



**Politechnika
Śląska**

PROJEKT INŻYNIERSKI

System wizyjny do śledzenia ruchu gałek ocznych

Bartosz WUWER

Nr albumu: 296949

Kierunek: Automatyka i Robotyka

Specjalność: Technologie informacyjne w automatyce i robotyce

PROWADZĄCY PRACĘ

Dr inż. Krzysztof Jaskot

KATEDRA Automatyki i Robotyki

Wydział Automatyki, Elektroniki i Informatyki

Gliwice 2025

Tytuł pracy

System wizyjny do śledzenia ruchu gałek ocznych

Streszczenie

(Streszczenie pracy – odpowiednie pole w systemie APD powinno zawierać kopię tego streszczenia.)

Słowa kluczowe

(2-5 słów (fraz) kluczowych, oddzielonych przecinkami)

Thesis title

Vision system for tracking the movement of the eyeballs

Abstract

(Thesis abstract – to be copied into an appropriate field during an electronic submission – in English.)

Key words

(2-5 keywords, separated by commas)

Spis treści

1	Wstęp	1
1.1	Wprowadzenie i osadzenie w dziedzinie	1
1.2	Cel oraz zakres pracy	2
2	Śledzenie ruchu gałek ocznych – analiza tematu	3
2.1	Sformułowanie problemu	3
2.1.1	Uwaga wzrokowa	3
2.1.2	Poszukiwanie wzrokowe	4
2.1.3	Analiza ruchów oczu	4
2.2	Osadzenie tematu w kontekście aktualnego stanu wiedzy	7
2.2.1	Metody śledzenia ruchu oczu: przegląd historyczny i alternatywy . .	7
2.2.2	Analiza wideo z wykorzystaniem źrenicy i odbicia rogówkowego . .	8
2.3	Studia literaturowe	9
2.3.1	Metoda źrenicy i odbicia rogówkowego w badaniach	10
2.3.2	Porównanie z video-okulografią z użyciem kamerki internetowej . .	11
2.3.3	Podsumowanie metod w literaturze	12
3	Wymagania i narzędzia	13
3.1	Wymagania funkcjonalne	13
3.1.1	Limitacje sprzętowe i środowiskowe	13
3.1.2	Wykorzystanie platform i narzędzi programistycznych	14
3.1.3	Zakres działania systemu	14
3.2	Wymagania нефункционалне	15
3.3	Przypadki użycia	15
4	[Właściwy dla kierunku – np. Specyfikacja zewnętrzna]	17
5	[Właściwy dla kierunku – np. Specyfikacja wewnętrzna]	19
6	Weryfikacja i walidacja	21
7	Podsumowanie i wnioski	23

Bibliografia	26
Spis skrótów i symboli	29
Źródła	31
Lista dodatkowych plików, uzupełniających tekst pracy	33
Spis rysunków	35
Spis tabel	37

Rozdział 1

Wstęp

1.1 Wprowadzenie i osadzenie w dziedzinie

Śledzenie ruchu gałek ocznych (ang. eye-tracking), zwane również jako okulografia, jest techniką, która badana jest od ponad stu lat. Istotność tej techniki wynika z faktu, że ruchy gałek ocznych są ściśle związane z procesami poznawczymi, takimi jak uwaga, percepcja, pamięć, czy procesy decyzyjne. Skupiając wzrok na danym punkcie, umieszczamy go w centralnym obszarze naszego pola widzenia które charakteryzuje się największą rozdzielczością, co pozwala na dokładne analizowanie szczegółów. Ten fakt wpływa także na proces skupienia – gdy koncentrujemy się na danym obiekcie lub obszarze, skupiamy na nim wzrok (często wystarczy jedynie krótki moment).

Możliwość rejestrowania ruchów oczu pozwala na zrozumienie w jaki sposób obserwator eksploruje otaczający go świat. Posiadając tę wiedzę możliwe jest wyciągnięcie wniosków na temat tego co jest interesujące lub istotne dla obserwatora, jakie emocje się z tym wiążą, czy nawet jakie procesy poznawcze zachodzą w jego umyśle, czy rozumie on to co widzi. Nie trudnym jest zauważyć jak cenne mogą być te informacje w szerokim spektrum dziedzin.

Okulografia odgrywa kluczową rolę w psychologii poznawczej, psychologii społecznej, neurobiologii, marketingu, czy medycynie. W psychologii poznawczej ruch oczu jest ściśle związany z pamięcią, podejmowaniem decyzji, obciążeniem poznawczym i uczeniem się asocjacyjnym. W psychologii społecznej eye-tracking pozwala na wgląd w zachowania społeczne i ich analizę, co pozwala badać empatię, prospołeczność, czy fobie społeczne [16]. W neurobiologii bada się powiązania ruchu oczu ze szlakami neuronowymi odpowiedzialnymi za podejmowane akcje i procesy myślowe, dając możliwość w diagnozach i wsparciu osób dotkniętych chorobą Parkinsona [6], Alzheimer [5], a także autyzmem czy łagodnym upośledzeniem funkcji poznawczych [15]. W medycynie okulografia pozwala na diagnozę między innymi oczopląsu, który może być objawem np. stwardnienia rozsianego, uszkodzenia śródmózgowia lub urazu ucha wewnętrznego [1]. Niezaprzeczalnie technika

ta jest niezwykle cenna i efektywna na wielu płaszczyznach naukowych i praktycznych.

1.2 Cel oraz zakres pracy

Celem niniejszej pracy jest wykonanie narzędzia które pozwoli na uniwersalne śledzenie ruchu gałek ocznych, bez zdefiniowanej docelowej grupy użytkowników. System ten powinien być prosty w obsłudze i niewymagający zaawansowanej technologii, ale nie ograniczający w razie potrzeby bardziej zaawansowanych użytkowników, prezentując użyteczność zarówno dla naukowców, studentów, jak i hobbystów. Tego typu narzędzie pozwoli na eksplorację danych eye-trackingowych, otwierając możliwości na dalszą analizę i interpretację wyników, a także zapewni solidną podstawę w razie potrzeby modyfikacji lub rozbudowy systemu w bardziej ukierunkowany sposób.

Praca ta skupia się przede wszystkim na samym procesie wykrywania źrenic i śledzenia ruchu gałek ocznych na ich podstawie oraz implementacji systemu wizyjnego, który pozwoli na zapis, wizualizację i potencjalną dalszą analizę danych zebranych przez kamerę, w tym także kamerę internetową. Całość zrealizowana jest w języku Python, z wykorzystaniem bibliotek takich jak OpenCV, NumPy, Pandas oraz Matplotlib, a wszelkie prezentowane dane zostały uchwycone przy użyciu wbudowanej kamery laptopa. Praca nie będzie obejmować rozległej analizy zebranych danych, ani eksperymentów przeprowadzonych z użyciem owego systemu. Nie została także stworzona grupa testowa, więc większość testów przeprowadzone zostały na autorze pracy.

- wprowadzenie w problem/zagadnienie
- osadzenie problemu w dziedzinie
- cel pracy
- zakres pracy
- zwięzła charakterystyka rozdziałów
- jednoznaczne określenie wkładu autora, w przypadku prac wieloosobowych – tabela z autorstwem poszczególnych elementów pracy

Rozdział 2

Śledzenie ruchu gałek ocznych – analiza tematu

2.1 Sformułowanie problemu

Widzeniem plamkowym (fovealnym) nazywamy wspomniane już wcześniej zjawisko, że widzimy w największej rozdzielczości jedynie centrum pola widzenia. Wynika to z budowy siatkówki, która posiada plamkę żółtą (fovea) – jest to niewielki obszar, który charakteryzuje się największym zagęszczeniem czopków. Taka budowa siatkówki sprawia, że nieustannie poruszamy oczami. Skupiając wzrok na danym punkcie, musimy być gotowi na to, że cała reszta pola widzenia straci dla nas na ostrości. Jest to proces, który przypomina filtrację wielu otaczających nas informacji, do paru, które mają w tym momencie znaczenie.

2.1.1 Uwaga wzrokowa

Uwaga wzrokowa może zostać opisana przez idiomy „gdzie” i „co”, które pozwalają na zobrazowanie jej selektywnej natury. „Gdzie” to proces wizualnego wyszukiwania i wyboru lokalizacji, która zwróciła naszą uwagę, w celu dokładniejszego zbadania. Istotnym aspektem tego procesu selekcji jest widzenie peryferyjne, czyli takie obejmujące obszar oddalony od naszego punktu skupienia, ale wciąż pozwalający na wyodrębnienie kształtów i ruchu, co w pewien sposób prowadzi nasze centralne spojrzenie. Przykładem może być spojrzenie przez okno – w pierwszym momencie nasz wzrok kierowany jest na wyraźne kształty, światła czy nagłe ruchy jak np. lądujący ptak.

„Co” można nazwać odwrotnością „gdzie” – jest to proces szczegółowego badania danego obszaru, charakteryzują go takie zjawiska jak fiksacja, czyli stabilizacja wzroku na danym punkcie oraz ruchy sakadowe, czyli szybkie mimowolne ruchy oczu pomiędzy kolejnymi punktami, które pozwalają na obserwacje. Całość tworzy kanał percepcyjny o ograniczonym zasięgu przestrzennym, obejmujący interesujący obszar, budując w ten

sposób naszą świadomość i zrozumienie celu naszego spojrzenia. Wracając do poprzedniego przykładu, po zauważeniu ptaka, nasze spojrzenie skupi się na nim (fiksacja), a ruchy sakadowe oczu pozwolą na dokładne zbadanie jego kształtu, koloru i detali. To właśnie połączenie „gdzie” i „co” pozwala na głębsze zrozumienie otaczającego nas świata, a owe ścieżki skanowania dają nam wgląd w proces poznawczy badanej osoby, jako że pomiar widzenia fovealnego w czasie odzwierciedla chwilowe i jawne skupienie uwagi wzrokowej obserwatora.

2.1.2 Poszukiwanie wzrokowe

Tutaj warto wspomnieć o poszukiwaniu wzrokowym, jest to proces aktywnego przeszukiwania pola widzenia w celu znalezienia konkretnego celu, a częścią tego procesu jest etap przeduwagowy. Etap ten równolegle analizuje duży obszar widzenia jednocześnie, dzieje się to automatycznie, nieświadomie i nie wymaga widzenia fovealnego. Co charakteryzuje ten proces to umiejętność rozpoznania czterech podstawowych cech: kolor, rozmiar, orientacja oraz obecność i/lub kierunek ruchu. Etap przeduwagowy jest pierwszym krokiem poszukiwania wzrokowego i obejmuje dużą część poszukiwania typu równoległego (np. próba zauważenia nagłego ruchu spadającej gwiazdy) i relatywnie niewielką część poszukiwania typu seryjnego (np. przeszukiwanie obiektów na stole w poszukiwaniu kluczy), które wymaga uwagi i przenoszenia wzroku od obiektu do obiektu.

2.1.3 Analiza ruchów oczu

Powyższe rozważania pozwalają przejść do głównego nurtu tematu śledzenia ruchu gałek ocznych, czyli analiza ruchów oczu. Rozróżniamy pięć podstawowych typów ruchów oczu: wspomniane już wcześniej sakadyczne, wergencyjne, przedsionkowe, płynne podążanie (smooth pursuit) oraz oczopląs fizjologiczny, który jest naturalny dla zdrowej osoby i często występuje podczas fiksacji. Ruchy te można podzielić na dobrowolne, mimowolne i odruchowe, a sygnały je kontrolujące pochodzą z obszarów korowych mózgu. Do opisu tych procesów można posłużyć się modelowaniem matematycznym. Do zrozumienia jawnej uwagi wzrokowej wystarczy modelowanie trzech typów ruchów: fiksacji, która pokazuje chęć utrzymania wzroku na stacjonarnym obiekcie, sakad, które mogą wskazywać na chęć zmiany punktu uwagi oraz płynnych pościgów, które podobnie do fiksacji, pozwalają na śledzenie obiektu, ale ruchomego.

Ruchy sakadowe

Sygnał sakad można opisać jako funkcja impuls/skok, gdzie impuls na wejściu reprezentuje prędkość, a skok pozycję, impuls jest przepuszczany przez filtr, który przekształca go w skok. Prostą reprezentacją ruchu sakadowego jest filtr liniowy różniczkujący, który

dokonyje potrzebnej konwersji informacji prędkość na przemieszczenie. Wzór owego filtra w dziedzinie czasu można zapisać równaniem 2.1.

$$x_t = g_0 \cdot s_t + g_1 \cdot s_{t-1} + g_2 \cdot s_{t-2} + \dots \quad (2.1)$$

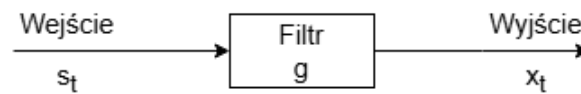
Wyjście x_t zależy od bieżącego wejścia s_t oraz jego poprzednich wartości (s_{t-1}, s_{t-2}, \dots) ważonych odpowiednimi współczynnikami filtra (g_0, g_1, g_2, \dots). Wzór 2.1 można przedstawić sumarycznie równaniem 2.2.

$$x_t = \sum_{k=0}^{\infty} g_k \cdot s_{t-k} \quad (2.2)$$

Filtr Haara jest jednym z przykładów filtrów który przybliża różniczkowanie, jest to filtr o długości 2, czyli operujący jedynie na dwóch kolejnych próbkach sygnału wejściowego (s_t, s_{t-1}), a przybliża on pierwszą pochodną, przyjmując, że prędkość zmian sygnału jest stała w czasie trwania dwóch próbek. Współczynniki filtra są równe $g_0 = 1$ oraz $g_1 = -1$, co oznacza że różniczkowanie jest zrealizowane poprzez różnicę między dwiema kolejnymi próbkami sygnału. W związku z tym transmitancję filtra Haara można zapisać jako równanie 2.3.

$$\begin{aligned} x_t &= g_0 \cdot s_t + g_1 \cdot s_{t-1} \\ x_t &= 1 \cdot s_t - 1 \cdot s_{t-1} \\ x_t &= s_t - s_{t-1} \\ \mathcal{Z}\{x_t\} &= \mathcal{Z}\{s_t - s_{t-1}\} \\ X(z) &= (1 - z) \cdot S(z) \\ \frac{X(z)}{S(z)} &= 1 - z \end{aligned} \quad (2.3)$$

Diagram omawianego modelu przedstawiony jest na rysunku 2.1.



Rysunek 2.1: Diagram modelu ruchów sakadowych z filtrem liniowym różniczkującym.

Płynne podążanie

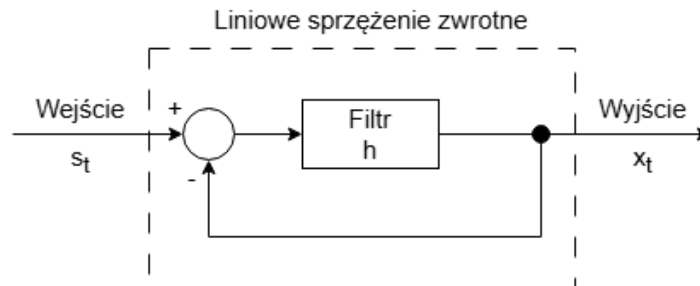
Płynne podążanie występuje gdy obserwator śledzi obiekt w ruchu, ruch ten oczywiście nie może być zbyt gwałtowny. Tego typu śledzenie jest przykładem systemu sterowania z ujemnym sprzężeniem zwrotnym. Do modelowania owych ruchów używana jest prosta pętla, którą można opisać następującym równaniem 2.4 w dziedzinie czasu.

$$h \cdot (s_t - x_t) = x_{t+1} \quad (2.4)$$

Wejście s_t to pozycja celu, wyjście x_t to pożądana pozycja oka. Można zauważyć, że w tym przypadku nie jest wymagana transformacja informacji wejściowej na wyjściową, a jedynie zmodyfikowanie jej wartości, w związku z tym h jest liniowym niezmiennym w czasie filtrem, czyli po prostu wzmocnieniem systemu. Przeprowadźmy teraz transformację Z na powyższym równaniu 2.4 by uzyskać transmitancję systemu przedstawioną wzorem 2.5.

$$\begin{aligned} \mathcal{Z}\{h \cdot (s_t - x_t)\} &= \mathcal{Z}\{x_{t+1}\} \\ H(z) \cdot (S(z) - X(z)) &= X(z) \\ H(z) \cdot S(z) - H(z) \cdot X(z) &= X(z) \\ H(z) \cdot S(z) &= (H(z) + 1) \cdot X(z) \\ \frac{X(z)}{S(z)} &= \frac{H(z)}{H(z) + 1} \end{aligned} \quad (2.5)$$

W ten sposób sygnał z receptorów oka służy za błąd, który następnie jest kompensowany w celu utrzymania obrazu w przestrzeni widzenia plamkowego. Diagram modelu przedstawiony jest na rysunku 2.2.



Rysunek 2.2: Diagram modelu płynnego podążania z liniowym sprzężeniem zwrotnym.

Fiksacja

Można zauważyć, że fiksacja jest procesem podobnym do płynnego podążania, z tą różnicą, że w tym przypadku obiekt jest nieruchomy. Jednakże proces ten nie da się bezpośrednio z sobą porównać i najprawdopodobniej nie ma wspólnego obwodu neuronalnego. Nasze komórki systemu wzrokowego są fizjologicznie wrażliwe na ruch, gdyby dany obiekt został unieruchomiony względem siatkówki, po krótkim czasie widzenie zaniknie. Powoduje to konieczność mikrosakad i innych mimowolnych, drobnych ruchów oczu. Można więc uznać, że model fiksacji jest podobny do modelu płynnego podążania, który próbuje utrzymać pozycję oka na danym punkcie, a mikrosakady i inne ruchy można uznać za szum w systemie kontrolnym, który można wyrazić jako $e_t = x_t - s_t$. Jest to losowa fluktuacja wokół punktu fiksacji, a jej wartość średnia pozostaje stała.

Ruch gałek ocznych i związane z nim widzenie jest bardzo szerokim tematem interdyscyplinarnym, na potrzeby tego projektu powyższe sformułowanie problemu, chociaż uproszczone i nie wyczerpujące, pozwala na zrozumienie w pełni istoty prezentowanego systemu.

2.2 Osadzenie tematu w kontekście aktualnego stanu wiedzy

Pomiar ruchu oczu był realizowany już w latach 70. XX wieku, nic dziwnego więc, że od tego momentu powstały różne techniki pozwalające na zebranie owych danych. Należy podkreślić, że metody te można podzielić na mierzące ruch oczu względem głowy, oraz takie które mierzą orientację oczu w przestrzeni, czyli „punkt spojrzenia”.

2.2.1 Metody śledzenia ruchu oczu: przegląd historyczny i alternatywy

Najstarszą z metod śledzenia ruchu oczu jest elektro-okulografia (EOG ang. *electrooculography*), wciąż wykorzystywana np. w badaniach klinicznych. Metoda ta polega na umieszczeniu elektrod na skórze twarzy wokół oczu i pomiarze różnicy potencjałów. Metoda ta z założenia mierzy ruch oczu względem głowy, ale można rozszerzyć ją o pomiar ruchu głowy dając możliwość wyliczenia punktu spojrzenia.

Jedna z najdokładniejszych metod pomiaru ruchu oczu polega na umieszczeniu soczewki kontaktowej bezpośrednio na gałce ocznej. Soczewka ta powinna być odpowiednio duża, by objąć zarówno rogówkę jak i twardówkę, a na jej powierzchni umieszczany jest albo obiekt optyczny, którego zadaniem jest dokładne odbijanie światła lub dostarczenie wyraźnych kształtów potrzebnych do śledzenia, albo cewka wykonana z drutu, która poruszając się w polu elektromagnetycznym pozwala na wykonanie pomiaru. Metoda ta jest

bardzo dokładna, ale niezwykle inwazyjna i nieprzyjemna dla badanego, a także, podobnie do elektro-okulografii, mierzy ruch oczu względem głowy.

Wideo-okulografia i foto-okulografia tworzą razem szeroką grupę metod, które opierają się na analizie wyróżniających się cech oka podczas jego obrotu, takich jak: kształt źrenicy, pozycja granicy tęczówki i twardówki, czy odbicie światła (często podczerwonego) od rogówki. Metody te są nieinwazyjne, ale same w sobie nie pozwalają na określenie punktu spojrzenia, więc w celu wyznaczenia owego punktu często stosuje się unieruchomienie głowy osoby badanej, wyznaczenie punktu odniesienia np. przez odbicia światła od powierzchni oka, lub odpowiednią kalibrację np. prosząc osobę badaną o utrzymanie wzroku na danym punkcie ekranu.

2.2.2 Analiza wideo z wykorzystaniem źrenicy i odbicia rogówkowego

Opisane metody wymagają unieruchomienia głowy (przez różnego rodzaju podpórki pod brodę lub głowę, albo nawet belki nagryzowe) lub zastosowanie dodatkowego sprzętu mierzącego ruch głowy, by wyznaczyć punkt spojrzenia. Jest to główna wada tych metod, jako że miejsce skupienia uwagi daje istotne informacje, które są najczęściej pożądane przez użytkowników okulografów. Dlatego metoda należąca do grupy wideo-okulografii, opierająca się na analizie obrazu wideo z kombinacją wykrycia źrenicy i odbicia rogówkowego, która pozwala na określenie punktu spojrzenia z dużą dokładnością, jest jedną z najpopularniejszych wyborów zarówno w świecie naukowym jak i komercyjnym. System prezentowany w tej pracy można zaliczyć do wideo-okulografii, z możliwością włączenia trybu z pomiarem odbicia rogówkowego, więc metoda ta zostanie omówiona szerzej od pozostałych.

Wykrywanie punktu spojrzenia metodą detekcji źrenicy i odbicia rogówkowego wymaga kilku kluczowych elementów. Potrzebna jest kamera wideo, która zarejestruje aktualny stan gałek ocznych oraz sprzęt, który pozwoli na przetworzenie obrazu i ostatecznie wykrycie interesujących cech, najlepiej w czasie rzeczywistym. Aktualnie oba wymienione elementy stają się coraz tańsze i bardziej wydajne, tak naprawdę większość osób posiada je w swoim smartfonie, co sprawia, że technika ta staje się bardziej dostępna dla szerokiego grona użytkowników. Aparatura ta jest także coraz bardziej miniaturyzowana, dzięki czemu aktualnie dostępne są systemy montowane na głowie, jak i montowane na biurku, które działają na tej samej zasadzie, różniąc się zasadniczo jedynie wielkością. Warto jednak zaznaczyć, że systemy montowane na głowie zazwyczaj posiadają także dodatkową kamerę wyznaczającą tak zwany POV (ang. point of view, czyli punkt widzenia), a spojrzenie jest monitorowane względem tego obrazu, w odróżnieniu do montowanych na biurku, które zazwyczaj mierzą spojrzenie względem pewnej powierzchni np. monitora. Należy omówić także źródło światła, którego odbicie rejestrowane jest podczas detekcji.

Możliwe jest wykorzystanie punktowego źródła światła białego o odpowiednim natężeniu, jednakże może to prowadzić do dyskomfortu przy użytkowaniu lub nawet chwilowego oślepienia osoby badanej, wytrącając ją z stanu skupienia. W związku z tym najczęściej wykorzystywanym są źródła światła podczerwonego, które znajdują się w większości poza widzialnym spektrum, nie irytując oka, nie zakłócając procesu badania i pozwalając na użycie wyższego natężenia, operowanie na takim zakresie częstotliwości światła wymaga użycia specjalnych kamer, które są w stanie rejestrować światło IR (ang. infrared).

Omawiane odbicie rogówkowe nazywane jest obrazem Purkiniego, budowa oka sprawia, że pojawiają się cztery odbicia, pierwsze to odbicie od zewnętrznej powierzchni rogówki (najbardziej widoczne), drugie od wewnętrznej powierzchni rogówki, trzecie od zewnętrznej powierzchni soczewki, a czwarte od wewnętrznej powierzchni soczewki. By poprawnie wyznaczyć punkt spojrzenia, wymagane jest wyznaczenie dwóch punktów odniesienia, pierwszy określający obrót oczu w oczodole, a drugi określający stałe położenie względem oczu. Pierwszym punktem zazwyczaj jest środek źrenicy, a drugim najczęściej jest pierwszy obraz Purkiniego, powstały przez źródło światła ustabilizowane względem głowy, lub powierzchni badanej. Przy nieruchomej głowie, ruch oczu odczytywany będzie jako zmieniająca się różnica pomiędzy tymi dwoma punktami, w odwrotnym przypadku, gdy głowa poruszy się, a oczy pozostaną w fiksacji na danym punkcie, dla systemu montowanego na biurku różnica ta pozostanie stała, a dla systemu montowanego na głowie, różnica będzie zmieniać się proporcjonalnie do zmiany punktu widzenia. Istnieją także okulografy piątej generacji z technologią DPI (ang. Dual Purkinje Image), mierzą one dodatkowo czwarty obraz Purkiniego, dzięki temu, mogą rozróżniać ruchy translacyjne oka od ruchów rotacyjnych, jako że przy translacji obrazy Purkiniego poruszają się o tą samą odległość, a przy rotacji zmieniają swoje rozdzielanie, co zwiększa precyzję pomiaru, ale wymaga bardziej skomplikowanego oprogramowania, a także unieruchomienie głowy może okazać się konieczne.

2.3 Studia literaturowe

Powyższe sformułowanie problemu i jego osadzenie w kontekście aktualnego stanu wiedzy zostało przedstawione w oparciu o książkę „Eye Tracking Methodology: Theory and Practice” autorstwa Andrew T. Duchowskiego [2]. Książka ta dogłębnie omawia temat śledzenia ruchu oczu, pozwalając na zgłębienie zarówno podstaw procesów wzrokowych w kontekstach różnych dziedzin naukowych, a także szeroko omawia praktykę pomiaru ruchu oczu. Powyższa analiza tematu skupia się jedynie na najistotniejszych aspektach, które pozwolą na zrozumienie istoty prezentowanego systemu, a wiedzę na ich temat można poszerzyć lekturą wspomnianej książki.

Eye-tracking staje się coraz bardziej dostępny, co sprawia że technologia ta znajduje zastosowanie w wielu obszarach, zarówno w badaniach naukowych jak i do użytku konsu-

menckiego czy komercyjnego. Część producentów owych systemów udostępnia publikacje naukowe, w których zostały one wykorzystane, a ich analiza przedstawia różne rozwiązania omawianego tematu.

2.3.1 Metoda źrenicy i odbicia rogówkowego w badaniach

Jednym z przedstawicieli firm skupiających się na analizie zachowań ludzkich jest iMotions, wyróżniając się modularnym systemem z centralnym hubem. Firma ta oferuje biosensory w formie modułów, pozwalające na np. śledzenie oczu, analizę wyrazu twarzy, rejestrowanie aktywności elektrycznej mózgu i rejestrowanie aktywności serca. Dane zebrane przez różne czujniki są integrowane w jednym oprogramowaniu co pozwala badaczom na prostą synchronizację i analizę danych w jednym systemie [13]. Jednak mimo tak obszernego asortymentu biosensorów, to właśnie moduł śledzenia wzroku na ekranie (montowany na biurku) był wykorzystywany najczęściej w badaniach naukowych, z czego eye-tracking był najczęściej wykorzystywany w systemach jednomodułowych [11]. Pokazuje to jak uniwersalnym i podstawowym narzędziem jest eye-tracker w badaniach behawioralnych i nie tylko.

Jednym z przeprowadzonych badań z użyciem technologii iMotions było „Using Psychophysiological Data to Facilitate Reflective Conversations with Children about their Player Experiences”. Podobne badania były przeprowadzane wielokrotnie w stosunku do osób dorosłych, jednak w przypadku grupy badanej składającej się z dzieci, badacze mogą napotkać znaczny problem z odpowiednim zrozumieniem wywiadu, w którym uczestnicy dzielą się swoimi wrażeniami z przeprowadzonej rozgrywki. Omawiane badanie miało głównie na celu sprawdzenie czy zebrane dane z biosensorów są w stanie wspomóc refleksje dzieci na temat ich doświadczenia w grze i czy dzieci napotkają problemy z interpretacją prezentowanych danych, jako że odnoszenie się do stanu psychofizjologicznego może zminimalizować problem z komunikacją pomiędzy dzieckiem, które posiada ograniczone umiejętności werbalne, a badaczem.

Do śledzenia ruchu gałek ocznych na ekranie zastosowano okuloграф Smart Eye AI-X o częstotliwości próbkowania 60 Hz, posługujący się techniką detekcji źrenicy i odbicia rogówkowego, pozwalając na sporą swobodę ruchu głowy w przestrzeni 35 cm na 30 cm, jednocześnie posiadając dokładność (różnica pomiędzy rzeczywistą pozycją spojrzenia a pozycją spojrzenia zarejestrowaną przez okuloграф) $0,5^\circ$ i precyzję (średnia kwadratowa punktów mierzonych w jednej pozycji oka) $0,1^\circ$, zwracając dane wyjściowe binokularowe z wskaźnikiem jakości zawierające: punkt spojrzenia, średnicę źrenicy oraz znacznik czasowy [12]. Oprócz tego zastosowano moduł rejestrujący aktywność sercową, analizujący wyrazy twarzy oraz mierzący reakcję skórno-galwaniczną.

Badanie składało się z dwóch etapów, w pierwszym zebrano dane w laboratorium podczas gdy badane dzieci grały w dwie gry oraz podczas wywiadów bezpośrednio po zakończonej

rozgrywce, a następnie prezentowano badanym momenty z pierwszego etapu z uwzględnieniem danych z biosensorów i zadawano pytania o ich doświadczenia. Badania wskazały, że dzieci są w stanie zrozumieć i odpowiednio odnieść się do prezentowanych danych psychofizjologicznych, a oparcie się o nie może pomóc w zwerbalizowaniu swojego doświadczenia i informacji zwrotnej przez badane dzieci. Dzieci nie miały problemu z zrozumieniem większości pomiarów za wyjątkiem reakcji skórno-galwanicznej, jednakże dane eye-trackingowe były swego rodzaju wyjątkiem, jako że nie prezentowały bezpośrednio doświadczenia i uczuć związanych z danym wydarzeniem w grze, a raczej dawały wgląd w strategię objętą przez dziecko zmagające się z aktualnym wyzwaniem, mimo tego prezentowane śledzenie punktu spojrzenia było naturalne w interpretacji przez badanych [4]. Publikacja ta pokazuje jak okulografia może zostać skutecznie użyta w badaniach naukowych, ale przede wszystkim przedstawia istotność tej techniki w zastosowaniach komercyjnych np. testując doświadczenia graczy przy produkcji gier wideo dla różnych grup wiekowych.

2.3.2 Porównanie z video-okulografią z użyciem kamierki internetowej

Warto omówić także przykłady systemów opierających się jedynie na analizie wideo z kamierki internetowej, jako że system prezentowany w tej pracy działa przede wszystkim na tej zasadzie. Ten typ okulografii opartej na analizie wideo w czasie rzeczywistym cieszy się dużą popularnością wśród konsumentów, jako że sprzętowo wymaga od użytkownika posiadanie jedynie komputera z wbudowaną lub zewnętrzną kamerą, a całość przetwarzania obrazu, wyświetlania i zbierania danych oraz wyliczania punktu spojrzenia realizowane jest przez software producenta eye-trackera. Istotność takiego rozwiązania pojawia się także gdy przeprowadza się badania na dużej grupie, jako że osoba badana sama dostarcza sprzęt w formie laptopa czy nawet smartfona, a zbierane dane przesyłane są do badaczy przez internet, eliminując potrzeby stwarzania przestrzeni laboratoryjnej i planowania dogodnych terminów eksperymentów. Dlatego firmy specjalizujące się w okulografii, często oferują także oprogramowanie do śledzenia ruchu gałek ocznych przez kamerkę internetową, jak wspomniana już wcześniej firma iMotions. Często także dostępne są rozwiązania od niezależnych dostawców oprogramowania, które za subskrypcją, opłatę, lub całkowicie bezpłatnie oferują śledzenie punktu spojrzenia przez kamerę komputera, jednym z takich programów, który zostanie omówiony jest darmowy GazeRecorder.

W publikacji „WebET 3.0 -Validation Study Report” przeprowadzono badania sprawdzające efektywność śledzenia ruchu gałek ocznych przez oprogramowanie WebET 3.0 firmy iMotions z wykorzystaniem kamerek internetowych uczestników, bez względu na ich oświetlenie, rozdzielczość kamery czy prędkość internetu. Grupa badana została utworzona tak by reprezentować różne pochodzenie, wiek, płeć i kolor oczu, uczestnicy mogli także posiadać owłosienie twarzy oraz nosić okulary. By stworzyć punkty odniesienia po-

kazano badanym gify z kotami oraz emotikony w tych samych miejscach, a następnie obliczono dokładność mierząc odległości pomiędzy zarejestrowanymi punktami spojrzenia, a ustalonymi punktami odniesienia.

Wyniki dokładności wskazały, że system iMotions jest niezależny na zmienne wynikające z pochodzenia, koloru oczu, wieku, płci oraz owłosienia twarzy, jednak posiadanie okularów mogło wpływać na wynik śledzenia, rodzaj i natężenie oświetlenia także nie wpływał znacząco na działanie WebET 3.0. Ponad 90% uczestników posiadało dokładność mniejszą od 5° dokładności, 70% badanych posiadało dokładność mniejszą od 3° dokładności, a mediana dokładności wynosiła 2,08° [14]. Dokładność alternatywnego darmowego oprogramowania GazeRecorder została zaprezentowana w przeglądzie systematycznym i wynosiła 1,43° z możliwością poprawy do 1,3° przez unieruchomienie głowy, pozwalając na noszenie okularów za wyjątkiem okularów antyrefleksyjnych [3].

2.3.3 Podsumowanie metod w literaturze

Metoda śledzenia ruchu gałek ocznych przy użyciu źrenicy i odbicia rogówkowego w podczerwieni znacznie przewyższa – pod względem dokładności – metodę polegającą na kamerce internetowej, w związku z tym przy badaniach skupiających się na analizie gęsto ułożonych informacji wyspecjalizowany okulograf może być bardziej pożądanym. Jednakże aktualna dostępność sprzętu i oprogramowania potrzebnego do śledzenia z użyciem kamery internetowej daje możliwość na bardziej obszerne zastosowania, a ich dokładność pozwala na efektywne śledzenie punktu spojrzenia w przypadkach gdy informacje są ułożone z odpowiednimi dystansami.

- sformułowanie problemu
- osadzenie tematu w kontekście aktualnego stanu wiedzy (*state of the art*) o poruszonym problemie
- studia literaturowe [9, 10, 8, 7] - opis znanych rozwiązań (także opisanych naukowo, jeżeli problem jest poruszany w publikacjach naukowych), algorytmów,

Rozdział 3

Wymagania i narzędzia

3.1 Wymagania funkcjonalne

Wymagania funkcjonalne definiują konkretne działania i funkcje, które system powinien zapewniać. W przypadku tego projektu są one związane z obsługą danych z eye-trackera, ich analizą oraz wizualizacją wyników. Niniejsza podsekcja przedstawia szczegółowy opis funkcjonalności, jakie musi oferować system, aby spełniać założenia pracy. Postawienie konkretnych granic funkcjonalnych projektu jest kluczowym elementem tworzenia rozwiązania. Działają one na zasadzie zabezpieczenia przed usprawnianiem i roszszerzaniem projektu bez końca, a także pozwalają na skupienie uwagi inżyniera na istotnych aspektach. Można zauważyć korelację pomiędzy rolą wymagań funkcjonalnych w tworzeniu projektu, a plamką żółtą w oku omawianej w sekcji 2.1, oba elementy dbają o to by wiele informacji zostało przefiltrowane do tych, które są najważniejsze.

3.1.1 Limitacje sprzętowe i środowiskowe

Do śledzenia ruchu gałek ocznych używany jest bardzo szeroki wachlarz sprzętów i związanych z nimi rozwiązań technicznych omówionych w sekcjach 2.2 i 2.3, jednakże w założeniu tego projektu było opierać się na systemie wizyjnym, co pozwala na ograniczenie znanych metod do video-okulografii. Kolejną ustaloną limitacją było opieranie się jedynie na aktualnie dostępnym sprzęcie, którym był laptop ASUS Vivobook 15 z wbudowaną kamerką internetową USB2.0 HD UVC o częstotliwości próbkowania 30 Hz, podczas korzystania z kamery automatycznie włączane było także niewielkie źródło światła przylegające z jej prawej strony. Tego typu laptop z wbudowaną kamerą reprezentuje jakość sprzętu często używaną do zastosowań biurowych i użytku własnego, dlatego też założono, że źródła światła powinny zostać ograniczone do podstawowego oświetlenia wewnętrznego, tworząc środowisko typowe do pokoju mieszkalnego, w którym korzysta się z komputera.

3.1.2 Wykorzystanie platform i narzędzi programistycznych

By uprościć użycie komputera wspomnianego w sekcji 3.1.1 w pracy pozostano na już zainstalowanym systemie operacyjnym, którym jest Windows 11. Podobnie postawiono wymagania względem edytora kodu źródłowego i systemu kontroli wersji, jako że użycie znajomego środowiska, które już jest zainstalowane i było wcześniej używane na wykorzystanym laptopie, pozwala na szybkie zamknięcie przygotowań do pracy. W związku z tym skupiono się na rozwiązaniach, które są kompatybilne z edytorem tekstu Visual Studio Code i klientem Git z graficznym interfejsem użytkownika GitHub Desktop, jako że narzędzia te ściśle współpracują z sobą i serwisem hostingowym GitHub tworząc intuicyjną i znajomą w użyciu całość. Stawianym wymaganiem do języka programowania, oprócz wspomnianej kompatybilności, jest jego popularność, uniwersalność, klarowność i szeroki wybór bibliotek. Język o takich cechach znacznie zwiększy produktywność, uprości stosowanie bardziej skomplikowanych narzędzi i pozwoli na efektywne uczenie się i rozwiązywanie problemów opierając się na społeczności użytkowników. Zalety te wspomogą uzyskanie natychmiastowych rezultatów pracy oraz skoncentrowanie uwagi na poprawie i rozwój funkcjonalności systemu. Łatwo dostępnym i popularnym językiem o wysokim poziomie abstrakcji jest Python, który jest oficjalnie wspierany przez Visual Studio Code. Język ten posiada możliwość korzystania z biblioteki OpenCV, która daje dostęp do wielu rozwiązań przetwarzających obraz. Ostatnim, bardziej personalnym atutem proponowanego języka, jest niewielkie ówczesne doświadczenie programowania w nim, a co za tym idzie szansa na zyskanie cennych umiejętności w trakcie pracy.

3.1.3 Zakres działania systemu

Nakreślone wymagania w sekcji 3.1.1 i 3.1.2 odpowiednio ograniczyły możliwe rozwiązania, pozwalając na ustalenie realnych funkcjonalności programu. Aspektem, który przede wszystkim determinuje stopień skomplikowania stawianych wymagań, jest ograniczony czas na wykonanie omawianego systemu. Skąplikowanie

Przetwarzanie obrazu z kamery odbywa się w czasie rzeczywistym Wykrywanie rejonu twarzy, w tym rejonie wykrywanie rejonu oczu, a w rejonie oczu wykrywa źrenicę Pozwala na wykrycie i zidentyfikowanie obu oczu z rozróżnieniem lewego i prawego Wykrywanie źrenic jest uniwersalne względem oświetlenia otoczenia pozwala na wizualizację procesu śledzenia, bez zaciemniania obrazu z kamery wyznacza umowny punkt odniesienia z możliwością prostej kalibracji Nie ogranicza bardziej zaawansowanych użytkowników, oferując wyznaczenie punktu odniesienia przy zastosowaniu dodatkowego sprzętu Zebranie danych o pozycji źrenicy z indeksem wskazującym odpowiednie oko i możliwość przedstawienia ich w formie wykresów czasowych

3.2 Wymagania niefunkcjonalne

3.3 Przypadki użycia

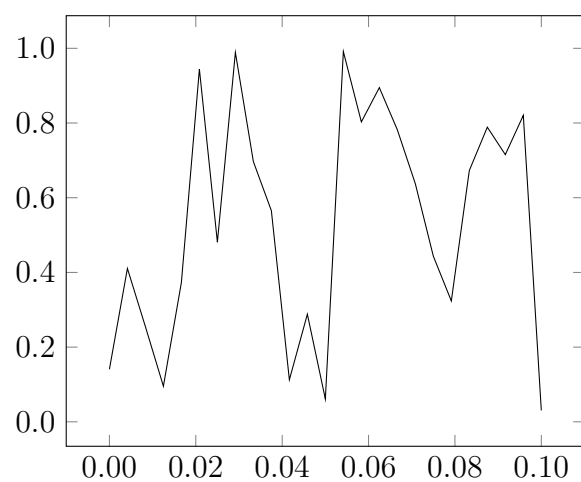
- wymagania funkcjonalne i niefunkcjonalne
- przypadki użycia (diagramy UML) – dla prac, w których mają zastosowanie
- opis narzędzi, metod eksperymentalnych, metod modelowania itp.
- metodyka pracy nad projektowaniem i implementacją – dla prac, w których ma to zastosowanie

Rozdział 4

[Właściwy dla kierunku – np. Specyfikacja zewnętrzna]

Jeśli „Specyfikacja zewnętrzna”:

- wymagania sprzętowe i programowe
- sposób instalacji
- sposób aktywacji
- kategorie użytkowników
- sposób obsługi
- administracja systemem
- kwestie bezpieczeństwa
- przykład działania
- scenariusze korzystania z systemu (ilustrowane zrzutami z ekranu lub generowanymi dokumentami)



Rysunek 4.1: Podpis rysunku po rysunkiem.

Rozdział 5

[Właściwy dla kierunku – np. Specyfikacja wewnętrzna]

Jeśli „Specyfikacja wewnętrzna”:

- przedstawienie idei
- architektura systemu
- opis struktur danych (i organizacji baz danych)
- komponenty, moduły, biblioteki, przegląd ważniejszych klas (jeśli występują)
- przegląd ważniejszych algorytmów (jeśli występują)
- szczegóły implementacji wybranych fragmentów, zastosowane wzorce projektowe
- diagramy UML

Krótką wstawka kodu w linii tekstu jest możliwa, np. **int a;** (biblioteka `listings`). Dłuższe fragmenty lepiej jest umieszczać jako rysunek, np. kod na rys 5.1, a naprawdę długie fragmenty – w załączniku.

```
1 class test : public basic
2 {
3     public:
4         test (int a);
5         friend std::ostream operator<<(std::ostream & s,
6                                         const test & t);
7     protected:
8         int _a;
9
10 };
```

Rysunek 5.1: Pseudokod w `listings`.

Rozdział 6

Weryfikacja i walidacja

- sposób testowania w ramach pracy (np. odniesienie do modelu V)
- organizacja eksperymentów
- przypadki testowe zakres testowania (pełny/niepełny)
- wykryte i usunięte błędy
- opcjonalnie wyniki badań eksperymentalnych

Tabela 6.1: Nagłówek tabeli jest nad tabelą.

ζ	metoda						
	alg. 1	alg. 2	alg. 3			alg. 4, $\gamma = 2$	
			$\alpha = 1.5$	$\alpha = 2$	$\alpha = 3$	$\beta = 0.1$	$\beta = -0.1$
0	8.3250	1.45305	7.5791	14.8517	20.0028	1.16396	1.1365
5	0.6111	2.27126	6.9952	13.8560	18.6064	1.18659	1.1630
10	11.6126	2.69218	6.2520	12.5202	16.8278	1.23180	1.2045
15	0.5665	2.95046	5.7753	11.4588	15.4837	1.25131	1.2614
20	15.8728	3.07225	5.3071	10.3935	13.8738	1.25307	1.2217
25	0.9791	3.19034	5.4575	9.9533	13.0721	1.27104	1.2640
30	2.0228	3.27474	5.7461	9.7164	12.2637	1.33404	1.3209
35	13.4210	3.36086	6.6735	10.0442	12.0270	1.35385	1.3059
40	13.2226	3.36420	7.7248	10.4495	12.0379	1.34919	1.2768
45	12.8445	3.47436	8.5539	10.8552	12.2773	1.42303	1.4362
50	12.9245	3.58228	9.2702	11.2183	12.3990	1.40922	1.3724

Rozdział 7

Podsumowanie i wnioski

- uzyskane wyniki w świetle postawionych celów i zdefiniowanych wyżej wymagań
- kierunki ewentualnych danych prac (rozbudowa funkcjonalna ...)
- problemy napotkane w trakcie pracy

Bibliografia

- [1] Redakcja Diagnostyki. *Oczopląs - co może oznaczać? Jakie badania warto wykonać?* 2023. URL: <https://diag.pl/pacjent/artykuly/oczoplas-co-moze-oznaczac-jakie-badania-warto-wykonac/#2-0-czym-swiadczy-oczoplas> (term. wiz. 10.01.2025).
- [2] Andrew T. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Londyn: Springer, 2007. ISBN: 978-1-84628-609-4.
- [3] Melanie Heck, Christian Becker i Viola Deutscher. „Webcam Eye Tracking for Desktop and Mobile Devices: A Systematic Review”. W: *Proceedings of the Annual Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences, 2023. DOI: 10.24251/HICSS.2023.825.
- [4] Janelle E. MacKenzie, Madison Klarkowski, Ella M. Horton, Maryanne Theobald, Susan Danby, Lisa Kervin, Lance Barrie, Philippa K. Amery, Manesha Andradi, Simon S. Smith, Regan L. Mandryk i Daniel Johnson. „Using Psychophysiological Data to Facilitate Reflective Conversations with Children about their Player Experiences”. W: *Proc. ACM Hum.-Comput. Interact.* 8.CHI PLAY (2024). DOI: 10.1145/3677112. URL: <https://doi.org/10.1145/3677112>.
- [5] Ieva Miseviciute. *4 methods to assess Alzheimer’s with eye tracking*. 2025. URL: <https://www.tobii.com/resource-center/learn-articles/4-methods-to-assess-alzheimers-with-eye-tracking> (term. wiz. 09.01.2025).
- [6] Ieva Miseviciute. *How does Parkinson’s disease alter visual search?* 2025. URL: <https://www.tobii.com/resource-center/scientific-publications/how-does-parkinsons-disease-alter-visual-search> (term. wiz. 09.01.2025).
- [7] Imię Nazwisko i Imię Nazwisko. *Tytuł strony internetowej*. 2021. URL: <http://gdzies/w/internecie/internet.html> (term. wiz. 30.09.2021).
- [8] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. „Tytuł artykułu konferencyjnego”. W: *Nazwa konferencji*. 2006, s. 5346–5349.
- [9] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. „Tytuł artykułu w czasopiśmie”. W: *Tytuł czasopisma* 157.8 (2016), s. 1092–1113.

- [10] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. *Tytuł książki*. Warszawa: Wydawnictwo, 2017. ISBN: 83-204-3229-9-434.
- [11] iMotions Science Team. *iMotions Publications: 2023 Report*. 2024. URL: <https://imotions.com/wp-content/uploads/brochures/Research%20report%202024.pdf> (term. wiz. 20.01.2025).
- [12] iMotions Science Team. *Smart Eye AI-X*. 2025. URL: <https://imotions.com/products/hardware/smart-eye-ai-x/#product-specifications> (term. wiz. 20.01.2025).
- [13] iMotions Science Team. *Who we are: about us*. 2025. URL: <https://imotions.com/about-us/> (term. wiz. 26.01.2025).
- [14] Divya Seernani, Morten Mosbaek Pedersen i Kerstin Wolf. *WebET 3.0 - Validation Study Report*. Spraw. tech. 2023. DOI: 10.13140/RG.2.2.32959.07849.
- [15] Tobii. *Eye tracking in neurology and psychiatry research*. 2025. URL: <https://www.tobii.com/resource-center/reports-and-papers/eye-tracking-neurology-and-psychiatry-research> (term. wiz. 09.01.2025).
- [16] Tobii. *Psychology and neuroscience use cases*. 2025. URL: <https://www.tobii.com/products/eye-trackers/screen-based/psychology-neuroscience-use-cases> (term. wiz. 09.01.2025).

Dodatki

Spis skrótów i symboli

DNA kwas deoksyrybonukleinowy (ang. *deoxyribonucleic acid*)

MVC model – widok – kontroler (ang. *model-view-controller*)

N liczebność zbioru danych

μ stopnień przyleżności do zbioru

\mathbb{E} zbiór krawędzi grafu

\mathcal{L} transformata Laplace’a

Źródła

Jeżeli w pracy konieczne jest umieszczenie długich fragmentów kodu źródłowego, należy je przenieść w to miejsce.

```
1 if (_nClusters < 1)
2     throw std::string ("unknown_number_of_clusters");
3 if (_nIterations < 1 and _epsilon < 0)
4     throw std::string ("You_should_set_a_maximal_number_of_
        iteration_or_minimal_difference_epsilon.");
5 if (_nIterations > 0 and _epsilon > 0)
6     throw std::string ("Both_number_of_iterations_and_minimal_
        epsilon_set—you_should_set_either_number_of_iterations_
        or_minimal_epsilon.");
```

Lista dodatkowych plików, uzupełniających tekst pracy

W systemie do pracy dołączono dodatkowe pliki zawierające:

- źródła programu,
- dane testowe,
- film pokazujący działanie opracowanego oprogramowania lub zaprojektowanego i wykonanego urządzenia,
- itp.

Spis rysunków

2.1	Diagram modelu ruchów sakadowych z filtrem liniowym różniczkującym. .	5
2.2	Diagram modelu płynnego podążania z liniowym sprzężeniem zwrotnym. .	6
4.1	Podpis rysunku po rysunkiem.	18
5.1	Pseudokod w <code>listings</code>	20

Spis tabel

6.1	Nagłówek tabeli jest nad tabelą.	22
-----	--	----