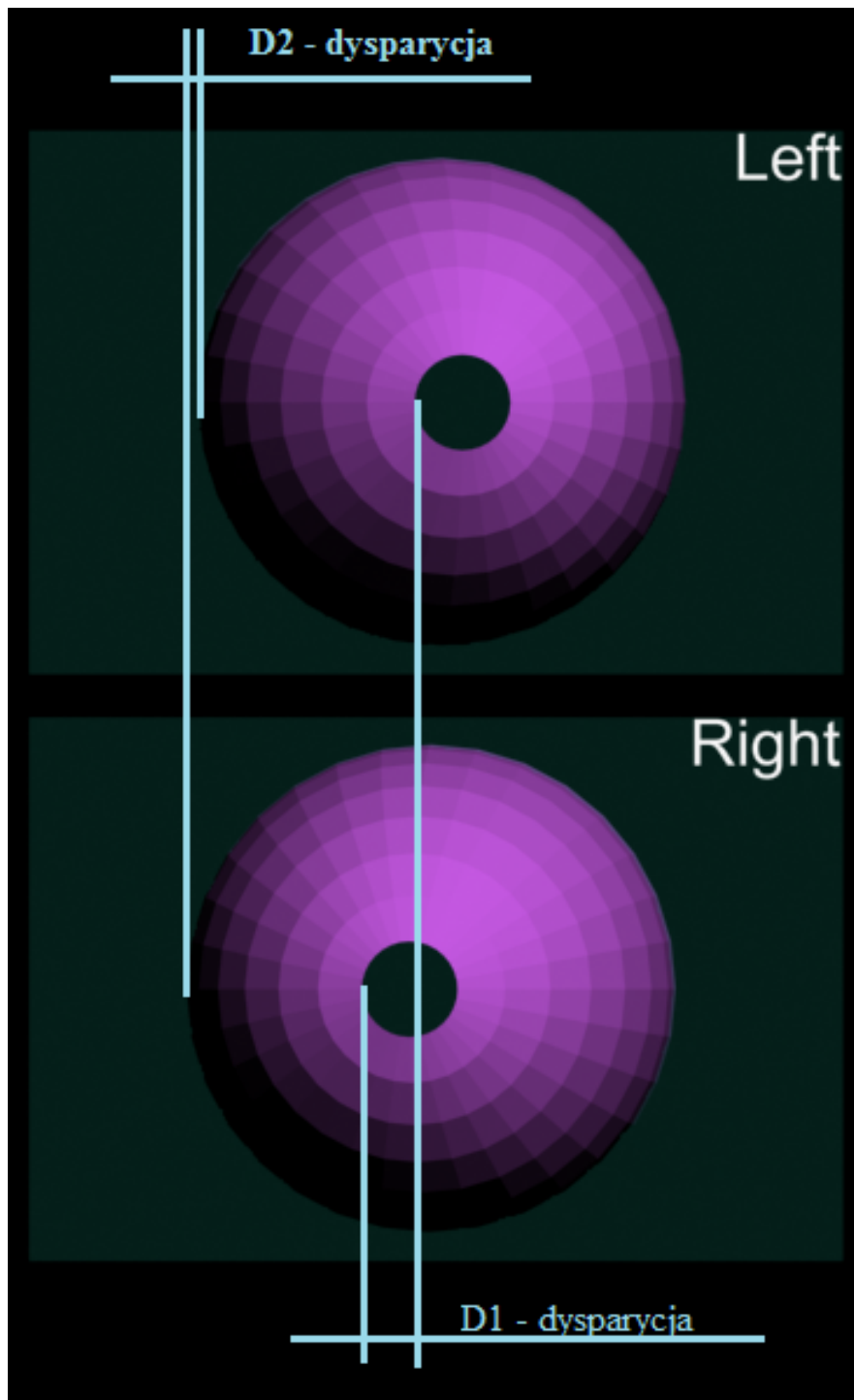
Taki sposób wyświetlania napisów podczas filmu jest zdecydowanie przyjemniejszy dla oka. Kolejną zaletą jest utrzymanie ciągłości obiektów, żaden z obiektów w scenie nie jest przecinany. Nie jest także naruszana ciągłości głębi obiektów w scenie.

Wszystkie te zalety są osiągnięte przez odpowiednie sterowanie głębokością umieszczenia napisów. Jest możliwe do wykonania na takiej samej zasadzie jak osiągnięta jest głębokość w scenach stereoskopowych. Obiekty mają swoją głębokość, która jest uzależniona od odległości poziomego przesunięcia danego przedmiotu na lewym obrazie, w stosunku do obrazu prawego. Tak samo można sztucznie sterować głębokością napisów. Aby zmieniać głębokość wyświetlanego tekstu, należy przesuwać go o taką samą wartość, w kierunku poziomym, ale o przeciwnych zwrotach dla obrazów lewego i prawego.

W skrócie program ma za zadanie zrobić wstępny przebieg przez całość materiału i ustalenie mapy dysparycji (w pewnym sensie mapy głębi) dla każdej z klatek sekwencji wideo. Następnie dla całego zakresu trwania poszczególnych napisów, wybierana jest maksymalna wartość dysparycji, która występuje w czasie wyświetlania danego tekstu. Mając tą wartość, w odpowiadających sobie kolejnych klatkach dla lewego i prawego oka, możemy rozsunąć umieszczany tekst o odpowiednią liczbę pikseli, która równa jest właśnie wyznaczonej maksymalnej wartości dysparycji. Dzięki temu tekst który będzie wyświetlany na ekranie, zawsze będzie przed najbliższym widzowi obiektem.

Materiały i dane aplikacji **WEJŚCIE:**

* Sekwencja obrazów dla lewego i prawego oka:



Są to kolejne klatki klipu stereoskopowego oddzielnie dla lewego i prawego oka. Tak jak widać na rysunku różnią się one przesunięciami w poziomie poszczególnych pikseli, gdzie im większa odległość pomiędzy pikselami tym większa głębia danych pikseli.

* Plik z napisami w formacie SRT

Do programu należy jeszcze dostarczyć plik z napisami w formacie SRT. Poniżej podany jest przykładowy listing pliku w takim formacie:

```
1 1
2 00:00:00,000 --> 00:00:00,080
3 Subtitles, subtitles..
4
5 2
6 00:00:00,120 --> 00:00:00,200
7 Subtitles, subtitles..
8
9 3
0 00:00:00,240 --> 00:00:00,320
1 Subtitles, subtitles..
2
3 4
4 00:00:00,360 --> 00:00:00,440
5 Subtitles, subtitles..
6
7 5
8 00:00:00,480 --> 00:00:00,560
9 Subtitles, subtitles..
```

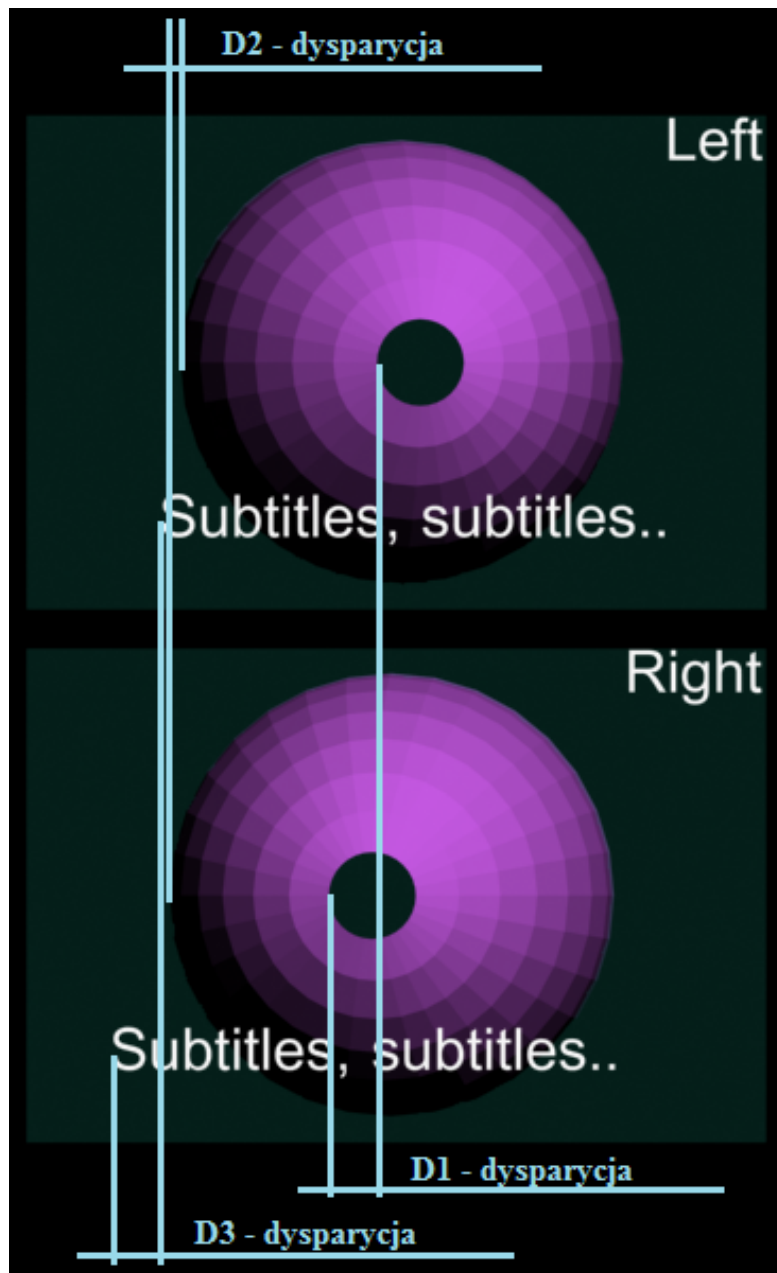
Licznik kolejnych napisów
Zakres napisów
(hh:mm:ss,ms)
Kolejne linie napisów

Program ma za zadanie liczyć mapy dysparycji dla poszczególnych zakresów napisów. Natomiast po ich policzeniu wyznaczana jest maksymalna wartość dysparycji dla całego zakresu czasu trwania napisu, który jest określony w pliku SRT.

Materiały i dane aplikacji

WYJŚCIE:

* Sekwencja obrazów dla lewego i prawego oka, z dodanymi napisami, umieszczonymi w wyliczonej głębokości:



D1 – dysparycja najbliższego widzowi miejsca w obrazie stereoskopowym – jest to maksymalna znaleziona dysparycja

D2 – dysparycja miejsca na obrazie stereoskopowym, które jest w głębi

D3 – dysparycja napisów wyliczona z obrazu stereoskopowego równa dysparycji D1

$D2 < D1 = D3$

Głównym wynikiem działania programu jest sekwencja klatek dla lewego i prawego oka wzbogacona o napisy stereoskopowe. Napisy te są rozsunięte w poziomie o taką wartość aby w obrazie 3D, dla widza były zawsze przed najbliższym wyświetlanym obiektem.

Algorytm programu

1. Załadowanie danych wejściowych:
 - pliku napisów SRT – przetworzenie tego pliku w celu otrzymania: ramki rozpoczęcia i końca napisów, tekstu napisów
 - sekwencji klatek dla lewego oka
 - sekwencji klatek dla prawego oka
2. Wybór ilości klatek/s załadowanego klipu – potrzebne ze względu na przeliczenie z poszczególnych czasów napisów na klatki klipu
3. Wybór miejsca zapisu wyników
4. Wybór rodzaju danych na wyjściu programu:
 - do pliku XML
 - wklejenie napisów bezpośrednio do kopii sekwencji obrazów stereoskopowych podawanych na wejście
5. Wybór rozmiaru bloku analizy (Block size) algorytmu liczącego mapę dysparycji – sterowanie wielkością minimalnego obiektu który będzie brany pod uwagę podczas obliczeń
6. Wybór pozostałych parametrów algorytmu do wyliczania mapy dysparycji: Prefilter intensity size, Min disparity, Search length disparity, Texture threshold, Uniqueness ratio, Speckle window size, Speckle range Scale factor depth, Shift factor depth, Filter size, Gaussian blur filter, Median filter (wszystkie wymienione parametry i reszta, opisane w dalszej części pracy)
7. Przycisk „Start” - rozpoczęcie obliczeń
8. Sprawdzanie różnych warunków zezwalających na przystąpienie do obliczeń (czy są załadowane sekwencje obrazów dla lewego i prawego oka, ilość klatek dla lewego i prawego oka musi być równa, musi być większa od zera, wymiary obrazów dla lewego i prawego oka muszą być równe, musi być załadowany plik z napisami, chyba że jest to uruchomienie w trybie testowym)
9. Pobieranie pierwszej klatki w celu sprawdzenia parametrów obrazu (rozdzielczość, ilości kanałów) aby móc utworzyć tymczasowe obrazy pośrednie wykorzystywane w dalszych obliczeniach
10. **Główna pętla** – do końca napisów pobieranych z pliku *.srt:
 - sprawdzenie czy program ma być uruchomiony w normalnym trybie czy może jest uruchomiony tryb testowy
 - *Tryb normalny:
 - sprawdzenie miejsca początku i końca napisów do wpisania - program wykonuje obliczenia jedynie w miejscach gdzie mają być umieszczone napisy
 - pobieranie kolejnych klatek z zakresu napisu
 - zmniejszanie rozmiaru obrazów i przetworzenie na czarno-białe w celu skrócenia obliczeń
 - nałożenie filtrów Median filter lub/i Gaussian blur filter w zależności od ustawień przez użytkownika

- algorytm liczący mapę dysparycji dla każdej klatki (jego opis jest pod algorytmem programu)
- zapamiętanie maksymalnej wartości dysparycji z dotychczas policzonych klatek dla danego napisu
- po zakończeniu klatek dla danego napisu – wklejanie, w źródłowe klatki sekwencji, napisów o rozsunięciu równym maksymalnej zapamiętanej dysparycji dla danego napisu lub zapis wartości dysparycji i tekstu napisów do pliku XML
- przejście do kolejnego napisu i dalej pętla się powtarza od punktu 10

*Tryb testowy:

- pobieranie kolejnych klatek z całej sekwencji obrazów stereoskopowych
- zmniejszanie rozmiaru obrazów i przetworzenie na czarno-białe w celu skrócenia obliczeń
- nałożenie filtrów Median filter lub/i Gaussian blur filter w zależności od ustawień przez użytkownika
- algorytm liczący mapę dysparycji dla danej pary obrazów stereoskopowych (jego opis jest pod algorytmem programu)
- zapamiętanie maksymalnej wartości dysparycji dla danej pary obrazów stereoskopowych
- wklejanie, w kopie źródłowych klatek sekwencji, napisów i grafiki testowej o rozsunięciu równym maksymalnej zapamiętanej dysparycji dla danej pary obrazów stereoskopowych
- przejście do kolejnej klatki i dalej pętla się powtarza od punktu 10

11. **Koniec pętli** w momencie policzenia ostatniego napisu(dla trybu normalnego) lub policzenia ostatniej pary obrazów stereoskopowych (dla trybu testowego)

12. Zwalnianie zasobów

13. Koniec programu

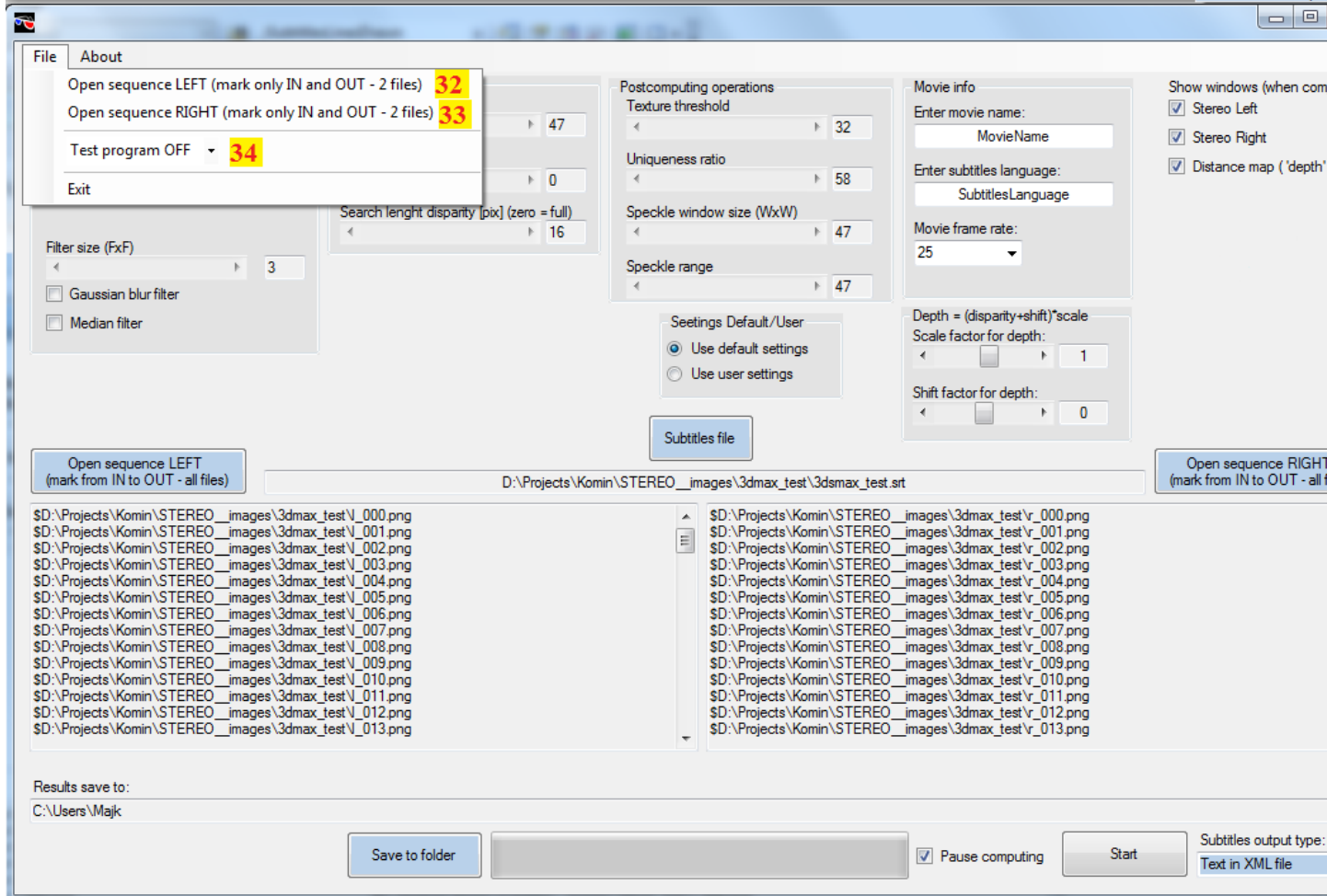
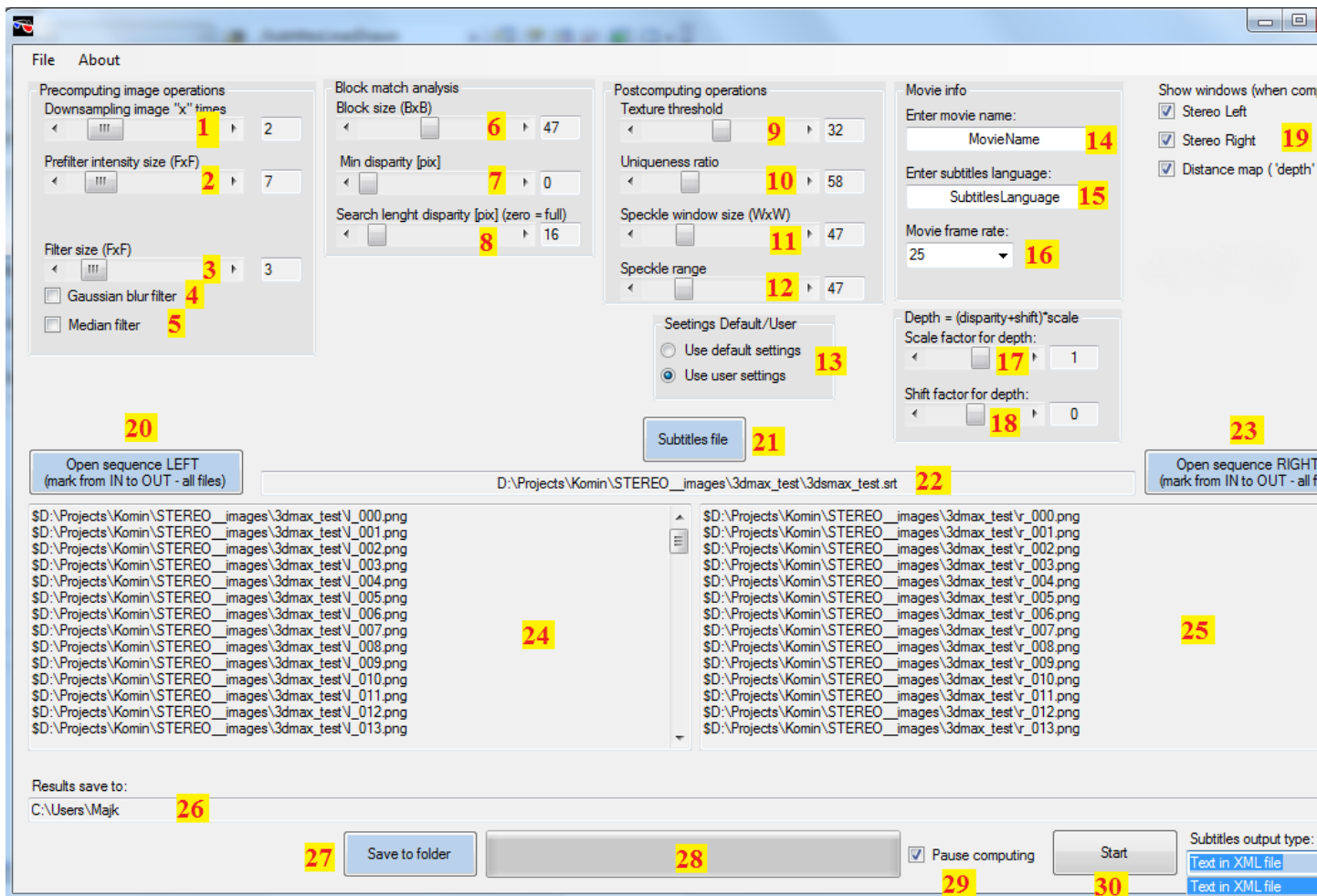
Do liczenia mapy dysparycji został wykorzystany algorytm Stereo Correspondence Block Matching udostępniony w bibliotece do przetwarzania obrazów OpenCV v2.0. Liczy on absolutne sumy intensywności w obrębie odpowiadających sobie okien dla obrazów lewego i prawego oka. Następnie porównuje te sumy i jeżeli różnica pomiędzy nimi jest mniejsza niż określony próg, wtedy uznawane są jako odpowiadające sobie i liczona jest dysparycja pomiędzy nimi, której wartość jest wstawiana w środek liczonego okna.

OpenCV oferuje jeszcze drugi algorytm Graph Cuts-Based Stereo Correspondence, który wylicza dokładniejszą mapę dysparycji lecz w znacznie dłuższym czasie.

Wybrany został jednak algorytm pierwszy gdyż w tym przypadku nie potrzeba nam wysokiej jakości mapy dysparycji, gdyż chodzi nam jedynie zbadanie głębokości dużych obiektów.

Dodatkowo przy tym wyborze znacząco zyskujemy na czasie obliczeń.

Interfejs programu



Precomputing image operations - Parametry operacji na obrazach stereoskopowych występujących przed algorytmem liczącym mapę dysparycji:

1. Downsampling image „x” times – liczba przez jaką zostanie podzielona szerokość i wysokość obrazów wejściowych – ma to na celu zmniejszenie czasu obliczeń.

Należy ją tak ustawiać aby obraz do obliczeń był rozmiarów ok 400x300pix

Domyślna wartość = 2

2. Prefilter intensity size (FxF) – rozmiar okna filtra o kształcie kwadratu i długości boku „F” - algorytm ma za zadanie wyrównanie jasności obu obrazów stereoskopowych w oknie o zadanym wymiarze

Domyślna wartość = 7

3. Filter size (FxF) – rozmiar okna filtrów z pkt 4 i 5 – w zależności o ilości ziarna i szumów w materiale wejściowym należy zwiększać rozmiar okna filtrów. W pierwszym momencie należy używać filtru medianowego, który najlepiej usuwa szumy typu „pieprz i sól”, które to generują najwięcej błędów w obliczeniach

Domyślna wartość = 3

4. Gaussian blur filter – włącz/wyłącz rozmycie filtrem gaussowskim

Domyślna wartość = wyłączony

5. Median filter – włącz/wyłącz filtr medianowy

Domyślna wartość = wyłączony

Block match analysis – Parametry algorytmu które wykorzystywane są bezpośrednio w trakcie szukania odpowiadających sobie bloków na obrazach dla lewego i prawego oka:

6. Block size (BxB) – rozmiar okna algorytmu porównującego obrazy dla lewego i prawego oka – sterując tą wartością można sterować wielkością pomijanych w obliczeniach obiektów, takich jak owady, ptaki w oddali itp. czyli obiekty na które nie chcemy aby cały algorytm zwracał uwagę. Im większy rozmiar okna tym większe obiekty będą pomijane w obliczeniach mapy dysparycji

Domyślna wartość = 47 (dla obrazów o rozmiarze ok 400x300pix)

7. Min disparity [pix] – najmniejsza wartość dysparycji brana pod uwagę w czasie obliczeń.

Wartość domyślna = 0

8. Search length disparity [pix] (zero = full) – wartość którą należy dodać do „Min disparity” aby otrzymać maksymalną wartość dysparycji jakiej będzie szukał algorytm pomiędzy obrazami lewego i prawego oka – z uwagi na wykorzystany algorytm wartość może być tylko wielokrotnością liczby 16, dla wartości zerowej algorytm przeszukuje wszystkie możliwe wartości dysparycji (zależne od wielkości wejściowego obrazka). Im większa wartość tym dłuższy czas obliczeń

Wartość domyślna = 16 (dla obrazów 400x300)

Postcomputing operations – Parametry części algorytmu dotyczącej operacji wykonywanych po policzeniu wstępnej mapy dysparycji:

9. Texture threshold – im większa wartość tego parametru tym fragmenty obrazów o większej tekstuowości są pomijane w obliczeniach dysparycji. Zmienia się ten parametr gdy chcemy odciać od obliczeń mapy dysparycji fragmenty powierzchni obrazu gdzie nie ma dużej ilości szczegółów, ponieważ w takich miejscach mogą powstawać błędy dopasowań

Wartość domyślna = 32

10. Uniqueness ratio – określa jak często algorytm ma odrzucać w obliczeniach piksele których prawdopodobieństwo dopasowania (lewego obrazu z prawym) jest podobne do prawdopodobieństwa dopasowania z innym pikselem ale o innej dysparycji. Podsumowując, jak bardzo algorytm ma szukać jednego prawidłowego dopasowania dwóch odpowiadających sobie pikseli z lewego i prawego obrazu.

Wartość domyślna = 58

11. Speckle window size (WxW) – rozmiar okna do szukania „plam” – przy obliczaniu map dysparycji częstym miejscem pojawiania się błędów są krawędzie obiektów. Występują tam duże przekłamanie zależne od wielkości okna i krawędzi obiektów. Przekłamanie te objawiają się jako bardzo różnie rozmieszczone, bardzo małe obszary o jasności nieadekwatnej do dysparycji.

Zastosowany algorytm pozwala w dużej mierze pozbyć się takich plam o „losowej” jasności, a w połączeniu z resztą algorytmów z „Postcomputing operations” praktycznie eliminuje się takie przekłamanie

Wartość domyślna = 47

12. Speckle range – wartość określająca maksymalną możliwą zmianę wartości jasności wewnątrz okna „Speckle window” która jest uznawana jako niepożądana

Wartość domyślna = 47

13. Settings Default/User:

- Use default settings – użycie domyślnych nastaw jakie zostały zapisane w programie; przełączenie na tę opcję zawsze skutkuje przywróceniem nastaw domyślnych
- Use user settings – pozwala użytkownikowi sterować parametrami obliczeń

Movie info – informacje o filmie potrzebne do obliczeń lub do uzupełnienia pliku XML z napisami i parametrem określającym głębokość napisów w filmie

14. Enter movie name – nazwa filmu zapisana do pliku XML
15. Enter subtitles language – język napisów zapisany do pliku XML
16. Movie frame rate – ilość ramek/s podawanego na wejście filmu, potrzebna do obliczeń w programie

Depth = (disparity + shift) * scale – ustawienia współczynników manipulujących parametrem głębokości zapisywanym w pliku XML. Współczynniki zmieniają parametr głębokości tak jak to jest pokazane we wzorze:

17. Scale factor for depth – współczynnik skalujący dysparycję
18. Shift factor for depth – współczynnik przesuwający dysparycję

Pozostała część interfejsu:

19. Show windows (when computing) – pozwala włączać/ wyłączać widoczne podczas obliczeń okien wyświetlających klatki:
 - obrazu dla lewego oka
 - obrazu dla prawego oka
 - mapy dysparycji
20. Open sequence LEFT (mark from IN to OUT – all files) – ładowanie sekwencji obrazów dla lewego oka – sposób umożliwiający ładowanie niedużej ilości plików, natomiast o nazwie nie składającej się tylko z cyfr, nazwa może być dowolna (np. lewy_001.png); w oknie dialogowym należy zaznaczyć klatkę początkową a następnie odnaleźć klatkę końcową i z klawiszem SHIFT kliknąć na ostatnią klatkę, następnie kliknąć Open
21. Subtitles file – ładowanie pliku napisów w formacie *.srt (darmowy konwerter napisów on-line KonwerterNapisow.pl)
22. Ścieżka załadowanego pliku napisów
23. Open sequence RIGHT (mark from IN to OUT – all files) – ładowanie sekwencji obrazów dla prawego oka – sposób tak jak w pkt. 20
24. Ścieżki do plików dla obrazów lewego oka
25. Ścieżki do plików dla obrazów prawego oka
26. Results save to – miejsce zapisania wyników działania programu (np. pliku XML, obrazów testowych, obrazów dla lewego i prawego oka z wklejonymi napisami)

27. Save to folder – wybór folderu do zapisania wyników programu z pkt. 26

28. Pasek postępu obliczeń

29. Pause computing – wstrzymanie obliczeń umożliwiające ich kontynuację po ponownym wciśnięciu tego pola

30. Start – rozpoczęcie głównego algorytmu programu

31. Subtitles output type:

- Text in XML file – zapisanie informacji o głębokości napisów i samego tekstu napisów w pliku XML

- Draw subs into sequence – wklejenie napisów bezpośrednio w kopię sekwencji obrazów dla lewego i prawego oka; wklejany tekst jest z odpowiednim przesunięciem w lewym i prawym oku, tak aby była zachowana żądana głębokość napisów

32. Open sequence LEFT (mark only IN and OUT – 2 files) – ładowanie sekwencji obrazów dla lewego oka – sposób ładowania umożliwiający ładowanie dużej ilości plików, ale z zastrzeżeniem że nazwa składa się tylko z liczb określających numery kolejnych klatek (np. 0001111.png); w oknie dialogowym należy zaznaczyć klatkę początkową a następnie odnaleźć klatkę końcową i z klawiszem CTRL kliknąć na ostatnią klatkę, następnie kliknąć Open

33. Open sequence RIGHT (mark only IN and OUT – 2 files) – ładowanie sekwencji obrazów dla lewego oka – sposób ładowania tak jak w pkt. 32

34. Test program (ON/OFF) – włączenie lub wyłączenie trybu testowego aplikacji.

Tryb ten pozwala na pokazanie poprawności działania programu dla sekwencji obrazów lewego i prawego oka.

- OFF – normalny tryb pracy programu

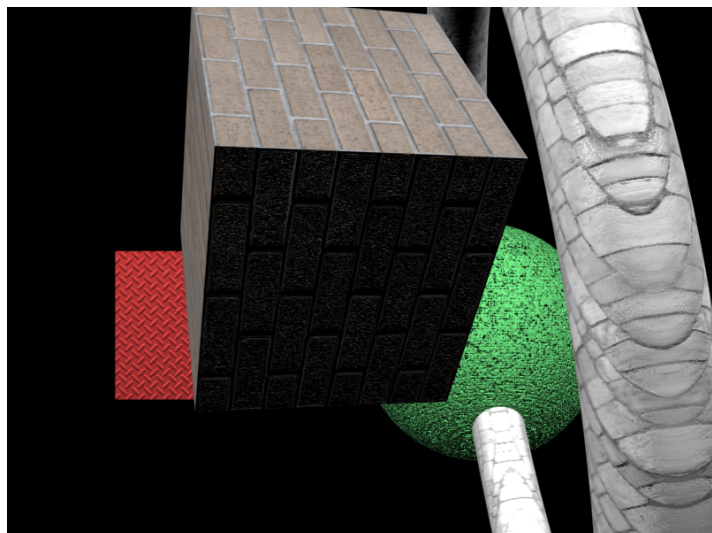
- ON – tryb testowy – w folderze na wyjściu programu pojawia się sekwencja dla lewego i prawego oka oraz sekwencja map dysparycji dla każdej z odpowiadających sobie par obrazów stereoskopowych. W tym trybie do programu nie trzeba ładować pliku napisów, gdyż do sekwencji stereoskopowej wklejane są napisy testowe i grafika testowa. Tryb testowy wkleja napisy testowe, tak jakby dla każdej klatki napisów były oddzielne napisy. Efektem tego po zakończeniu obliczeń w katalogu wyjściowym mamy sekwencję obrazów dla lewego i prawego oka z wstawioną grafiką testową i napisami testowymi o głębokości wyliczonej tylko dla pojedynczej pary stereoskopowej. Podsumowując, w tym trybie, głębokość napisów i grafiki zmienia się z klatki na klatkę i jest taka sama jak głębokość najbliższego dla widza obiektu.

Przykładowa sekwencja obrazów wejściowych

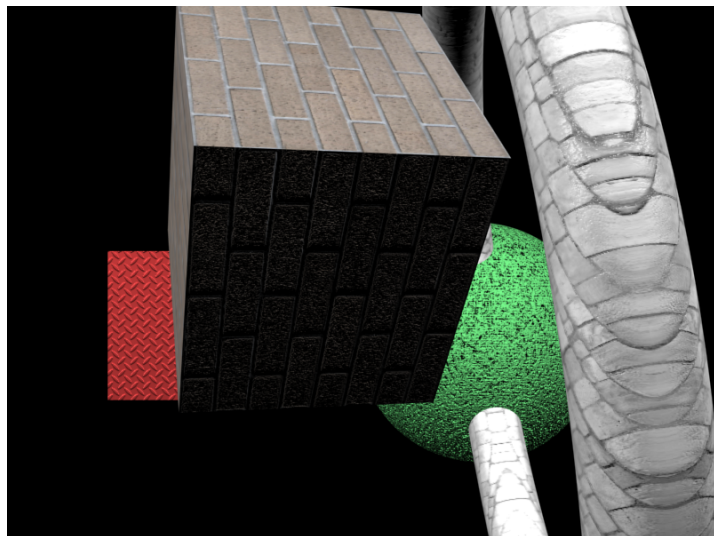
WEJŚCIE:

Wybrana para obrazów stereoskopowych z sekwencji:

Obraz dla lewego oka



Obraz dla prawego oka



Plik z napisami w formacie *.srt:

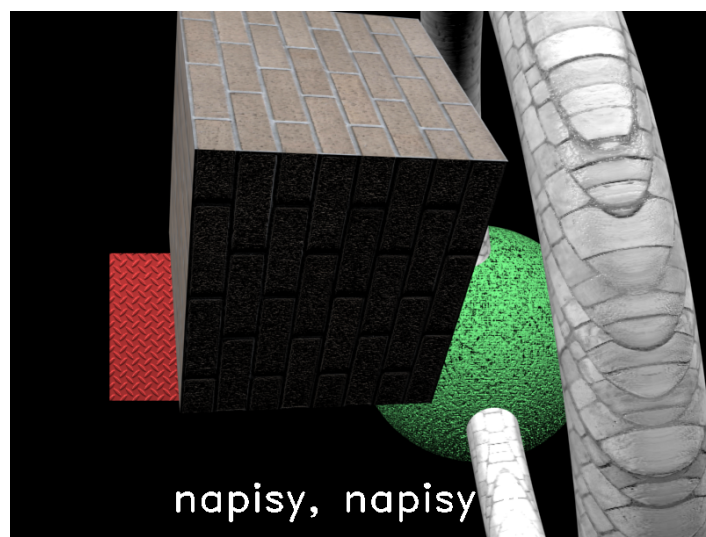
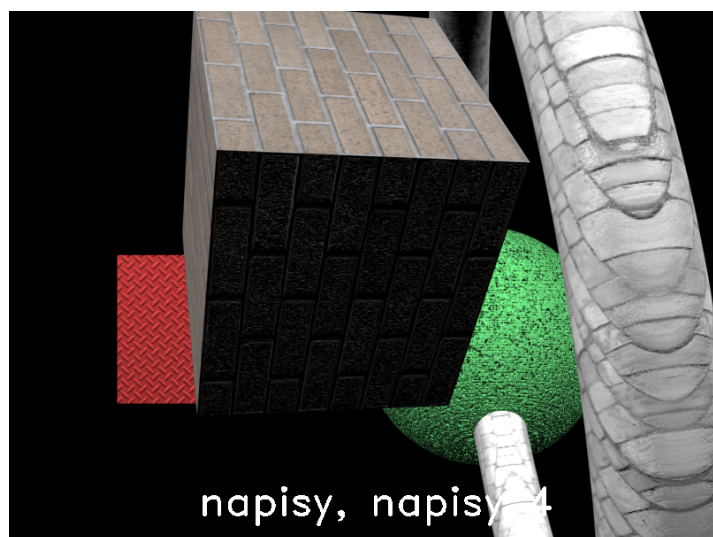
```
1
00:00:00,000 --> 00:00:00,960
napisy, napisy 1

2
00:00:01,000 --> 00:00:01,960
napisy, napisy 2

3
00:00:02,000 --> 00:00:02,960
napisy, napisy 3

4
00:00:03,000 --> 00:00:04,000
napisy, napisy 4
```


WYJŚCIE:
(dla napisów wklejanych bezpośrednio w sekwencji)

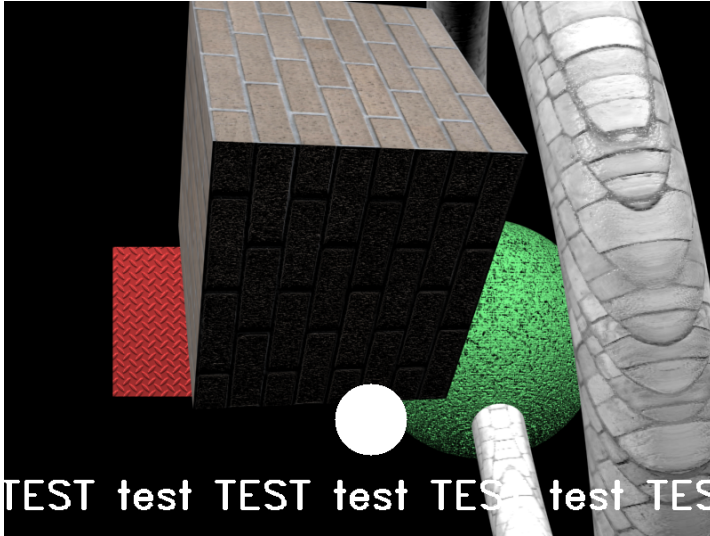


WYJŚCIE:
(dla pliku XML)

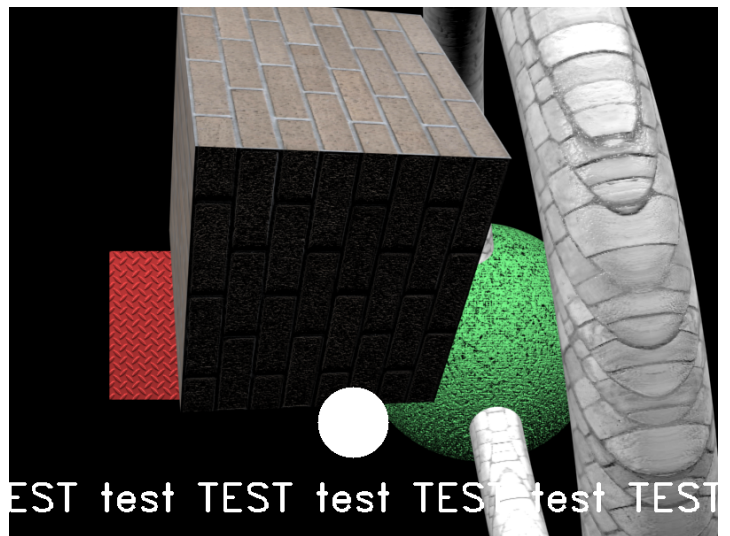
```
<?xml version="1.0" encoding="UTF-8"?>
<SubtitleID>1c93837c-d504-48bd-b54b-387036103799</SubtitleID>
<MovieTitle>MovieName</MovieTitle>
<ReelNumber>0</ReelNumber>
<Language>SubtitlesLanguage</Language>
<Subtitle SpotNumber="1" TimeIn="0:0:0:0" TimeOut="0:0:0:960">
  <Text HPosition="0" VAlign="bottom" VPosition="10" ZPosition="0,921875">napisy, napisy 1</Text>
</Subtitle>
<Subtitle SpotNumber="2" TimeIn="0:0:1:0" TimeOut="0:0:1:960">
  <Text HPosition="0" VAlign="bottom" VPosition="10" ZPosition="1,125">napisy, napisy 2</Text>
</Subtitle>
<Subtitle SpotNumber="3" TimeIn="0:0:2:0" TimeOut="0:0:2:960">
  <Text HPosition="0" VAlign="bottom" VPosition="10" ZPosition="1,1171875">napisy, napisy 3</Text>
</Subtitle>
<Subtitle SpotNumber="4" TimeIn="0:0:3:0" TimeOut="0:0:4:0">
  <Text HPosition="0" VAlign="bottom" VPosition="10" ZPosition="1,875">napisy, napisy 4</Text>
</Subtitle>
```

WYJŚCIE:
(dla trybu testowego)

Obraz dla lewego oka



Obraz dla prawego oka



Mapa dysparycji:
(po pierwszym uruchomieniu)



Mapa dysparycji:
(po drugim uruchomieniu)



Tryb testowy został stworzony do pokazania (sprawdzenia) mapy dysparycji oraz sprawdzenia algorytmu wklejającego napisy do sekwencji obrazów stereoskopowych w wyliczonej głębokości.

Uruchamiając ten tryb za pierwszym razem na wyjściu programu otrzymujemy parę obrazów stereoskopowych takie jak zamieszczone na początku poprzedniej strony oraz mapę dysparycji „po pierwszym uruchomieniu”. Para stereoskopowa posiada wklejone napisy testowe i koło testowe.

Ten tryb różni się od zwykłego trybu liczenia głębokości napisów ponieważ liczy on maksymalną dysparycję dla pojedynczej pary stereoskopowej i z taką dysparycją są wklejane napisy testowe. Efektem tego mamy sekwencję par stereoskopowych dla których napisy testowe są wstawiane na takiej głębokości, na jakiej znajduje się najbliższy dla widza obiekt w danej parze stereoskopowej. Oznacza to że głębokość napisów uaktualnia się razem z kolejnymi parami stereoskopowymi.

Dodatkowo można program uruchomić ponownie w trybie testowym tylko jako wejście podać materiał wyjściowy z pierwszego uruchomienia (czyli razem z napisami i kołem testowym). Po obliczeniach otrzymamy mapę dysparycji „po drugim uruchomieniu”. Można w niej odnaleźć informacje o głębokościach obiektów w scenie (oraz przy takich materiałach jakie został podane na wejście – razem z testową grafiką i napisami) a także głębokość napisów w scenie. Głębia odzwierciedla jasność pikseli, im jaśniejszy piksel tym w oglądanym obrazie stereoskopowym jest bliżej widza. Na mapie dysparycji „po drugim uruchomieniu” widać że jasność pikseli napisów (pkt 1) jest identyczna jak jasność najbliższego obiektu, w tym przypadku rogu sześcianu (pkt 2). Oznacza to, że dla widza, sześcian i napisy będą na tej samej głębokości, co jest zgodne z oczekiwanym efektem.

Opis funkcji z kodu źródłowego

- `AllocateImages` – funkcja przygotowująca pamięć na obrazy źródłowe i pośrednie wykorzystywane w obliczeniach
- `DeallocateImages` – funkcja zwalniana powyższą pamięć
- `convertSysStrToStdStr` – konwersja z zmiennej typu `System::String` na `std::string`
- `drawInStereoStuff` – wkleja napisy z pliku SRT lub napisy testowe, bezpośrednio do kopii wejściowej sekwencji stereoskopowej
- `getPtrToCharFromSysString` – zwraca nazwę pliku do wskaźnika typu `CHAR`
- `getTimecodesAndSubtitles` – przetwarza plik napisów SRT i zapisuje w tablicach czas pojawienia się i wygaszenia napisów zmieniając go na numery klatek oraz pobieranie tekstu tych napisów
- `makingXML` – utworzenie pliku XML z wyliczonymi w programie informacjami (opis pliku XML powyżej)
- `startStereo` – główna funkcja programu, zawierająca główną pętlę

W kodzie źródłowym przyjęty został następujący sposób oznaczania zmiennych:

`zasiegzmiennej_typzmiennejNazwazmiennej`

`np.:`

`g_iLicznik` – zmienna globalna typu `int`

`_uiLicznik` – zmienna lokalna typu `unsigned int`

LITERATURA

- dokumentacja bibliotek przetwarzania obrazu OpenCV
- „Image processing and recognition” - Robert Sitnik
- *Cyfrowe przetwarzanie obrazu* – M.Kujawińska
- dokumentacja środowiska programistycznego C++
- „OpenCV learning” - O'Reily
- dokumentacja środowiska programistycznego Microsoft Visual C++ 2008 Express Edition