

Sprawozdanie z Pracowni Dyplomowej Inżynierskiej 1

Łukasz Żmuda

Temat pracy: [Wykorzystanie aktywnej wizji do identyfikacji obiektów](#)

Opiekun pracy: mgr inż. Tomasz Kornuta

1. Motywacja pracy.

W życiu codziennym identyfikacja obiektów za pomocą odczytów z kamery w trybie on-line przydaje się bardzo często. Może mieć zastosowanie np. w armii (namierzanie obiektów do zestrzelenia), ale również bardziej przyziemne np. robot może odczytywać położenie ludzkiej dłoni i w zależności od tego co zostało odczytane, wykonać odpowiednie zadanie. Stwarza to bardzo duże możliwości i umożliwia komunikację z robotem na odległość tylko i wyłącznie za pomocą gestów. Dzięki temu, znika całkowicie bariera językowa, robot rozumie polecenia bez względu na język, jakim posługuje się wydający komendy.

2. Cele i zakres pracy dyplomowej.

Docelowo robot na którym ma być umieszczona kamera, powinien znajdować ludzką dłoń i odczytywać jej kształt. Należy więc skupić się na dwóch aspektach.

2.1. Odczyt i odpowiednia obróbka obrazu z kamery. Ta część będzie realizowana za pomocą aplikacji Discode.

2.2. Ruchu ramienia robota w zależności od wyniku przetwarzania obrazu odczytanego przez kamerę. Robot będzie sterowany za pomocą aplikacji MRROC++.

Chcąc nieco ułatwić sobie zadanie na początku, skupię się na rozpoznawaniu jednokolorowej piłki. Następnie, gdy ten etap zostanie zrealizowany, skupię się na rozpoznawaniu kształtu ludzkiej dłoni.

3. Zrealizowane zadania.

W mijającym semestrze skupiłem się przede wszystkim na zapoznaniu się z aplikacjami MRROC++, oraz Discode, a także Żbik. Zainstalowałem i skonfigurowałem odpowiednio aplikację MRROC++ i Eclipse. Pozwala on na zarządzaniem pracą robotów zarówno za pomocą graficznego interfejsu użytkownika, jak i za pomocą modułów napisanych w C++ i wczytanego pliku konfiguracyjnego. Aby móc przeprowadzać testy zainstalowałem aplikację Żbik. Na początku wykorzystywałem tylko GUI, dzięki temu mogłem szybko poznać budowę i zasadę działania poszczególnych modułów mrroc++. Następnie napisałem szkielek własnego modułu, który uruchomiłem w trybie testowym. Tryb testowy pozwala na uruchamianie zadań bez zaangażowania rzeczywistych robotów co zwłaszcza na początku pracy jest bardzo wskazane. Jednak testy wykonywałem jeszcze bez wykorzystania Żbik'a. Kolejną aplikacją, którą poznałem był Discode. Jest ona stworzona do przetwarzania obrazów odczytanych bezpośrednio z kamery. Składa się z czterech podstawowych komponentów: źródła, procesory danych, proxy i komponent odpowiedzialny za prezentację danych i ich zapis do pliku. Discode pozwala na sprawne wyszukiwanie obiektu znajdującego się w kadrze kamery. Aby uruchomić zadanie należy napisać odpowiednią

funkcjonalność w języku C++ i dodatkowo wczytać plik konfiguracyjny (w nim są umieszczone m.in. ścieżki do wejściowych plików). Po instalacji Discode'a, wykonałem wszystkie zadania związane z przetwarzaniem filmów i obrazów znajdujących się w tutorialu i dotyczących poziomu podstawowego.

4. Dalszy plan pracy.

Teraz, kiedy dokładnie poznałem wszystkie narzędzia, które będą mi potrzebne, mogę zabrać się do realizacji właściwego zadania.

Żeby nie uszkodzić robota będę na początku wykorzystywał tryb testowy. Zatem konieczne będzie uruchamianie MRROC++'a z aplikacją Żbik.

Kolejnym etapem będzie napisanie szkieletu programu w MRROC++. Chodzi tutaj o napisanie odpowiednich modułów odpowiedzialnych za określone zachowania.

4.1. Jeżeli poszukiwany obiekt nie jest widoczny w kamerze to robot powinien szukać tego obiektu, tzn. obracać ramieniem na którym umieszczona jest kamera.

4.2. Jeżeli znajdzie obiekt to powinien do niego się zbliżyć.

4.3. Jeśli się zbliżył powinien wykonać ruchy z góry na dół i na boki, aby jak najdokładniej zidentyfikować obiekt.

4.4. Kolejny moduł odpowiedzialny będzie za podążanie za obiektem w przypadku, kiedy ten będzie zmieniał swoje położenie.

Przetworzenie obrazu z kamery odbędzie się oczywiście za pomocą aplikacji Discode. W zależności od odczytu MRROC++ będzie wykonywał jeden z opisanych wyżej modułów.

Docelowo rozpoznawanym obiektem ma być dłoń, a dokładniej jej kształt. Jednak w związku z tym, że nie mam jak dotąd żadnego doświadczenia praktycznego w przetwarzaniu obrazów, postanowiłem podzielić ten problem na dwie fazy.

W pierwszej fazie moim celem będzie rozpoznanie jednokolorowej piłki. Jeżeli osiągnę swój cel tzn. robot będzie znajdował piłkę i podążał za nią, dopiero wtedy skupię się na rozpoznawaniu dłoni i jej kształtu. Taki podział znacząco ułatwi mi realizację zadań.

Najważniejsza i najtrudniejsza będzie dla mnie faza pierwsza. Żeby osiągnąć efekt końcowy będzie trzeba dodać tylko funkcjonalność związaną z rozpoznawaniem kształtu dłoni.