

# **Sztuczna inteligencja i inżynieria wiedzy**

**Sprawozdanie nr 4  
Rozpoznawanie obrazów**

**Przemysław Pietrzak, 238083  
Środa, 17:05**

# Spis treści

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Opis implementacji</b>   | <b>2</b>  |
| <b>2</b> | <b>Prezentacja działania algorytmu</b>                            | <b>2</b>  |
| 2.1      | Kubek . . . . .   | 3         |
| 2.2      | Portfel . . . . .   | 3         |
| 2.3      | Mysz komputerowa . . . . .  | 3         |
| 2.4      | Książka . . . . .   | 4         |
| 2.5      | Kaktus . . . . .  | 4         |
| 2.6      | Omówienie wyników . . . . .                                       | 5         |
| <b>3</b> | <b>Badanie wpływu parametrów na algorytm spójności sąsiedztwa</b> | <b>6</b>  |
| 3.1      | Liczba sąsiadów . . . . .   | 6         |
| 3.2      | Próg spójności . . . . .  | 6         |
| 3.3      | Omówienie wyników . . . . .                                       | 6         |
| <b>4</b> | <b>Badanie wpływu użytej transformaty na wyniki metody RANSAC</b> | <b>7</b>  |
| 4.1      | Transformata afiniczna . . . . .                                  | 7         |
| 4.2      | Transformata perspektywiczna . . . . .                            | 7         |
| 4.3      | Omówienie wyników . . . . .                                       | 7         |
| <b>5</b> | <b>Badanie wpływu parametrów na wyniki metody RANSAC</b>          | <b>8</b>  |
| 5.1      | Liczba iteracji . . . . .   | 8         |
| 5.2      | Maksymalny błąd . . . . .   | 8         |
| 5.3      | Omówienie wyników . . . . .                                       | 8         |
| <b>6</b> | <b>Badanie wpływu heurystyk na wyniki metody RANSAC</b>           | <b>9</b>  |
| 6.1      | Random . . . . .  | 9         |
| 6.2      | Heurystyka 1 . . . . .  | 9         |
| 6.3      | Heurystyka 2 . . . . .  | 9         |
| 6.4      | Heurystyka 3 . . . . .  | 9         |
| 6.5      | Heurystyka 4 . . . . .  | 9         |
| 6.6      | Heurystyka 5 . . . . .  | 9         |
| 6.7      | Omówienie wyników . . . . .                                       | 9         |
| <b>7</b> | <b>Podsumowanie</b>   | <b>10</b> |

# 1 Opis implementacji

Program składa się z kilku modułów odpowiadających za wyznaczanie par punktów kluczowych, wyszukiwanie par spójnych, wyznaczanie transformat, uruchomienie algorytmu RANSAC oraz usprawnianie procesu odczytu i zapisu obrazów oraz plików .json i .harriff.sift. Po podaniu dwóch plików źródłowych (w formacie .png) program korzystając z przekazanego nam skryptu wyznacza punkty kluczowe (regiony kluczowe) na obu obrazach. Następnie rozpoczęty jest proces łączenia punktów w pary na podstawie ich podobieństwa wizualnego (wartości features). W kolejnym kroku zbiór par punktów kluczowych jest filtrowany. W dalszych punktach brane pod uwagę są tylko te punkty, które spełniają kryterium spójności. Pary spójnych punktów kluczowych są następnie przekazywane do następnej części programu, która uruchamia metodę RANSAC, w każdym kroku algorytmu na podstawie wybranej transformaty jest wyznaczany model. Otrzymany model jest następnie ewaluowany. Ocena modelu jest liczba par punktów kluczowych, których odległość między punktem z obrazu drugiego, a transformowanym punktem z obrazu pierwszego jest mniejsza niż zadana wartość dozwolonego błędu. Na sam koniec najlepszy model jest używany do odfiltrowania par punktów kluczowych, które nie spełniają podanego kryteriu (odległość między punktami jest większa od dozwolonego błędu).

# 2 Prezentacja działania algorytmu

Algorytm uruchomiono na pięciu parach zdjęć przedstawiających przedmioty takie jak mysz komputerowa, książka, portfel i tym podobne. Przedmioty na zdjęcia znajdują się w możliwie różnych sceneriach, pozycjach oraz widoczne są z różnych perspektyw. Program został uruchomiony z następującymi parametrami.

Tabela 1: Parametry metody badania spójności par punktów kluczowych

| Liczba sąsiadów | Ustalony próg spójności |
|-----------------|-------------------------|
| 50              | 0,5                     |

Tabela 2: Parametry metody RANSAC

| Liczba iteracji | Maksymalny błąd | Transformata    | Heurystyka                 |
|-----------------|-----------------|-----------------|----------------------------|
| 5000            | 20              | Perspektywiczna | Pominiecie najgorszych par |

Na następnych stronach znajdują się porównania wyników w kolejnych etapach działania programu.

## 2.1 Kubek

W poniższym porównaniu zostały użyte dwa zdjęcia kubka przy różnych źródłach światła, na drugim zdjęciu kubek jest obrócony o 90 stopni (w płaszczyźnie XZ). Dodatkowo kubek został lekko obrócony wokół osi X, aby widoczne były różne części nadruku. Niebieskie (ciemne) linie oznaczają par punktów przed odfiltrowaniem (badanie spójności lub RANSAC), a żółte (jasne) po.

Rysunek 1: Zdjęcia porównawcze



Rysunek 2: Wynik badania spójności



Rysunek 3: Wynik metody RANSAC

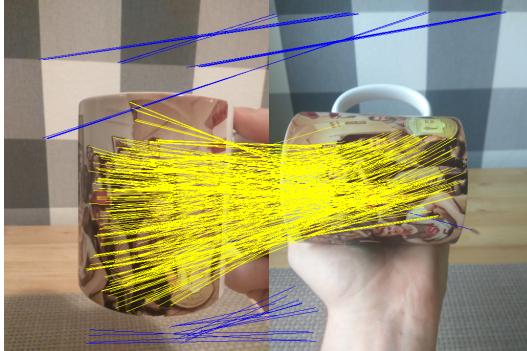


Tabela 3: Zestawienie wyników par w etapach

| Punkty (obraz 1) | Punkty (obraz 2) | Pary punktów | Pary spójne | Ostateczne pary |
|------------------|------------------|--------------|-------------|-----------------|
| 3090             | 2821             | 758          | 451         | 410             |

## 2.2 Portfel

W poniższym porównaniu zostały użyte dwa zdjęcia portfela w różnych sceneriach, na pierwszym zdjęciu portfel leży na blacie jest on trzymany w ręce. Dodatkowo zdjęcia zostały dokonane pod innym kątem i w różnej pozycji portfela. Niebieskie (ciemne) linie oznaczają par punktów przed odfiltrowaniem (badanie spójności lub RANSAC), a żółte (jasne) po.

Rysunek 4: Zdjęcia porównawcze



Rysunek 5: Wynik badania spójności



Rysunek 6: Wynik metody RANSAC



Tabela 4: Zestawienie wyników par w etapach

| Punkty (obraz 1) | Punkty (obraz 2) | Pary punktów | Pary spójne | Ostateczne pary |
|------------------|------------------|--------------|-------------|-----------------|
| 3986             | 9905             | 1075         | 265         | 261             |

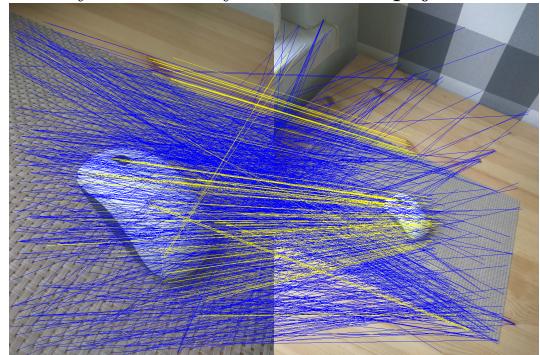
### 2.3 Mysz komputerowa

Zdjęcia zostały zrobione z różnych odległości oraz pod lekko innym kątem, tło pozostało takie samo dlatego nadal występują pary łączące podkładkę zamiast myszki. Niebieskie (ciemne) linie oznaczają par punktów przed odfiltrowaniem (badanie spójności lub RANSAC), a żółte (jasne) po.

Rysunek 7: Zdjęcia porównawcze



Rysunek 8: Wynik badania spójności



Rysunek 9: Wynik metody RANSAC

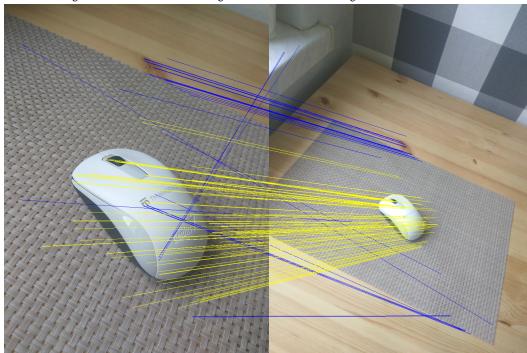


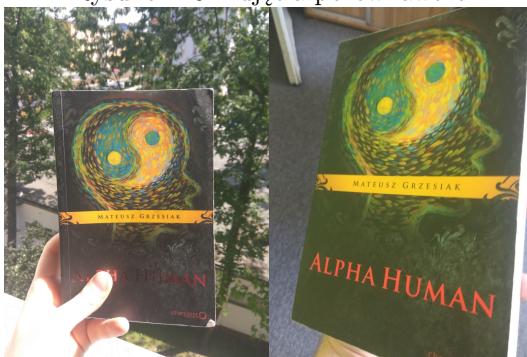
Tabela 5: Zestawienie wyników par w etapach

| Punkty (obraz 1) | Punkty (obraz 2) | Pary punktów | Pary spójne | Ostateczne pary |
|------------------|------------------|--------------|-------------|-----------------|
| 15806            | 5939             | 851          | 149         | 112             |

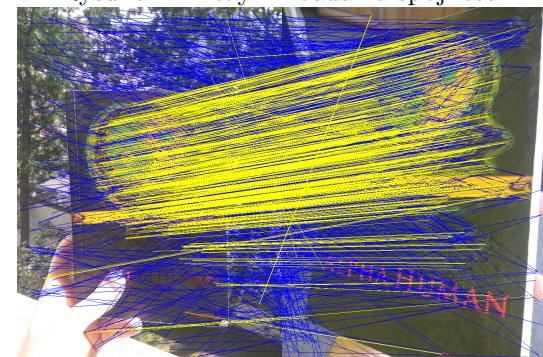
## 2.4 Książka

Zdjęcia książki zostały zrobione w różnej scenerii i pod różnymi kątami. Niebieskie (ciemne) linie oznaczają par punktów przed odfiltrowaniem (badanie spójności lub RANSAC), a żółte (jasne) po.

Rysunek 10: Zdjęcia porównawcze



Rysunek 11: Wynik badania spójności



Rysunek 12: Wynik metody RANSAC

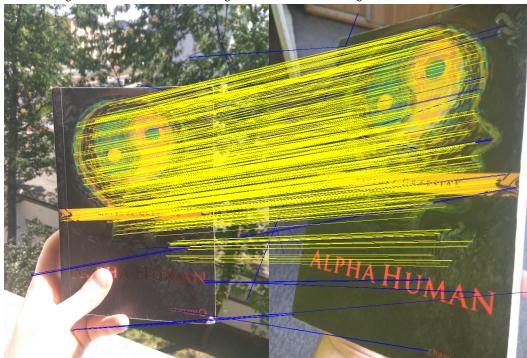


Tabela 6: Zestawienie wyników par w etapach

| Punkty (obraz 1) | Punkty (obraz 2) | Pary punktów | Pary spójne | Ostateczne pary |
|------------------|------------------|--------------|-------------|-----------------|
| 10654            | 5212             | 1548         | 717         | 647             |

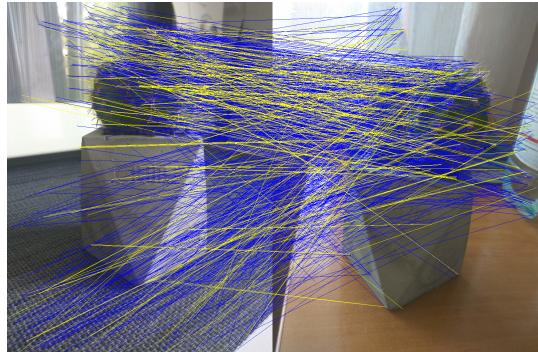
## 2.5 Kaktus

Zdjęcia zostały zrobione w różnych pomieszczeniach i odrobinę różnym oświetleniu. Niebieskie (ciemne) linie oznaczają pary punktów przed odfiltrowaniem (badanie spójności lub RANSAC), a żółte (jasne) po. W przypadku tej pary zdjęć konieczne była modyfikacja parametrów algorytmu spójności sąsiedztwa, liczba sąsiadów została zredukowana do 10, natomiast próg spójności został zmniejszony do wartości 0.35. W przypadku większych parametrów algorytm nie potrafił znaleźć żadnej pary punktów spójnych.

Rysunek 13: Zdjęcia porównawcze



Rysunek 14: Wynik badania spójności



Rysunek 15: Wynik metody RANSAC

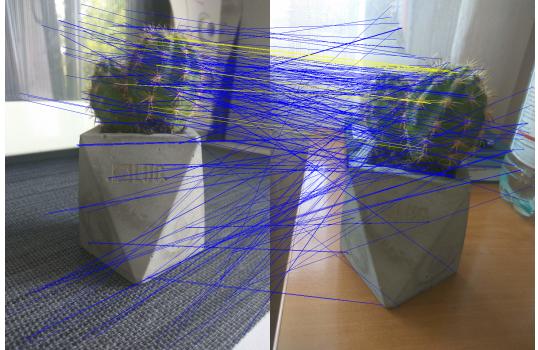


Tabela 7: Zestawienie wyników par w etapach

| Punkty (obraz 1) | Punkty (obraz 2) | Pary punktów | Pary spójne | Ostateczne pary |
|------------------|------------------|--------------|-------------|-----------------|
| 12967            | 3569             | 946          | 90          | 18              |

## 2.6 Omówienie wyników

Wynik w przypadku większości badanych par obrazów jest zadowalający, widoczna jest redukcja par zarówno po sprawdzeniu spójności jak i uruchomieniu metody RANSAC. Najlepsze rezultaty uzyskano podczas badania zdjęć kubka oraz książki. W przypadku myszki i kaktusa skrypt wyznaczający punkty kluczowe skupiał się nie na samym przedmiocie, a powierzchni na której leżały przedmioty. Z tego powodu liczba par spójnych była bardzo mała w stosunku do liczby punktów kluczowych na obu obrazach i została ponownie zredukowana po użyciu metody RANSAC. Pary punktów kluczowych były wyznaczane na podstawie błędu średniokwadratowego z cech punktów, być może inna metoda pozwoliłaby na uzyskanie lepszych rezultatów. Liczba spójnych par mogłaby także zostać zwiększoa poprzez obniżenie progu spójności, który w przypadku badanych obrazów wynosił 0.4.

### 3 Badanie wpływu parametrów na algorytm spójności sąsiedztwa

Główym zadaniem algorytmu spójności sąsiedztwa jest odfiltrowanie nieznaczących par. Pary w poprzednim kroku są wyliczane na podstawie cech punktu kluczowego, jednak takie podejście może generować dużą ilość szumu w przypadku punktów kluczowych, które wyglądają tak samo, ale nie odnoszą się do tego samego regionu na drugim obrazie. Sprawdzając sąsiedztwo każdego punktu z pary możemy odizolować odosobnione regiony obrazu, w których znajdują się punkty kluczowe, otrzymując przy tym zbiór sensownych par.

#### 3.1 Liczba sąsiadów

Badanie ma na celu wyznaczenie zależności wynikowej liczby par spójnych punktów kluczowych oraz czasu przetwarzania od liczności sąsiedztwa. Algorytm został uruchomiony na każdym z pięciu obrazów używanych w poprzednim etapie. Badania przeprowadzono dla liczności sąsiedztwa o rozmiarach: 5, 20, 50, 100, 300. Brany pod uwagę próg spójności wynosi 0,5 (50%). Czasy przetwarzania są średnią z pięciu uruchomień.

Tabela 8: Porównanie liczby par spójnych i czasu przetwarzania w zależności od liczności sąsiedztwa

| Obiekt  | Pary | Liczność sąsiedztwa |     |     |     |     |     |                        |      |      |      |      |       |
|---------|------|---------------------|-----|-----|-----|-----|-----|------------------------|------|------|------|------|-------|
|         |      | 1                   | 5   | 20  | 50  | 100 | 300 | 1                      | 5    | 20   | 50   |      |       |
|         |      | Pary spójne         |     |     |     |     |     | Czas przetwarzania [s] |      |      |      |      |       |
| Kubek   | 758  | 330                 | 382 | 421 | 432 | 424 | 531 | 1,01                   | 1,03 | 1,10 | 1,27 | 1,71 | 5,58  |
| Myszka  | 851  | 152                 | 82  | 73  | 40  | 30  | 428 | 1,40                   | 1,41 | 1,47 | 1,67 | 2,17 | 6,61  |
| Kaktus  | 946  | 102                 | 12  | 0   | 0   | 0   | 0   | 1,73                   | 1,89 | 1,81 | 2,01 | 2,87 | 7,68  |
| Portfel | 1075 | 242                 | 223 | 237 | 228 | 228 | 238 | 2,33                   | 2,38 | 2,41 | 2,53 | 2,70 | 9,51  |
| Książka | 1548 | 549                 | 520 | 636 | 650 | 708 | 825 | 5,31                   | 4,73 | 5,48 | 6,08 | 6,97 | 14,20 |

Przeprowadzone badania jasno wskazują na to, że czas przetwarzania jest zależny zarówno od liczby par punktów spójnych na wejściu algorytmu jak i liczności sąsiedztwa. Zwiększenie liczby badanych par powoduje wzrost wymaganej liczby iteracji, ponieważ koniecznie jest zmierzenie odległości między większą liczbą punktów. Podobna sytuacja występuje w przypadku liczności sąsiedztwa, gdzie większa liczba sąsiadów powoduje, że algorytm musi zbadać przynależność do sąsiedztwa dla większej liczby przypadków.

Liczba par spójnych jest najbardziej stabilna przy rozmiarze sąsiedztwa z przedziału 20-50. Warto zaznaczyć, że większa liczba punktów spójnych niekoniecznie oznacza większą skuteczność algorytmu, ponieważ punkty te mogą łączyć ze sobą całkowicie niepotrzebne regiony. Dla mniejszej liczby algorytm wyłapuje dużą liczbę bezsensownych, ponieważ nie ma on wtedy wystarczającej liczby danych wymaganych do odpowiedniego odfiltrowania zbioru par. Najbardziej widoczne było to w przypadku zdjęć przedstawiających myszkę komputerową gdzie algorytm, oprócz par łączących regiony myszki, przepuszczał także te łączące elementy stołu czy podkładki.

Zbyt duża liczność sąsiedztwa z kolei powodowała, że przestawało mieć ono jakiekolwiek znaczenie. Algorytm uznawał za spójną prawie każdą parę ze zbioru wejściowego. Jest to całkowicie zrozumiałe, ponieważ w przypadku większej liczności sąsiedztwa szansa na to, że regiony z sąsiedztwa punktu X (na pierwszym obrazie) tworzą pary z sąsiadami punktu Y na drugim obrazie i odwrotnie.

### 3.2 Próg spójności

Badanie ma na celu wyznaczenie zależności wynikowej liczby par spójnych punktów kluczowych od progu dedcyującego o spójności pary. Algorytm został uruchomiony na każdym z pięciu obrazów używanych w poprzednim etapie. Badania przeprowadzono progów spójności o wartościach: 0,00, 0,25, 0,50, 0,75 oraz 1,00. W przypadku tego badania liczba sąsiadów została ustalona na 50, ponieważ jest to wartość optymalna (dla badanych obrazów) według poprzedniego badania. Czasy przetwarzania są średnią z pięciu uruchomień.

Tabela 9: Porównanie liczby par spójnych i czasu przetwarzania w zależności od progu spójności

| Obiekt  | Pary | Próg spójności |      |      |      |      |                        |      |      |      |      |
|---------|------|----------------|------|------|------|------|------------------------|------|------|------|------|
|         |      | 0,00           | 0,25 | 0,50 | 0,75 | 1,00 | 0,00                   | 0,25 | 0,50 | 0,75 |      |
|         |      | Pary spójne    |      |      |      |      | Czas przetwarzania [s] |      |      |      |      |
| Kubek   | 758  | 758            | 481  | 432  | 266  | 0    | 1,29                   | 1,32 | 1,31 | 1,31 | 1,32 |
| Myszka  | 851  | 851            | 256  | 40   | 0    | 0    | 1,78                   | 1,75 | 1,79 | 1,82 | 1,79 |
| Kaktus  | 946  | 946            | 0    | 0    | 0    | 0    | 2,51                   | 2,58 | 2,57 | 2,54 | 2,61 |
| Portfel | 1075 | 1075           | 301  | 228  | 76   | 0    | 2,71                   | 2,72 | 2,71 | 2,73 | 2,71 |
| Książka | 1548 | 1548           | 788  | 650  | 108  | 0    | 5,82                   | 5,84 | 5,81 | 5,83 | 5,84 |

Badania jasno wskazują na to, że wartość progu, który decyduje o spójności pary nie ma żadnego wpływu na czas przetwarzania. Ewaluacja i sprawdzenie warunku musi nastąpić dla każdej pary, niezależnie od wartości progu. Jednak liczba uzyskanych par punktów spójnych może wpływać na działanie algorytmu RANSAC w późniejszych etapach.

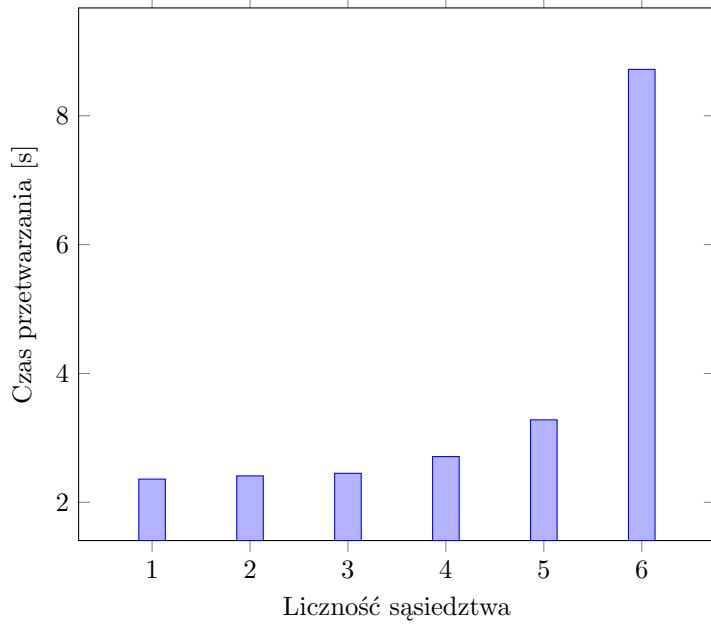
Przy przypadku progu o wartości 0,00 (0%) każda para uznawana jest za spójną, odwrotnie sytuacja wygląda w przypadku progu spójności 1,00 (100%), gdzie wymagane jest, aby każdy z sąsiadów punktu X był w parze z sąsiadami odpowiadającego punktu Y na drugim obrazie. Sytuacja ta jest praktycznie niemożliwa do wystąpienia, dlatego też liczba wyznaczonych punktów spójnych wynosi 0.

W przypadku pozostałych wartości widać spadek liczby par spójnych wraz ze wzrostem progu spójności, jest to zrozumiałe, ponieważ większy próg spójności oznacza bardziej rygorystyczne wybieranie par regionów kluczowych.

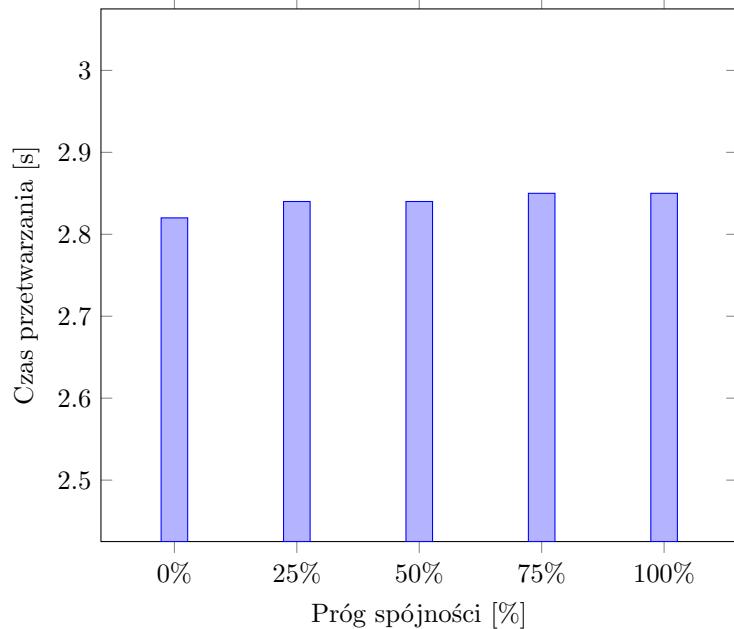
Najlepsze wyniki, zgodnie z instrukcją, można otrzymać dla wartości progu spójności 0,50 (50%) i 0,75 (75%).

### 3.3 Omówienie wyników

Rysunek 16: Wpływ liczności sąsiedztwa na czas przetwarzania



Rysunek 17: Wpływ wartości progu spójności na czas przetwarzania



Z przeprowadzonych badań można wywnioskować, że w przypadku badanych obrazów najbardziej skuteczne parametry to liczność sąsiedztwa znajdująca się w zakresie (20,50) oraz próg

spójności o wartościach około 0,50 (50%). Nieodpowiedni dobór parametrów może doprowadzić do zbyt rygorystycznej selekcji par punktów spójnych lub do występowania dużej ilości szumu w zbiorze par spójnych. Tak jak wspomniano w badaniach wartość progu spójności nie ma żadnego wpływu na czas przetwarzania, większa liczność sąsiedztwa jednak może znacznie wpłynąć na czas potrzebny do uzyskania wyniku.

## 4 Badanie wpływu użytej transformaty na wyniki metody RANSAC

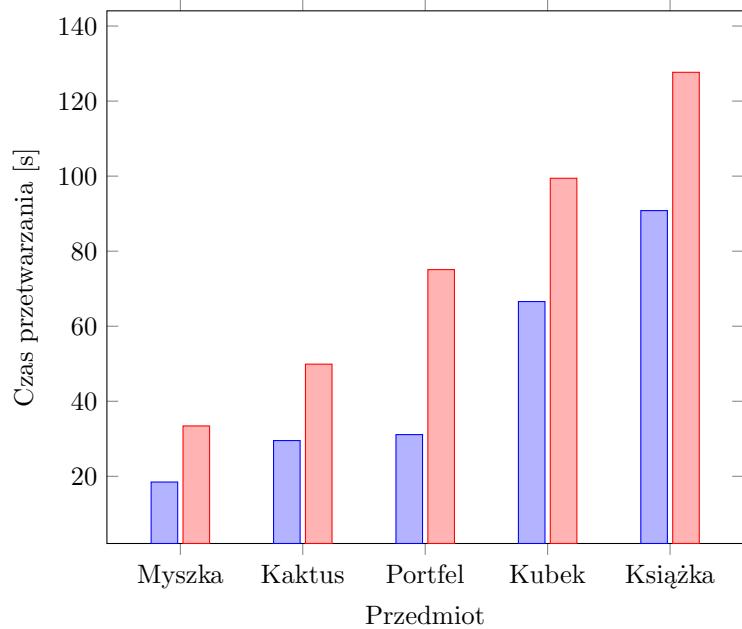
W przypadku kiedy obiekt na zdjęciach znajduje się w różnych ustawieniach, czyli np. jest obrócony lub leży pod kątem, podobnieństwo może nie zostać stwierdzone. Rozwiązaniem tego problemu są transformaty, które odpowiadają za przekształcenie jednego obrazu w drugi. Badanie skupia się na dwóch transformatach: afincznej i perspektywicznej. Transformata afinczna odpowiada za przekształcenia typu przesunięcie, skalowanie i obrót. Transformata perspektywiczna natomiast rozszerza ją i bierze pod uwagę przekształcenie polegające na spłaszczeniu i rozciąganiu obrazu. Dzięki nim jesteśmy w stanie porównać obiekty nawet w przypadku kiedy oba obrazy pokazują je pod innym kątem.

Badania przeprowadzone zostały dla liczby iteracji równej 5000, dodatkowo próg błędu wynosi 25. Przedstawione rezultaty są średnimi wynikami z pięciu uruchomień. Kolumny Para i Iteracja oznaczają kolejno liczbę otrzymanych w wyniku działania metody RANSAC par oraz, w której iteracji zostało znalezione to rozwiązanie. Czas przetwarzania natomiast oznacza liczbę sekund potrzebną na ukończenie wszystkich pięciu tysięcy iteracji.

Tabela 10: Porównanie wyników dla transformaty afincznej i perspektywicznej

| Przedmiot | Transformata afinczna |          |                        | Transformata perspektywiczna |          |                        |
|-----------|-----------------------|----------|------------------------|------------------------------|----------|------------------------|
|           | Pary                  | Iteracja | Czas przetwarzania [s] | Pary                         | Iteracja | Czas przetwarzania [s] |
| Kubek     | 406                   | 12       | 66,57                  | 406                          | 14       | 99,43                  |
| Portfel   | 214                   | 231      | 31,11                  | 239                          | 760      | 75,09                  |
| Myszka    | 36                    | 1131     | 18,47                  | 49                           | 638      | 33,42                  |
| Książka   | 595                   | 4000     | 90,81                  | 648                          | 102      | 127,67                 |
| Kaktus    | 13                    | 3573     | 29,52                  | 15                           | 3582     | 49,89                  |

Rysunek 18: Wpływ wyboru transformaty na czas przetwarzania



Głównym wnioskiem wysuwającym się po przeprowadzeniu badań jest to, że czas przetwarzania w przypadku transformaty perspektywicznej jest dłuższy. Spowodowane jest to tym, że uwzględnienie transformacji rozciągania i zwężania obrazu wymaga większej liczby obliczeń. W większości przypadków transformata perspektywiczna osiągała lepsze wyniki, ponieważ większość par zdjęć źródłowych różniła się kątem nachylenia przedmiotu. Można także zauważyć, że transformata perspektywiczna zwykle znajdowała lepsze wyniki już w początkowych iteracjach. Jednak mogła na to mieć także wpływ losowość metody RANSAC.

## 5 Badanie wpływu parametrów na wyniki metody RANSAC

Metoda RANSAC jako parametry przyjmuje dwie wartości, liczbę iteracji oraz maksymalną wartość błędu, której przekroczenie powoduje odfiltrowanie pary ze zbioru wynikowego. Do obliczeń wykorzystano transformatę afiniczą z heurystyką wybierającą trzy losowe pozycje ze zbioru par punktów kluczowych.

Tabela 11: Wyniki pomiarów czasu w zależności od liczby iteracji

| Przedmiot | Liczba iteracji |       |       |        |
|-----------|-----------------|-------|-------|--------|
|           | 10              | 500   | 1000  | 5000   |
|           | Czas [s]        |       |       |        |
| Myszka    | 0,04            | 2,70  | 4,57  | 29,97  |
| Kaktus    | 0,05            | 2,41  | 4,82  | 32,48  |
| Portfel   | 0,12            | 3,31  | 6,84  | 41,89  |
| Kubek     | 0,15            | 9,36  | 19,01 | 102,29 |
| Książka   | 0,21            | 10,84 | 26,92 | 124,54 |

Tabela 12: Wyniki działania algorytmu w zależności o wartość błędu maksymalnego

| Przedmiot | Wartość maksymalnego błędu |     |     |     |
|-----------|----------------------------|-----|-----|-----|
|           | 1                          | 20  | 40  | 200 |
|           | Największy konsensus       |     |     |     |
| Kubek     | 80                         | 407 | 407 | 445 |
| Portfel   | 59                         | 222 | 225 | 242 |
| Myszka    | 6                          | 37  | 40  | 52  |
| Książka   | 60                         | 645 | 645 | 650 |
| Kaktus    | 11                         | 12  | 12  | 12  |

Jak można zauważyć w badaniach liczba iteracji, zaraz obok rozmiaru list par punktów spójnych jest jednym z najważniejszych czynników wpływających na czas przetwarzania. Przy zbyt małej liczbie iteracji algorytm nie jest w stanie znaleźć optymalnego rozwiązania duża liczba iteracji jednak znaczco wpływa na czas działania programu, podczas gdy rozwiązanie mogło zostać znalezione już na samym początku.

Maksymalny błąd natomiast jest jednym z czynników definiujących skuteczność metody RANSAC, zbyt mały błąd powoduje odfiltrowanie znacznie liczby właściwych par, natomiast zbyt duży pozostawia w danych wiele szumów. Przeprowadzone badania wskazują na to, że dla badanych par obrazków, najbardziej optymalny wartość błędu wynosi od 20 do 40.

## 6 Badanie wpływu heurystyk na wyniki metody RANSAC

W badanej metodzie transformaty wyliczają model na podstawie 3 lub 4 par punktów kluczowych. Sposób wyboru tych wartości może znacząco wpływać na jakość działania metody RANSAC. Poniższe heurystyki mają za zadanie dokonać odpowiedniego wyboru tych punktów celem utworzenia poprawnego modelu. Ze względu na dużą liczbę obliczeń badania zostały przeprowadzone na zdjęciach przedstawiających przedmiot Książka. Wykorzystana została transformata perspektywiczna. Przedstawione wyniki są średnią z pięciu uruchomień.

Badane heurystyki:

- **losowa** - pary punktów wybierane są całkowicie losowo
- **modyfikacja rozkładu** - pary, które prowadzą do poprawy wybierane są częściej
- **odległości par punktów** - wybierane są pary, których punkty leżą blisko siebie na obrazach
- **usuwanie niepoprawnych par** - pary, które doprowadzają do znalezienia najgorszego dotychczas rozwiązania nie są brane pod uwagę w kolejnych iteracjach
- **rozmiar sąsiedztwa** - wybierane są pary, których rozmiar sąsiedztwa jest największy

Rysunek 19: Badanie wpływu heurystyk na wynik metody RANSAC

| Heurystyka            | Najlepsze rozwiązanie |          |          | Całkowite przejście |          |
|-----------------------|-----------------------|----------|----------|---------------------|----------|
|                       | Wynik                 | Iteracja | Czas [s] | Iteracje            | Czas [s] |
| Losowa                | 618                   | 1587     | 39,76    | 2000                | 45,23    |
| Modyfikacja rozkładu  | 619                   | 1579     | 47,03    | 2000                | 59,54    |
| Odległość r=4, R=400  | 628                   | 1317     | 32,92    | 2000                | 48,48    |
| Odległość r=4, R=1000 | 640                   | 1642     | 39,55    | 2000                | 48,43    |
| Usuwanie              | 610                   | 1401     | 34,76    | 2000                | 49,09    |
| Sąsiedztwo - 5        | 604                   | 404      | 33,55    | 2000                | 167,09   |
| Sąsiedztwo - 20       | 615                   | 409      | 35,63    | 2000                | 182,62   |

Wszystkie badane metody uzyskały zadowalające wyniki. Najwyższy wynik uzyskała prosta heurystyka wyboru punktów na podstawie odległości  $r=4$  oraz  $R=1000$ . Głównym założeniem tej heurystyki jest to, że prawdopodobieństwo wylosowania punktów leżących blisko siebie jest większe niż wylosowania takich, które znajdują się w większej odległości. Dolna granica odległości jest odpowiedzialna za zminimalizowanie możliwości wystąpienia błędów obliczeniowych. Parametry  $r$  oraz  $R$  należy dobrać odpowiednio do badanej przez nas pary obrazów. Dodatkowo warto wspomnieć, że heurystyka ta wymaga wielu obliczeń, dlatego jej czas wykonania może być znacznie większy. Jednak jedną z jej zalet jest to, że pozwala ona znaleźć bardzo dobre rozwiązania w stosunkowo małej liczbie iteracji. Łącząc ją z badaną w późniejszych etapach metodą estymacji wymaganej liczby iteracji można uzyskać satysfakcyjujące wyniki w niedługim czasie. Pod względem iteracji wyróżnia się heurystyka polegająca na wybieraniu par, których punkty mają największą liczebność sąsiedztwa. Obie z badanych wartości parametru  $r$  uzyskały satysfakcyjny wynik już przy ok. 400 iteracjach. Czas obliczeniowy potrzebny na wyznaczenie sąsiedztwa

dla każdego z punktów w parach jest jednak ogromny dlatego rezultaty czasowe, szczególnie dla całkowitego przejścia), nie są zadowalające. Usprawnienie implementacji, poprzez przykładowo, wyznaczanie sąsiedztwa na wielu wątkach jednocześnie mogłoby doprowadzić do lepszych rezultatów czasowych. Najgorzej wypadła heurystyka wyznaczająca pary punktów kluczowych 'Modyfikacja rozkładu', która polega na wybieraniu częściej tych par, które wcześniej zwracały dobre rezultaty. Otrzymany wynik jednak może być wynikiem wadliwej implementacji.

## 7 Badanie metody ograniczania liczby iteracji

Metoda ograniczenia liczby iteracji polega na wyznaczeniu odpowiedniej liczby obiegów metody RANSAC na podstawie prawdopodobieństwa tego czy po określonej liczbie iteracji model jest wystarczająco dobry oraz prawdopodobieństwa określającego jaka jest szansa na to, że losowo wybrana para ze zbioru ni jest szumem.

Badania zostały przeprowadzone na wartościach prawdopodobieństwa  $p$  równych 0,10, 0,50, 0,75, oraz 0,9. Wartości parametru  $p$  zostały estymowane na podstawie stosunku liczby par przed i po zbadaniu spójności. W badaniach wykorzystano transformatę perspektywiczną oraz heurystykę opierającą się o odległości punktów kluczowych. Wyniki są wartościami uśrednionymi uzyskanyimi na podstawie pięciu uruchomień.

Rysunek 20: Wyniki dla pary obrazów przedstawiających książkę

| p    | Najlepsze rozwiązanie |          |          | Całkowite przejście |          |
|------|-----------------------|----------|----------|---------------------|----------|
|      | Wynik                 | Iteracja | Czas [s] | Iteracje            | Czas [s] |
| 0,10 | 400                   | 2        | 0,05     | 3                   | 0,05     |
| 0,50 | 550                   | 9        | 0,31     | 21                  | 0,60     |
| 0,75 | 581                   | 19       | 0,37     | 43                  | 0,81     |
| 0,90 | 600                   | 48       | 0,79     | 72                  | 1,42     |

Rysunek 21: Wyniki dla pary obrazów przedstawiających kubek

| p    | Najlepsze rozwiązanie |          |          | Całkowite przejście |          |
|------|-----------------------|----------|----------|---------------------|----------|
|      | Wynik                 | Iteracja | Czas [s] | Iteracje            | Czas [s] |
| 0,10 | 52                    | 1        | 0,02     | 1                   | 0,02     |
| 0,50 | 340                   | 5        | 0,07     | 6                   | 0,8      |
| 0,75 | 398                   | 6        | 0,08     | 12                  | 0,14     |
| 0,90 | 403                   | 16       | 0,15     | 21                  | 0,40     |

Już dwie wybrane pary obrazów pozwalają na zaletę metody estymacji iteracji. Kosztem utracenia kilku par punktów kluczowych czas i liczba iteracji zmniejszyły się prawie stukrotnie. Z przeprowadzonych badań można jasno wywnioskować, że wartość parametru  $p$  nie powinna być mniejsza niż 0,5, ponieważ redukuje to liczbę iteracji do bardzo małych wartości, które nie pozwalają na znalezienie satysfakcyjującego wyniku. Wraz ze wzrostem wartości parametru  $p$  liczba iteracji wzrasta, jednak nawet dla wartości nieskoniecznie bliskim 1 jest ona stosunkowo mała. Metoda ta miałaby zastosowanie w systemach, gdzie wymagana jest szybka, nawet jeśli nie w pełni poprawna, odpowiedź. Niestety mała liczba iteracji prowadzi do dużej losowości otrzymywanych wyników. Rezultaty znajdujące się w powyższych tabelach zostały uzyskane poprzez uśrednienie wyników otrzymanych po pięciu uruchomieniach metody, jednak nie jest do wystarczające do właściwej oceny tego sposobu na przyśpieszenie metody RANSAC. W celu dokładnego zweryfikowania tej metody konieczne byłoby uśrednienie wyników otrzymanych z minimum kilkudziesięciu prób.

## 8 Podsumowanie

Wykrywanie obiektów na obrazach jest w pewnym stopniu skomplikowanym procesem, składającym się z wielu etapów. Po ekstrakcji punktów kluczowych charakteryzujących każdy z obrazów, należy na podstawie ich podobieństwa (cech) połączyć je w odpowiednie pary. To właśnie ten proces zajmuje najwięcej czasu, ponieważ znalezienie odpowiednich par ma bardzo dużą złożoność obliczeniową. Przypuśćmy, że obraz pierwszy posiada 6000 punktów kluczowych, natomiast obraz drugi 5000. W celu wyznaczenia par punktów kluczowych należy dla każdego punktu z obrazu pierwszego obliczyć podobieństwo do każdego z punktów na obrazie numer dwa i odwrotnie. Następnie dla każdego punktu należy sprawdzić, który z punktów z drugiego obrazu jest do niego najbardziej podobny. Ostatecznie nawet dla niewielkich obrazów musi sprawdzić 30000000 przypadków. Dla większych obrazów liczba ta może wzrosnąć nawet kilkadziesiąt czy kilkaset razy. Właśnie z tego powodu w mojej implementacji wykorzystałem moduł ‘multiprocessing’ dostępny w środowisku Python w celu prowadzenia równoległych obliczeń na wielu rdzeniach. Niezwykle przydatny okazał się też moduł ‘NumPy’ pozwalający na przeprowadzanie obliczeń na macierzach i wektorach z niewiarygodną szybkością.

Usunięcie szumów poprzez badanie spójności sąsiedztwa punktów kluczowych znacznie zmniejsza liczbę branych pod uwagę par. Poprzez wybieranie tych par, które łączą ze sobą zbliżone regiony można usunąć szумy łączące losowe regiony zdjęć. Niezwykle ważny jest jednak odpowiedni dobór parametrów. Zbyt duża liczебność sąsiedztwa lub za duży próg spójności mogą spowodować odfiltrowanie poprawnych par punktów. W przypadku małego sąsiedztwa natomiast, lub bardzo niskiego progu spójności, ryzykujemy umieszczenie dużej ilości szumu w zbiorze wynikowym. Warto wspomnieć, że liczenie odległości między punktami można łatwo przekształcić na operacje na macierzach, co po wykorzystaniu modułu ‘NumPy’ może znacznie zwiększyć wydajność programu.

W kolejnym etapie spójne pary punktów kluczowych wykorzystywane są do znalezienia odpowiedniego model transformacji obrazu, który pozwoli na pozbycie się błędów wynikających z obrotu czy różnych pozycji obiektu na obrazach. Transformaty są obliczane na podstawie losowej próbki danych ze zbioru par punktów kluczowych. W klasycznej wersji metody RANSAC próbka ta jest wybierana całkowicie losowo, jednak jak się okazało w badaniach, tworzenie modelu na podstawie określonych punktów prowadzi do lepszych wyników. Trzeba jednak pamiętać, że metoda estymacji RANSAC nie jest deterministyczna, a wyliczenie odpowiednich parametrów modelu może w niektórych wypadkach zająć dużo czasu. Warto więc ją usprawnić poprzez implementację metod pozwalających na przykładowo estymację liczby wymaganych do uzyskania odpowiedniego modelu iteracji. Podsumowując, RANSAC jest świetnym narzędziem do estymacji parametrów modeli na podstawie zbiorów zawierających dużą liczbę danych odstających, w celu uzyskania lepszych rezultatów jednak warto się także przyjrzeć jego rozszerzeniom MSAC oraz MLESAC.