دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

تمرین سری پنج

درس بینایی ماشین

فرهاد دلیرانی

۹۶۱۳۱۱۲۵

dalirani@aut.ac.ir

dalirani.1373@gmail.com

فهرست

ابزارهای استفاده شده ۱
تمرین ۱ ..... ۲
تمرین ۲ ..... ۷
تمرین ۳ ..... ۲۱

ابزارهای استفاده شده

زبان برنامه نویسی: Python 3.6

محیط توسعه: PyCharm

سیستم عامل:Windows 10

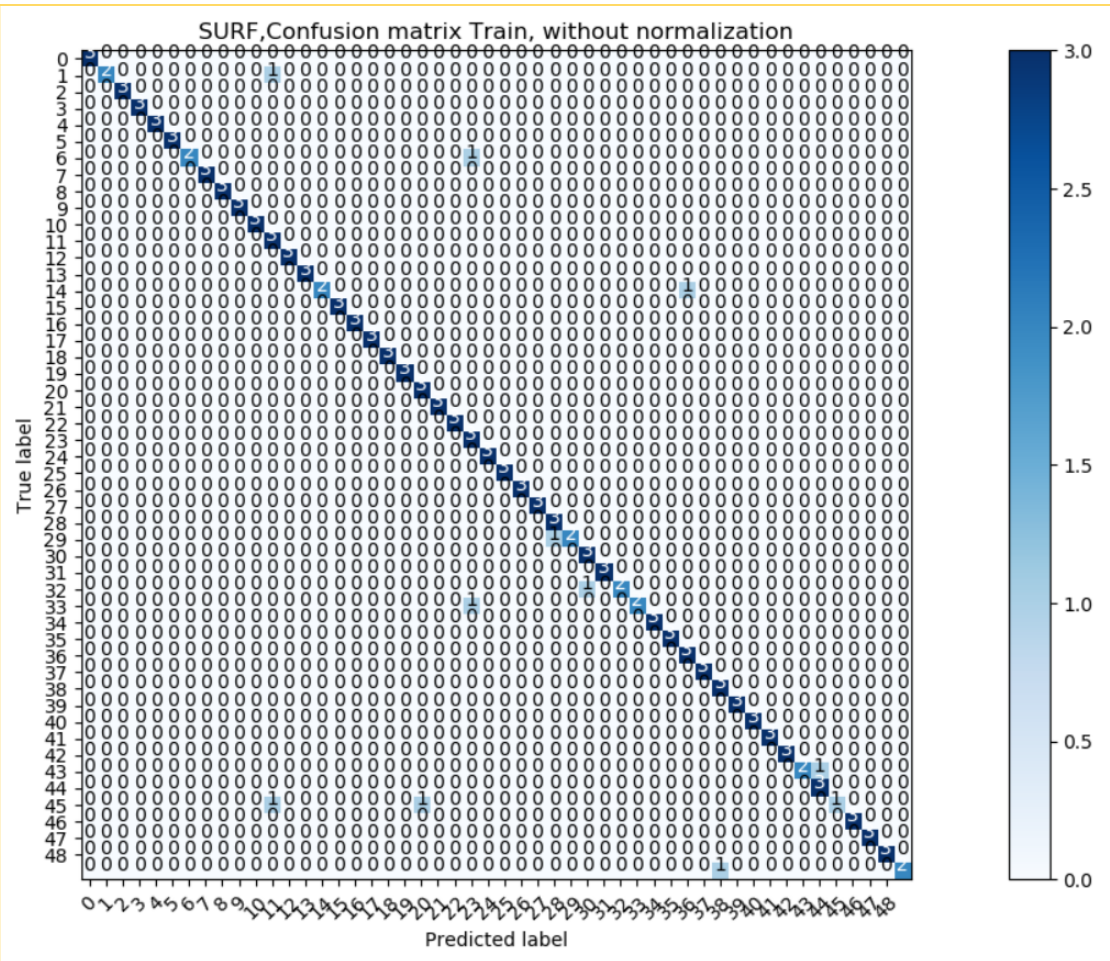
تمرین ۱

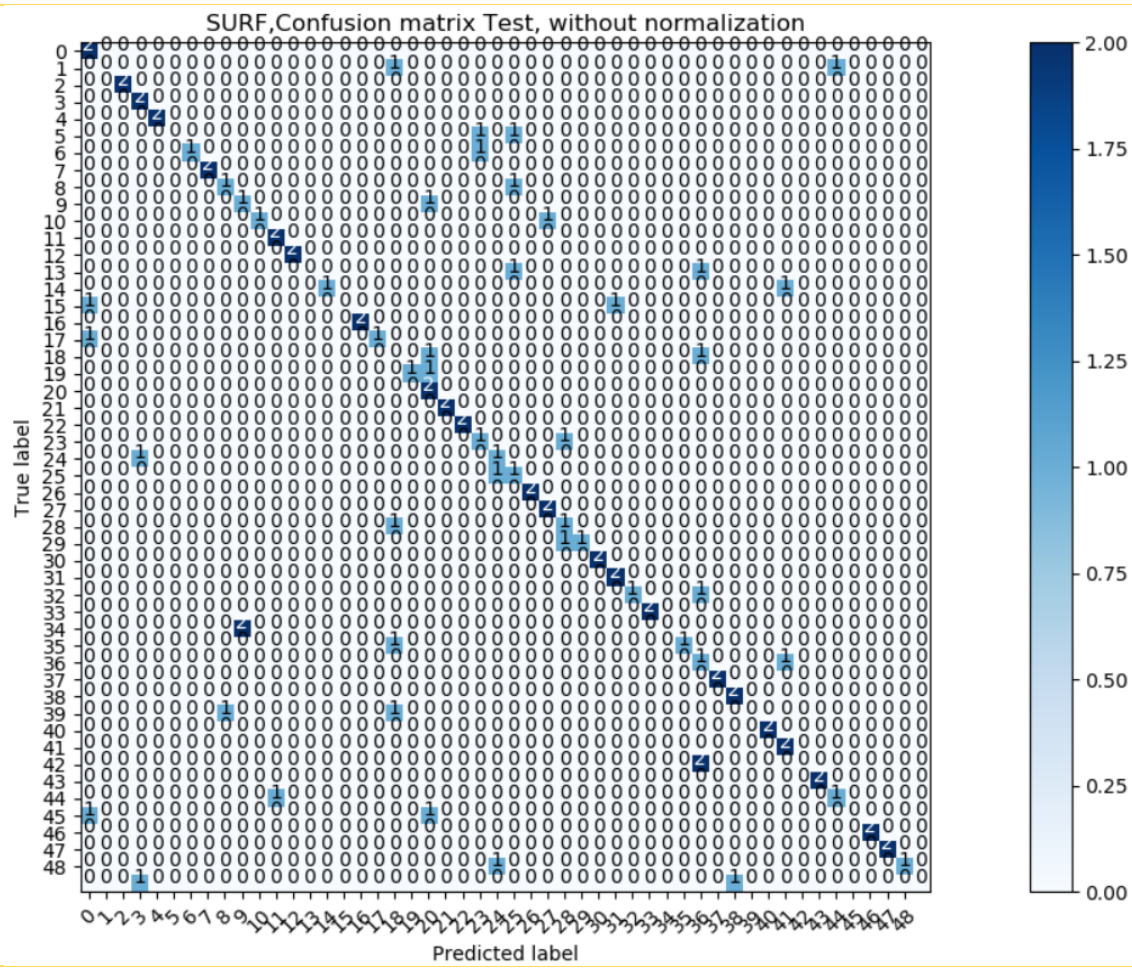
کدهای این سوال در فایل‌های problem-1-surf.py، problem-1-LBP.py, problem-1-freak.py classifier.py، read_dataset.py و visual_bag_of_words.py قرار دارد. برای دسته‌بندی با استفاده از ویژگی‌های Surf، ابتدا برای تصویرهای موجود در مجموعه‌ی آموزش تمام ویژگی‌های سورف استخراج شده است. سپس ویژگی‌های سورف به mini-batch kmeans داده شده است و ۱۰۰ خوشه استخراج می‌شود. بعد از آن باید برای هر تصویر یک بردار ویژگی استخراج کنیم، ابتدا برای تصویر تمام ویژگی‌های سورف را استخراج می‌کنیم و سپس مشخص می‌کنیم هر ویژگی استخراج شده معادل کدام یک از مرکز خوشه‌ها است. بردار ویژگی معادل یک تصویر برابر با بردار نرمال شده، میزان تکرار ویژگی‌های مربوط به خوشه‌ها می‌شود. سپس با استفاده از این بردارها یک دسته‌بند ایجاد می‌کنیم، از تنظیمات پیش فرض SVM کتابخانه‌ی Sklearn برای دسته‌بند استفاده شده است.

برای ویژگی freak دقیقا مانند بالا عمل می‌کنیم با این تفاوت که نقاط کلیدی را با سورف پیدا می‌کنیم و توصیفگر را با Freak ایجاد می‌کنیم.

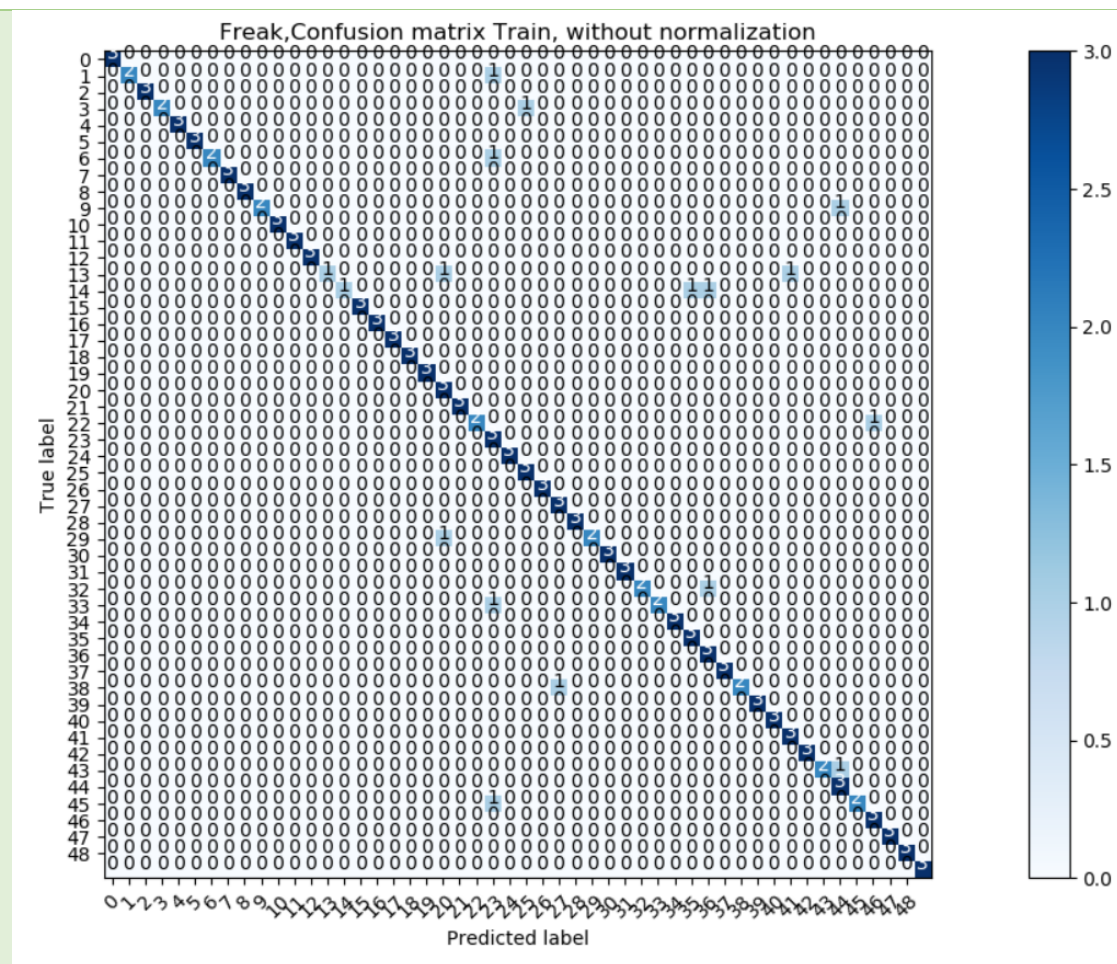
برای LBP از شعاع ۳ و تعداد نقاط در همسایگی برابر با ۹ استفاده می‌کنیم و بافت هر تصویر مجموعه آموزش را به دست می‌آوریم. سپس با مجموعه آموزش یک دسته‌بند را آموزش می‌دهیم.

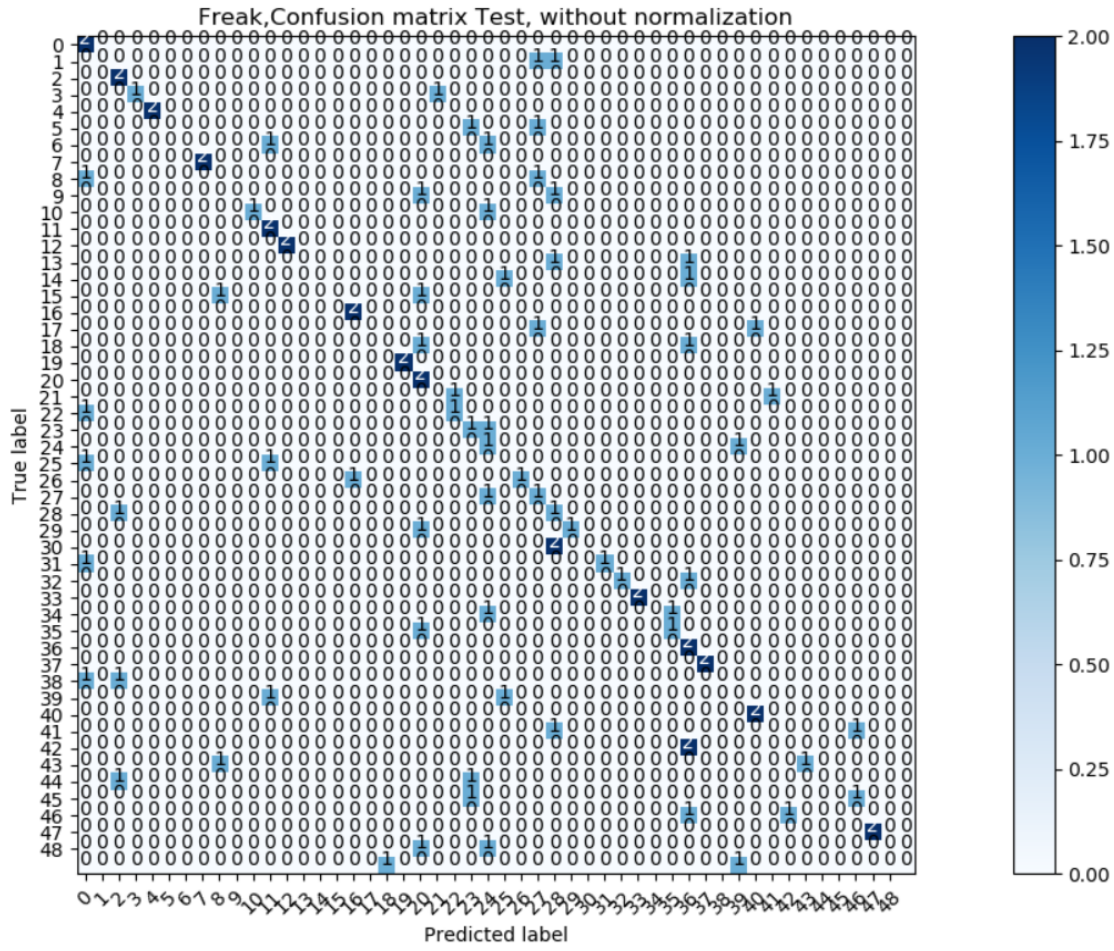
با استفاده از SURF نتایج زیر حاصل شد:

	SURF	
	Train Accuracy: 0.9333333333333333 Train F1-score: 0.9308571428571427 Train Precision-score: 0.9333333333333333 Train Recall-score: 0.9333333333333333	**Accuracy, F1-Score, Precision, Recall For Training Set**
	Test Accuracy: 0.63 Test F1-score: 0.5957619047619047 Test Precision-score: 0.63 Test Recall-score: 0.63	**Accuracy, F1-Score, Precision, Recall For Test Set**
		**Confusion matrix for Training Set**

	**Confusion matrix for Test Set**

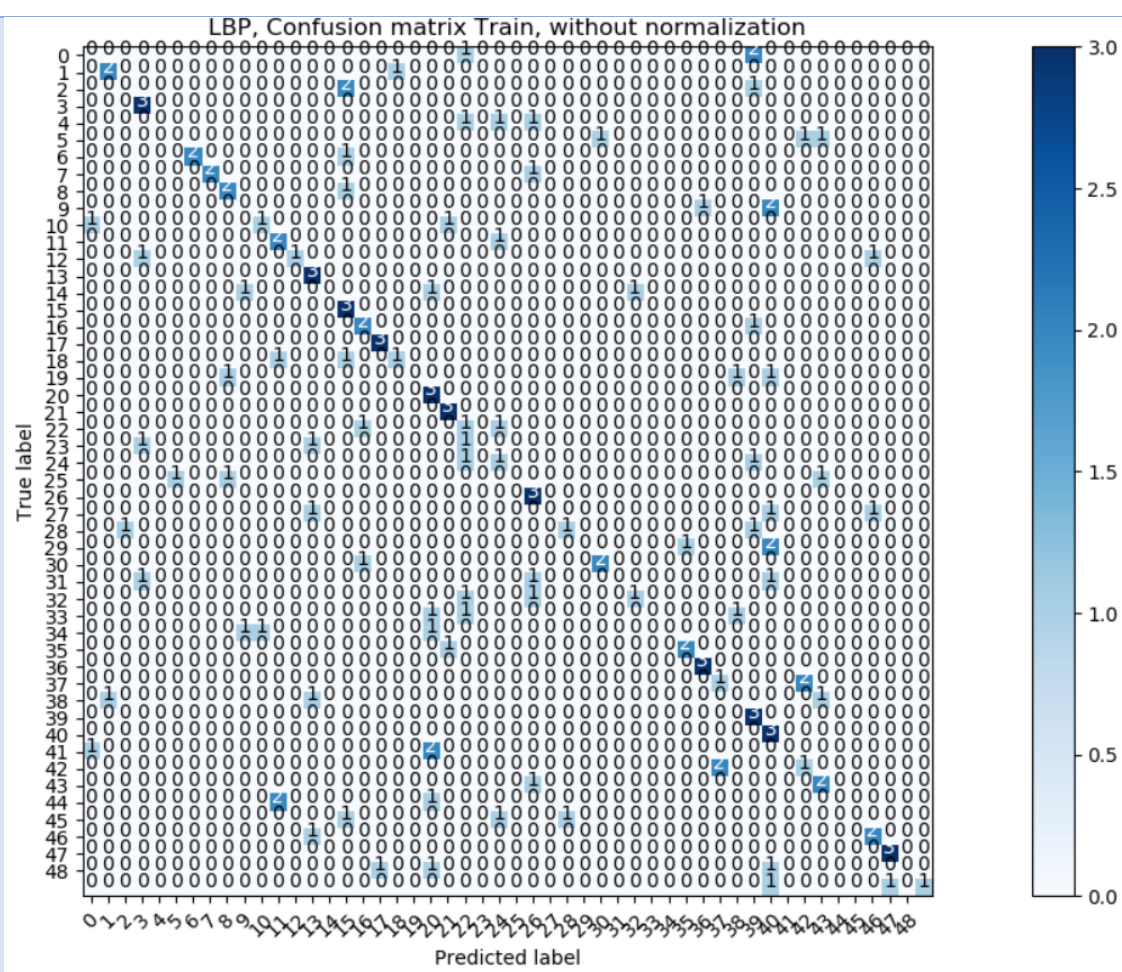
با استفاده از Freak نتایج زیر حاصل شد:

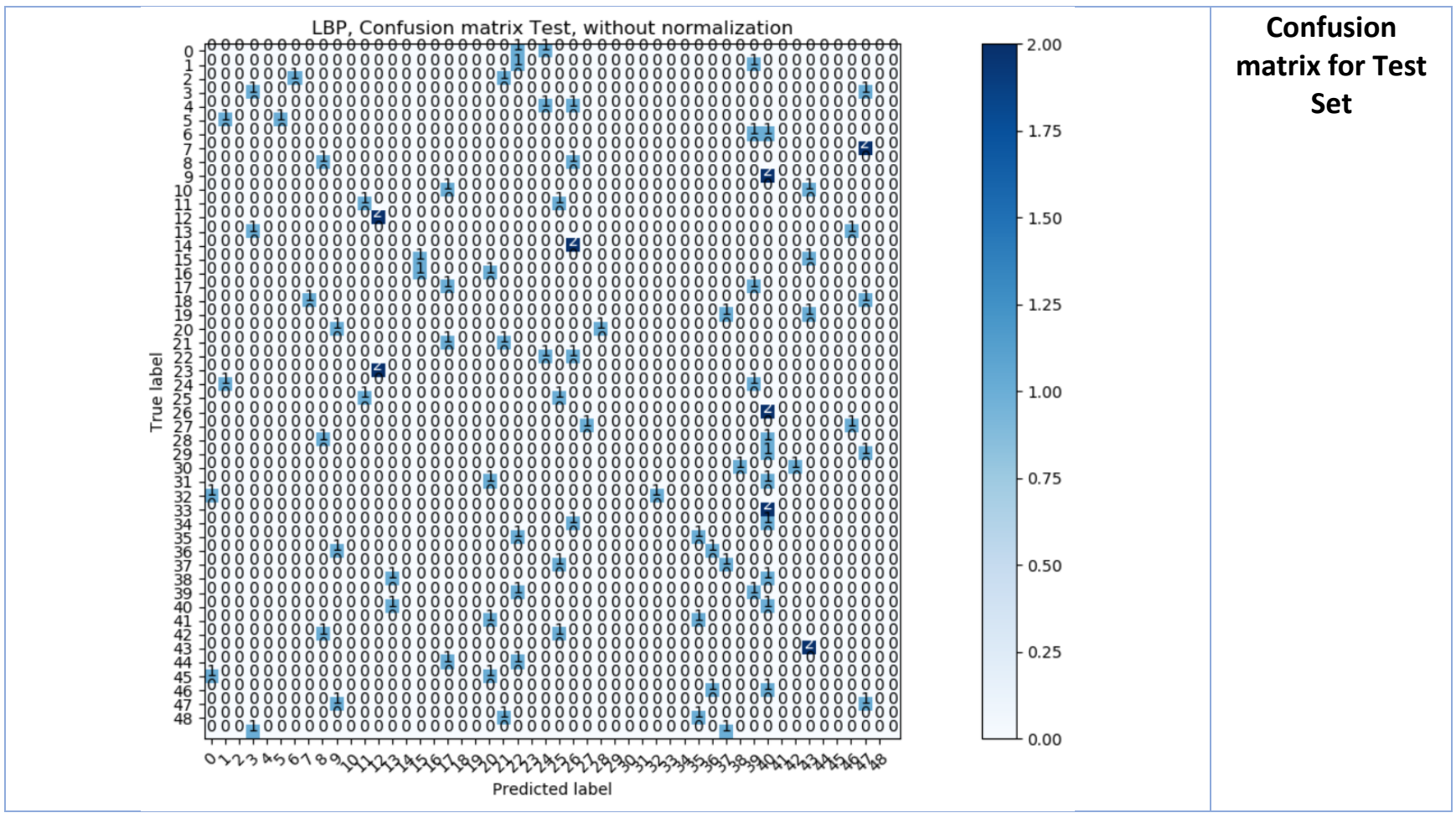
Freak	
Train Accuracy: 0.9 Train F1-score: 0.8987142857142857 Train Precision-score: 0.9 Train Recall-score: 0.9	**Accuracy, F1-Score, Precision, Recall For Training Set**
Test Accuracy: 0.41 Test F1-score: 0.35362193362193367 Test Precision-score: 0.41 Test Recall-score: 0.41	**Accuracy, F1-Score, Precision, Recall For Test Set**
	**Confusion matrix for Training Set**



Confusion matrix for Test Set

با استفاده از LBP نتایج زیر حاصل شد:

LBP	
Train Accuracy: 0.42 Train F1-score: 0.3395684315684316 Train Precision-score: 0.42 Train Recall-score: 0.42	**Accuracy, F1-Score, Precision, Recall For Training Set**
Test Accuracy: 0.2 Test F1-score: 0.16130952380952382 Test Precision-score: 0.2 Test Recall-score: 0.2	**Accuracy, F1-Score, Precision, Recall For Test Set**
	**Confusion matrix for Training Set**

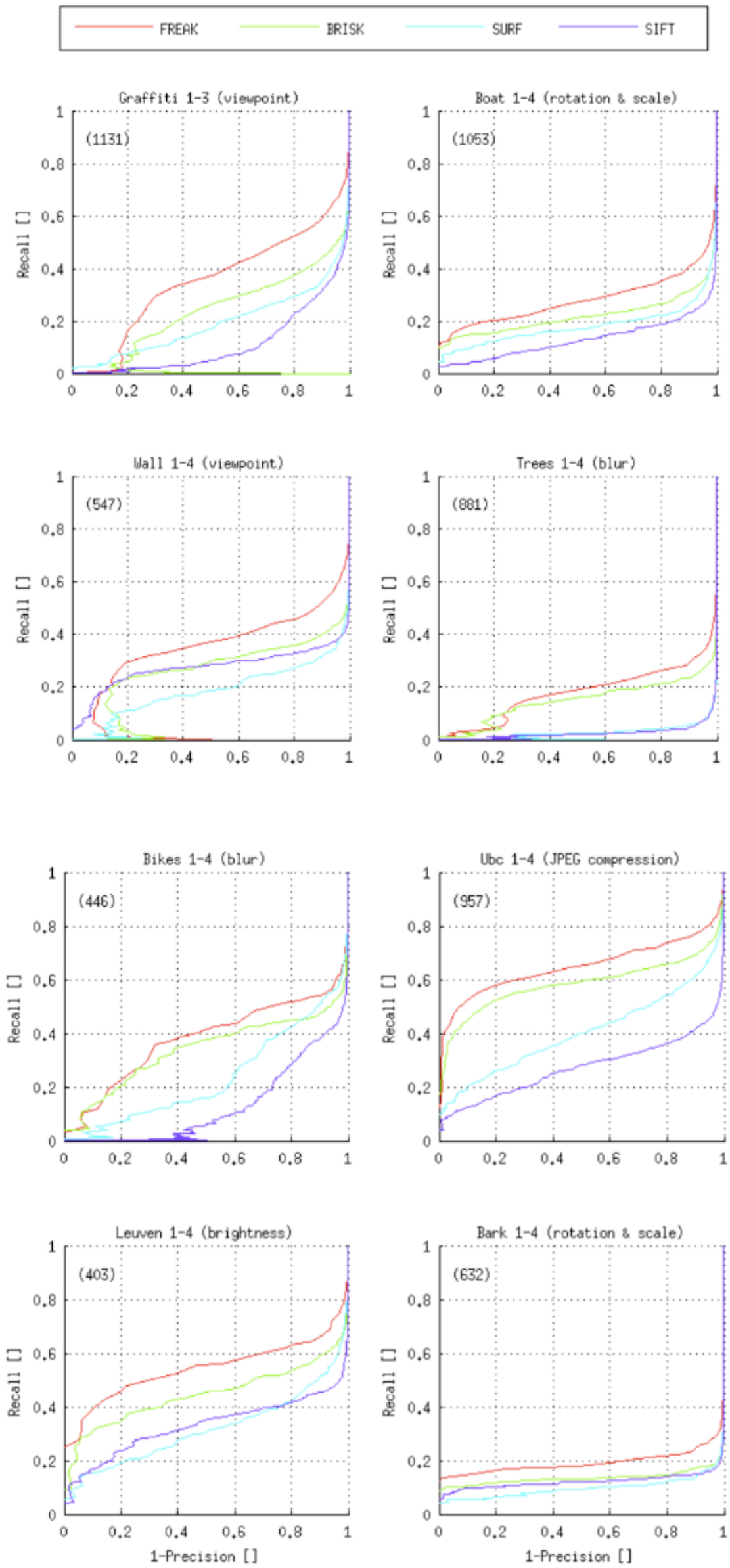


سرعت انجام محاسبات در سورف، فریک و lbp بسیار بیشتر از سیفت است.

در جدول‌های بالا نتایج حاصل از دسته‌بندی به کمک Freak،SURF و LBP ارائه شد. عملکرد SURF چه از نظر precision و recall و چه از نظر confusion matrix بسیار خوب است و بهتر از سایر موارد است. اگر به confusion matrix مجموعه تست و آموزش دسته‌بندی که با SURF کار می‌کند دقت کنیم، مشاهده می‌شود که مقادیر روی قطر اصلی غیر صفر است و در بیشتر خانه‌های خارج از قطر اصلی مقادیر صفر هستند، که نشان دهنده‌ی این هست که دسته‌بندی که با سورف کار می کنند به خوبی مجموعه تست را دسته‌بندی کرده است. مجموعه تست شامل تصاویری است که نسبت به تصاویر درون مجموعه آموزش دچار چرخش بزرگ نمایی و جابجایی شده اند. در نتیجه می‌توان نتیجه گرفت عملگر سیفت نسبت به چرخش، مقیاس و جابه‌جایی به خوبی مقاوم است.

بعد از سورف، دو روش freak و sift عملکرد مشابه و بسیار نزدیکی داشته‌اند، هر دو برای مجموعه تست عملکردی بین ۴۱ تا ۴۲ درصد داشته‌اند، فریک مقداری بهتر عمل کرده است. بعد از این دو، روش LBP عملکرد قرار دارد.

سورف و فریک از نظر سرعت و دقت بهتر از سیفت عمل می‌کنند و در آزمایش‌هایی که انجام داده‌ایم این موضوع مشاهده شد، تصویر زیر هم از مقاله فریک برداشته شده است که در تایید نتایج گرفته شده است. سیفت از تخمین DoG برای پیدا کردن نقاط کلیدی استفاده می‌کند در حالی که سورف از Fast Hessian برای پیداکردن نقاط کلیدی استفاده می‌کند. توصیفگر سورف پیچیده‌تر از توصیفگر سیفت است که همه‌ی این‌ها موجب عملکرد بهتر سورف شده است.

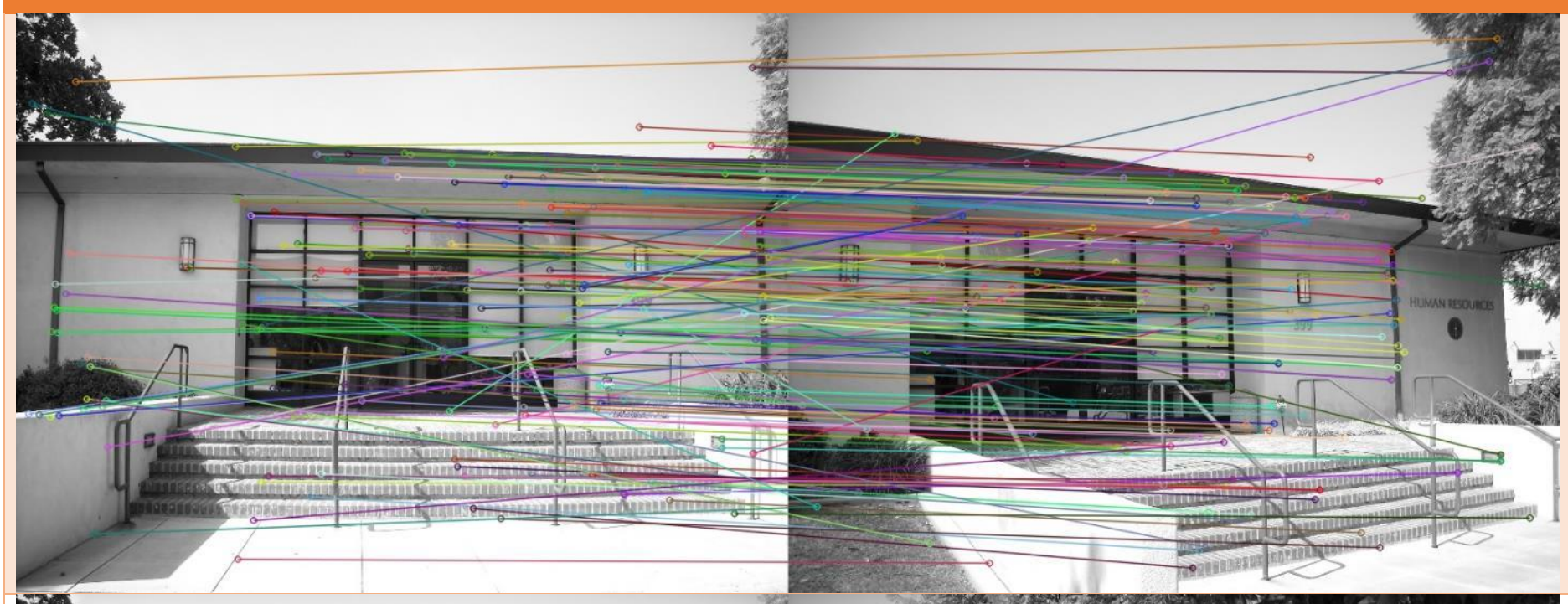
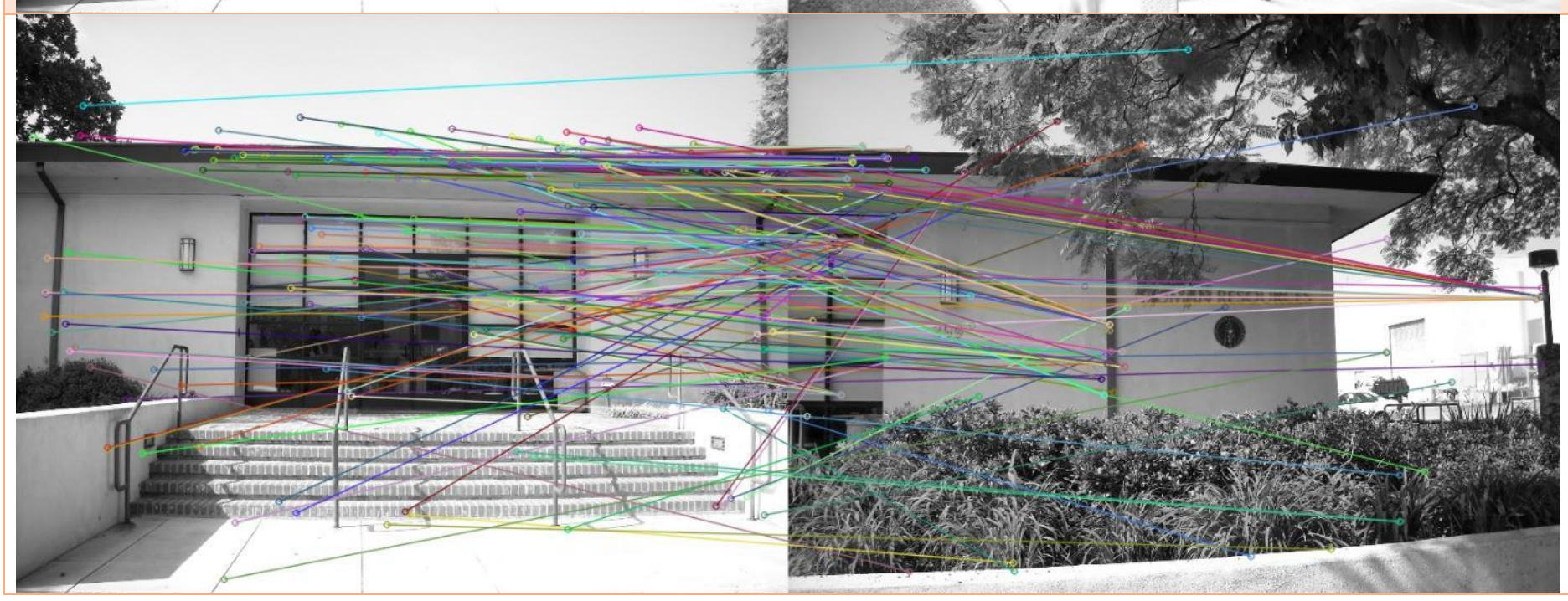

FREAK
BRISK
SURF
SIFT
Graffiti 1-3 (viewpoint)
(1131)
Boat 1-4 (rotation & scale)
(1053)
Wall 1-4 (viewpoint)
(547)
Trees 1-4 (blur)
(881)
Bikes 1-4 (blur)
(446)
Ubc 1-4 (JPEG compression)
(957)
Leuven 1-4 (brightness)
(403)
Bark 1-4 (rotation & scale)
(632)
Recall []
1-Precision []

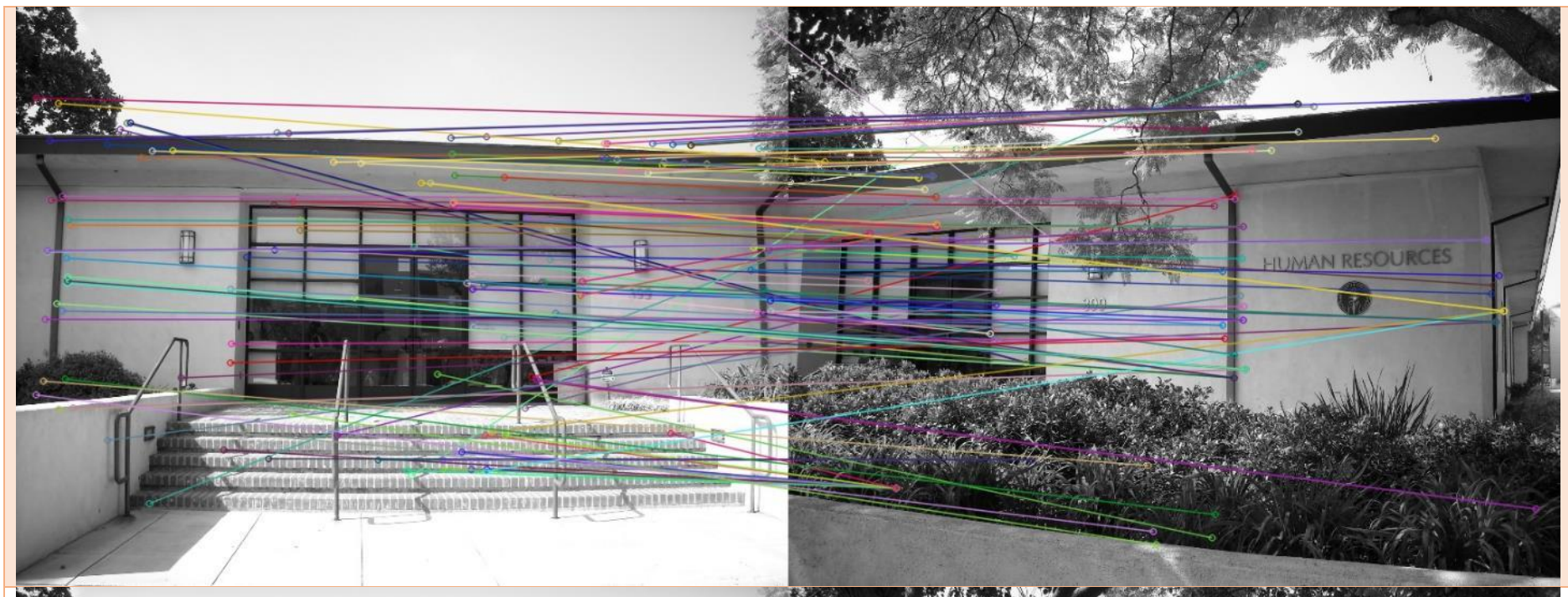
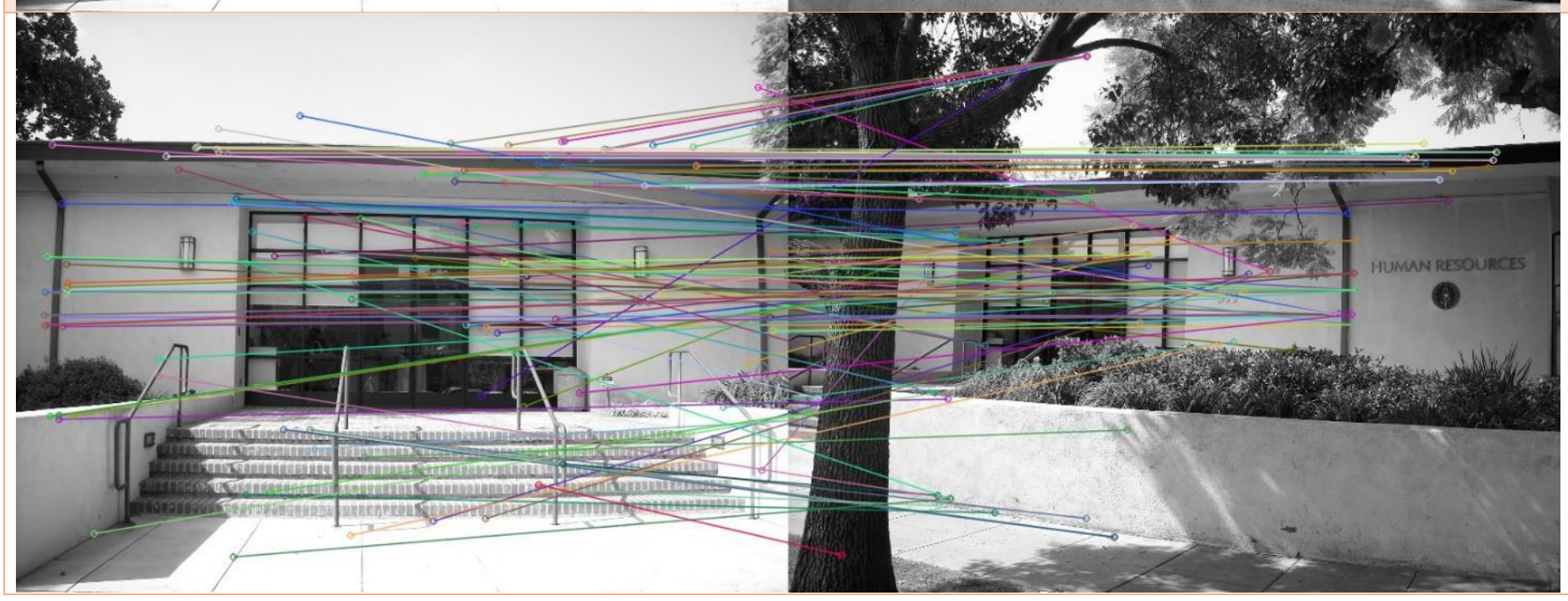
تمرین ۲

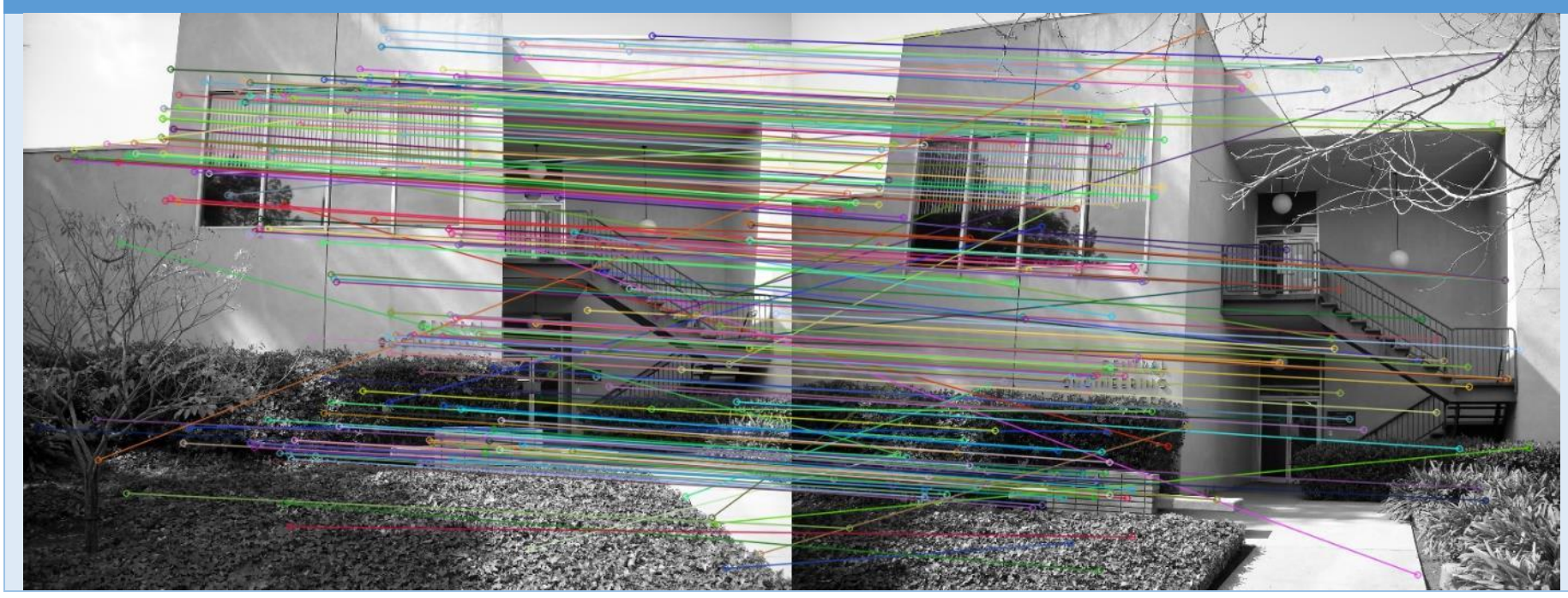
کدهای این سوال در problem-2.py قرار دارد. در این سوال از freak و surf استفاده شده است. در این سوال برای ساختمان‌های مختلف ویژگی‌های مختلف را استخراج می‌کنیم و سپس نمایش می‌دهیم. در نمایش انطباق‌ها فقط ویژگی‌هایی که به خوبی انطباق داده شده اند را نمایش می‌دهیم و از نمایش تمام انطباق‌ها دوری می‌کنیم. در ادامه ۳۹ خروجی این سوال را مشاهده می‌کنید:

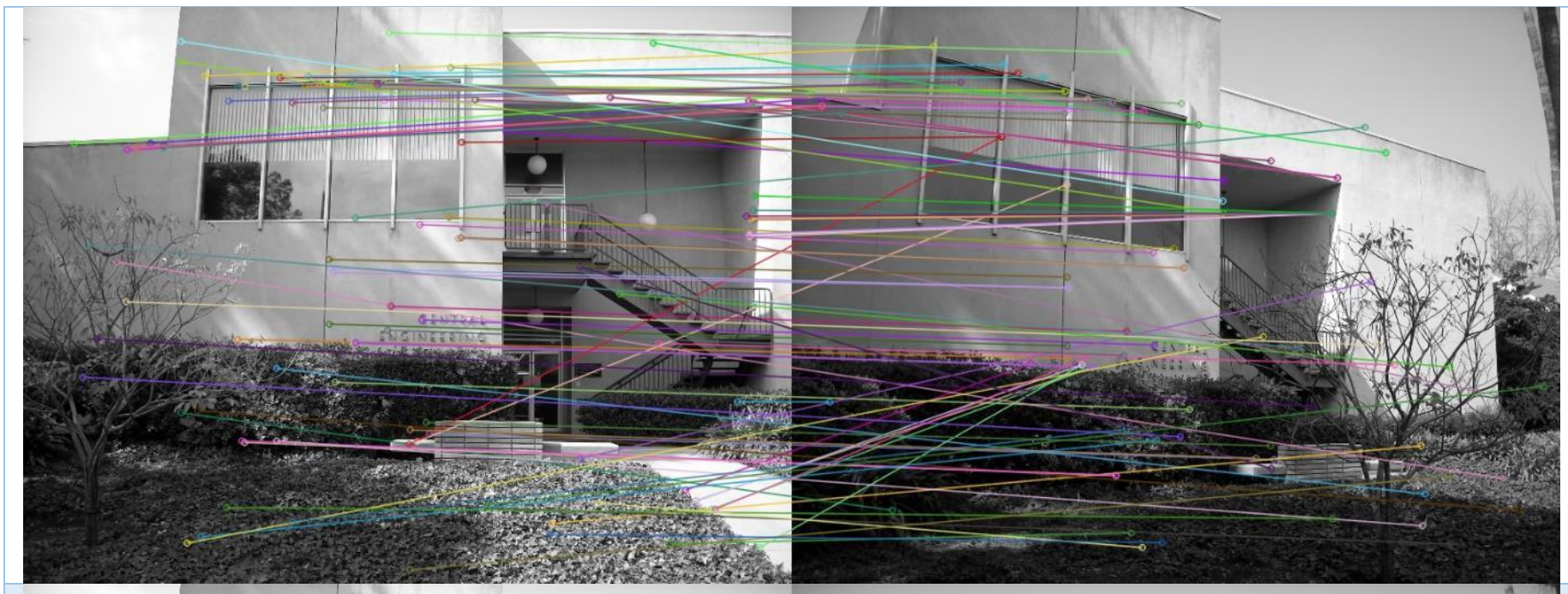
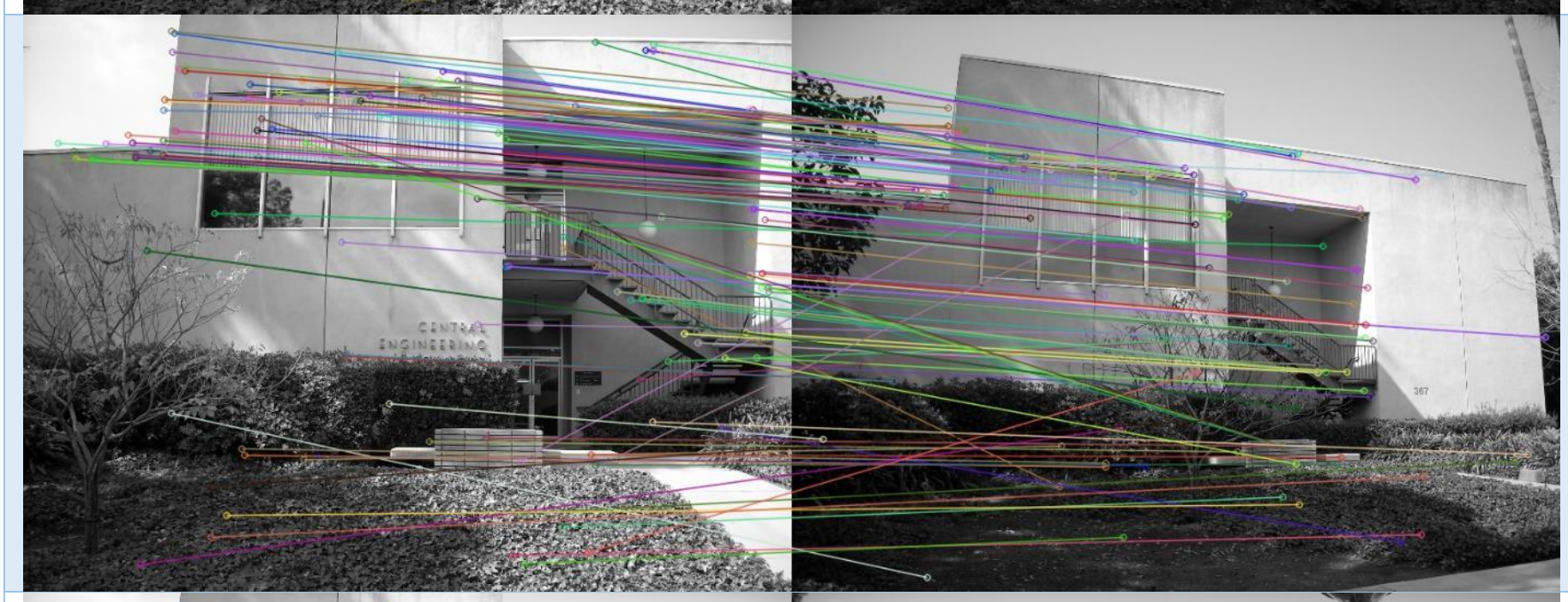
انطباق ساختمان ۱ با ۴ تصویر دیگرش SURF	N.O Of good matches
	۱۸۳
	۱۰۱
	۱۳۰

	۱۵۶

انطباق ساختمان ۲ با ۴ تصویر دیگرش SURF	N.O Of good matches
	۲۰۶
	۲۴۹

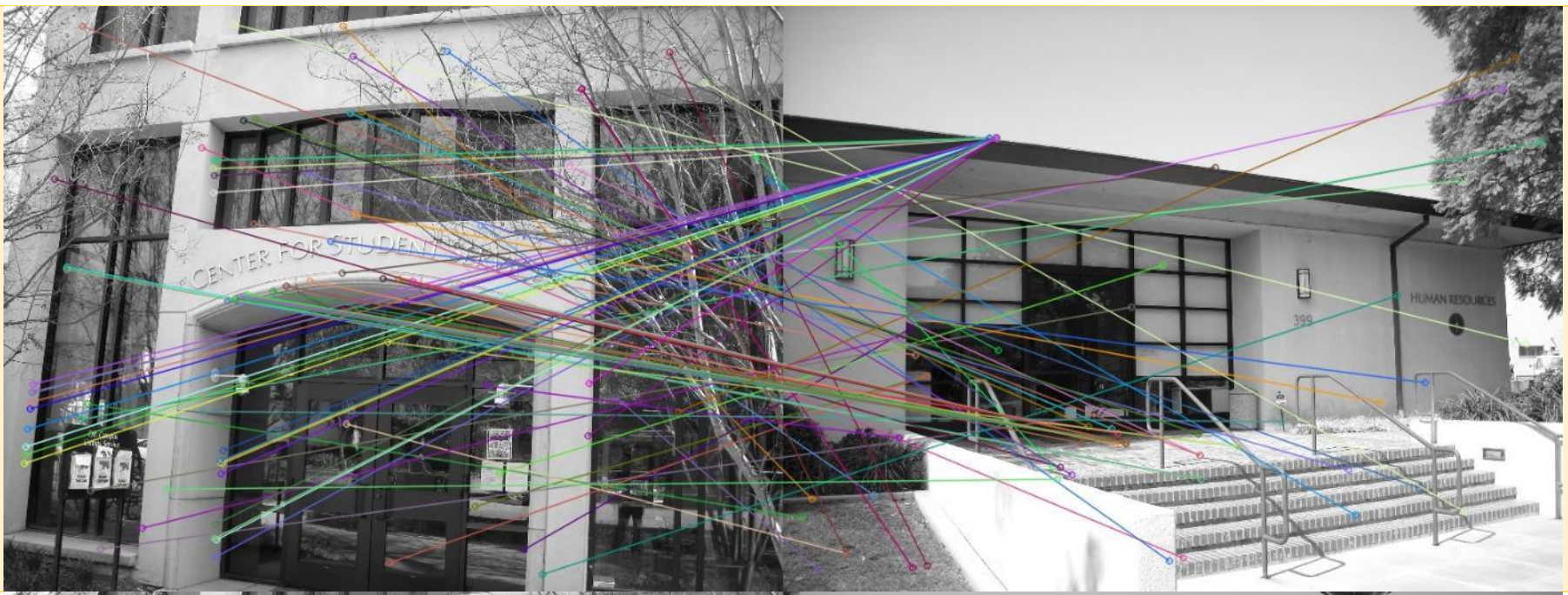
	۱۳۵
	۱۳۳

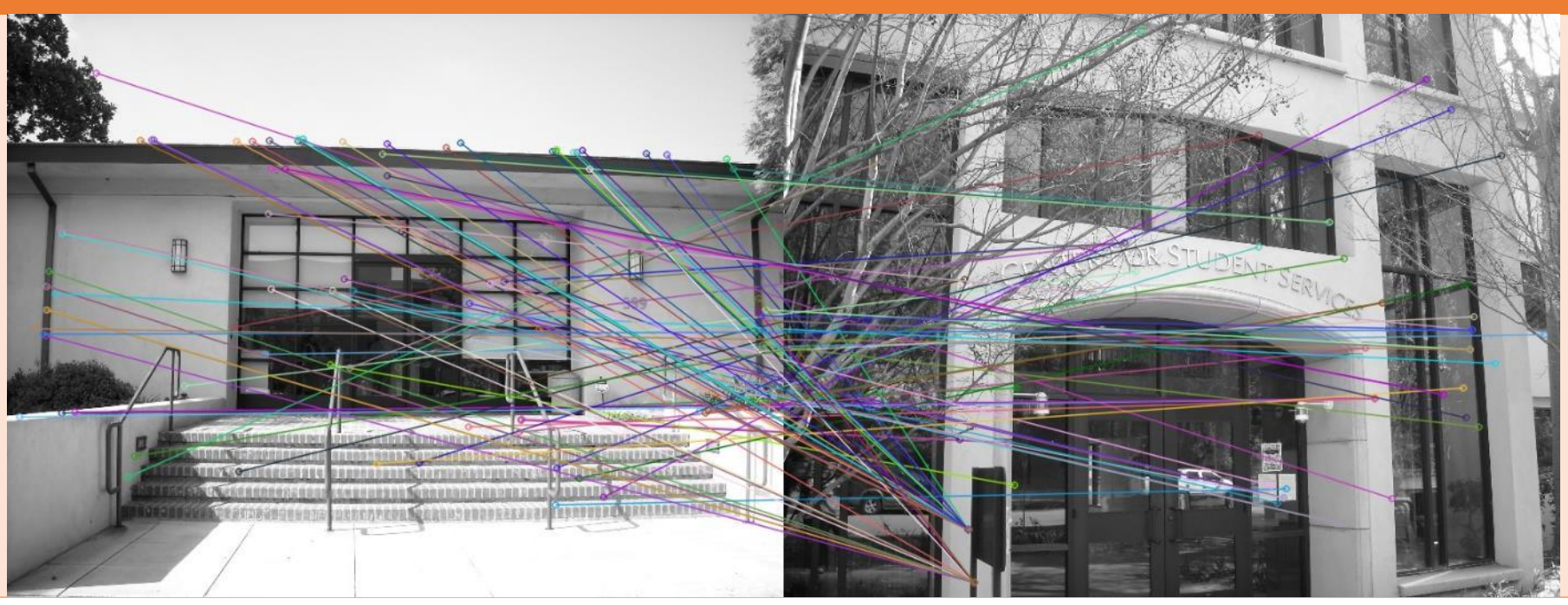
انطباق ساختمان ۳ با ۴ تصویر دیگرش SURF	N.O Of good matches
	۲۳۸

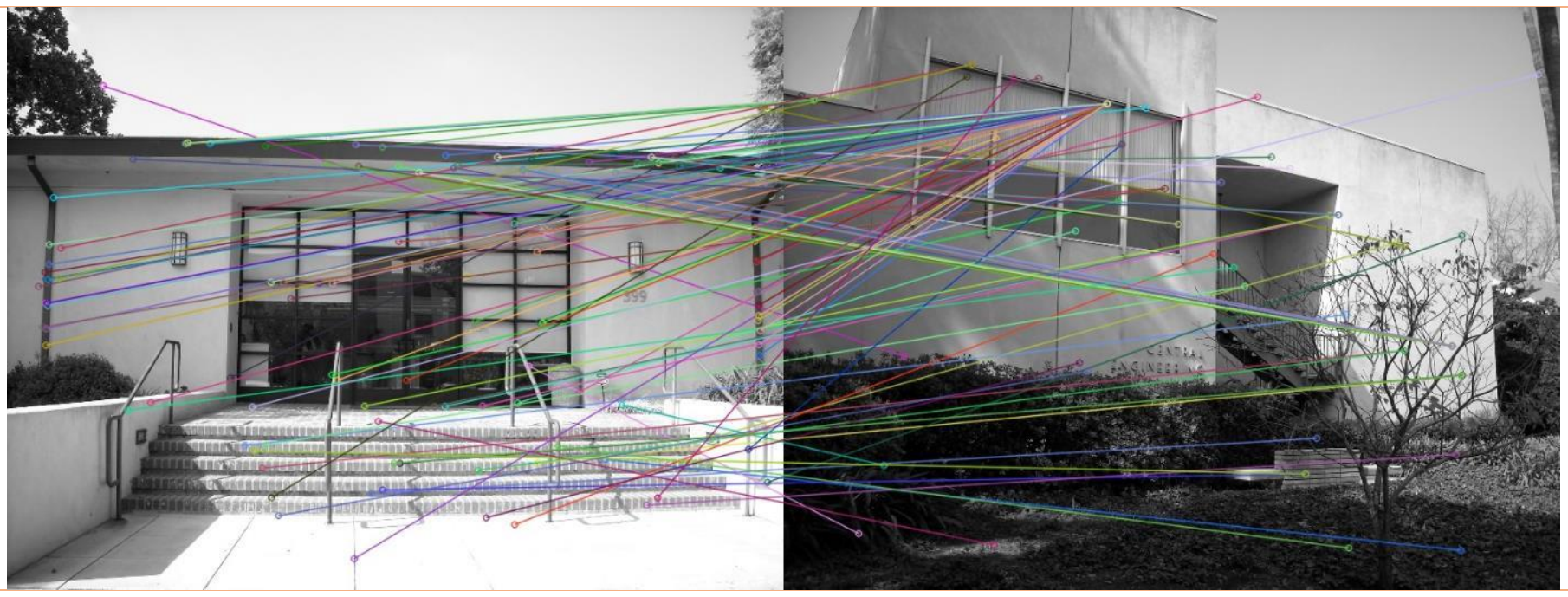
	۱۱۰
	۱۴۵
	۲۲۷

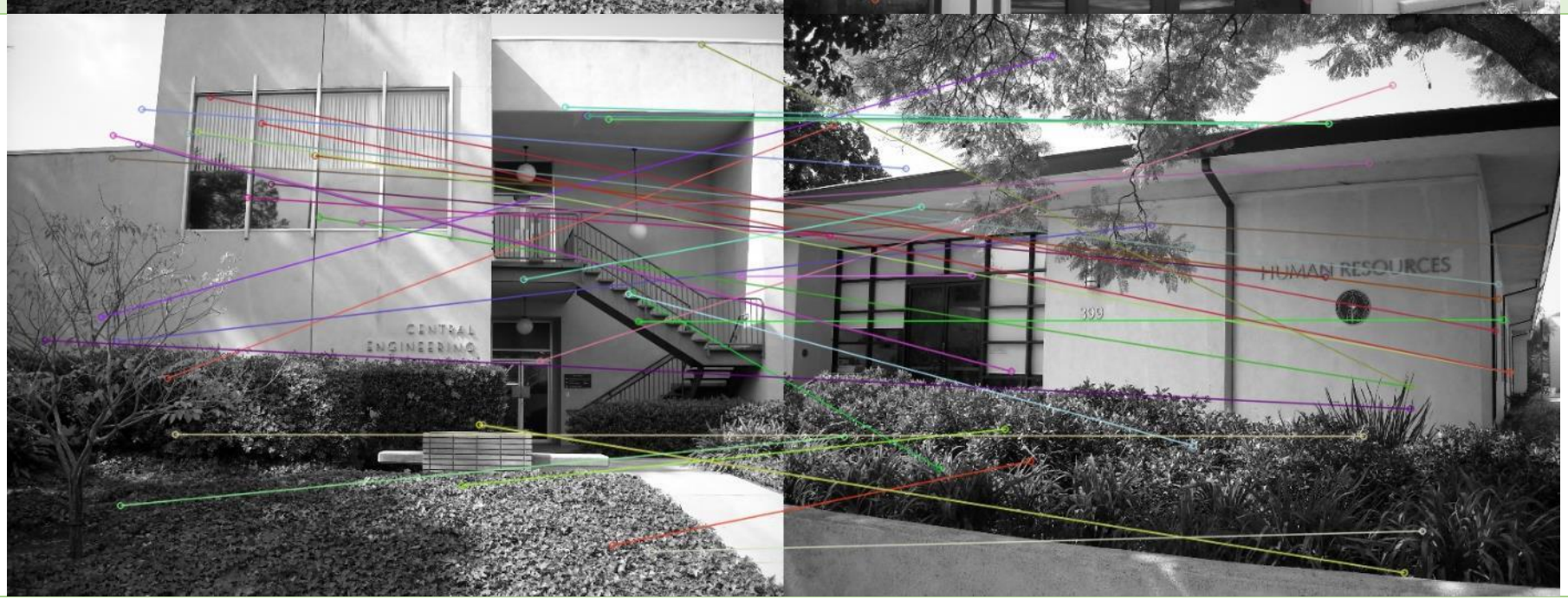
انطباق تصویر یک ساختمان با دو تصویر از دو ساختمان دیگر

N.O Of good matches	انطباق ساختمان ۱ با دو تصویر دیگر از دو ساختمن دیگرSURF

	۹۸
	۶۶

انطباق ساختمان ۲ با دو تصویر دیگر از دو ساختمان دیگر SURF	N.O Of good matches
	۹۳

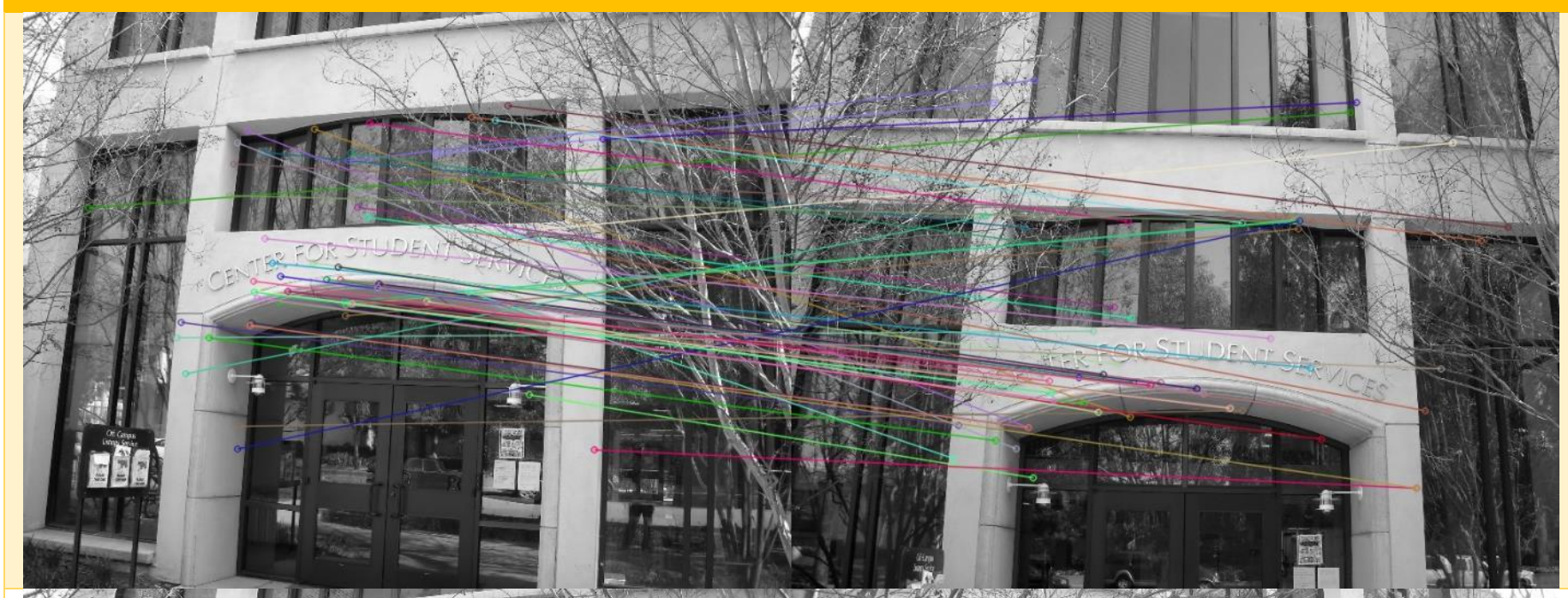
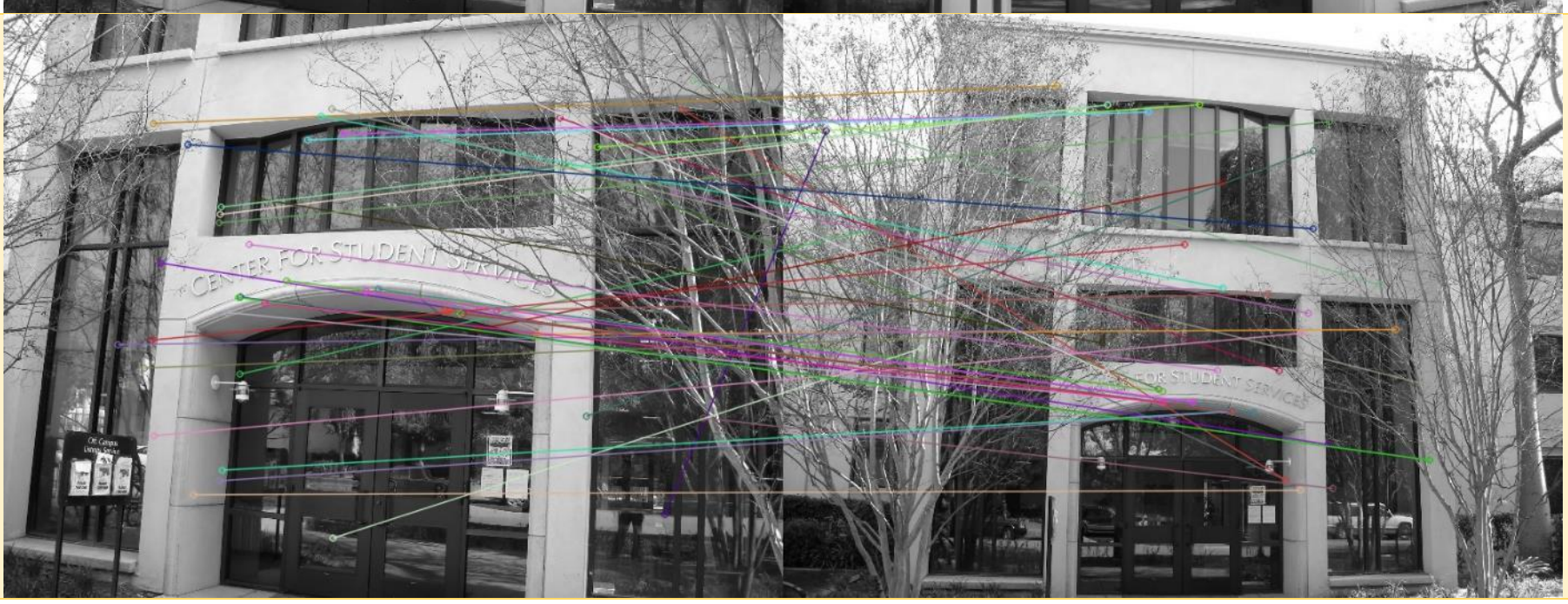
	۱۱۴

انطباق ساختمان ۳ با دو تصویر دیگر از دو ساختمان دیگرSURF	N.O Of good matches
	۶۵
	۳۵

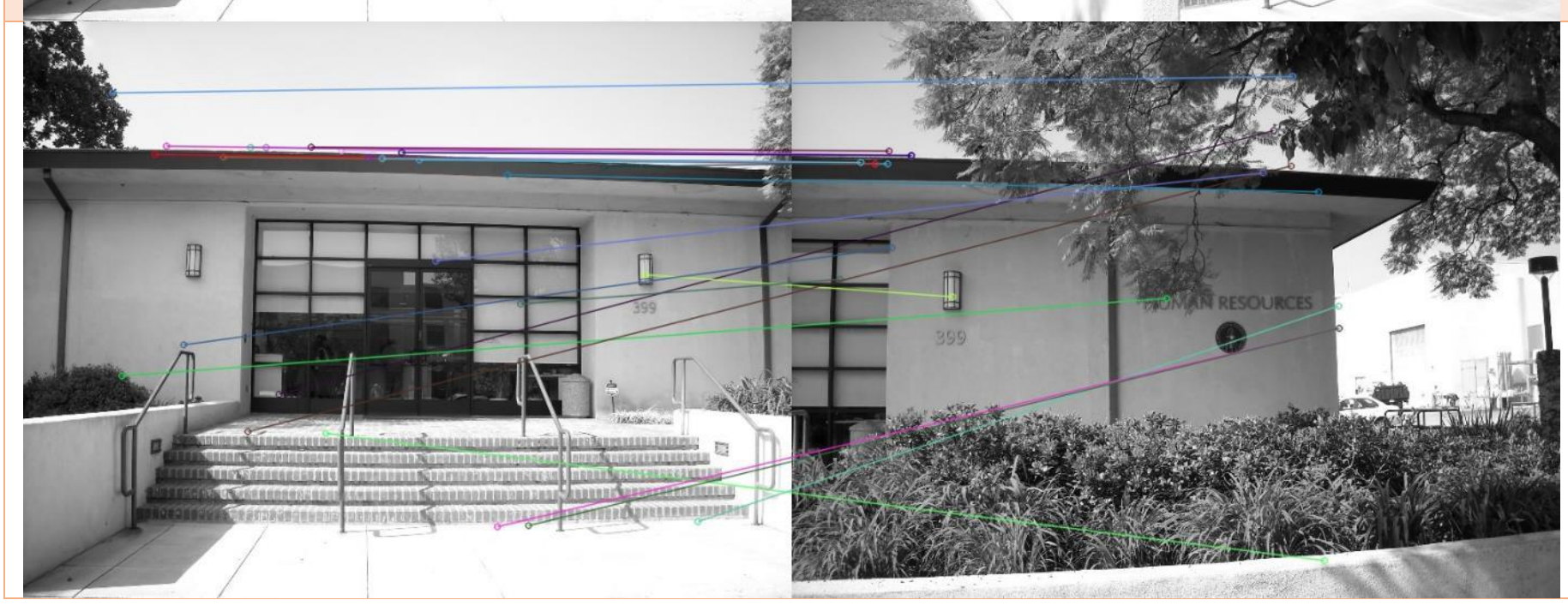
در ۱۸ تصویر بالا نتایج حاصل از انطباق بین ساختمان‌های مختلف با تصویرهایی از همان ساختمان که دچار چرخش تغییرمقیاس و جابه‌جایی شده اند را مشاهده می‌کنید(از ویژگی surf استفاده شده است). همچنین انطباق بین ساختمان‌ها با ساختمان‌های دیگر را مشاهده می‌کنید. در سمت راست هر تصویر تعداد انطباق‌های خوب نمایش داده شده است.

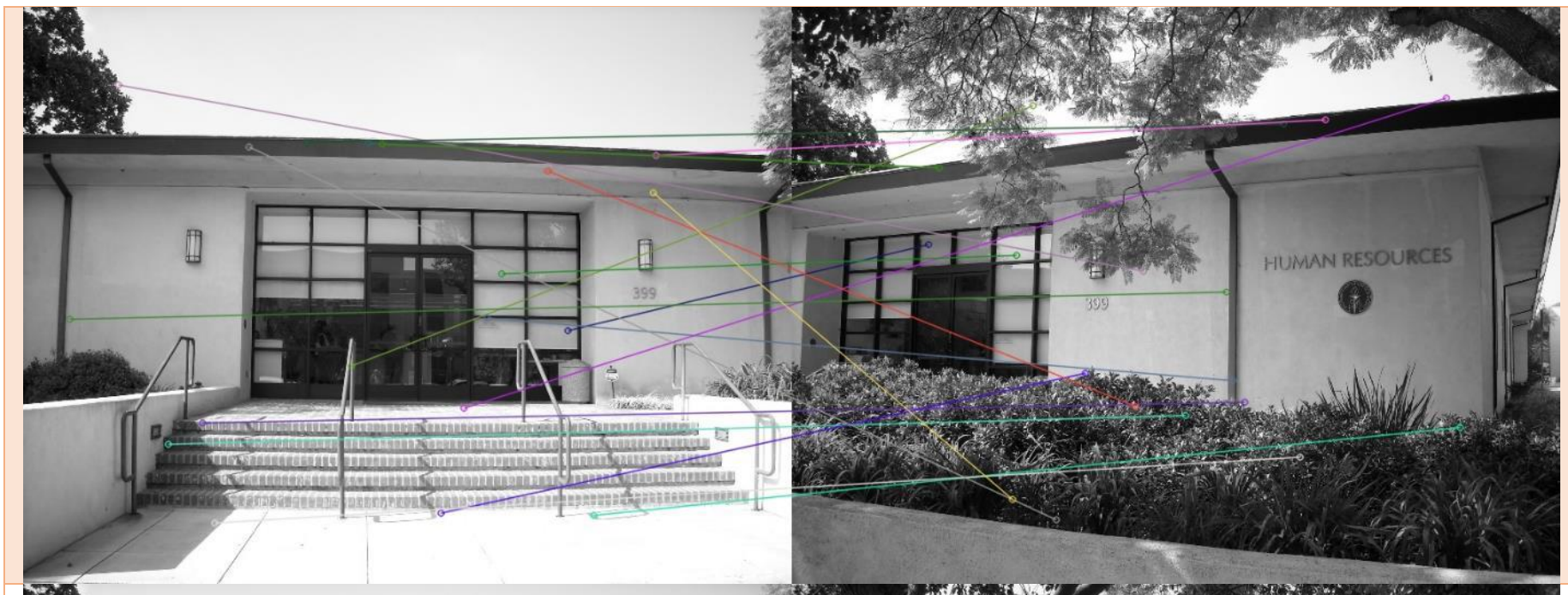
در انطباق یک ساختمان با تصاویر دیگر همان ساختمان تعداد انطباق‌های خوب بسیار زیاد است که بیشتر آن‌ها نیز درست هستند، در انطباق یک ساختمان با تصاویر دیگر همان ساختمان نکته‌ی جالبی مشاهده می‌شود: هر چه میزان مقیاس، چرخش و جابه‌جایی شدیدتر بوده است تعداد انطباق‌های خوب کاهش پیدا کرده است و هر چه میزان تغییرات کمتر بوده است تعداد انطباق‌های خوب بیشتر شده است.

در انطباق بین یک ساختمان با ساختمان‌های دیگر، تعداد انطباق‌های خوب به شدت کاهش یافته است، همچنین بیشتر انطباق‌ها در نقاطی مثل برگ‌ها، شیشه‌ها رخ داده است.

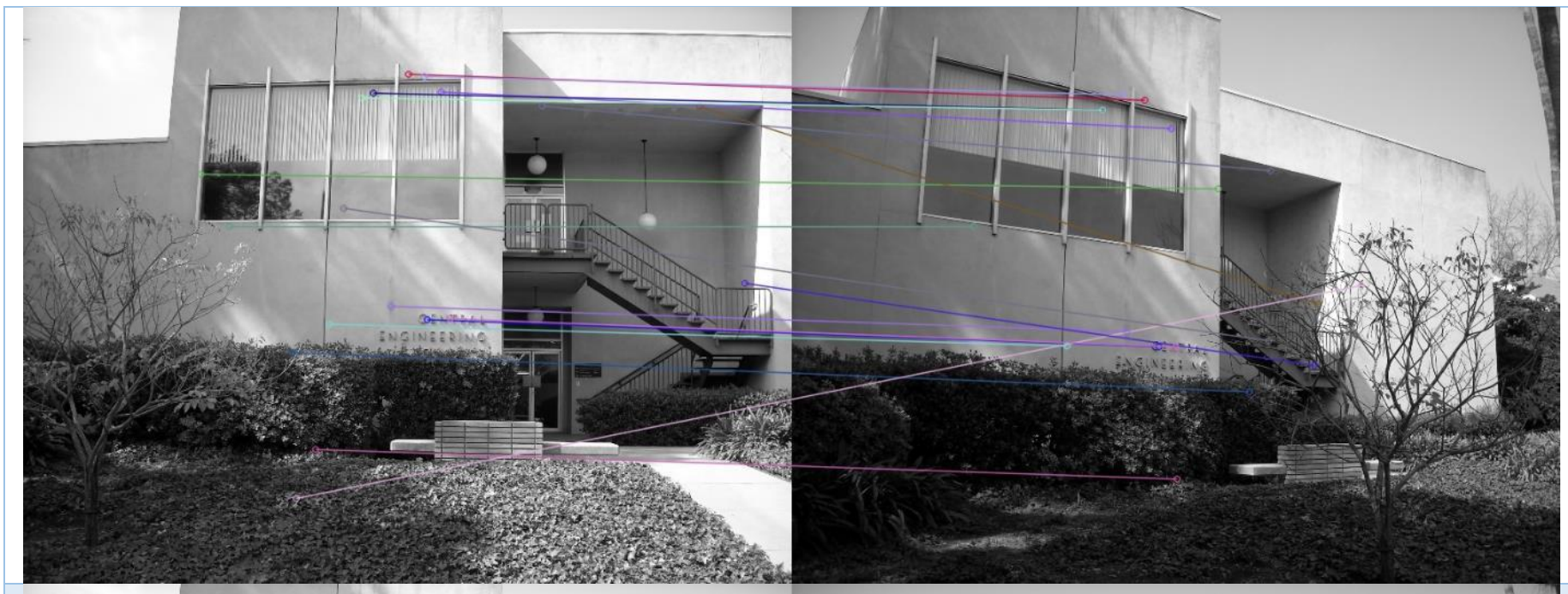
انطباق ساختمان ۱ با ۴ تصویر دیگرش Freak	N.O Of good matches
	56
	69
	48

	81

انطباق ساختمان ۲ با ۴ تصویر دیگرش Freak	N.O Of good matches
	97
	25

	19
	27

انطباق ساختمان ۳ با ۴ تصویر دیگرش Freak	N.O Of good matches
	74

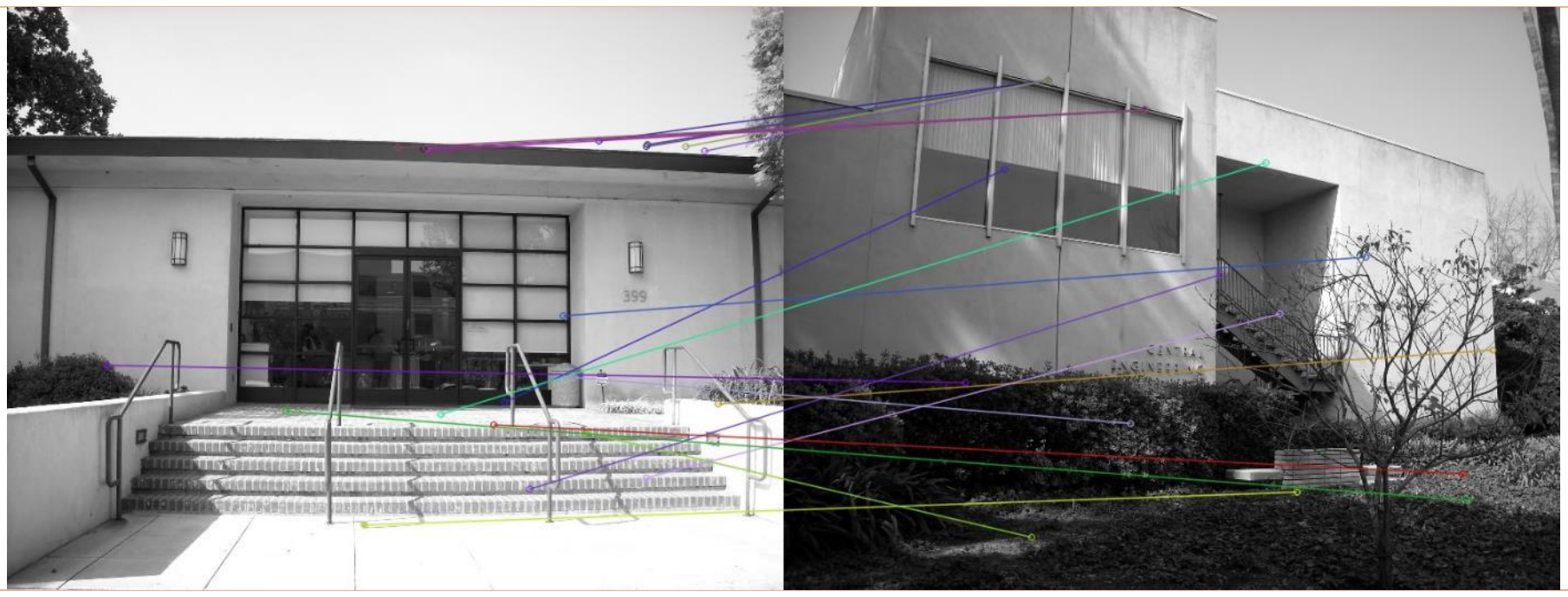
	**18**
	29
	47

انطباق تصویر یک ساختمان با دو تصویر از دو ساختمان دیگر

N.O Of good matches	انطباق ساختمان ۱ با دو تصویر دیگر از دو ساختمن دیگرFreak

	23
	30

انطباق ساختمان ۲ با دو تصویر دیگر از دو ساختمان دیگرFreak	N.O Of good matches
	17

	22

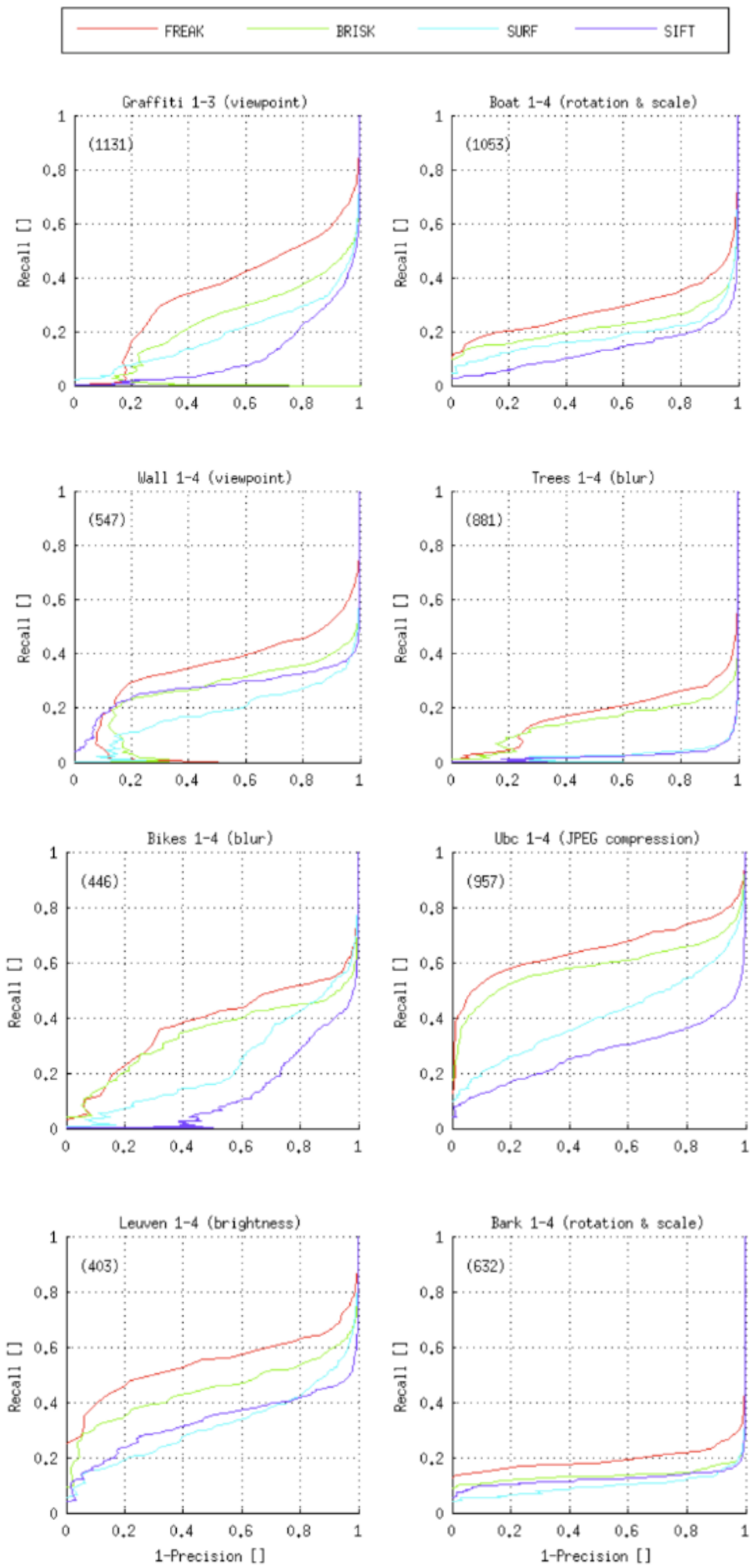
انطباق ساختمان ۳ با دو تصویر دیگر از دو ساختمان دیگرFreak	N.O Of good matches
	20
	13

در ۱۸ تصویر بالا نتایج حاصل از انطباق بین ساختمان‌های مختلف با تصویرهایی از همان ساختمان که دچار چرخش تغییرمقیاس و جابه‌جایی شده اند را مشاهده می‌کنید(از ویژگی Freak استفاده شده است). همچنین انطباق بین ساختمان‌ها با ساختمان‌های دیگر را مشاهده می‌کنید. در سمت راست هر تصویر تعداد انطباق‌های خوب نمایش داده شده است.

در انطباق یک ساختمان با تصاویر دیگر همان ساختمان تعداد انطباق‌های خوب بسیار زیاد است که بیشتر آن‌ها نیز درست هستند، در انطباق یک ساختمان با تصاویر دیگر همان ساختمان نکته‌ی جالبی مشاهده می‌شود: هر چه میزان مقیاس، چرخش و جابه‌جایی شدیدتر بوده است تعداد انطباق‌های خوب کاهش پیدا کرده است و هر چه میزان تغییرات کمتر بوده است تعداد انطباق‌های خوب بیشتر شده است.

در انطباق بین یک ساختمان با ساختمان‌های دیگر، تعداد انطباق‌های خوب به شدت کاهش یافته است، همچنین بیشتر انطباق‌ها در نقاطی مثل برگ‌ها، شیشه‌ها رخ داده است.

مقایسه بین Freak و SURF: در نمایش انطباق‌ها، زیر مجموعه‌ای از انطباق‌ها را نشان دادیم که نام آن‌ها را انطباق‌های خوب گذاشتیم. در انطباق‌ها با استفاده از SURF تعداد انطباق‌های بین دو تصویر بیشتر از تعداد انطباق‌ها در Freak است، **ولی** کیفیت انطباق‌ها در فریک بهتر است و تعداد بیشتر از انطباق‌ها درست هستند. فریک از سیستم بینایی انسان الگوبرداری شده است. به گونه‌ای که ابتدا به صورت کلی انطباق‌ها بررسی می‌شوند و در ادامه در صورت شباهت با جزیئات بیشتری به نقاط کلیدی جهت شباهت پرداخته می‌شود. نتایجی که در مقاله فریک آمده هم در تایید نتایج به دست آمده آزمایش انجام شده است. فریک تحت تغییر روشنایی، جابه‌جایی، مقیاس و چرخش بسیار بهتر از همه روش‌ها عمل کرده است.



عملکر فریک و سورف بهتر از دو روش قبلی (سیفت و اچ.او.جی) است زیرا در سورف از روش Fast Hessian برای پیدا کردن نقاط کلیدی استفاده می‌شود در حالی که در سیفت از روش تخمین DOG برای پیدا کردن نقاط کلیدی پیدا می‌شود، روش مورد استفاده سورف قوی‌تر است در نتیجه نقاط کلیدی بهتری به دست می‌آید. همچنین توصیفگر سورف پیچیده‌تر از توصیف‌گر سیفت است به همین دلیل تطبیق هایی با کیفیتی بهتر به دست می‌آید. سورف و فریک از نظر سرعت بسیار بهتر از سیفت عمل می‌کنند. نتایج به دست آمده در مقاله سورف، نتایج به دست آمده را تایید می‌کنند:

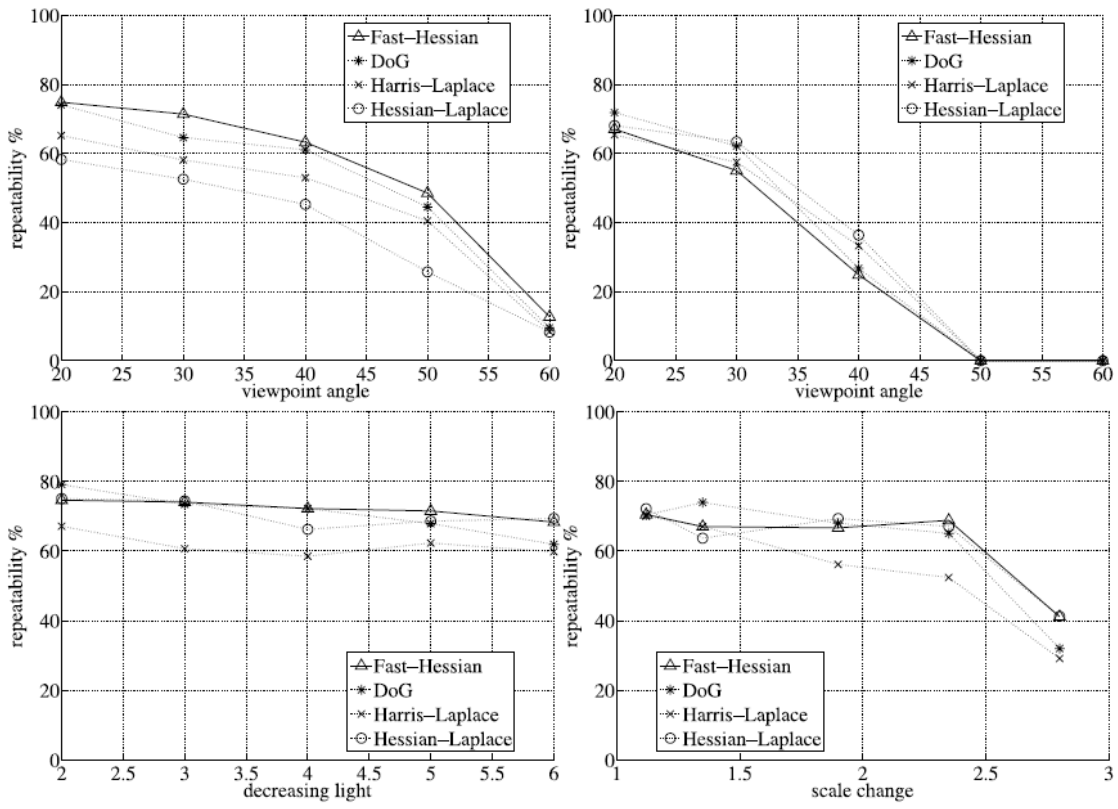
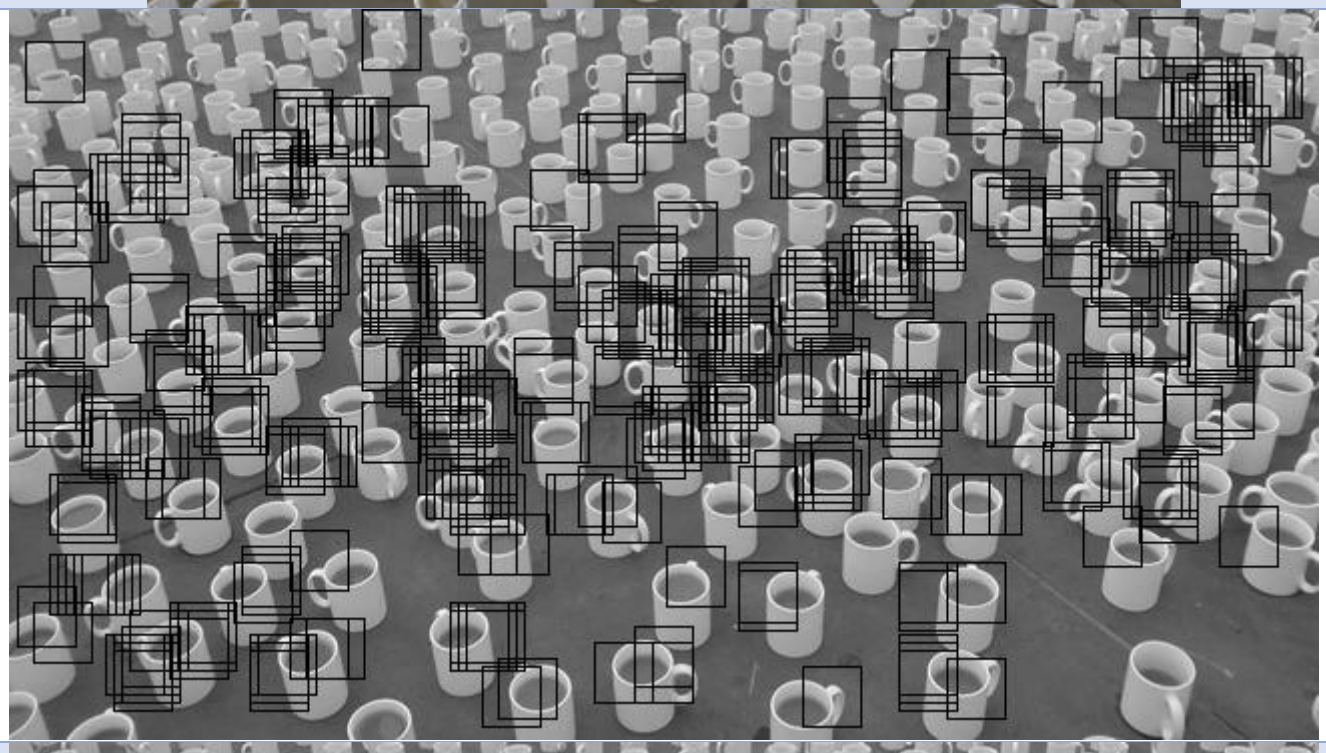


Fig. 6. Repeatability score for image sequences, from left to right and top to bottom, Wall and Graffiti (Viewpoint Change), Leuven (Lighting Change) and Boat (Zoom and Rotation).

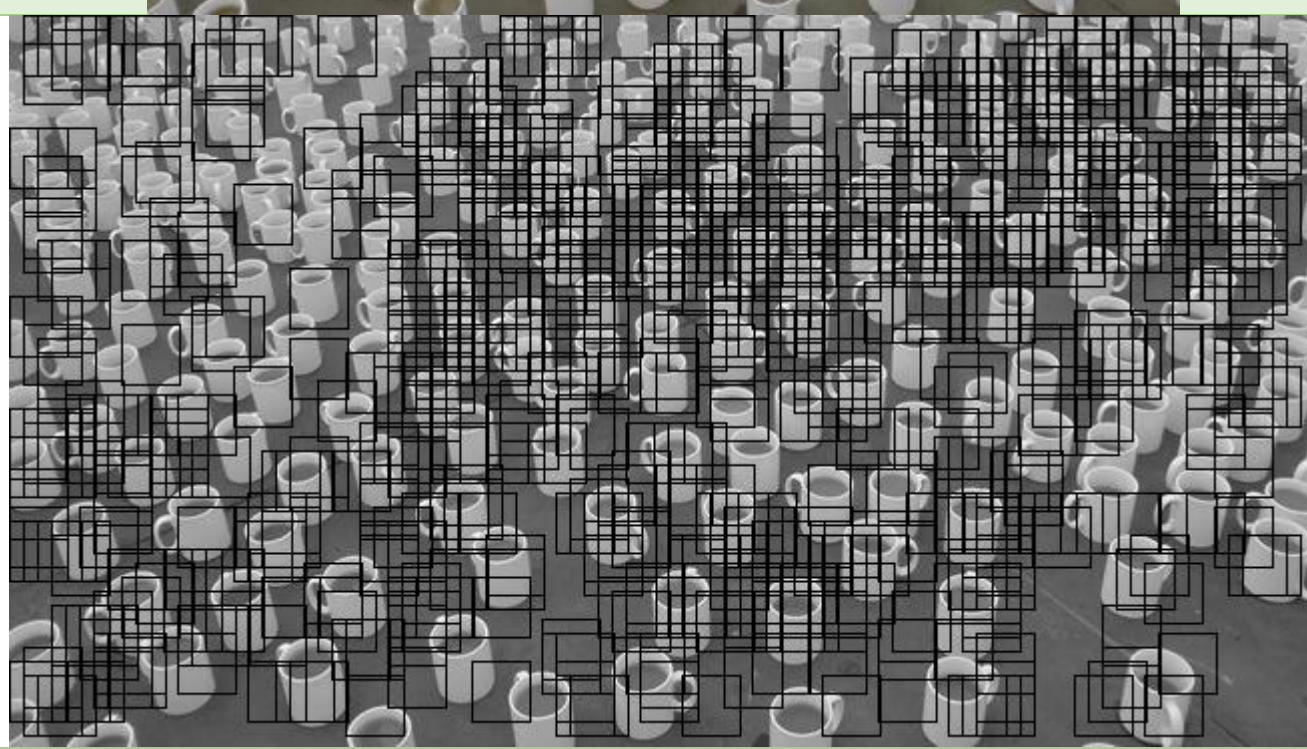
تمرین ۳

کدهای این قسمت در problem-3-surf.py و problem-3-freak.py قرار دارند. برای انجام این قسمت مقداری کلیشه را بزرگ‌تر کرده‌ایم، سپس یک پنجره را بر روی تصویر حرکت می‌دهیم و انطباق نقاط کلیدی را به دست می‌آوریم، اگر تعداد انطباق‌های خوب بیش از حد مشخصی باشد، آن محل را لیوان تشخیص می‌دهیم.

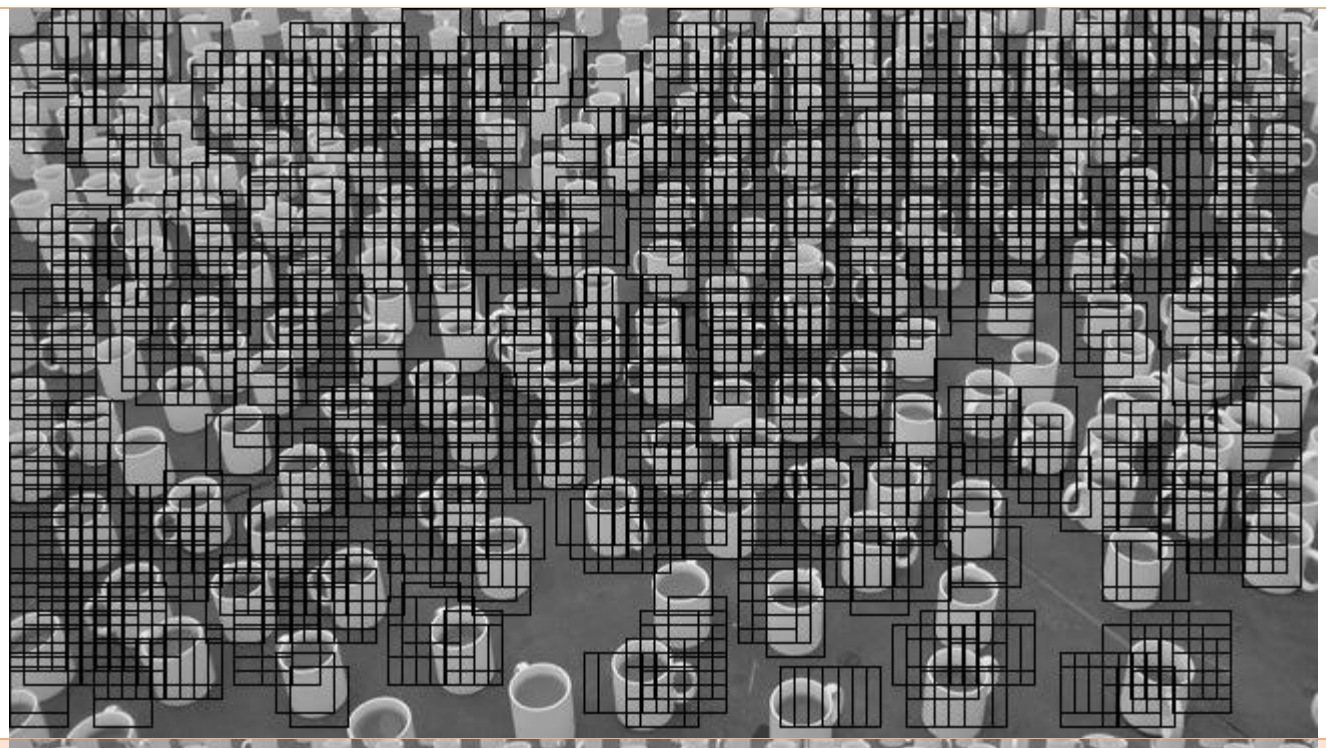
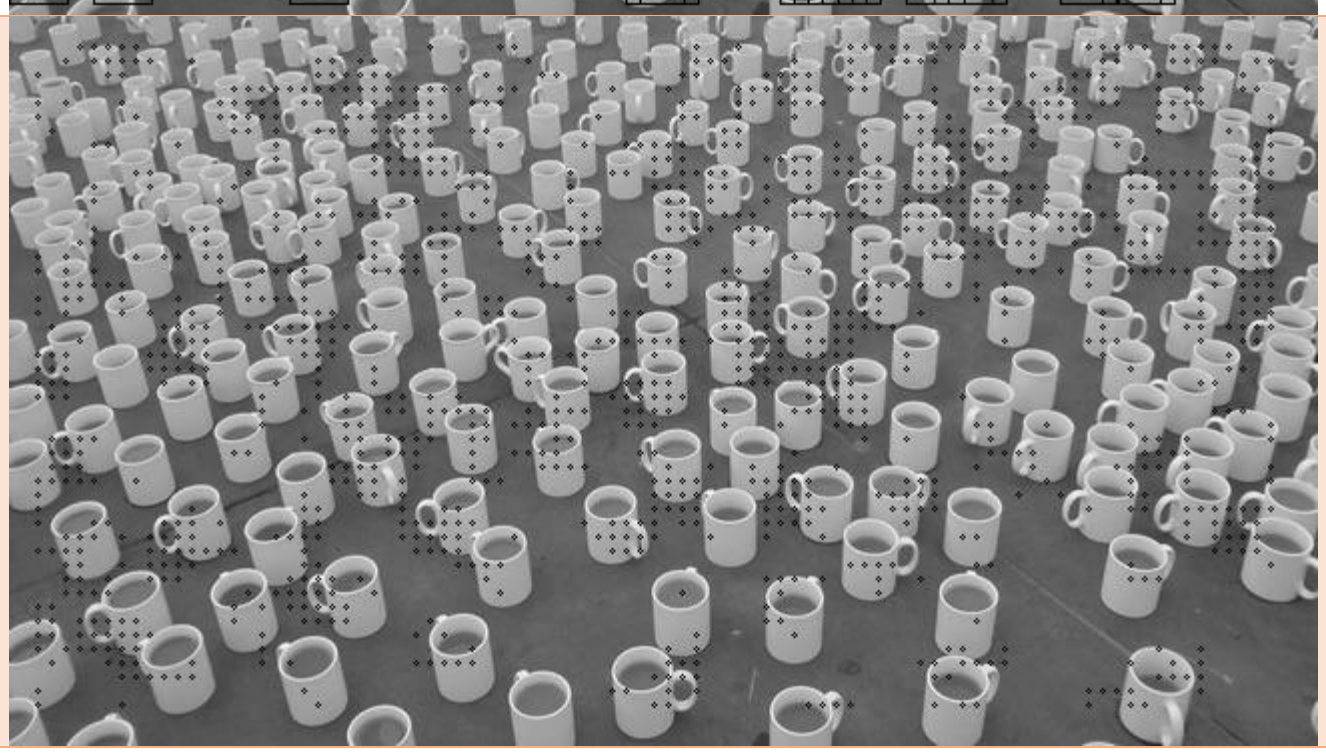
در زیر خروجی این روش را مشاهده می‌کنید:

	روش SURF
تصویر ورودی	
محل لیوان‌ها با استفاده از ویژگی‌های SURF	
محل لیوان‌ها با استفاده از ویژگی‌های SURF	

	روش FREAK

تصویر ورودی	
محل لیوان‌ها با استفاده از ویژگی‌های FREAK	
محل لیوان‌ها با استفاده از ویژگی‌های FREAK	

	SIFT
تصویر ورودی	

		محل لیوان‌ها با استفاده از ویژگی‌های سیفت
		محل لیوان‌ها با استفاده از ویژگی‌های سیفت

در سوال‌های قبل مقایسه هایی انجام شد. در اینجا به خوبی سرعت بیشتر فریک و سورف نسبت به سیفت را مشاهده می‌کنیم. از نظر بهتر بودن ابتدا سیفت بهتر بوده است بعد فریک بعد سورف. دلیل این امر در مقاله زیر با رنگ آبی مشخص شده است، SURF از نظر چرخش و تاری و ... بهتر از سیفت است ولی SIFT از نظر Scale بهتر عمل می‌کند، از آنجایی که لیوان‌ها با اسکیل‌های متفاوت داریم، این فاکتور بسیار تعیین کننده شده است و باعث شده است که سیفت عملکرد بهتری از سورف داشته باشد. از آنجایی که توصیفگر فریک را برای نقاط کلیدی حاصل از سورف استفاده کرده‌ایم، عملکرد ضعیف‌تری در مقابل سیفت داشته است ولی چون توصیفگری قوی‌تر از توصیفگر سورف است، عملکرد فریک از سورف بهتر بوده است.

GRD Journals- Global Research and Development Journal for Engineering | Volume 2 | Issue 4 | March 2017
ISSN: 2455-5703

Comparison of Feature Detection and Matching Approaches: SIFT and SURF

Darshana Mistry
PhD Student
Department of Computer Engineering
Technical Associate, eInfochips Training and Research Academy, Ahmedabad

Asim Banerjee
Professor
Department of Electrical Engineering
Dhirubhai Ambani Institue of Information and Communication Technology, Gandhinagar

Abstract

Feature detection and matching are used in image registration, object tracking, object retrieval etc. There are number of approaches used to detect and matching of features as SIFT (Scale Invariant Feature Transform), SURF (Speeded up Robust Feature), FAST, ORB etc. SIFT and SURF are most useful approaches to detect and matching of features because of it is invariant to scale, rotate, translation, illumination, and blur. In this paper, there is comparison between SIFT and SURF approaches are discussed. SURF is better than SIFT in rotation invariant, blur and warp transform. SIFT is better than SURF in different scale images. SURF is 3 times faster than SIFT because using of integral image and box filter. SIFT and SURF are good in illumination changes images.
Keywords- SIFT (Scale Invariant Feature Transform), SURF (Speeded up Robust Feature), invariant, integral image, box filter