

Sztuczna inteligencja i inżynieria wiedzy

Podobieństwo obrazów

1) Wstęp teoretyczny

Algorytmy służące do analizy zdjęć pełnią kluczową rolę w dzisiejszym świecie. 18 czerwca 2013 NASA ogłosiło Asteroid Grand Challenge, mający za zadanie pomóc usprawnić algorytm do wyszukiwania na zdjęciach asteroid, które mogą stanowić potencjalne zagrożenie. Każdego dnia na całym świecie lekarze wykorzystują podobne algorytmy w celu np. identyfikacji komórek rakowych. Niemal każdy z nas posiada telefon zdolny do wykrycia, kiedy fotografowana osoba się uśmiecha. Istnieje wiele zadań, jakie dany algorytm może spełniać - od wykrywania obecności jakiegoś obiektu, szukania powtarzających się wzorców, klasteryzacji aż do określania różnic między zdjęciami. Tak samo jak różne są cele, różne są też sposoby ich osiągania. Jedno z prostszych zadań to określanie podobieństwa między obrazami. Dokonano implementacji takiego rozwiązania z użyciem algorytmu RANSAC, mającego za zadanie znaleźć przekształcenie geometryczne między zdjęciami.

2) Podstawy teoretyczne

Pierwszym krokiem w celu analizy zdjęcia jest często wyszukanie charakterystycznych dlań punktów kluczowych. Istnieje wiele algorytmów mających za zadanie odszukanie pikseli / grup pikseli, które wyróżniają zdjęcie. Jednym z nich jest algorytm o nazwie Harris affine region detector. Polega on na pomiarze cech charakterystycznych dla kolejnych fragmentów obrazu. Dla każdego obszaru mierzona jest korelacja między pikselami, wyrażona za pomocą następującej macierzy.

$$A(x) = \sum_{p,q} w(p,q) \begin{bmatrix} I_x^2(x) & I_x I_y(x) \\ I_x I_y(x) & I_y^2(x) \end{bmatrix}$$

Gdy λ_1 λ_2 , wyprowadzone za pomocą wzoru:

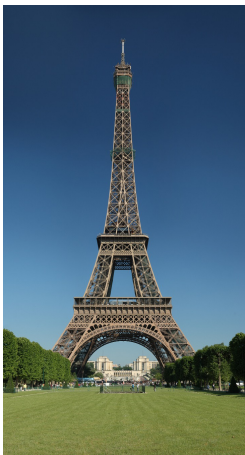
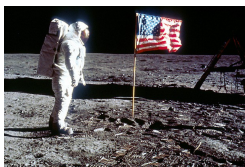


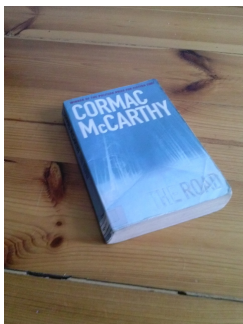
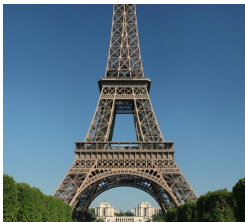
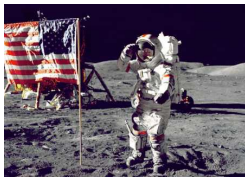



$$R = \det(A) - \alpha \operatorname{trace}^2(A) = \lambda_1 \lambda_2 - \alpha (\lambda_1 + \lambda_2)^2$$

przyjmując duże wartości pozytywne dany fragment obrazu zawiera punkt kluczowy.

Dla każdego opisanego punktu należy wybrać interesujące cechy (opisać go). Do tego można zastosować algorytm SIFT - Scale-invariant feature transform. Każdy punkt jest opisywany jako wektor cech niezależnych od jego położenia - w ten sposób ten sam obiekt opisany podobnymi punktami może zostać zidentyfikowany niezależnie od zdjęcia. Należy wspomnieć, że algorytm SIFT ma w gruncie rzeczy dużo szersze zastosowanie.

3) Realizacja ćwiczenia i wyniki pomiarów

Użyte obrazy:

	 moon1 4227 punktów kluczowych	 sintel1 1240 punktów kluczowych	 CAM00257 13299 punktów kluczowych	 TheRoad1 3447 punktów kluczowych
Eiffel_1 5915 punktów kluczowych				
	 moon2 2803 punktów kluczowych	 sintel2 747 punktów kluczowych	 CAM00258 6245 punktów kluczowych	 TheRoad2 3233 punktów kluczowych
Eiffel_2 8454 punktów kluczowych				

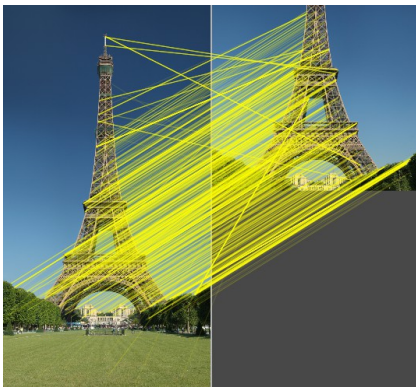
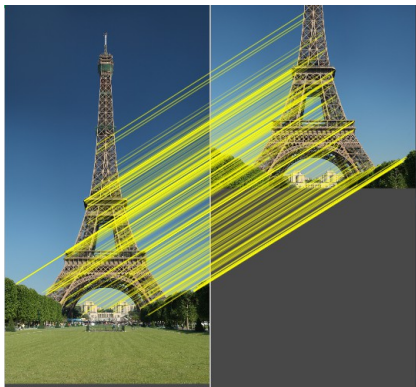
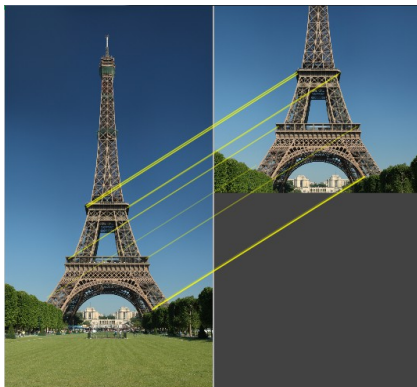
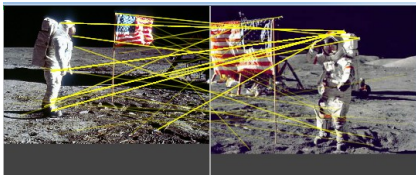
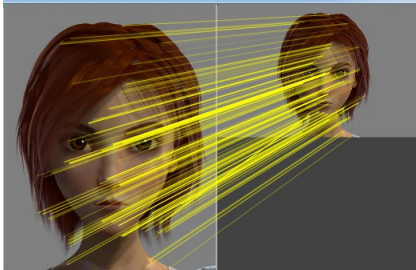
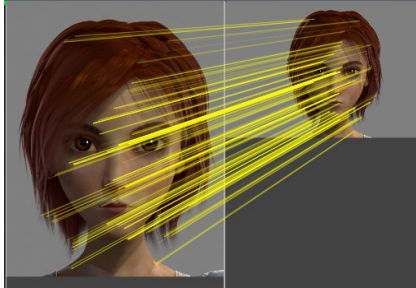

Obrazy zastosowane do analizy algorytmu

Pierwszym krokiem było wczytanie danych punktów kluczowych porównywanych obrazów a następnie wyznaczenie par odpowiadających sobie punktów kluczowych. Warto zaznaczyć, że jest to zdecydowanie najbardziej skomplikowana pod względem obliczeniowym część ćwiczenia. Dla każdego punktu z obrazu 1 należy wyznaczyć punkt najbardziej doń podobny pod względem odczytanych cech z obrazu 2. Dobór par dla dwóch obrazów po 3,300 punktów kluczowych każdy trwa ~7 sekund.

Pierwszy algorytm do zaimplementowania to algorytm do analizy spójności sąsiedztwa pary punktów kluczowych. Liczbę par z poprzedniego etapu można ograniczyć gdy zauważymy, że punkty, które na pierwszym obrazie były blisko siebie po przekształceniu na odpowiadające punkty z pary także powinny znaleźć się w niewielkiej odległości.

l.p.	Zdjęcie 1	Zdjęcie 2	Ilość par				
			Algorytm doboru par	N: 20 Minimalna zgodność: 10%	N: 20 Minimalna zgodność: 40%	N: 40 Minimalna zgodność: 10%	N: 40 Minimalna zgodność: 40%
1	Eiffel_1	Eiffel_2	2133	1216	76	1174	15
2	moon1	moon2	583	0	0	0	0
3	sintell	sintell2	505	423	73	462	12
4	CAM00257	CAM00258	3819	377	0	203	0
5	TheRoad1	TheRoad2	1235	128	0	55	0

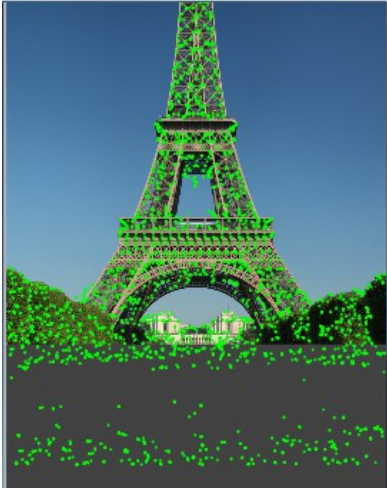
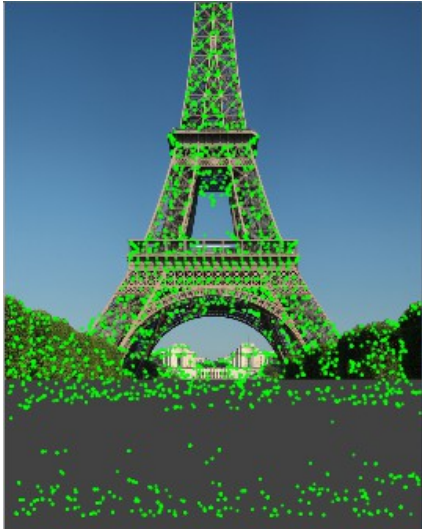
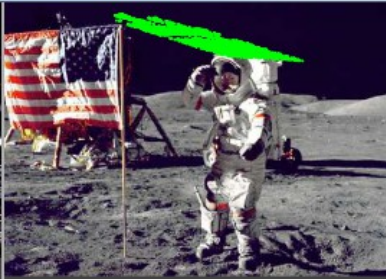
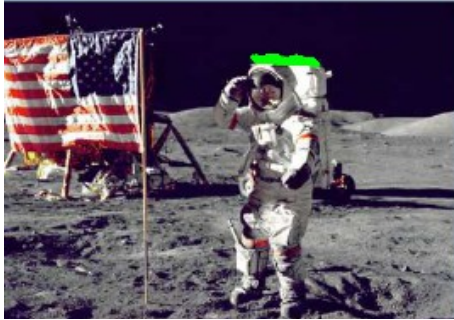
Wyniki algorytmu redukcji par


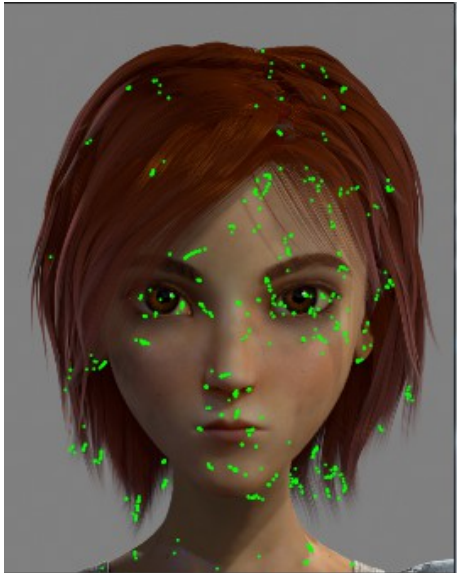
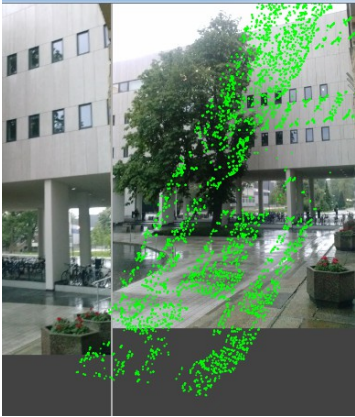
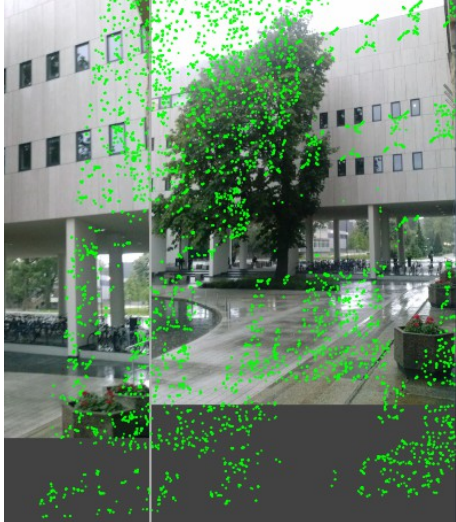


Algorytm doboru par	N: 20 Min percentage: 10%	N: 20 Min percentage: 40%
		
	Brak znalezionych par	Brak znalezionych par
		
Błąd rysunku	Błąd rysunku	Brak znalezionych par

		<p>Brak znalezionych par</p>
---	--	------------------------------

Przykładowe wyniki uzyskane za pomocą algorytmu

Algorytm RANSAC (RANDOM Sample Consensus) ma za zadanie znaleźć model, taki, że będzie on dobrze dopasowywał się do zaobserwowanych danych. Algorytm jest niedeterministyczny. Wraz ze wzrostem iteracji rośnie prawdopodobieństwo, że błąd dopasowania modelu zmaleje. W naszym przypadku zastosowaliśmy ten algorytm do znalezienia przekształcenia z jednego obrazu na drugi. Przykładem może być książka sfotografowana pod innymi kątami. Punkty kluczowe znajdujące się np. na obrazku na okładce powinny umożliwić znalezienie transformacji. Błąd przekształcenia możemy znaleźć porównując położenie punktów po przekształceniu z ich właściwą pozycją.

l.p.	Zdjęcie 1	Zdjęcie 2	Maksymalny błąd	
			5%	10%
1	Eiffel_1	Eiffel_2		
2	moon1	moon2		

3	sintel1	sintel2		
4	CAM00257	CAM00258		
5	TheRoad1	TheRoad2		

Wyniki uzyskane z użyciem algorytmu RANSAC

4) Czas wykonywania algorytmu

l.p.	Zdjęcie 1	Zdjęcie 2	Czas [s]				
			Algorytm doboru par	Redukcja par		RANSAC	
				N=20	N=40	Ilość iteracji = 5000	Ilość iteracji = 15000
1	Eiffel_1	Eiffel_2	12.8	22.2	38.0	4.0	12.0
2	moon1	moon2	3.5	0.7	0.7	0.9	2.6
3	sintel1	sintel2	0.8	0.9	1.8	0.9	2.2
4	CAM00257	CAM00258	36.6	20.5	23.6	6.5	19.5
5	TheRoad1	TheRoad2	6.5	2.5	3.6	2.2	6.0

Czas wykonywania każdego z algorytmów dla każdej z par

Możemy zauważyć, że czas w algorytmach redukcji par oraz RANSAC rośnie liniowo w stosunku do odpowiednio: ilości najbliższych branych pod uwagę punktów (parametr N, liczności sąsiedztwa) oraz ilości iteracji.

5) Wnioski

Wpływ parametrów na efektywność działania metody zgodności sąsiedztwa:

- parametr liczności sąsiedztwa

Najważniejszym skutkiem zwiększenia wartości tego parametru było wydłużenie czasu obliczeń. Zaobserwowano niemal liniową korelację. Dodatkowym skutkiem było zmniejszenie liczby znalezionych par. Jest to oczekiwane – gdy na zdjęciach brak oczywistej korelacji (np. prostego przesunięcia) szanse, że dopasowana para będzie miała podobną translację jest niewielka. Para obrazów TheRoad1 i TheRoad2 idealnie pokazuje jak nawet niewielka zmiana perspektywy może drastycznie wpłynąć na znalezione podobieństwo.

- parametr spójności sąsiedztwa

Zwiększenie parametru miało dosyć intuicyjny skutek – mniej par zostało znalezionych, ale te, które zostały wybrane są znacznie wyraźniejsze. Gdy porównamy obrazki Eiffel_1 oraz Eiffel_2 dla wartości parametru 10% i 40% nie trudno zauważyć, że linie uzyskane przy wartości parametru 40% pokrywają się z zagęszczeniami na obrazku 20%.

Należy zauważyć, że grupa obrazów, która miała dużą ilość par (ze względu na dużą ilość punktów kluczowych) po zastosowaniu algorytmu uzyskała dużo mniejszy procent par, które przeszły test sąsiedztwa.

Wpływ parametrów na efektywność działania metody RANSAC:

- parametr progu akceptacji danych

Nie zaobserwowano szczególnych różnic między obrazami wynikowymi dla różnych wartości parametru. Pary Eiffel_1 oraz Eiffel_2 oraz sintel1 i sintel2 uzyskały znakomite rezultaty niezależnie od wartości progu.

- ilość iteracji

Większa ilość iteracji powoduje niemal liniowy wzrost czas obliczeń i daje większe szanse na uzyskanie akceptowalnych rezultatów. Jednak nawet 50000 iteracji nie gwarantuje, że uzyskane wyniki będą w jakikolwiek sposób użyteczne.

Jak już wspomniano, pary Eiffel_1 oraz Eiffel_2 oraz sintel1 i sintel2 uzyskały idealne wyniki. Możliwa jest nawet rekonstrukcja niewidocznego fragmentu obrazu na podstawie punktów wynikowych. Jest to spowodowane bardzo dużym podobieństwem oby obrazów. Rozczarowuje wynik pary TheRoad1 i TheRoad2 – uzyskanie jakiegokolwiek podobieństwa wiązało się z parametrem szczęścia osoby wywołującej algorytm. Zaprezentowane wyniki dla tej pary wymagały ~60000 iteracji. Możemy stwierdzić, że im mniejsza liczba punktów kluczowych tym większe prawdopodobieństwo uzyskania dobrego wyniku.