



**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	[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
3	Wymagania i narzędzia	9
4	[Właściwy dla kierunku – np. Specyfikacja zewnętrzna]	11
5	[Właściwy dla kierunku – np. Specyfikacja wewnętrzna]	13
6	Weryfikacja i walidacja	15
7	Podsumowanie i wnioski	17
	Bibliografia	19
	Spis skrótów i symboli	23
	Źródła	25
	Lista dodatkowych plików, uzupełniających tekst pracy	27
	Spis rysunków	29
	Spis tabel	31

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 [9]. 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 [3], Alzheimerem [2], a także autyzmem czy łagodnym upośledzeniem funkcji poznawczych [8]. 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 stworzenie 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

[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 dokonuje potrzebnej konwersji informacji prędkość na przemieszczenie. Wzór owego filtra w dziedzinie czasu można zapisać następująco

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

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). W zapisie sumarycznym wzór prezentuje się następująco

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

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

$$x_t = g_0 \cdot s_t + g_1 \cdot s_{t-1}$$

$$x_t = 1 \cdot s_t - 1 \cdot s_{t-1}$$

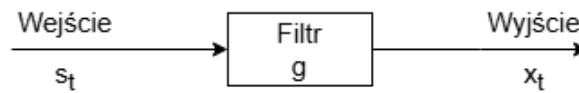
$$x_t = s_t - s_{t-1}$$

$$\mathcal{Z}\{x_t\} = \mathcal{Z}\{s_t - s_{t-1}\}$$

$$X(z) = (1 - z) \cdot S(z)$$

$$\frac{X(z)}{S(z)} = 1 - z$$

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 ruchu jest przykładem systemu sterowania z ujemnym sprzężeniem zwrotnym. Do modelowania owych ruchów używana

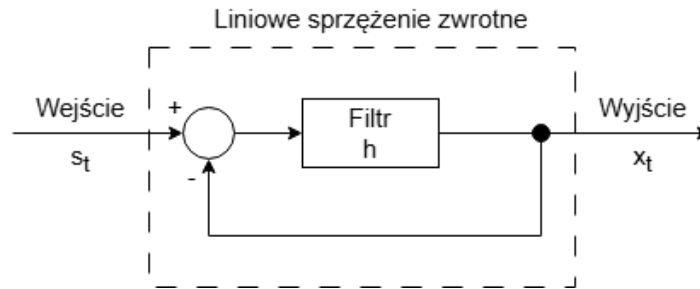
jest prosta pętla, którą można opisać następującym równaniem w dziedzinie czasu

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

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 by uzyskać transmitancję systemu

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

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 zebranie owych danych. Poniżej przedstawiono kilka z nich, większość

- sformułowanie problemu
- osadzenie tematu w kontekście aktualnego stanu wiedzy (*state of the art*) o poruszonym problemie
- studia literaturowe [6, 7, 5, 4] - opis znanych rozwiązań (także opisanych naukowo, jeżeli problem jest poruszany w publikacjach naukowych), algorytmów,

Wzory

$$y = \frac{\partial x}{\partial t} \tag{2.1}$$

jak i pojedyncze symbole x i y składa się w trybie matematycznym.

Rozdział 3

Wymagania i narzędzia

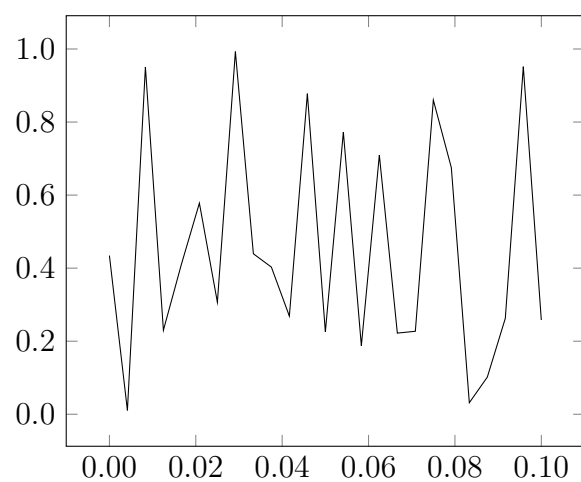
- wymagania funkcjonalne i нефункционалне
- przypadki użycia (diagramy UML) – dla prac, w których mają zastosowanie
- opis narzędzi, metod eksperymentalnych, metod modelowania itp.
- metodyka pracy nad projektowaniem i implementacją – dla prac, w których ma to zastosowanie

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] 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).
- [3] 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).
- [4] Imię Nazwisko i Imię Nazwisko. *Tytuł strony internetowej*. 2021. URL: <http://gdzies/w/internecie/internet.html> (term. wiz. 30.09.2021).
- [5] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. „Tytuł artykułu konferencyjnego”. W: *Nazwa konferencji*. 2006, s. 5346–5349.
- [6] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. „Tytuł artykułu w czasopiśmie”. W: *Tytuł czasopisma* 157.8 (2016), s. 1092–1113.
- [7] Imię Nazwisko, Imię Nazwisko i Imię Nazwisko. *Tytuł książki*. Warszawa: Wydawnictwo, 2017. ISBN: 83-204-3229-9-434.
- [8] 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).
- [9] 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.	12
5.1	Pseudokod w <code>listings</code>	14

Spis tabel

6.1	Nagłówek tabeli jest nad tabelą.	16
-----	--	----