

# Podobieństwo obrazów

Szymon Woźniak, 235040

16.05.2019

# 1 Wstęp teoretyczny

## 1.1 Cel zadania

Celem zadania jest określania podobieństwa obrazów poprzez poszukiwanie obiektów występujących na obu.

## 1.2 Punkty kluczowe

Punkty kluczowe opisują lokalną charakterystykę wizualną danego obrazu. Każdy punkt kluczowy opisany jest wektorem cech. Do celów zadania zostanie użyty gotowe narzędzie do ekstrakcji cech, korzystający z detektora Harris-Affine i deskryptora SIFT.

## 1.3 Wyznaczanie par punktów kluczowych

Do wyznaczania par punktów kluczowych zastosowany zostanie użyty algorytm wzajemnego najbliższego sąsiada. Dla każdego punktu z jednego z obrazów zostaje wyznaczony punkt z drugiego obrazu będący najbliższym do niego w sensie podobieństwa cech. Ta sama procedura jest przeprowadzana dla punktów z drugiego obrazu. Para punktów ( $P, Q$ ) zostaje utworzona tylko wtedy, gdy najbliższym punktem dla punktu  $P$  jest punkt  $Q$ , a najbliższym dla punktu  $Q$  jest punkt  $P$ . Funkcja odległości używana do porównywania punktów kluczowych ma postać:

$$distance(P, Q) = \sum_{i=1}^n |features(P, i) - features(Q, i)|, \quad (1)$$

gdzie:

- $n$  jest liczbą cech punktu kluczowego,
- $features(R, i)$  jest wartość  $i$ -tej cechy z punktu  $R$ .

## 1.4 Algorytm analizy spójności sąsiedztwa

Pary punktów kluczowych wyznaczanych przy pomocy algorytmu wzajemnego najbliższego sąsiada nie są wystarczające do wyznaczania podobieństwa obrazów. Jest tak ponieważ znacząca liczba znajdowanych par stanowi szum informacyjny.

Do odfiltrowania nieznaczących par można wykorzystać algorytm analizy spójności sąsiedztwa. Algorytm ten na wejściu przyjmuje zbiór par znalezionych na obrazach. Zachowywane są tylko te pary, które spełniają założone kryterium spójności. Najpierw dla każdego punktu na obu obrazach wyznaczane zostaje jego sąsiedztwo. Sąsiedztwem nazywany jest zbiór punktów o wielkości  $k$ , z tego samego obrazu, których odległość od zadanego punktu jest najmniejsza i większa od 0.

Para punktów kluczowych  $(Q_i, Q_j)$  sąsiaduje z parą punktów  $(P_i, P_j)$  jeżeli punkt  $Q_i$  należy do sąsiedztwa punktu  $P_i$ , a punkt  $Q_j$  należy do sąsiedztwa punktu  $P_j$ . Sąsiedztwo pary punktów kluczowych jest nazywane **spójnym** jeżeli, para punktów kluczowych sąsiaduje z co najmniej  $n$  innymi parą punktów kluczowych.

## 1.5 Metoda Random Sample Consensus

Metoda Random Sample Consensus ("RANSAC") jest metodą pseudolosową służącą do estymacji parametrów modelu uczenia maszynowego. Jest ona zdolna wyznaczać parametry modelu nawet dla danych o dużej liczbie danych odstających. Na początku metoda losuje z całego zbioru danych  $D$  minimalną próbę danych pozwalającą na przeprowadzenie uczenia. W kolejnym kroku przeprowadza się uczenie. Na koniec oceniana jest jakość modelu. Każda obserwacja sprawdzana jest pod kątem zgodności z modelem. Jeżeli wartość błędu nie przekracza zadanego progu, to mówi się że dana obserwacja jest zgodna z modelem. Obserwacje zgodne tworzą **konsensus**, a jego rozmiar oddaje ogólną jakość modelu. W metodzie RANSAC przeprowadzanych jest  $i$  iteracji opisanych wcześniej, a wynikiem działania jest model posiadający największy konsensus. W opisywanym jako modele użyte zostaną dwa opisane poniżej przekształcenia geometryczne.

### 1.5.1 Model - transformata afiniczna

**Transformata afiniczna** jest złożeniem trzech przekształceń elementarnych - obrót, skalowanie i translacja. Może zostać wyznaczona przy użyciu trzech par punktów kluczowych w następujący sposób:

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad (2)$$

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

### 1.5.2 Model - transformata perspektywiczna

Transformata perspektywiczna rozszerza transformatę afiniczną o efekt rozcięcia/gniesienia/spłaszczenia wynikający z perspektywy. Można ją wyznaczyć przy po-

mocy 4 par punktów kluczowych według poniższych równań:

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -u_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -u_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4x_4 & -u_4y_4 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1x_1 & -v_1y_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2x_2 & -v_2y_2 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3x_3 & -v_3y_3 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4x_4 & -v_4y_4 \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \quad (4)$$

$$\mathbf{H} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \quad (5)$$

## 1.6 Heurystyki przyspieszające metodę RANSAC

### 1.6.1 Heurystyki odległości par punktów

Ta heurystyka mówi, że jako próbka do algorytmu RANSAC powinny być wybierane punkty leżące stosunkowo blisko siebie. Ograniczeniem może być na przykład to, że odległość pomiędzy każdą parą wybranych punktów powinna być większa niż 1% i mniejsza niż 30% wielkości obrazu.

### 1.6.2 Heurystyka modyfikacji rozkładu

Ta heurystyka mówi, że rozkład prawdopodobieństwa wyboru par punktów kluczowych powinien być modyfikowany w trakcie działania metody. Pary punktów, które uzyskiwały większe rozmiary konsensusu powinny być losowane z większym prawdopodobieństwem.

### 1.6.3 Heurystyka liczby iteracji

W podstawowej wersji metody RANSAC liczba iteracji jest stała i musi zostać podana przez użytkownika. Jeżeli jednak podana liczba iteracji będzie zbyt duża w stosunku do trudności problemu, to od pewnego momentu metoda może nie generować już lepszych modeli. Liczbę iteracji można jednak oszacować zgodnie ze wzorem:

$$k = \frac{\log(1-p)}{\log(1-w^n)} \quad (6)$$

, gdzie:

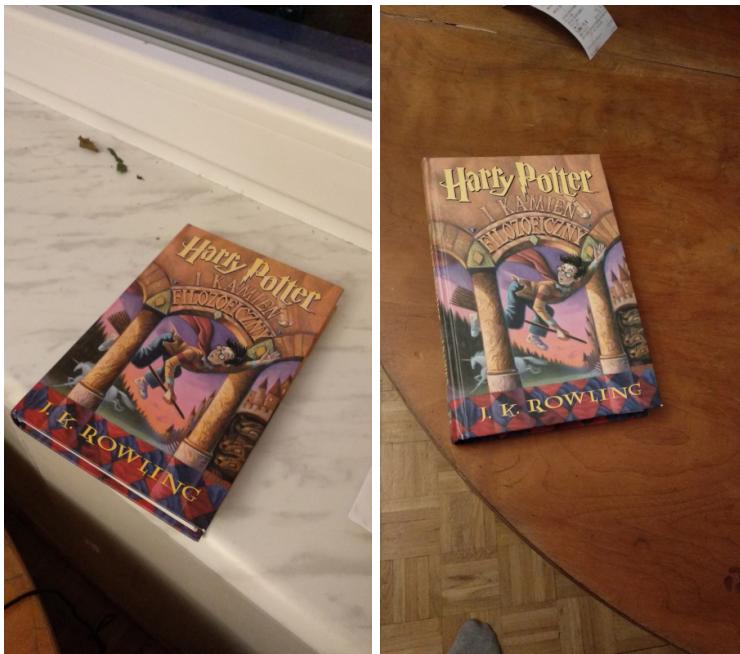
- $p$  - prawdopodobieństwo, że po oszacowanej liczbie iteracji model jest dostatecznie dobrym przybliżeniem danych,
- $w$  - prawdopodobieństwo, że wylosowana para nie jest szumem,
- $n$  - liczba par potrzebnych do wyznaczenia parametrów transformaty (3 lub 4),

- $k$  - szacowana liczba iteracji.

Prawdopodobieństwo  $p$  musi zostać zadane przez użytkownika. Prawdopodobieństwo  $w$  może zostać oszacowane jako stosunek liczby par przed i po zastosowaniu algorytmu analizy spójności sąsiedztwa.

## 2 Badania

### 2.1 Wykorzystane pary obrazów



Rysunek 1: Książka "Harry Potter i Kamień Filozoficzny" w dwóch różnych ujęciach.



Rysunek 2: Książka Harry Potter i Przeklęte Dziecko w dwóch różnych ujęciach.



Rysunek 3: Książka "Brudne Gry" w dwóch różnych ujęciach.



Rysunek 4: Kubek w dwóch różnych ujęciach.



Rysunek 5: Ossolineum w dwóch różnych ujęciach.

## 2.2 Skuteczność algorytmu analizy spójności sąsiedztwa

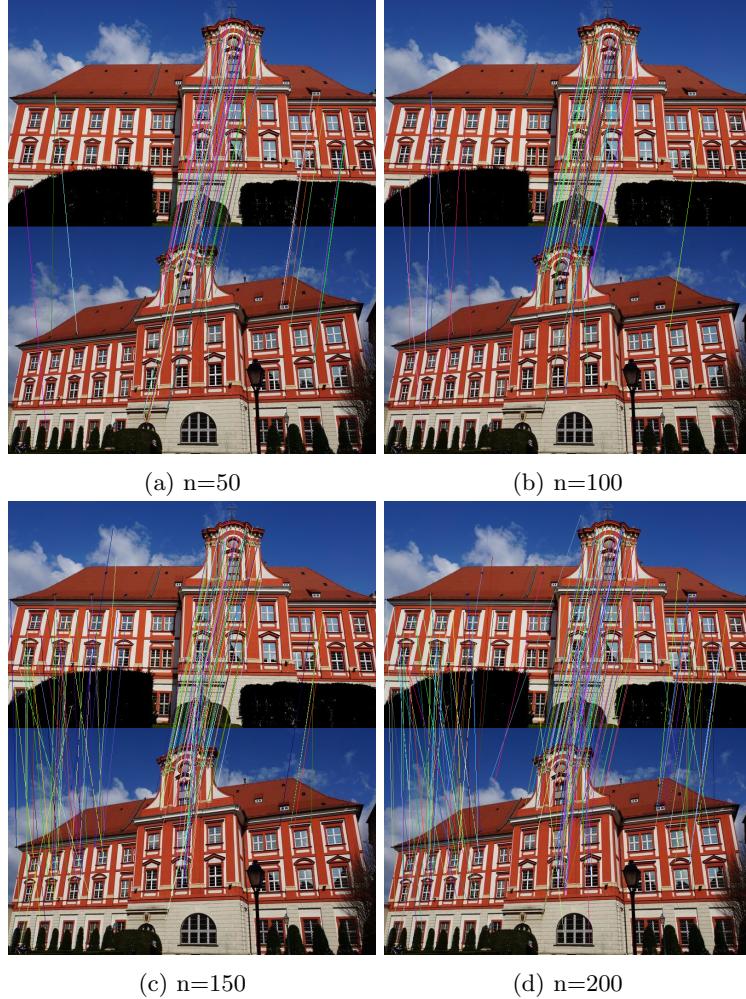
Wyniki badań skuteczność algorytmu spójności sąsiedztwa zostaną przedstawione jako porównanie dwóch par obrazów z naniesionymi parami punktów kluczowych. Jedna para będzie reprezentować wszystkie znalezione pary, a druga pary po zastosowaniu algorytmu analizy spójności sąsiedztwa. Do przeprowadzenia tych badań wykorzystane zostaną zdjęcia z rysunku 5.



Rysunek 6: Wszystkie pary punktów kluczowych znalezione na obrazach Osolineum

### 2.2.1 Wpływ wielkości sąsiedztwa na skuteczność algorytmu spójności

Wielkość sąsiedztwa została zbadana przy wartości progu spójności ustawionej na 0.5.



Rysunek 7: Pary punktów kluczowych odfiltrowane za pomocą algorytmu spójności sąsiedztwa dla różnych wartości wielkości sąsiedztwa.

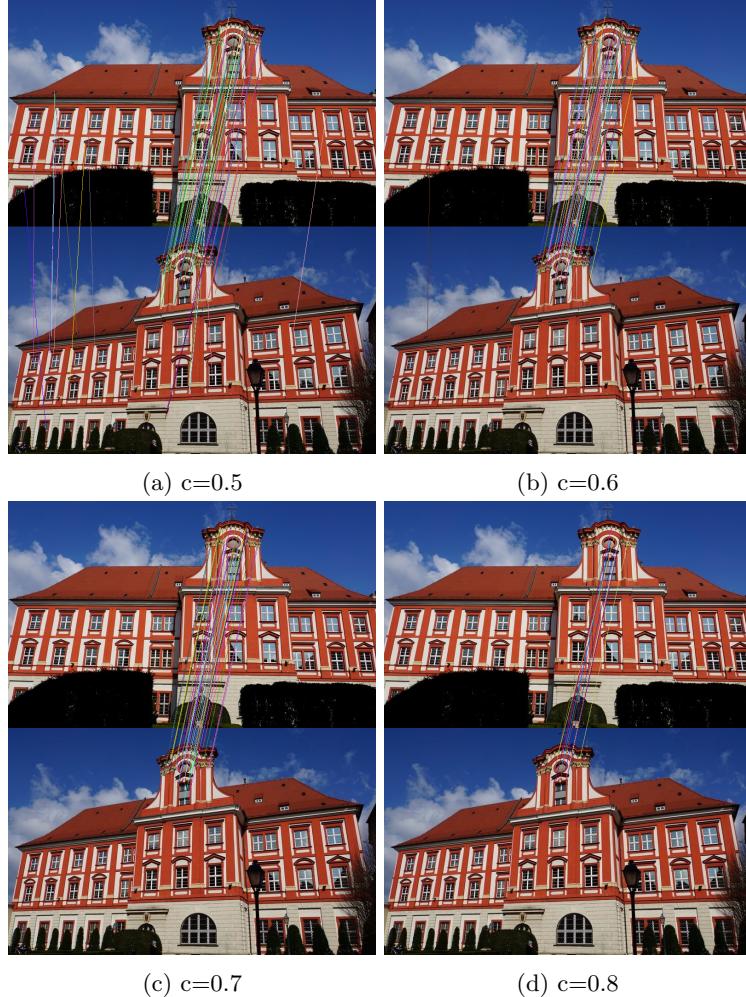
Tabela 1: Porównanie liczby zaakceptowanych par punktów kluczowych i czasu przetwarzania w zależności od wielkości sąsiedztwa

Wielkość sąsiedztwa	Liczba par spójnych	Czas[ms]
50	97	180
100	88	190
150	141	200
200	188	215

**Wnioski** Ustalenie dobrej wartości parametru wielkości sąsiedztwa jest istotnym wymogiem poprawnego działania algorytmu analizy spójności sąsiedztwa. Przy dobieraniu wartości tego parametru można kierować się gęstością punktów kluczowych na analizowanych obrazach. Im jest ona większa tym większa powinna być ustalona wielkość sąsiedztwa. Jak widać na rysunku 7 przy zbyt małych wartościach tego parametru algorytm może pominać istotne punkty. Z kolei jeżeli jest on za duży, to mogą zostać znalezione pary będące szumem. Dla analizowanego obrazu wartości pomiędzy 100 a 150 wydają się być odpowiednie. Jak widać w tabeli 1, wpływ ustalonej wartości tego parametru na czas działania algorytmu istnieje, natomiast jest stosunkowo mały i prawdopodobnie może zostać pominięty.

### 2.2.2 Wpływ progu spójności na skuteczność algorytmu

Wpływ progu spójności sąsiedztwa został zbadany przy wartości wielkości sąsiedztwa ustawionej na  $n=100$ .



Rysunek 8: Pary punktów kluczowych odfiltrowane za pomocą algorytmu spójności sąsiedztwa dla różnych wartości progu spójności.

Tabela 2: Porównanie liczby zaakceptowanych par punktów kluczowych w zależności od wartości progu spójności sąsiedztwa

Próg spójności	Liczba par spójnych
0.5	88
0.6	62
0.7	54
0.8	16
0.9	0

**Wnioski** Ustalenie dobrej wartości parametru wielkości sąsiedztwa jest kolejną istotną kwestią przy dobieraniu parametrów do algorytmu analizy spójności sąsiedztwa. Wartość tego parametru jest powiązana z wielkością sąsiedztwa więc wartości dla obu parametry powinny być ustalane równocześnie. Jak widać na rysunku 8 zbyt małe wartości tego parametru mogą skutkować akceptowaniem par stanowiących szum. Z kolei zbyt duże wartości mogą powodować zgubienie istotnej, pod kątem analizy podobieństwa, informacji.

Wpływ wielkości tego parametru na czas działania algorytmu okazał się nieznaczący.

## 2.3 Skuteczność działania metody RANSAC

### 2.3.1 Wpływ rodzaju transformaty

Badania w tej sekcji zostaną przedstawione jako zestawienia wyników działania metody RANSAC na dwóch różnych obrazach przy wykorzystaniu transformaty afinicznej i transformaty perspektywicznej.

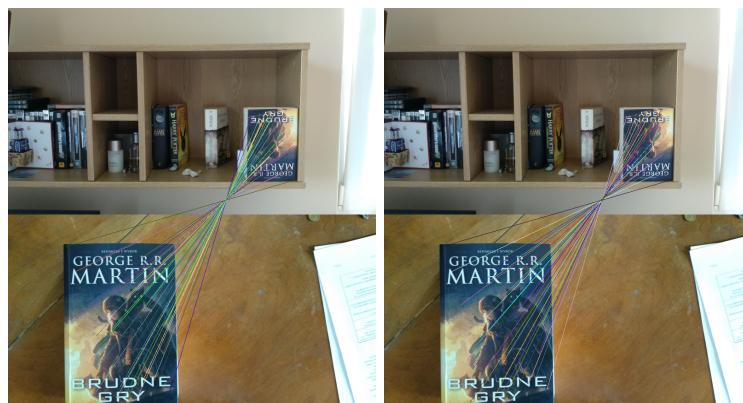
Wszystkie badania w tej sekcji zostały przeprowadzone bez użycia heurystyk, z wartościami parametrów:

- liczba iteracji = 10000,
- próg akceptacji = 1.

**Para obrazów bez zmiany perspektywy** Na tej parze obrazów, obiekt jest w obu przypadkach widoczny od frontu. Liczba wszystkich par punktów kluczowych na obrazie wynosi 371.



Rysunek 9: Wszystkie pary punktów kluczowych znalezione na zdjęciach książki "Brudne gry".



(a) Transformata afiniczna

(b) Transformata perspektywiczna

Rysunek 10: Pary punktów kluczowych znalezione przez metodę RANSAC na zdjęciach książki "Brudne gry" przy użyciu obu przekształceń geometrycznych.



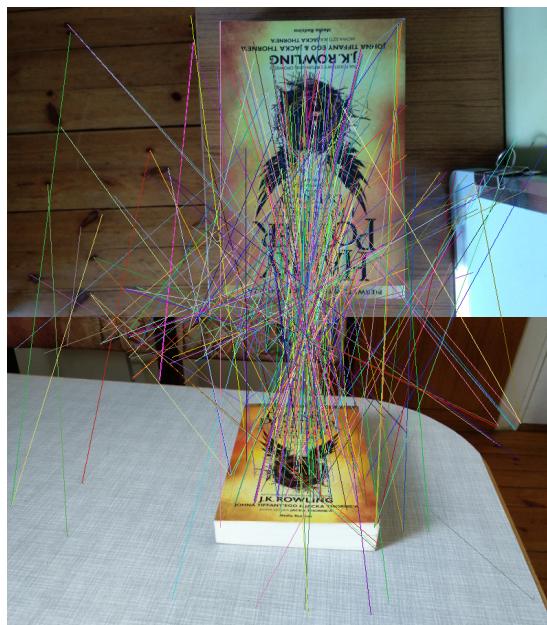
(a) Transformata afinicza



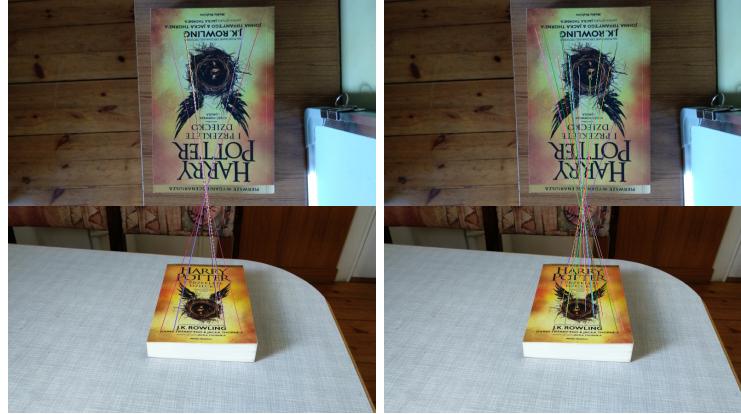
(b) Transformata perspektywiczna

Rysunek 11: Najlepsze znalezione przekształcenia geometryczne zaaplikowane do całego pierwszego zdjęcia.

**Para obrazów ze zmianą perspektywy** Na tej parze obrazów, kąt patrzenia na obiekt się zmienia. Liczba par punktów kluczowych na obrazach wynosi 328.



Rysunek 12: Wszystkie pary punktów kluczowych znalezione na zdjęciach książki "Harry Potter i przeklęte dziecko".



(a) Transformata afinicza      (b) Transformata perspektywiczna

Rysunek 13: Pary punktów kluczowych znalezione na zdjęciach książki "Harry Potter i przeklęte dziecko" przez metodę RANSAC przy użyciu obu przekształceń geometrycznych.



(a) Transformata afinicza      (b) Transformata perspektywiczna

Rysunek 14: Najlepsze znalezione przekształcenia geometryczne zaaplikowane do całego pierwszego zdjęcia.

Tabela 3: Liczba par punktów kluczowych znalezionych przez metodę RANSAC na obrazach z perspektywą i bez perspektywy (średnia z 10 przebiegów)

	Transformata afinicza	Transformata perspektywiczna
Bez perspektywy	63	60
Z perspektywą	17	35

**Wnioski** Jak widać na rysunku 10 i w tabeli 3 w przypadku pierwszego obrazu, obie transformacje poradziły sobie podobnie. Widać również jednak, że rozwiązanie z transformacją perspektywiczną znajduje średnio mniej par, co może być związane z większą przestrzenią poszukiwań i stałą liczbą iteracji. Na rysunkach 13 14 widać jednak, że w momencie kiedy obiekt na jednym z obrazów znajduje się pod innym kątem, transformacja perspektywiczna radzi

sobie zdecydowanie lepiej od afanicznej. Jest to spowodowane ograniczeniami wynikającymi wprost z modelu - transformacja afaniczna nie jest w stanie reprezentować perspektywy.

### **2.3.2 Wpływ parametrów algorytmu**

### **2.3.3 Wpływ heurystyk metody RANSAC**