

*Automatyczne tworzenie trójwymiarowego
planu pomieszczenia z zastosowaniem
metod stereowizyjnych*

autor: **Robert Drab**
opiekun naukowy: **dr inż. Paweł Rotter**

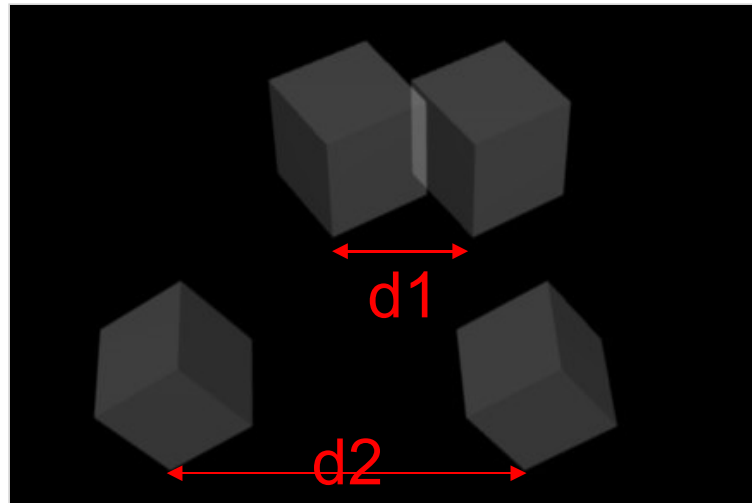
1. Wstęp

Zagadnienie generowania trójwymiarowego planu pomieszczenia z wykorzystaniem metod stereowizyjnych jest interesujące ze względu na możliwości wykorzystania tego typu informacji tak w sterowaniu robotów jak i innych obszarach, nie koniecznie związanych z automatyką.

Główny nacisk został położony na same algorytmy stereowizyjne, w szczególności na zastosowanie większej liczby kamer w celu poprawienia dokładności odtwarzania przestrzeni 3D. Do realizacji projektu zostało wykorzystane środowisko Matlab oferujące szeroką gamę narzędzi oraz algorytmów wspomagających jak obsługa akwizycji danych, filtracja, itp.

2. Przetwarzanie informacji w systemie stereowizyjnym

Stereowizja bazuje na dysparycji, czyli odległości między obrazami tego samego punktu na płaszczyznach dwu kamer.



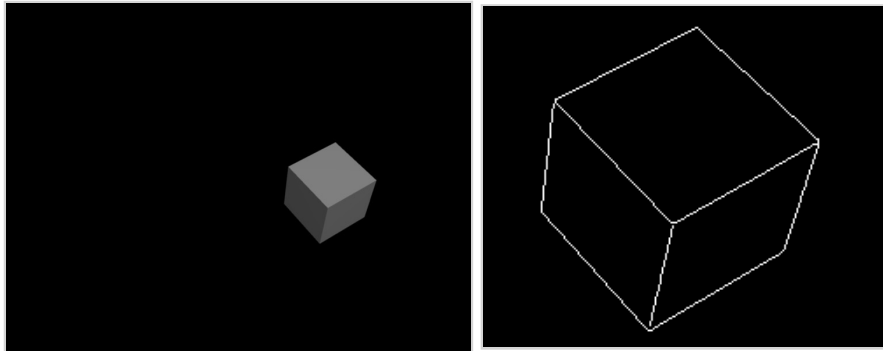
Rysunek 1: Dysparycja

Mniejsza dysparycja oznacza w rzeczywistości większą odległość danego obiektu od układu kamer. W powyższym przypadku widać, iż górny sześcian jest dużo dalej niż dolny.

W procesie tworzenia rzadkiej mapy dysparycji każda stereopara musi przejść przez trzy etapy przetwarzania:

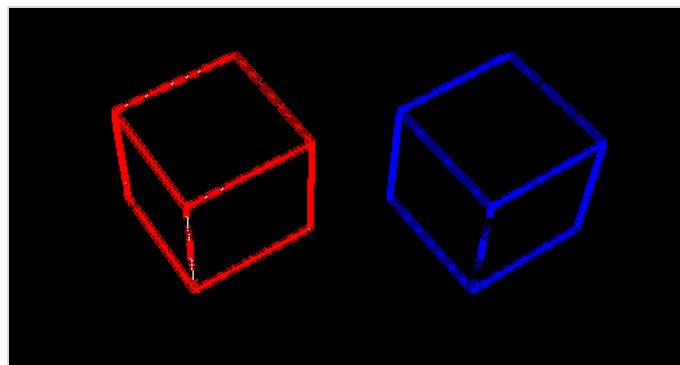
- detekcja punktów charakterystycznych dla których obliczana będzie wartość głębi, np. krawędzi
- dopasowywanie odpowiedników
- rekonstrukcja współrzędnych w przestrzeni 3D

W celu wyznaczenia punktów do dalszego przetwarzania należy przeprowadzić detekcję krawędzi. Poniżej znajduje się przykładowy obraz wejściowy oraz efekt filtracji:



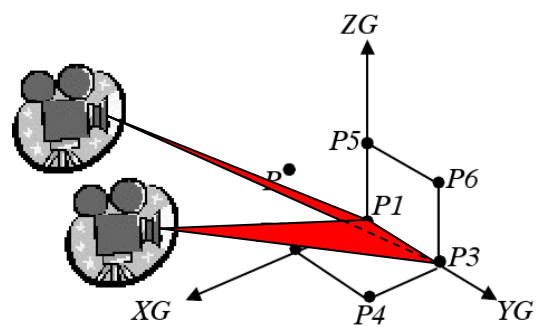
Rysunek 2: Obraz przed i po detekcji krawędzi

Na tak przygotowanych obrazach możemy przeprowadzić operację dopasowywania odpowiedników. Generalnie nie jest to zagadnienie proste, ze względu na możliwość występowania fałszywych dopasowań. W celu wyeliminowania tego zjawiska stosuje się miary dopasowania, w tym wypadku wykorzystywana była miara SSD (sum of squared difference). W ten sposób badany jest piksel wraz z sąsiedztwem, co pozwala stwierdzić, czy znaleziony piksel na drugim obrazie odpowiada poszukiwanemu. Kolejnym zabezpieczeniem jest wykorzystanie dopasowania zwrotnego, które polega na tym, iż najpierw np. lewy obraz jest dopasowywany do prawego, a następnie odwrotnie. Do wynikowej mapy dysparycji przechodzą tylko te dopasowania, które dla każdego z dwu dopasowań były zgodne.



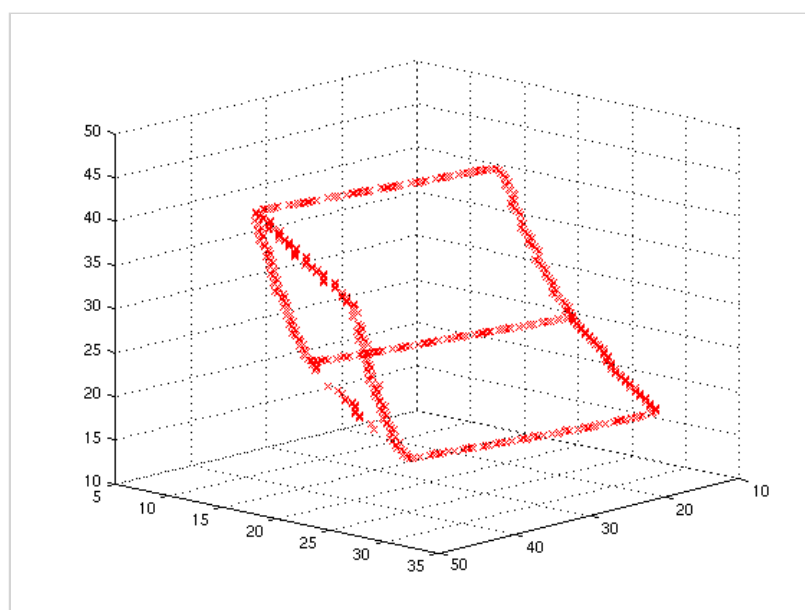
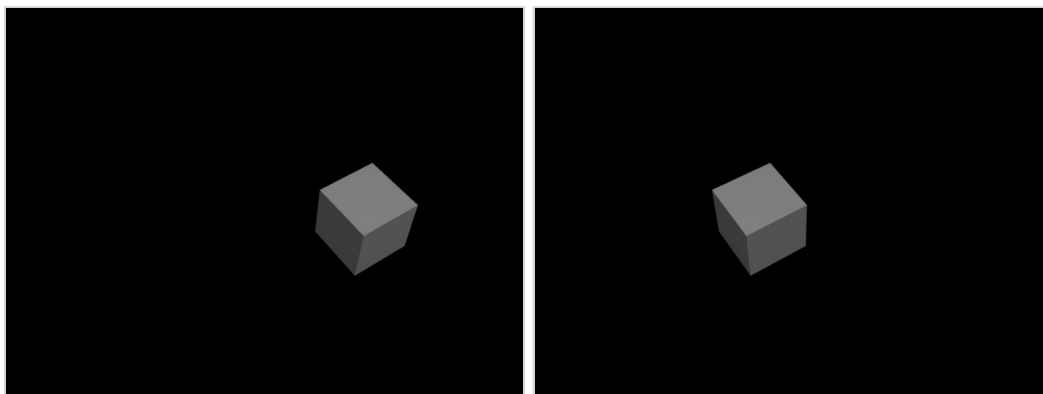
Rysunek 3: Przykładowa mapa dysparycji

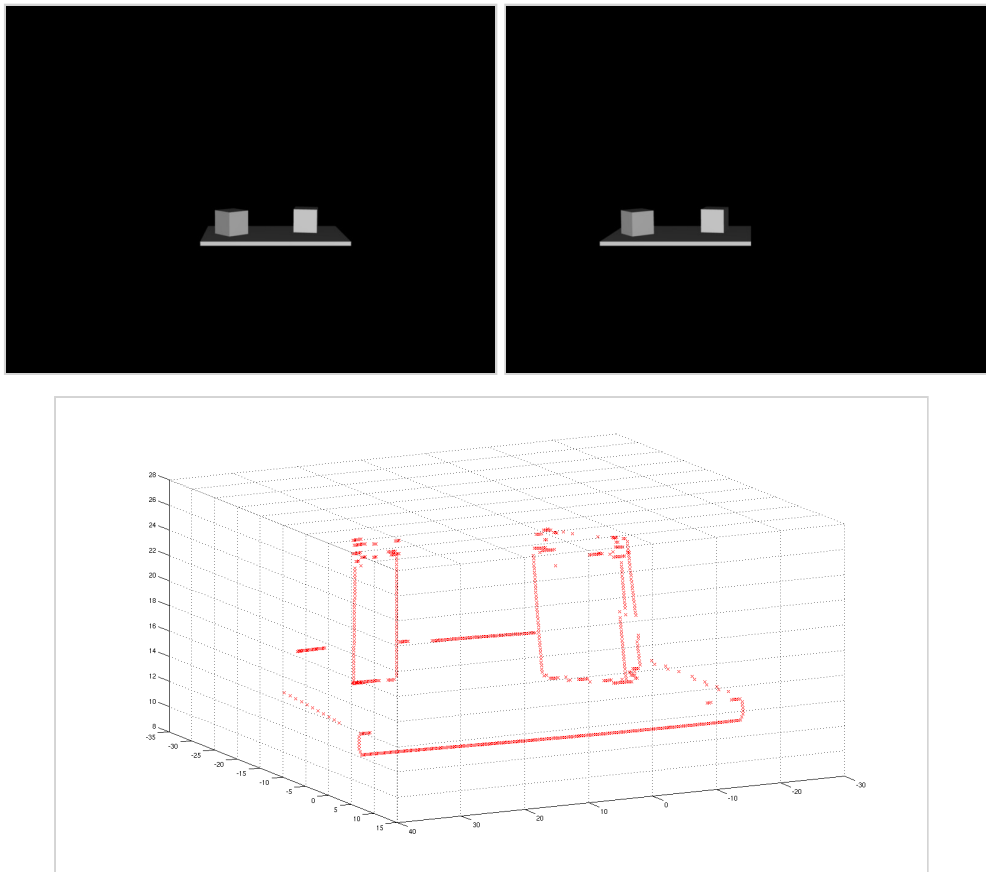
Ostatnim etapem przetwarzania jest rekonstrukcja przestrzeni 3D. W tym przypadku jest do tego wykorzystana transformacja płaska, bardzo efektywna i elegancka metoda odzyskiwania informacji trójwymiarowej, niemniej jednak nie pozbawiona wad – efekt przetwarzania uzależniony od numerycznego wyznaczania prostej między dwoma punktami, na czym bazuje cała metoda. W przypadku niewielkiej odległości między punktami, błędy numeryczne mogą prowadzić do nieprzewidywalnych wyników ze względu na złe uwarunkowanie zadania. Jest to istotne w bliskim otoczeniu płaszczyzny łączącej oś poziomą (YG) z osią optyczną kamery – rysunek poniżej.



Rysunek 4: Obszar silnej wrażliwości transformacji płaskiej na obliczenia numeryczne

Najciekawszą częścią są oczywiście rezultaty. Poniżej znajdują się przykładowe stereopary oraz wyniki przetwarzania.

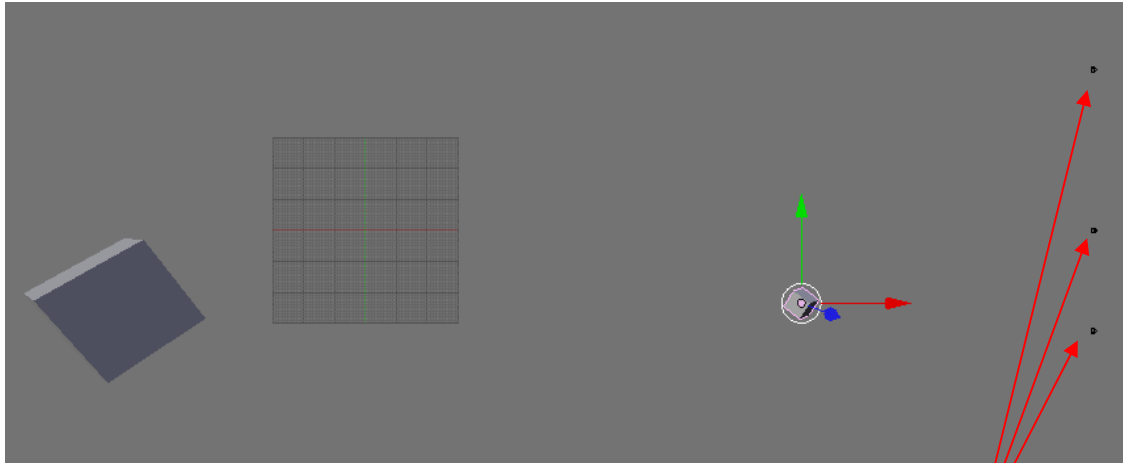




Jak widać na powyższym przykładzie nie wszystkie krawędzie zostały poprawnie odnalezione. Wynika to głównie z dużego wpływu oświetlenia sceny na jakość końcowego wyniku.

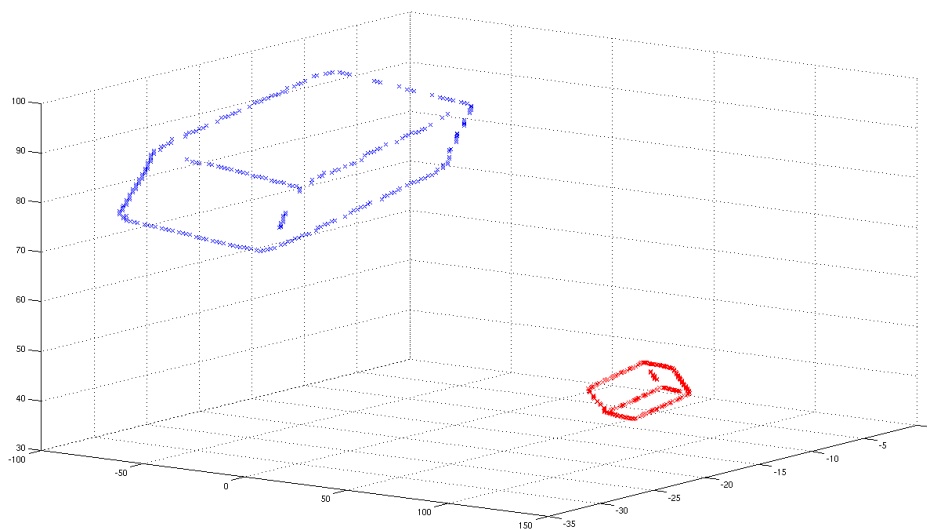
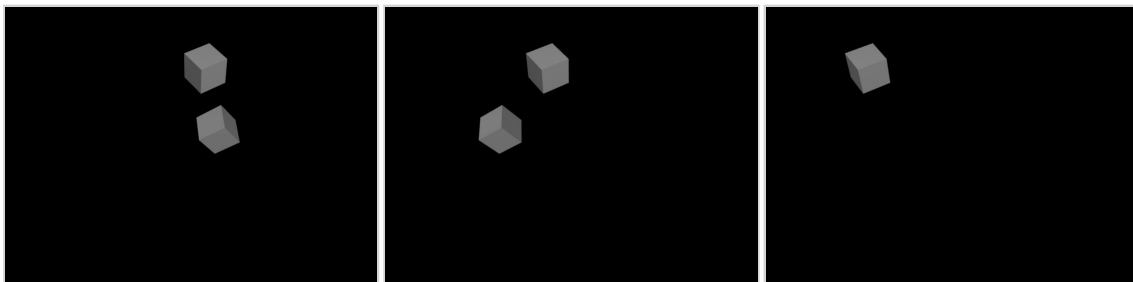
3. Stereowizja wielokamerowa

Stereowizja wielokamerowa jest wciąż mało popularna pomimo sporego potencjału tej metody. Jak wiadomo ilość poziomów głębi oraz precyzja ich wydobywania ze stereopary jest funkcją rozdzielczości kamer oraz odległości między osiami optycznymi kamer. Wyższa rozdzielczość oznacza większą ilość poziomów głębi, lecz również wyższą cenę zestawu stereowizyjnego. Z kolei większa odległość między kamerami skutkuje większą ilością poziomów głębi, jak i precyzją umiejscowienia w przestrzeni, jednak w przypadku zbyt szerokiego rozstawienia kamer bliski plan staje się niemożliwy do odtworzenia z powodu dysparycji przekraczającej rozdzielczość kamery. Pomysł jest więc następujący: zamiast dwu drogich kamer wysokiej klasy, które i tak nie gwarantują dobrej jakości przetwarzania, wykorzystana zostanie większa liczba zwykłych kamer. W tym wypadku badania były prowadzone na trzech kamerach rozstawionych w odległościach zgodnych z zasadą złotego podziału – rysunek poniżej.



Kamery

W ten sposób, korzystając niejako z trzech oddzielnych zestawów stereowizyjnych, jesteśmy w stanie, korzystając z najbardziej skrajnie rozstawionych kamer uzyskać dokładne informacje na temat najdalszej (najgłębszej) części sceny, następnie korzystając z dwu bliżej rozstawionych kamer przetwarzamy środkową część sceny, a najbliższy plan analizujemy wykorzystując dwie najbliższe położone kamery. W porównaniu do standardowego zestawu stereowizyjnego, który albo nie był w stanie przetworzyć bliskiej części sceny, albo dawał nieprawidłowe informacje na temat części odległej, tego typu zestaw daje bardzo dobre efekty, obejmując całą scenę oraz dostarczając dokładne dane na temat położenia obiektów w przestrzeni. Poniżej znajdują się przykładowe trzy zdjęcia z poszczególnych kamer oraz efekt przetwarzania.



Kolorem niebieskim zaznaczony został zaznaczony sześcian odtworzony z najbardziej skrajnej pary kamer, natomiast czerwony z pary najbardziej zbliżonych kamer. Jest to przykład z powyższego rysunku z zaznaczonymi kamerami. Sześciany zostały poprawnie odwzorowane, zarówno ich kształt, jak i umiejscowienie w przestrzeni dokładnie zgadza się ze współrzędnymi z programu do tworzenia grafiki 3D, w którym scena została przygotowana.

Jak widać metoda ta daje bardzo dobre rezultaty w przypadku przetwarzania rozległych, skomplikowanych scen. Daje również możliwość ograniczenia wpływu przesłonek obiektów.

4. Podsumowanie

Projekt jest ciągle rozwijany, gdyż do tej pory testy prowadzone były wyłącznie na obrazach syntetycznych. Kolejnym etapem będzie przetwarzanie obrazów rzeczywistych, które są dużo bardziej wymagające, przede wszystkim ze względu na filtrację i detekcję krawędzi. Docelowo projekt miałby działać w czasie zbliżonym do rzeczywistego, co umożliwiłoby tworzenie trójwymiarowego planu całych budynków, np. przez robota mobilnego. Na obecnym etapie projekt daje bardzo zadowalające efekty oraz przejawia duże możliwości dalszego rozwoju.