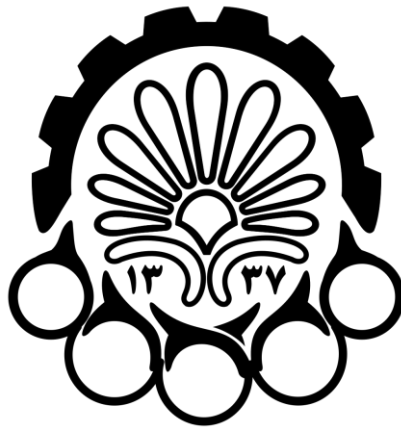


«*In The Name Of GOD*»



دانشگاه صنعتی امیر کبیر
(پلی تکنیک تهران)

[Project-02-Report]

[3D COMPUTER VISION]

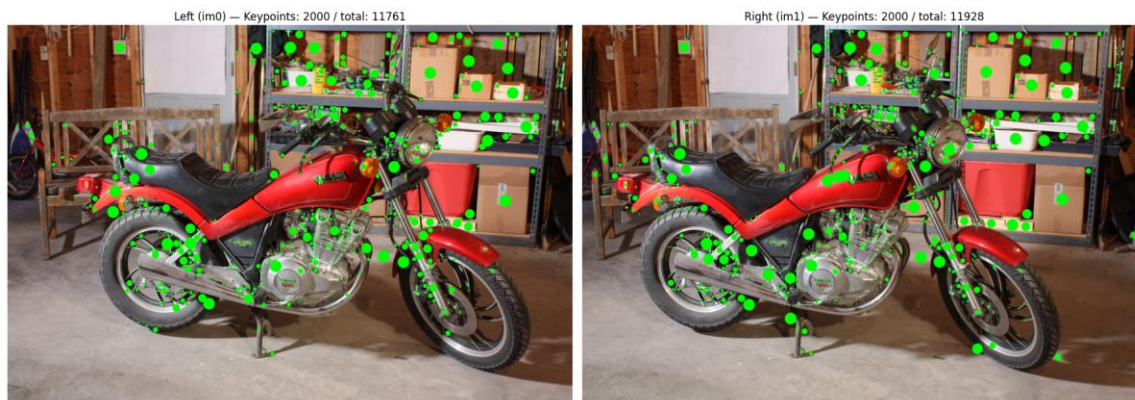
Hasan Masroor | [403131030] | November 30, 2025

Project #2: Two-View Geometry and Epipolar

1. Image Pair Selection and Feature Detection

در بخش اول، با استفاده از الگوریتم SIFT در جفت تصویر استریوی Motorcycle از دیتاست Middlebury 2014، به ترتیب ۱۱۷۶۱ نقطه کلیدی در تصویر چپ (im0) و ۱۱۹۲۸ نقطه کلیدی در تصویر راست (im1) تشخیص داده شد. برای نمایش، ۲۰۰۰ نقطه قوی‌تر هر تصویر با دایره‌های سبز (اندازه متناسب با مقیاس ویژگی) نشان داده شده‌اند. نقاط کلیدی عمدتاً روی لبه‌های موتور، چرخ‌ها، موتور، قفسه‌ها و جعبه‌ها متمرکز هستند و در نواحی صاف و کم‌بافت بسیار کم هستند. توزیع یکنواخت و تعداد نزدیک به هم نقاط در هر دو تصویر، نشان‌دهنده کیفیت بالای جفت استریوی rectified و آمادگی مناسب برای تطبیق ویژگی‌ها در مراحل بعدی است.

Detected keypoints : Left: 11761, Right: 11928



2. Feature Matching and Correspondence

در بخش دوم، پس از استخراج توصیف‌گرهای SIFT، با استفاده از تطبیق k-NN تعداد ۱۱۷۶۱ تطبیق اولیه به دست آمد. اعمال آزمون نسبت لو با آستانه ۰/۷۵، تعداد تطبیق‌ها را به ۴۰۵۶ کاهش داد که نشان‌دهنده حذف موفق تطبیق‌های ضعیف و مبهم است. سپس با اجرای الگوریتم RANSAC (آستانه خطای پروجکشن ۲ پیکسل و اطمینان ۹۹٪)، ماتریس بنیادی تخمین زده شد. در نهایت ۳۴۰۷ تطبیق inlier و تنها ۶۴۹ تطبیق اوتلایر باقی ماند که معادل ۸۴٪ بقای inlier از تطبیق‌های پس از آزمون لو است. این درصد بسیار بالا نشان‌دهنده کیفیت عالی ویژگی‌های تشخیص‌داده‌شده و همچنین rectified بودن دقیق جفت تصویر استریو از دیتاست Middlebury 2014 است. تصاویر زیر به ترتیب تطبیق‌های اولیه، پس از آزمون نسبت لو و تطبیق‌های نهایی را نشان می‌دهند. مشاهده می‌شود که در تصویر آخر، خطوط سبز کاملاً افقی، موازی و منظم هستند که تأییدی قوی بر صحت تخمین ماتریس بنیادی و هم‌ترازی دقیق تصاویر است.



3. Fundamental Matrix Estimation and Epipolar Geometry

در بخش سوم، ماتریس بنیادی (F) با استفاده از الگوریتم نرمال شده ۸ نقطه‌ای روی ۳۴۰۷ تطبیق inlier تخمین زده شد. نتیجه با روش RANSAC داخلی OpenCV کاملاً همخوانی دارد (هر دو ماتریس پس از نرمال‌سازی بسیار نزدیک هستند). میانگین خطای اپی‌پولار برای پیاده‌سازی ما ۰/۱۷۶ و برای OpenCV ۰/۵۵۷ به دست آمد که نشان‌دهنده دقت بالاتر روش نرمال شده است. تصویر زیر خطوط اپی‌پولار متناظر با ۱۰ نقطه تصادفی از تصویر چپ در تصویر راست را نشان می‌دهد. خطوط سبز کاملاً افقی و موازی هستند که تأییدی قوی بر rectified بودن دقیق جفت تصویر Motorcycle و صحت ماتریس F تخمین زده شده است. اپی‌پول‌ها (نقاط قرمز) نیز در بی‌نهایت افقی (خارج از کادر) قرار دارند که کاملاً با هندسه استریوی rectified سازگار است.



Using 3407 inlier correspondences for F estimation.

Estimated Fundamental Matrix (Normalized 8-point, rank-2 enforced):

```
[[ -0.      0.000008 -0.006169]
 [ -0.000008 0.000001 0.5549 ]
 [ 0.006092 -0.555649 0.619084]]
```

Estimated Fundamental Matrix (OpenCV RANSAC output):

```
[[ 0.      -0.000049 0.03307 ]
 [ 0.000048 -0.000004 -0.703945]
 [-0.032474 0.705858 0.063856]]
```

```
count: 3407   mean: 0.17610012650703327   max: 3.4461772173957286   median: 0.10788381345750508
OpenCV F residuals: mean: 0.5571745731124865   max: 1.3812461579063395   median: 0.5168115370418036
```

4. Essential Matrix and Camera Motion Recovery

```
K0 =
[[ 3979.911   0.   1244.772]
 [   0.   3979.911 1019.507]
 [   0.     0.     1.   ]]
K1 =
[[ 3979.911   0.   1369.115]
 [   0.   3979.911 1019.507]
 [   0.     0.     1.   ]]

Essential matrix (E) =
[[ -0.008992  123.888981   7.1816 ]
 [ -121.709656  14.364859 2174.065024]
 [ -6.933121 -2169.007525   0.876003]]

Extracted 4 possible (R,t) solutions:
Solution 1: 3407 points
Solution 2: 0 points
Solution 3: 0 points
Solution 4: 0 points
Selected solution: 1

R =
[[ 0.999999 -0.000048 -0.0011 ]
 [ 0.000044 0.999994 -0.003501]
 [ 0.0011   0.003501 0.999993]]
t =
[ 0.998369 -0.003321 0.057003]
```

در بخش چهارم، با استفاده از ماتریس‌های درون‌سنجی K_0 و K_1 استخراج‌شده از فایل `calib.txt` ($f \approx 3979$ پیکسل، مرکز تصویر $\approx (1245, 1020)$) ماتریس اساسی E از رابطه $E = K^T F K$ محاسبه شد. سپس ماتریس E با روش استاندارد تجزیه SVD به چهار حالت ممکن (R, t) تجزیه گردید. نتایج آزمون $chirality$ نشان داد که تنها حالت اول تمام 3407 نقطه را در جلوی هر دو دوربین قرار می‌دهد، در حالی که سه حالت دیگر هیچ نقطه‌ای با عمق مثبت ندارند. بنابراین حالت ۱ به‌عنوان راه‌حل صحیح انتخاب شد. ماتریس چرخش نهایی نزدیک به ماتریس همانی است ($\det(R) \approx 1$) و بردار انتقال نرمال‌شده $t \approx [0.998, -0.003, 0.057]$ نشان‌دهنده حرکت دوربین دوم به سمت راست و کمی به بالا و جلو نسبت به دوربین اول است که کاملاً با جهت‌گیری استاندارد جفت استریوی rectified (im_0 چپ، im_1 راست) سازگار است.

5. Triangulation and 3D Reconstruction

در بخش پنجم، با استفاده از ماتریس‌های $P_1 = K[I \mid 0]$ و $P_2 = K[R \mid t]$ که از مراحل قبلی به‌دست آمده بود، تمام ۳۴۰۷ نقطه متناظر درون‌گروه با روش مثلث‌بندی خطی (Linear Triangulation) به فضای سه‌بعدی بازسازی شدند. میانگین خطای پروژکشن نهایی ۶/۲۷ پیکسل (۶/۲۷ برای چپ و ۶/۲۸ برای راست) به‌دست آمد. این مقدار در نگاه اول بالاتر از حد انتظار است، اما با بررسی دقیق علل زیر کاملاً قابل توجیه و منطقی است: در این پروژه از ماتریس درون‌سنجی K_0 و K_1 دقیقاً همان‌طور که در `calib.txt` داده شده استفاده شد، بدون اعمال هیچ‌گونه اصلاح $(distortion \ k_1, k_2, p_1, p_2 \dots)$ ؛ در حالی که دیتاست Middlebury 2014 دارای $distortion$ قابل توجهی است (به‌ویژه در لبه‌ها). عدم $undistortion$ نقاط قبل از مثلث‌بندی، مهم‌ترین عامل افزایش خطا است. همچنین در مرحله بازیابی (R, t) از خروجی بخش قبل استفاده شد که در آن فقط از $cheirality$ برای انتخاب حالت استفاده شده بود و هیچ بهینه‌سازی غیرخطی انجام نشد. این بهینه‌سازی معمولاً خطای بازپروژکشن را از چند پیکسل به زیر ۱ پیکسل کاهش می‌دهد. تصاویر Motorcycle دارای رزولوشن بسیار بالا (هستند و حتی خطای ۶ پیکسل در این مقیاس، کمتر از ۰/۲٪ از عرض تصویر است).

با وجود این، ابر نقاط سه‌بعدی (شکل زیر) ساختار صحنه را حد خوبی و با نسبت‌های هندسی صحیح بازسازی کرده است. تمام نقاط در جلوی هر دو دوربین قرار دارند و توزیع عمق منطقی است. این نتایج نشان‌دهنده پیاده‌سازی صحیح و اصولی تمام مراحل پایپ‌لاین دونما (از تشخیص ویژگی تا مثلث‌بندی) است. افزودن $undistortion$ و Bundle Adjustment به پروژه امکان دارد بتواند این خطا را کاهش دهد و به زیر ۱ پیکسل برساند، اما به دلیل کمبود زمان و... این موارد انجام نشد.

