

Tomasz GRZEJSZCZAK, Natalia BARTOSIAK, Martyna WOJNAR,
Kamil SKOWROŃSKI, Eryka PROBIERZ
Politechnika Śląska

REGULACJA POZYCJI ROBOTA SPOŁECZNEGO W SPRZĘŻENIU ZWROTNYM Z SYSTEMEM WIZYJNYM

Streszczenie. Utrzymanie kontaktu wzrokowego jest ważnym aspektem w interakcji człowieka z robotem (HRI). Artykuł przedstawia implementację i procedurę strojenia regulatora PID. Jako sygnał sprzężenia zwrotnego użyty jest system wizyjny umożliwiający wykrywanie twarzy. Odległość wykrytej twarzy od środka obrazu traktowana jest jako uchyb. We wnioskach podsumowano praktyczne problemy napotkane podczas fizycznej implementacji.

POSITION REGULATION OF THE SOCIAL ROBOT IN FEEDBACK WITH THE VISION SYSTEM

Summary. Maintaining eye contact is an important aspect of Human-Robot Interaction (HRI). The paper presents the implementation and tuning procedure of the PID controller. A vision system with a face detection algorithm is used as a feedback signal. The distance between the detected face and the center of the image is an error. The conclusions summarize the practical problems encountered during physical implementation.

1. Wstęp

Robot społeczny to autonomiczny robot, który wchodzi w interakcje i komunikuje się z ludźmi lub innymi autonomicznymi obiektami, przestrzegając zachowań społecznych i reguł związanych z jego rolą [2,3]. Ważnym zadaniem podczas konstrukcji robotów społecznych jest implementacja odpowiednich algorytmów planowania trajektorii, czyli zmiany położenia, prędkości i przyspieszenia w celu sprawienia by interakcja człowieka z robotem (ang. *Human Robot Interaction - HRI*) przebiegała w sposób naturalny [4]. Jednym z aspektów HRI mających na celu polepszenie jakości interakcji jest utrzymanie kontaktu wzrokowego. Udowodniono, że uczestnicy wykazali bardziej skoordynowane i zsynchronizowane multimodalne zachowania między mową a spojrzeniem, gdy z powodzeniem nawiązano i utrzymano większy kontakt wzrokowy [8]. Utrzymanie kontaktu wzrokowego odbywa się najczęściej za pomocą systemu wizyjnego wyposażonego w algorytmy wykrywania twarzy. Najbardziej popularnym algorytmem dającym zadowalające wyniki jest algorytm bazujący na kaskadach Haara który z powodzeniem można zastosować do

HRI [5]. W poniższym opracowaniu zastosowano system wizyjny umożliwiający detekcję twarzy jako sygnał sprzężenia zwrotnego regulatora PID, który wyznacza pozycję robotycznej głowy. Głównym aspektem badawczym jest dostosowanie nastaw regulatora działającego w sposób dyskretny z próbkowaniem zależnym od czasu przetwarzania sygnału wizyjnego [6,7]. Dodatkowo układ posiada 2 wejścia i wyjścia: dla rotacji poziomej i pionowej. Nastawy regulatora dobrano zgodnie z metodą Zieglera Nicholasa [1]. Podobne zastosowania regulacji trajektorii z zastosowaniem PID można odnaleźć w przykładach robotów śledzących i podążających za użytkownikiem [9].

2. Opis stanowiska

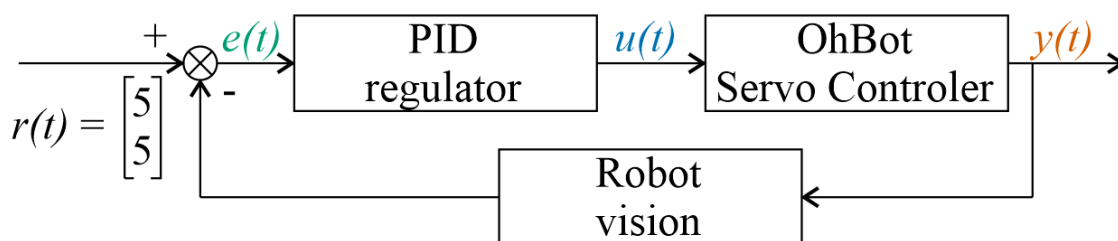
Stanowisko robota społecznego składa się z robotycznej głowy, czyli gotowego urządzenia o nazwie OhBot, do którego przymocowano kamerę USB. Stanowisko zaprezentowane jest na rysunku 1. Oprogramowanie robota wykonano w języku Python z pomocą dedykowanych bibliotek. Głównymi elementami stanowiska są:

- Robotyczna głowa w której skład wchodzi serwomotory, kontroler serwomotorów oparty o bibliotekę *ohbot*, oraz konstrukcja imitująca głowę.
- System wizyjny z kamerą Logitech c920 bazujący na bibliotece *OpenCV*
- Regulator PID zaimplementowany z pomocą biblioteki *simple_pid*



Rys. 1. Stanowisko robota społecznego *OhBot* z kamerą

Postawione zadanie regulacji polega na imitacji kontaktu wzrokowego z użytkownikiem. Celem systemu wizyjnego jest wykrycie twarzy użytkownika a celem regulatora jest korekcja pozycji robota tak aby wykryta twarz znalazła się na środku obrazu. W tym celu utworzono pętlę sprzężenia zwrotnego, zgodnie z rysunkiem 2.



Rys. 2. Schemat blokowy układu regulacji

Układ ma własności nadążne w związku z zastosowaniem serwomotorów. Zastępcza transmisja serwomotoru jest inercją z małym opóźnieniem wynikającą z taniej konstrukcji robotycznej głowy (np. luzy na plastikowych zębatkach). Zakłada się jednak, że opóźnienie jest pomijalne a obiekt dla którego będzie projektowany regulator PID jest inercją z opóźnieniem. W rozdziale 4 opisano sposób doboru nastawy regulatora zgodnie z syntezą Zieglera-Nicholsa, jednak alternatywą jest wyznaczenie parametrów dla obiektu zastępczego. Dokonując syntezy regulatora PID zgodnie z regułami Zieglera-Nicholsa zaobserwowaliśmy bardzo dobre własności układu sterowania. W związku z tym, że obiekt może być reprezentowany przez inercję z opóźnieniem, co spełnia założenia metody, wybrana procedura doboru parametrów PID jest metodą referencyjną. Zatem można wnioskować, że inercja z opóźnieniem jest dobrą aproksymacją obiektu sterowania.

Warto zauważyć, że zgodnie ze specyfikacją biblioteki *ohBot*, sygnałem sterującym serwomotorami jest ich pozycja docelowa w przedziale $[0,10]$, dlatego wartością zadaną jest środek przedziału, czyli $r(t)=[5,5]^T$. Każdy pozostały sygnał również jest rozpatrywany w przedziale $[0,10]$.

W związku z tym, że system wizyjny zwraca błąd w osi pionowej i poziomej a sygnał sterujący również przesyłany jest do 2 serwokontrolerów modyfikujących położenie głowy w pionie i poziomie, układ posiada 2 wejścia i 2 wyjścia. Jednak ze względu na brak zależności skrośnych, pętla regulacji jest układem typu 2 x SISO (ang. Single Input Single Output). Każdy z powyższych sygnałów jest wektorem 2 wymiarowym.

3. System wizyjny

Ważnym elementem umożliwiającym pomiar błędu i regulację pozycji robota jest system wizyjny. Opracowany algorytm przetwarzania obrazów składa się z algorytmu detekcji twarzy na obrazie za pomocą kaskad Haara. Testy prowadzono na domyślnych wagach dostarczonych wraz z biblioteką *OpenCV*. Algorytm składa się z następujących punktów:

1. Pobranie klatki z kamery
2. Detekcja twarzy na obrazie (pozycji prostokąta okalającego)
3. Dla twarzy o najwyższej pewności, detekcji wyznaczenie środka prostokąta okalającego (x_F, y_F)
4. Sprowadzenie wyniku do przedziału $[0,10]$

Należy pamiętać, że wynikiem układu sterowania jest pozycja serwomotoru, czyli $y(t)$. Pozycja ta wpływa bezpośrednio na pole widzenia kamery. Pozycja wykrytej twarzy jest w układzie współrzędnych kamery, więc przed wyznaczeniem uchybu należy proporcjonalnie zmienić wyjście systemu wizyjnego poprzez przemnożenie wartości przez 10 i dzieląc przez wymiary klatki obrazu.

Obliczenia w układzie kamery powinny uwzględniać skalowanie oraz zniekształcenia geometryczne wynikające z właściwości fizycznych soczewek. Są to jednak prace przewidziane w dalszych badaniach mających na celu poprawę jakości. Problem skalowania bezpośrednio przekłada się na wzmocnienie graniczne, co

zaobserwowano przy doborze parametrów. Dla wyeliminowania tego problemu założono stałą odległość użytkownika od kamery wynoszącą 1m.

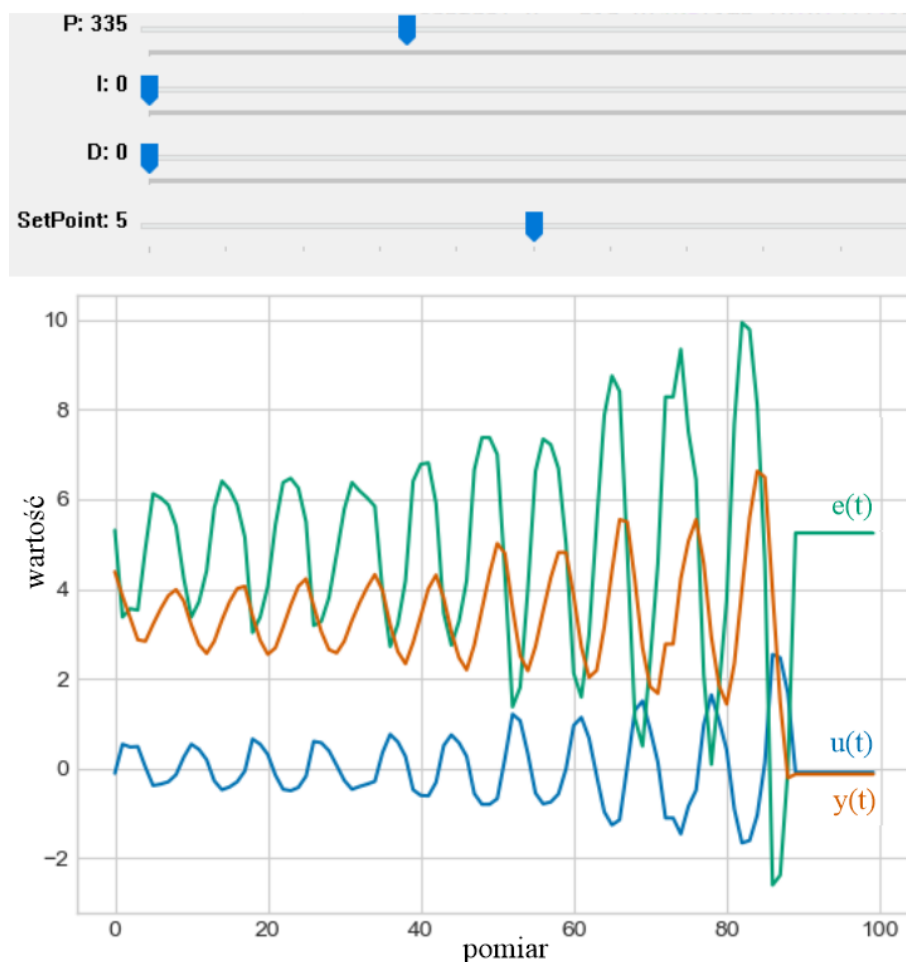
4. Nastawy regulatorów

W celu dobrania odpowiednich parametrów regulatora PID wybrano metodę Zieglera-Nicholsa. Biorąc pod uwagę wzór regulatora

$$K_r(s) = k_r \left(1 + \frac{1}{T_i s} + T_d s \right) \quad (1)$$

należy wyznaczyć wzmocnienie regulatora k_r , oraz stałe czasowe elementu całkującego T_i oraz różniczkującego T_d . Program obsługujący system wizyjny i kontroler wyposażono w suwaki, które umożliwiały na bieżąco modyfikacje parametrów regulatora oraz obserwację 100 ostatnich pomiarów. W związku z ograniczeniami implementacyjnymi, wartości PID widoczne na rysunku 3 są dzielone przez 1000, co definiuje dokładność pomiarową.

Pierwszym krokiem metody Zieglera-Nicholsa jest wyznaczenie wzmocnienia granicznego, przy którym w układzie występują nie gasnące oscylacje (układ na granicy stabilności).



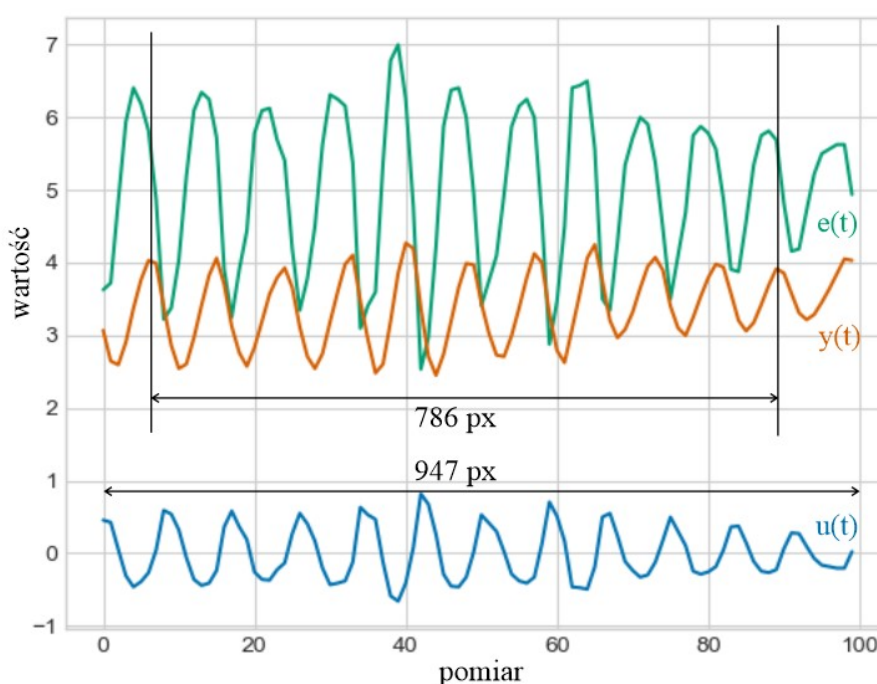
Rys. 3. Destabilizacja układu po próbie ustawienia wzmocnienia granicznego

Wartości graniczne to $k_{gr} = [0,301 \ 0,331]^T$, dla których przedstawiono przebiegi

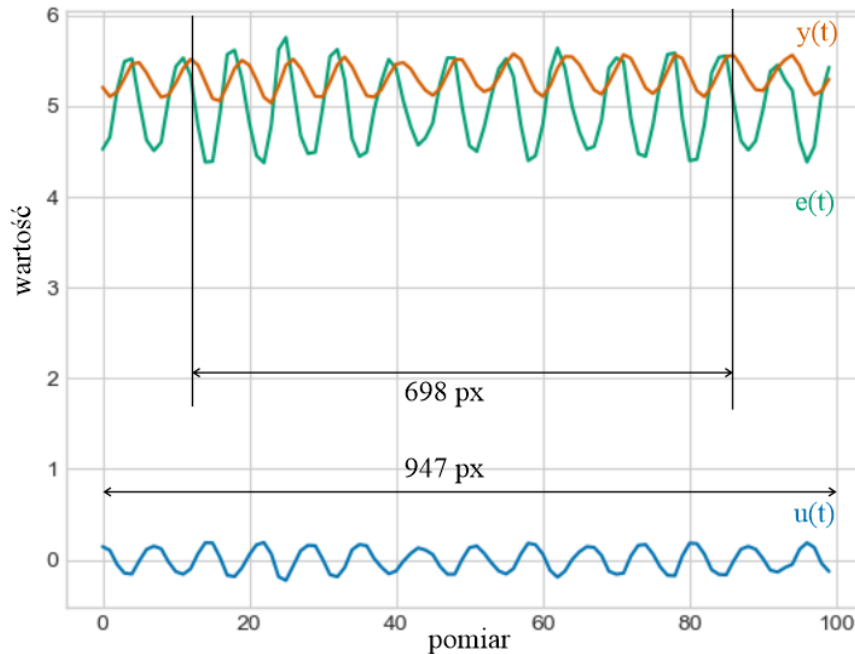
na rysunkach 4 i 5. W trakcie testów to zadanie okazało się trudne w związku z kilkoma efektami. Po pierwsze układ wymaga dwóch oddzielnych nastaw dla sterowania w pionie jak i poziomie ponieważ zakres ruchu robotycznej głowy w poziomie wynosi 180° natomiast w pionie 90° . Ponadto pole widzenia kamery również jest różne gdyż dostarczany obraz jest w proporcjach 16:9. Innym zaobserwowanym efektem była różna wartość wzmocnienia granicznego w zależności o odległości użytkownika od kamery. Czym dalej znajdowała się osoba tym układ był bardziej wrażliwy. Efekt spowodowany był różnicą między dostępnym ruchem a polem widzenia kamery oraz możliwym zakresem ruchu. Ostatecznie, wszystkie testy wykonywano w odległości 1m od kamery.

Kolejnym krokiem jest wyznaczenie stałych czasowych. W trakcie tego eksperymentu również zaobserwowano ciekawe wnioski wynikające z dyskretnego charakteru układu. Na prezentowanych przebiegach przedstawianych jest 100 ostatnich pomiarów. W zależności od mocy obliczeniowej komputera, pomiary te mogą zajmować zróżnicowany czas. Wykonano trzykrotny pomiar czasu przetwarzania pętli odpowiedzialnej za obsługę serwokontrolera, systemu wizyjnego, regulatora PID i modułu generującego przebiegi czasowe dla generacji 100 pomiarów i uzyskano średnią 9,06 s. Oznacza to, że program operuje średnio z częstotliwością 11 klatek na sekundę, a każdy pomiar na wykresie to około 0,09 s. Istnieje jednak obawa, że dla komputera o innej mocy obliczeniowej obliczenia należy powtórzyć.

Stałe czasowe wyznaczono dokonując pomiaru szerokości cykli na wyświetlanym obrazie. Zapisany obraz z przebiegami posiadał określoną rozdzielczość. Wyliczono, że 100 pomiarów zajmuje na obrazie 947 pikseli, które trwają 9,06 sekundy. Zatem skoro 10 okresów oscylacji ma szerokość 786 pikseli, to trwały one 7,52 sekundy. Zatem okres oscylacji dla osi pionowej wynosi 0,752 s. Podobnie obliczenia dla osi poziomej dają wynik 0,670 s. Zatem $T_{osc} = [0,670 \ 0,752]^T$.



Rys. 4. Układ granicznie stabilny w osi oy dla $k_{gr} = 331/1000$



Rys. 5. Układ granicznie stabilny w osi ox dla $k_{gr} = 302/1000$

Implementacja regulatora PID w bibliotece *simple_pid* odbywa się przez podanie trzech wzmoceń $\{k_P, k_I, k_D\}$, natomiast regulator przybiera postać

$$K_r(s) = k_P + \frac{k_I}{s} + k_D s. \quad (2)$$

Zatem w celu wyznaczenia zestawu 3 wzmoceń należy zastosować przekształcenia zgodnie z metodą Zieglera-Nicholsa

$$k_r = k_P = 0.6 k_{gr}, \quad (3)$$

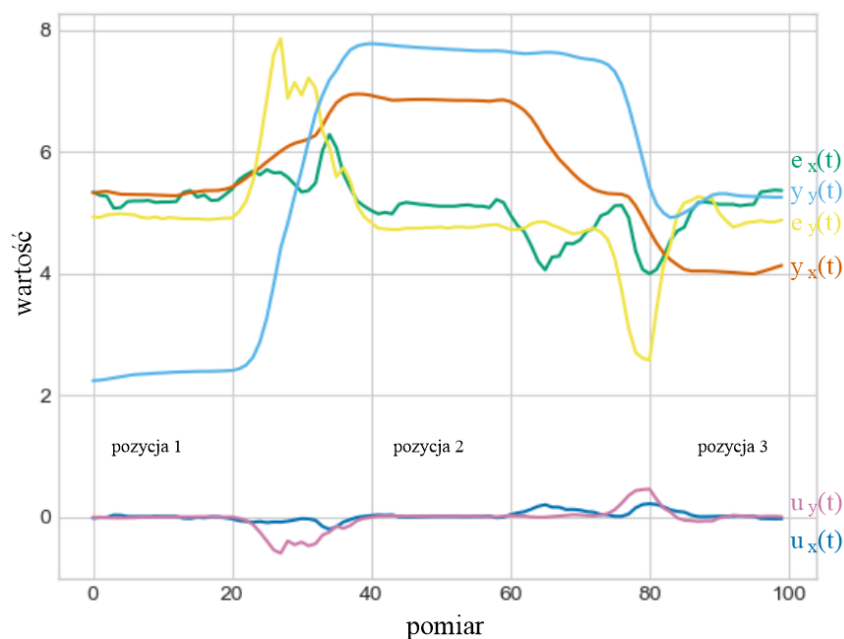
$$T_i = 0.5 T_{osc}, k_I = \frac{2 k_r}{T_{osc}}, \quad (4)$$

$$T_d = 0.125 T_{osc}, k_D = 0.125 k_r T_{osc}. \quad (5)$$

Po zastosowaniu przekształceń (3-5), dla przedstawionego przykładu, uzyskano nastawy regulatora $k_P = [0,180 \ 0,199]^T$, $k_I = [0,0090 \ 0,0088]^T$ oraz $k_D = [0,00056 \ 0,00055]^T$. Demonstracja układu regulowanego znajduje się na rysunku 6. W trakcie demonstracji użytkownik zmienił swoją pozycję 3 krotnie, zmieniając wartość wyjścia z układu $y(t)$ dla obu osi ox i oy .

5. Wnioski

Przedstawiona metoda z sukcesem umożliwia śledzenie użytkownika i imitację utrzymywania kontaktu wzrokowego przez robota społecznego. Robot podąża wzrokiem za twarzą użytkownika. Dobrane nastawy regulatora gwarantują szybką reakcję. Analizując przebiegi można zaobserwować szybką redukcję błędu przy zmianie pozycji trwającej około 1 sekundę (około 7-15 pomiarów), co wskazuje na dobre własności dynamiczne układu. W układzie nie występuje przesterowanie. Scenariusze symulacyjne zostały dobrane w taki sposób aby nie osiągały wartości granicznych $[0,10]$.



Rys. 6. Odpowiedzi układu regulowanego

Problem jednak pojawia się w przypadku gdy na obrazie znajduje się wiele twarzy lub nie wykryto żadnej. W przypadku wielu twarzy robot przełącza się losowo pomiędzy różnymi użytkownikami sprawiając wrażenie niestabilnego. Zastosowano algorytmy poprawiające występujące negatywne efekty takie jak anti-windup dla kumulacji błęd przy niewykryciu twarzy oraz ważenie pewności wykrytej twarzy i podążanie za twarzą o największej pewności wykrycia w przypadku wykrycia wielu.

Ostatni zauważony efekt wynika z geometrii kamery i został opisany w rozdziale 4 przy okazji wyznaczania k_{gr} . Efekt dotyczy wpływu odległości użytkownika od kamery na wzmocnienie graniczne.

6. Dalsze prace

Należy zaznaczyć, że zaproponowana metoda nie jest wolna od wad. Dalsze prace będą skupiać się na unieważnieniu systemu na zmienne niezależnie od bodźców zakłócających. Jedną ze zmiennych, która posiada niebagatelne znaczenie jest moc obliczeniowa jednostki. Należy zbadać wpływ osiągniętej liczby klatek na sekundę na stałą czasową T_{osc} . Jest to znaczące, ponieważ dalsze prace prowadzone będą w celu maksymalnego uproszczenia systemu, tak aby był on równie wydajny w jednostkach o ograniczonych zasobach obliczeniowych. Kolejnym równie istotnym aspektem jest analiza odległości użytkownika od ekranu. Roboty społeczne wchodzące w interakcje z człowiekiem analizują i przetwarzają dane osób będących w ich bliskiej odległości (w większości do 10m), niemniej analiza i rozróżnienie odległości stojącej osoby lub osób pozwoliłaby na odróżnianie aktywnego rozmówcy od innych uczestników spotkania, a także na wykrywanie osób zbliżających się lub odchodzących od robota.

W celu identyfikacji tych parametrów należy zmierzyć 2 zmienne w systemie i uzależnić parametry PID od ich wartości. Zmiennymi są czas przetwarzania pętli, czyli jednego punktu pomiarowego z wykresu oraz wielkość prostokąta okalającego wynikającego z wykrywania twarzy, świadczący o odległości użytkownika od kamery.

PODZIĘKOWANIA

Praca TG została częściowo wsparta przez Politechnikę Śląską poprzez dotację na utrzymanie i rozwój potencjału badawczego w 2022 roku oraz poprzez dotację na utrzymanie i rozwój potencjału badawczego w 2022 roku dla młodych naukowców numer 02/060/BKM22/0041. Praca NB i MW była wspierana przez Program Mentorski realizowany przez Politechnikę Śląską (Program Mentorski – „Rozwiń Skrzydła”) i opłacany z rezerwy Prorektora ds. Studenckich i Kształcenia: MPK: 60/ 001 GZF: SUBD. Praca PE była wspierana częściowo przez Unię Europejską poprzez Europejski Fundusz Społeczny jako stypendium w ramach Grantu POWR.03.02.00-00-I029, a częściowo przez Politechnikę Śląską poprzez dotację na utrzymanie i rozwój potencjału badawczego w 2022 roku dla młodych naukowców numer 02/060/BKM22/0036.

LITERATURA

1. Gessing R.: Control fundamentals. Silesian University of Technology, 2004.
2. Hegel F., Muhl C., Wrede B., Hielscher-Fastabend M., Sagerer G.: Understanding social robots. In 2009 Second International Conferences on Advances in Computer-Human Interactions, IEEE, 2009. p. 169-174.
3. Leite I., Martinho C., Paiva A.: Social robots for long-term interaction: a survey. International Journal of Social Robotics, 5(2), 2013, p. 291-308.
4. Nocentini O., Fiorini L., Acerbi G., Sorrentino A., Mancioffi G., Cavallo F.: A survey of behavioral models for social robots. Robotics, 8(3), 54. 2019.
5. Pan Y., Ge S. S., He H., Chen, L.: Real-time face detection for human robot interaction. In RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2009, p. 1016-1021.
6. Shanavas I. H., Reddy P. B., Doddegowda M. C.: A Personal Assistant Robot Using Raspberry Pi. In 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), IEEE, 2018, p. 133-136.
7. Wei L. H.: Design and Development of an Omnidirectional Mobile Base for a Social Robot, 2011.
8. Xu T., Zhang H., Yu C.: See you see me: The role of eye contact in multimodal human-robot interaction. ACM Transactions on Interactive Intelligent Systems (TiiS), 6(1), 2016, p. 1-22.
9. Yao N., Anaya E., Tao Q., Cho S., Zheng H., Zhang F.: Monocular vision-based human following on miniature robotic blimp. In 2017 IEEE International Conference on Robotics and Automation (ICRA) IEEE, 2017, p. 3244-3249.