

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Звіт

про виконання лабораторної роботи №2
з дисципліни “Розпізнавання образів”

Виконали:
Студенти 4 курсу
Групи КА-76 і КА-72
Борбела Артур
Фалілеєва Дар’я

Київ 2020

Теоретичні данні:

ORB представлений в 2011г. В його основі лежить комбінація таких алгоритмів як детектор FAST (Features from Accelerated Segment Test) і дескриптор BRIEF (Binary Robust Independent Elementary Features) з некоторими удосконаленнями. Зображення розбивається на патчі. Нехай патч P має розміри $S \times S$ пікселів. З патча вибирається некоторим образом множество пар пікселів $\{(X, Y), \forall X, Y \text{ в околиці}\}$, для яких строится набір бінарних тестів:

$$\tau(P, X, Y) = \begin{cases} 1, I(X) < I(Y) \\ 0, \text{иначе} \end{cases}$$

де $I(Y)$, $I(X)$ – інтенсивність пікселів Y , X відповідно. Для кожного патча вибирається множество, що містить n_d пар точок, які однозначно визначають набір бінарних тестів. Далі на основі цих тестів строится бінарна строка:

$$f_{n_d}(P) = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(P, X_i, Y_i)$$

SIFT (Scale Invariant Feature Transform). Будується простір змінного масштабу, в ньому обчислюються функції LoG з різним параметром згладжування. Точка вважається ключовою, якщо вона є локальним екстремумом різниці Гауссіанів. Після безліч передбачуваних ключових точок уточнюється (видаляються точки з малим контрастом і на кордоні об'єктів) і обчислюються їх орієнтації. Для цього будується зважена гістограма градієнтів в околиці, вибирається напрямок, що відповідає максимальній компоненті гістограми. Точки присвоюються всім напрямкам, яким відповідають значення компонент гістограми, великих заданого порогу. Інваріантний щодо зрушень, обертань, масштабу, що не зміщується локальні екстремуми

$G(x, y, \sigma)$ – гауссове ядро зі степінню розмиття σ

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

Для формування дескриптора SIFT (Scale Invariant Feature Transform) спочатку обчислюються значення магнітуди і орієнтації градієнта в кожному пікселі, що належить околиці особливої точки розміром $16 * 16$ пікселів. Магнітуди градієнтів при цьому враховуються з вагами, пропорційними значенню функції щільності нормального розподілу з математичним очікуванням в розглянутій особливій точці і стандартним відхиленням, рівним половині ширини околиці (ваги Гауссова розподілу використовуються для того, щоб зменшити вплив на підсумковий дескриптор градієнтів, обчислених в пікселях, які перебувають далі від особливої точки).

У кожному квадраті розміром $4 * 4$ пікселя обчислюється гістограма орієнтованих градієнтів шляхом додавання зваженого значення магнітуди градієнта до одного з 8 бінов гістограми. Щоб зменшити різні "граничні" ефекти, пов'язані з віднесенням схожих градієнтів до різних квадратів (що може виникнути внаслідок невеликого зсуву розташування особливої точки) використовується билинейна інтерполяція: значення магнітуди кожного градієнта додається не тільки в гістограму, відповідну квадрату, до якого даний піксель відноситься, але і до гістограми, відповідним сусіднім квадратах. При цьому значення магнітуди додається з вагою, пропорційним відстані від пікселя, в якому обчислений даний градієнт, до центру відповідного квадрата. Всі обчислені гістограми об'єднуються в один вектор, розміром, рівним 128

Отриманий дескриптор перетворюється, щоб зменшити можливі ефекти від зміни освітленості. Зміна контрасту зображення (значення інтенсивності кожного пікселя множиться на деяку константу) призводить до такого ж зміни в значеннях магнітуд градієнтів. Тому очевидно, що даний ефект може бути знівельовано шляхом нормалізації дескриптора таким чином, щоб його довжина стала дорівнює одиниці. Зміни яскравості зображення (до значення інтенсивності кожного пікселя додається деяка константа) не впливають на значення магнітуд градієнтів. Таким чином, SIFT-дескриптор є інваріантні по відношенню до афінних змін освітленості. Однак можуть виникати і нелінійні

зміни в освітленості внаслідок, наприклад, різної орієнтації джерела світла по відношенню до поверхонь тривимірного об'єкту. Дані ефекти можуть викликати велика зміна в ставленні магнітуд деяких градієнтів (при цьому мають незначний вплив на орієнтацію вектора градієнта). Щоб уникнути цього, використовують відсікання по деякому порогу (за результатами експериментів показано, що оптимальним є значення 0.2), яке застосовують до компонентів нормалізованого дескриптора. Після застосування порога дескриптор знову нормалізується. Таким чином, зменшується значення великих магнітуд градієнтів і збільшується значення розподілу орієнтацій даних градієнтів в околиці особливої точки.

Отримані дані :

Датасет 1

src/result/ORB/ result_correct src/result/ORB/ result_wrong

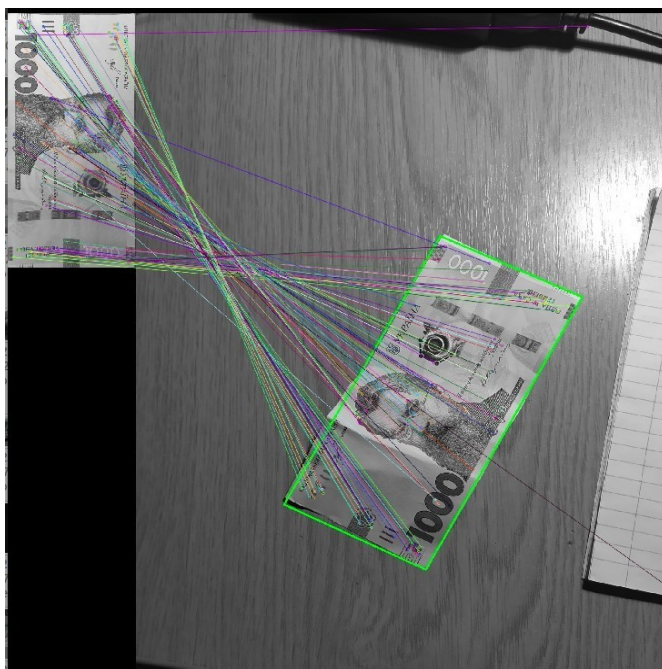
src/result/SIFT/ result_correct src/result/SIFT/ result_wrong

Датасет 2

src/result1/ORB/result_correct src/result1/ORB/result_wrong

src/result1/SIFT/result_correct src/result1/SIFT/result_wrong

Зразок



Пояснення, що до назв директорій та зображень :

Якщо директрія має шлях ./result/ORB — це результати роботи алгоритму ORB, аналогічно й з алгоритмом SIFT

Якщо в шляхові є

...../correct_data - це зображення на яких справді є предмет розпізнавання;

...../wrong_data — зображення на якому предмета немає;

Назва зображень в результаті роботи може бути :

3_jopa.jpg

12_undefined_jopa.jpg

15_good.jpg

resize32 -//- .jpg

Пояснення:

....._jopa - не пройшла перевірку, тобто в моєму випадку не розпізнано більше десяти характерних точок із-за цього не намагаємось визначити контури об'єкту.

....._good — характерних точок більше десяти, малюємо контур

об'єкта._undefined_jopa - все погано (не був згенерований дескриптор або некоректно знайдені matches що взаємо пов'язане)

Якщо починається на resize... зображення було масштабоване.

Дані метрики порівняння:

Зразок:

/result_correct/1

	Time	Good_points/matches	Avarage_distance
ORB	0.2508082389831543	0.004	349.496303527832
Sift	3.508439302444458	0.09834469328140215	233.6188332242882

Пояснення:

result_correct/1_*.jpg — назва фото

Time — час роботи алгоритму.

Good_points/matches — відношення good_points (відсіювання по distance) до всіх інших.

Avarage_distance — середня distance по всім точкам.

Датасет 1

src/compare_result/compare_correct_data.txt
src/compare_result/compare_wrong_data.txt

(correct_data - там де є зображений предмет)
(wrong_data предмета немає)

Датасет 2

src/compare_result1/compare_correct_data.txt
src/compare_result1/compare_wrong_data.txt

Висновки :

- 1) Зменшений розмір зображення позитивно впливає на швидкість відпрацювання алгоритму (аж в 3-рази) й майже не як не впливає на два інші показника(незначні зміни).

/result_correct/1

	Time	Good_points/matches	Avarage_distance
ORB	0.2508082389831543	0.004	349.496303527832
Sift	3.508439302444458	0.09834469328140215	233.6188332242882

/result_correct/resize1

	Time	Good_points/matches	Avarage_distance
ORB	0.07330846786499023	0.002	342.12963946533205
Sift	0.7891435623168945	0.11684518013631938	252.1168591388825

/result_correct/2

	Time	Good_points/matches	Avarage_distance
ORB	0.2384777069091797	0.0	345.4539998168945
Sift	2.7140140533447266	0.09834469328140215	255.73354266054326

/result_correct/resize2

	Time	Good_points/matches	Avarage_distance
ORB	0.07433652877807617	0.006	338.66166110229494
Sift	0.7490072250366211	0.10516066212268745	259.34689775297915

/result_correct/3

	Time	Good_points/matches	Avarage_distance
ORB	0.2540555000305176	0.0	389.83626132202147
Sift	2.998670816421509	0.0019474196689386564	268.9107511904827

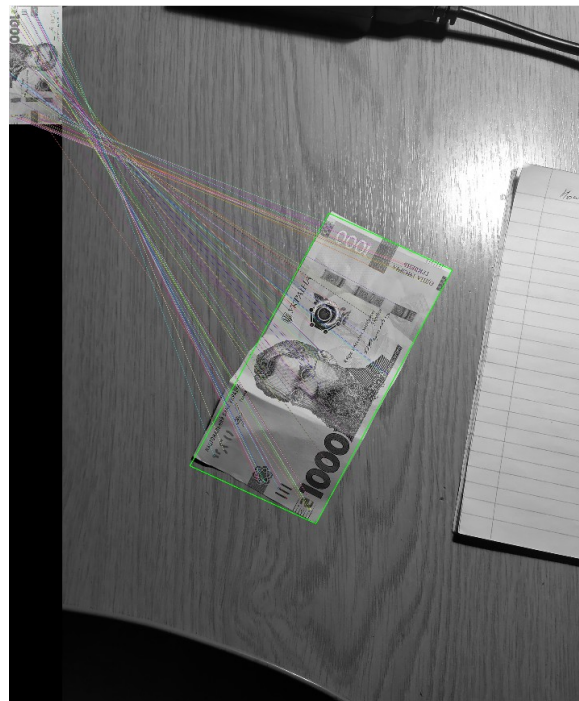
```
/result_correct/resize3
      Time      Good_points/matches  Avarage_distance
ORB  0.08049464225769043 0.006 372.15639205932615
Sift 0.8748307228088379 0.0038948393378773127 278.3393977832887
```

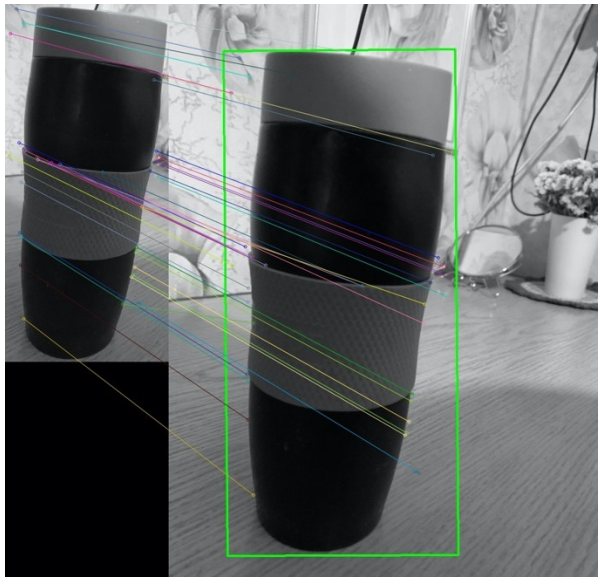
```
/result_correct/4
      Time      Good_points/matches  Avarage_distance
ORB  0.23672986030578613 0.006 359.762241394043
Sift 2.7215631008148193 0.007789678675754625 292.1632188167405
```

```
/result_correct/resize4
      Time      Good_points/matches  Avarage_distance
ORB  0.07669544219970703 0.01 359.8940180358887
Sift 0.7554645538330078 0.011684518013631937 308.05245503664247
```

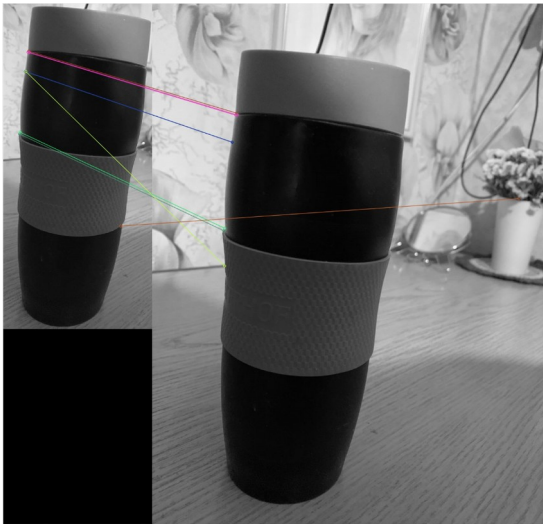
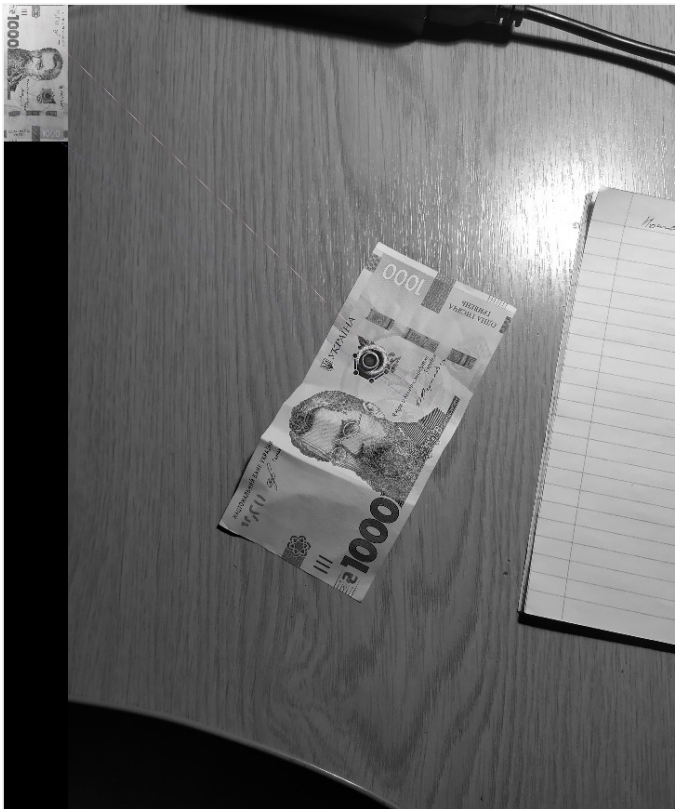
2) ORB працює швидше, але фатально програє в точності SIFT для двох Датасетів.

SIFT





ORB



1) предмет зображений під певним кутом

Датасет 1

/result_correct/17

	Time	Good_points/matches	Avarage_distance
ORB	0.24353408813476562	0.006	349.69171398925783
Sift	2.8482844829559326	0.06523855890944498	252.7604462160224

/result_correct/28

ORB	0.2415759563446045	0.004	356.27570602416995
Sift	2.7298333644866943	0.004868549172346641	283.9256147776514

/result_correct/12

ORB	0.24527549743652344	0.004	357.1088797607422
Sift	2.728064775466919	0.0019474196689386564	283.79237720820083

/result_correct/57

ORB	0.23975253105163574	0.004	349.61754217529295
Sift	2.65129017829895	0.007789678675754625	293.2202192193229

/result_correct/65

ORB	0.24764060974121094	0.006	351.9070592041016
Sift	2.780367374420166	0.008763388510223954	268.1016812422027

/result_correct/86

ORB	0.24127888679504395	0.008	343.84426052856446
Sift	2.859158992767334	0.04771178188899708	254.89923993847037

/result_correct/88

ORB	0.24094510078430176	0.008	354.88300802612304
Sift	2.602071762084961	0.006815968841285297	300.9023860645387

.....

Датасет 2

/result_correct/1

	Time	Good_points/matches	Avarage_distance
ORB	0.03208017349243164	0.04	78.1581004374736
Sift	0.30515027046203613	0.050955414012738856	241.4944540619091

/result_correct/2

ORB 0.03180408477783203 0.0 393.8182727710621
Sift 0.30513501167297363 0.006369426751592357 223.20531594829194

/result_correct/4

ORB 0.032943010330200195 0.013513513513513514 360.49138723837365
Sift 0.33025527000427246 0.03184713375796178 177.5539457965049

/result_correct/5

ORB 0.0330502986907959 0.0 363.8315786413244
Sift 0.31815028190612793 0.01910828025477707 180.33037380048424

/result_correct/10

ORB 0.0322115421295166 0.0 357.4829930485906
Sift 0.30649280548095703 0.012738853503184714 212.70452098481974

.....

(Зображення додавати не став тяжко)

ORB працює швидше але майже нічого не розпізнає. А SIFT саме краще підходить до цього.

2) зображено частина предмета

Датасет 1

/result_correct/30

Time	Good_points/matches	Avarage_distance
------	---------------------	------------------

ORB 0.2540926933288574	0.004	357.01111386108397
Sift 3.255985736846924	0.004868549172346641	264.2544801557888

/result_correct/31

ORB 0.2304368019104004	0.006	471.6197974243164
Sift 2.5571212768554688	0.0029211295034079843	378.219352640451

/result_correct/32

ORB 0.2663285732269287	0.0	367.76223794555665
Sift 3.2579658031463623	0.033106134371957155	262.18984364673946

/result_correct/33

ORB 0.2513868808746338	0.006	354.3928780517578
Sift 3.094437599182129	0.014605647517039922	265.2826874128451

/result_correct/60

ORB 0.23626112937927246 0.004 347.6864210510254

Sift 2.6240482330322266 0.0019474196689386564 290.8707513744202
/result_correct/61

ORB 0 0 inf
Sift 2.5617175102233887 0.012658227848101266 421.8635881195514

.....

Датасет 2

/result_correct/3

	Time	Good_points/matches	Avarage_distance
--	------	---------------------	------------------

ORB	0.030644893646240234	0.013513513513513514	347.22745905695734
Sift	0.2994227409362793	0.03821656050955414	208.5466129278681

/result_correct/31

ORB	0.030127763748168945	0.0	506.5220840041702
Sift	0.2931499481201172	0.29936305732484075	321.92253550146796

/result_correct/61

ORB	0.03044891357421875	0.0	396.39107513427734
Sift	0.2901492118835449	0.12738853503184713	255.97837504004218

/result_correct/45

ORB	0.03096771240234375	0.0	352.34680485081026
Sift	0.29912781715393066	0.050955414012738856	199.85883809958293

/result_correct/46

ORB	0.032050132751464844	0.0	360.1155519227724
Sift	0.29380130767822266	0.006369426751592357	251.61875959262727

.....

SiFT децю розпізнає , ORB взагалі дивно себе веде.

В деяких випадках SiFT вражав точністю



3) погане освітлення

Датасет 1

/result_correct/15

Time	Good_points/matches	Avarage_distance
ORB 0 0 inf		
Sift 2.5931808948516846	0.0019474196689386564	446.44620473057523

/result_correct/39

ORB 0 0 inf
Sift 0 0 inf

/result_correct/67

ORB 0 0 inf		
Sift 2.5343849658966064	0.0038948393378773127	442.9573506891089

/result_correct/78

ORB 0 0 inf
Sift 0 0 inf

.....

Датасет 2

/result_correct/83

Time	Good_points/matches	Avarage_distance
ORB 0.033084869384765625	0.013513513513513514	354.56633263665276
Sift 0.31195902824401855	0.006369426751592357	187.22244299141465

/result_correct/90

ORB 0.033660888671875	0.013513513513513514	384.0287970465583
Sift 0.3021411895751953	0.012738853503184714	201.80489888768287

/result_correct/94

ORB 0.03289341926574707	0.0	348.19909977268526
Sift 0.31155872344970703	0.0	201.91024644207803

.....

Обидва алгоритма справились дуже погано.

0 0 inf - означає що цю фотографію методи взагалі не сприймають.

Незнаю хто писав в оренсв цю частину бібліотеки (почерк індусов). Но коли в документації написано що метод повертає масив елементів, а тобі на певну фотографію прилітає None а не пустий масив. Це жахливо.

4) Розмите зображення

Датасет 1

/result_correct/7

	Time	Good_points/matches	Avarage_distance
ORB	0.23040294647216797	0.0	397.8114720458984
Sift	2.591653347015381	0.0029211295034079843	373.15341371500966

/result_correct/23

ORB	0.23286032676696777	0.002	390.0387359008789
Sift	2.5631625652313232	0.0	347.82016728149534

/result_correct/27

ORB	0.23228836059570312	0.002	410.7335396728516
Sift	2.6032397747039795	0.007789678675754625	357.46104296690766

/result_correct/44

ORB	0.2839233875274658	0.004	373.4208617248535
Sift	3.2188427448272705	0.0009737098344693282	303.53005109302256

/result_correct/66

ORB	0.2845597267150879	0.002	372.57111834716795
Sift	3.218555450439453	0.0009737098344693282	305.41866389407386

.....

Датасет 2

/result_correct/20

	Time	Good_points/matches	Avarage_distance
ORB	0.030204057693481445	0.013513513513513514	478.3088288178315
Sift	0.29671192169189453	0.05732484076433121	333.05068415745046

/result_correct/89

ORB 0.03030705451965332 0.0 426.6968726080817
Sift 0.2924668788909912 0.03184713375796178 271.85039199537533

/result_correct/32

ORB 0 0 inf
Sift 0.294522762298584 0.6050955414012739 357.4939579082902

/result_correct/16

ORB 0.03022170066833496 0.0 429.43729050095016
Sift 0.2912161350250244 0.006369426751592357 364.3568107459196

.....

Обидва алгоритми впорались дуже погано. (SIFT дещо краще)

5) на зображенні немає предмету

Датасет 1

/result_wrong/3

Time	Good_points/matches	Avarage_distance
------	---------------------	------------------

ORB 0 0 inf
Sift 0 0 inf
/result_wrong/4

ORB 0.23177766799926758 0.0 456.385321105957
Sift 2.6066970825195312 0.007789678675754625 365.9819351835167

/result_wrong/5

ORB 0.24315214157104492 0.004 351.2955401916504
Sift 2.68975567817688 0.006815968841285297 281.20740279111334

/result_wrong/6

ORB 0.23713350296020508 0.006 364.7546877746582
Sift 2.6116347312927246 0.0038948393378773127 322.7903240690445

/result_wrong/7

ORB 0.23581552505493164 0.0 370.6971237487793
Sift 2.929701566696167 0.0009737098344693282 298.56172236634785

/result_wrong/8

ORB 0.2582681179046631 0.0 344.14446420288084
Sift 3.024962902069092 0.0019474196689386564 282.1742162360885

/result_wrong/9

ORB 0.251662015914917 0.0 343.5658219604492
Sift 2.818023204803467 0.0009737098344693282 286.4381916005284

.....

Датасет 2

/result_wrong/8

Time	Good_points/matches	Avarage_distance
ORB 0.030458927154541016	0.0	426.2337852684227
Sift 0.2971165180206299	0.025477707006369428	274.1528598761103

/result_wrong/9

ORB 0.030729293823242188 0.0 388.13445756242083
Sift 0.298098087310791 0.16560509554140126 298.2715457017255

/result_wrong/10

ORB 0.030000925064086914 0.013513513513513514 423.2766971072635
Sift 0.29397082328796387 0.17197452229299362 288.175779743559

/result_wrong/11

ORB 0.030836820602416992 0.0 361.4008153966955
Sift 0.30238962173461914 0.050955414012738856 216.64384897803046

/result_wrong/12

ORB 0.030457019805908203 0.0 425.8359469851932
Sift 0.2924954891204834 0.006369426751592357 234.4510891665319

.....

Безумовна перемога ORB він раніше нічого не розпізнавав а саме зараз це йому допомогло.

Висновок з попередніх випадків:

Датасет1 краще розпізнається в любых випадках.

1) SIFT добре розпізнає, його й потрібно використовувати в таких випадках (предмет з різних ракурсів).

2) SIFT добре розпізнає , ORB дещо теж розпізнає не з такою точністю за те швидко.
(частина предмета).

3) , 4), 5) обидва алгоритма майже нічого не можуть розпізнати за те ORB робить це швидше.
(освітлення, розмитість, відсутність).

Узагальнюючі висновки:

1. На розгляд було два датасети дуже **не** схожі між собою.



Перший предмет більш унікальний ніж другий навіть для людини перший предмет легше розпізнати.

Із-за багатьох ярко виражених переходів кольору (високого контрасту) перший датасет був ідеальним кандидатом для розпізнавання дескриптором SIFT. З розпізнаванням даного предмету з різних ракурсів та частин його він впорався дуже добре. Але на це в нього пішло дуже багато часу 2~3 секунди, але це краще ніж нічого не розпізнати.

Перше (ракурс та масштаб) пов'язано з тим що це входило в одну з найважливіших причин створення даного дескриптора. Друге (частина предмету) тому що йому вдалось згенерувати багато ключових точок. Що неможна сказати про ORB.

ORB на цьому датасеті впорався дуже погано. Зміг генерувати деякі точки лише на зображенні трохи меншого масштабу. Жахливо впорався з ракурсом та частинами, але робив це принаймні швидко.

Із зміною освітлення або розмитістю все вийшло дуже погано. Оскільки SIFT втратив свою головну опору — контраст кольору й чіткі переходи. ORB в цьому випадку можна сказати впорався краще тому що робив це швидко й в деяких випадках знаходив більше ключових точок. Тому в даному випадку перемога ORB.

При відсутності даного предмета на зображенні теж краще впорався ORB (раніше нічого не розпізнавав а зараз йому це допомогло) + робив це швидше. Натомість SIFT прив'язаний до певних переходів контрасту помилявся й генерував багатенько точок. Особливо видно це на датасеті2 де предмет з менш унікальною кольоровою гамою.

З датасетом2 все вийшло більш печально. Предмет в обе'мі з різних ракурсів виглядав порізному + кольорова гама дуже однотипна. В цьому випадку робота SIFT була набагато гіршою. Коректно він зміг розпізнати лише при змінні масштабу, при змінні ракурсу не зміг тому, що предмет виглядав дещо інакше.

В загальному цей датасет виявся набагато важчим при нашій задачі.

При випадках зміни освітлення, розмитості, частини предмета обидва дескриптора були безсилі, а ORB робив швидше тому перемога ORB.

Відсутність предмета на зображенні добре збивало роботу SIFT (кольорова гама даного предмета дуже схожа з багатьма іншими предметами) що зумовлювало його генерувати багатенько ключових точок.