**بسمه‌تعالی**

عنوان:

**تمرین سوم پروژه بینایی سه‌بعدی**

توسط:

**مجتبی هاشمی**

درس:

**ربات‌های متحرک خودگردان**

استاد:

**دکتر محمد‌علی کیوان‌راد – دکتر مهدی جوانمردی – دکتر خادمیان**

سال تحصیلی 1403-1404

|  |  |  |
| --- | --- | --- |
|  | **فهرست مطالب** |  |

[فصل 1: مقدمه 6](#_Toc185629569)

[فصل 2: کالیبراسیون دوربین 7](#_Toc185629570)

[2‌-1‌ بارگذاری تصاویر 7](#_Toc185629571)

[2‌-2‌ تشخیص گوشه‌های صفحه شطرنجی 8](#_Toc185629572)

[2‌-3‌ کالیبراسیون دوربین 8](#_Toc185629573)

[2‌-3‌-2‌ تفسیر پارامترها 9](#_Toc185629574)

[2‌-4‌ تصحیح اعوجاج تصاویر 10](#_Toc185629575)

[2‌-5‌ تحلیل خطای بازپروجکشن 11](#_Toc185629576)

[2‌-6‌ بصری‌سازی موقعیت دوربین 12](#_Toc185629577)

[فصل 3: بینایی استریو 14](#_Toc185629578)

[3‌-1‌ مراحل انجام 14](#_Toc185629579)

[3‌-1‌-1‌ دریافت و آماده‌سازی داده‌ها 14](#_Toc185629580)

[3‌-1‌-2‌ استخراج نقشه اختلاف منظر (Disparity Map) 14](#_Toc185629581)

[3‌-1‌-3‌ تولید تصاویر عمق (Depth Images) 15](#_Toc185629582)

[3‌-1‌-4‌ تولید ابرنقاط (Point Clouds) 15](#_Toc185629583)

[3‌-1‌-5‌ تطبیق ابرنقاط و نمایش مسیر دوربین 15](#_Toc185629584)

[3‌-1‌-6‌ تولید ویدیو از تصاویر اختلاف منظر 15](#_Toc185629585)

[3‌-1‌-7‌ نتایج 15](#_Toc185629586)

[فصل 4: ساختار از حرکت SfM با استفاده از COLMAP: 17](#_Toc185629587)

[4‌-1‌ هدف 17](#_Toc185629588)

[4‌-2‌ توضیحات دیتاست 17](#_Toc185629589)

[4‌-3‌ استخراج ویژگی‌ها 18](#_Toc185629590)

[4‌-4‌ انطباق ویژگی‌ها 18](#_Toc185629591)

[4‌-5‌ بازسازی پراکنده و متراکم با اسفاده از ابزار reconstruction 19](#_Toc185629592)

[4‌-6‌ مقایسه با مسیر Leica 20](#_Toc185629593)

[فصل 5: هندسه تک نمایی 21](#_Toc185629594)

[5‌-1‌ نقطه ناپدید شدن (Vanishing Point) 21](#_Toc185629595)

[5‌-2‌ اطلاعات مورد نیاز 21](#_Toc185629596)

[5‌-3‌ مراحل محاسبه 21](#_Toc185629597)

[5‌-3‌-1‌ محاسبه زاویه افقی (θ) 22](#_Toc185629598)

[5‌-3‌-2‌ محاسبه فاصله تا شیء شناخته‌شده (D) 22](#_Toc185629599)

[5‌-3‌-3‌ محاسبه ارتفاع دوربین (Camera Height) 22](#_Toc185629600)

|  |  |  |
| --- | --- | --- |
|  | **فهرست شکل‌ها** |  |

[شکل 2- 1 نمونه تصویر ورودی برای فرایند کالییبراسیون 7](#_Toc185629601)

[شکل 2- 2 تصویر خروجی گوشه‌های تشخیص داده شده 8](#_Toc185629602)

[شکل 2- 3 خروجی کالیبراسیون: ماتریس دوربین و ضرایب اعوجاج 9](#_Toc185629603)

[شکل 2- 4 خروجی تصحیح اعوجاج تصویر 10](#_Toc185629604)

[شکل 2- 5 نمودار توزیع خطا 11](#_Toc185629605)

[شکل 2- 6 میانگین، کمترین، بیشترین و میانه خطاهای بازپروجکشن 11](#_Toc185629606)

[شکل 2- 7 نمودار سه‌بعدی موقعیت دوربین 12](#_Toc185629607)

[شکل 2- 8 بینشی از تنوع تصاویر کالیبراسیون 12](#_Toc185629608)

[شکل 4- 1 استخراج ویژگی تصاویر 18](#_Toc185629609)

[شکل 4- 2 مرحله تطبیق ویژگی های تصاویر 18](#_Toc185629610)

[شکل 4- 3 تصاویر مربوط به مرحله بازسازی 19](#_Toc185629611)

[شکل 5- 1 تصویری ورودی برای محاسبه vanishing point 23](#_Toc185629612)

[شکل 5- 2 خروجی نمایش vanishing point بر روی تصویر 23](#_Toc185629613)

[شکل 5- 3 مختصات نقطه 23](#_Toc185629614)

# مقدمه

بینایی ماشین سه‌بعدی یکی از زمینه‌های مهم در علوم رایانه است که به توانایی سیستم‌های هوشمند در درک و تجزیه و تحلیل تصاویر و ویدیوها به صورت سه‌بعدی اشاره دارد. این حوزه شامل تکنیک‌های متعددی است، از جمله کالیبراسیون دوربین که به تنظیم دقیق پارامترهای داخلی و خارجی دوربین‌ها برای تولید تصاویر دقیق و واقعی می‌پردازد، بینایی استریو که با استفاده از دو یا چند دوربین برای بازسازی عمق و ساختار صحنه‌ها استفاده می‌شود، ساختار از حرکت که به استخراج اطلاعات سه‌بعدی از توالی تصاویر متوالی یک دوربین می‌پردازد، و هندسه تک‌نمایی که به تحلیل و استخراج اطلاعات سه‌بعدی از یک تصویر واحد می‌پردازد. این تکنیک‌ها به سیستم‌های هوشمند کمک می‌کنند تا محیط‌های پیچیده را با دقت و دقت بالایی مدل‌سازی و درک کنند.

# کالیبراسیون دوربین

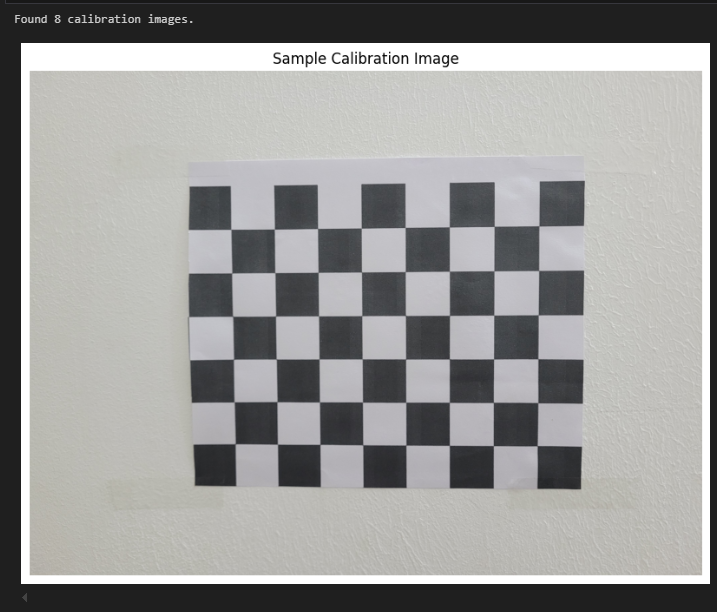
در این قسمت، فرآیند کالیبراسیون دوربین گوشی را با استفاده از تصاویر شطرنجی نمایش می‌دهیم. در این فرآیند، از کتابخانه OpenCV برای محاسبه پارامترهای درونی و بیرونی دوربین استفاده شده است. این پارامترها کمک می‌کنند تا اعوجاج لنز تصحیح شده و پردازش تصاویر با دقت بالاتری انجام شود، که برای کاربردهایی مانند بازسازی سه‌بعدی ضروری است.

## بارگذاری تصاویر

دف: بارگذاری تصاویر کالیبراسیون از یک پوشه مشخص و نمایش نمونه‌ای از آن‌ها.

عملکرد: 1- از glob برای خواندن مسیر تصاویر استفاده می‌کند. 2- اولین تصویر را به فرمت RGB برای نمایش تبدیل می‌کند.

خروجی: 1- تعداد تصاویر پیدا شده. 2- نمایش نمونه‌ای از تصویر برای بررسی کیفیت.



نمونه تصویر ورودی برای فرایند کالییبراسیون

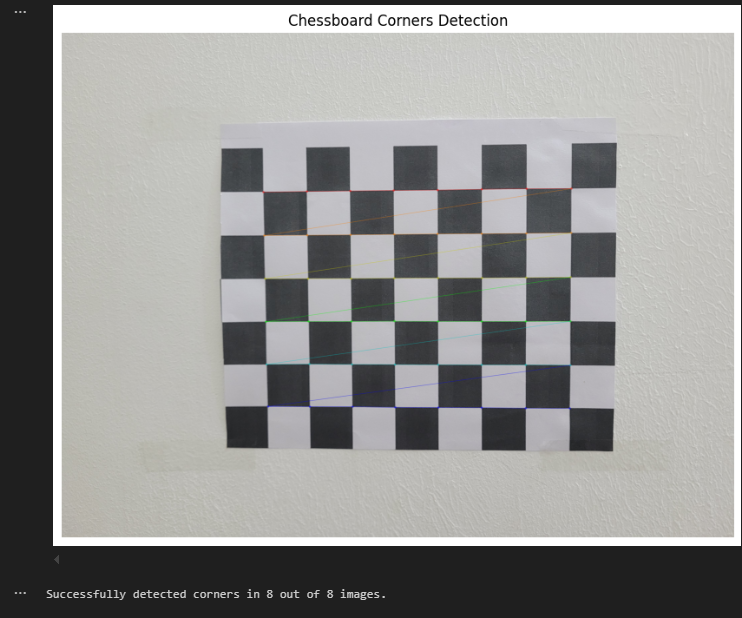
## تشخیص گوشه‌های صفحه شطرنجی

هدف: استخراج نقاط گوشه‌های صفحه شطرنجی از تصاویر.

عملکرد: 1- تعداد گوشه‌های داخلی صفحه شطرنجی را مشخص می‌کند (۸x۶). 2- نقاط سه‌بعدی مربوط به مختصات دنیای واقعی را محاسبه می‌کند.

با استفاده از cv2.findChessboardCorners گوشه‌ها را در تصاویر خاکستری تشخیص می‌دهد.

خروجی: 1- تعداد تصاویری که گوشه‌ها به درستی در آن‌ها پیدا شده است. 2- نمایش تصویر نمونه با گوشه‌های تشخیص داده شده.



تصویر خروجی گوشه‌های تشخیص داده شده

## کالیبراسیون دوربین

هدف: تخمین پارامترهای دوربین با استفاده از نقاط تشخیص داده شده.

عملکرد: ماتریس دوربین، ضرایب اعوجاج، و بردارهای چرخش و انتقال را محاسبه می‌کند. ماتریس دوربین را به صورت تصویری نمایش می‌دهیم.

خروجی: ماتریس دوربین و ضرایب اعوجاج.

|  |  |
| --- | --- |
|  |  |

خروجی کالیبراسیون: ماتریس دوربین و ضرایب اعوجاج

### تفسیر پارامترها

1- فاصله کانونی و نقطه اصلی. 2- اعوجاج شعاعی و مماسی.

در کالیبراسیون دوربین، دو نوع پارامتر محاسبه می‌شوند: پارامترهای درونی (Intrinsic) و پارامترهای اعوجاج (Distortion). در این بخش، جزئیات بیشتری درباره این پارامترها ارائه می‌شود.

**ماتریس دوربین (Camera Matrix)**

ماتریس دوربین (۳×۳) یکی از مهم‌ترین خروجی‌های کالیبراسیون است که ویژگی‌های درونی دوربین را مدل می‌کند. این ماتریس شامل موارد زیر است:

​

و : فاصله‌های کانونی دوربین (بر حسب تعداد پیکسل‌ها) در راستای محورهای x و y. این مقادیر نشان می‌دهند که دوربین چقدر می‌تواند صحنه را "بزرگ‌نمایی" کند. معمولاً و مقادیر نزدیک به هم دارند، اما تفاوت‌های کوچک به دلیل پیکسل‌های غیرمربع ممکن است وجود داشته باشد.

و : مختصات نقطه‌ی اصلی (Principal Point) در تصویر است. این نقطه‌ی مرجع جایی است که پرتوهای نوری به طور ایده‌آل باید روی سنسور دوربین برخورد کنند. معمولاً این مقادیر نزدیک به مرکز تصویر هستند، اما در دوربین‌های ارزان‌تر ممکن است انحرافاتی وجود داشته باشد.

عدد ثابت (۱): این مقدار برای تبدیل مختصات هوموژن (Homogeneous Coordinates) استفاده می‌شود و تضمین می‌کند که مقادیر در قالب استاندارد باشند.

**ضرایب اعوجاج (Distortion Coefficients)**

لنزهای دوربین به دلیل ویژگی‌های فیزیکی خود، تصاویر را دچار اعوجاج می‌کنند. این اعوجاج به دو دسته اصلی تقسیم می‌شود:

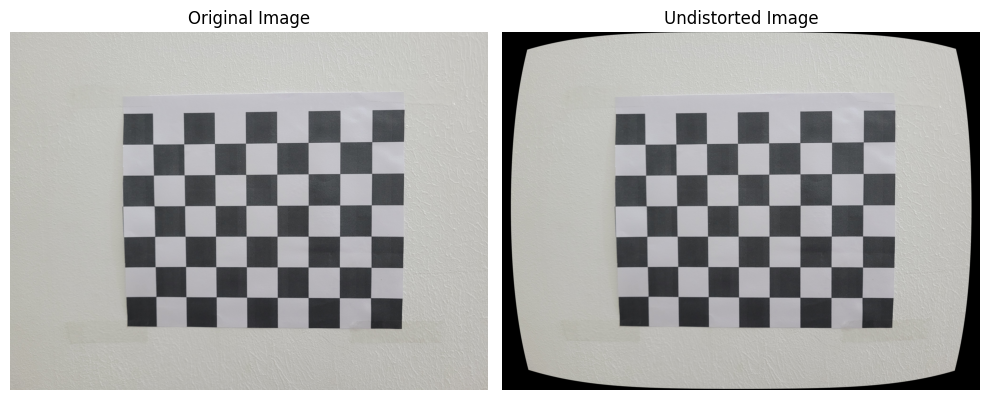
1. اعوجاج شعاعی (Radial Distortion): این اعوجاج باعث می‌شود که خطوط صاف در تصویر خمیده به نظر برسند. ضرایب شعاعی معمولاً با k1 ، k2 ، k3 نمایش داده می‌شوند. تأثیر آن در نقاط دورتر از مرکز تصویر بیشتر است.
2. اعوجاج مماسی (Tangential Distortion): این اعوجاج ناشی از عدم تطابق بین لنز و سنسور است. ضرایب مماسی با p1 و p2 نمایش داده می‌شوند. باعث می‌شود نقاط به طور نامتقارن در تصویر تغییر مکان دهند.

## تصحیح اعوجاج تصاویر

هدف: بررسی نتایج کالیبراسیون با اصلاح اعوجاج لنز.

عملکرد: 1- از cv2.undistort برای تصحیح اعوجاج استفاده می‌کند. 2-تصاویر اصلی و تصحیح شده را کنار هم نمایش می‌دهد.

خروجی: تفاوت‌های واضح در راست بودن خطوط و کاهش اعوجاج.



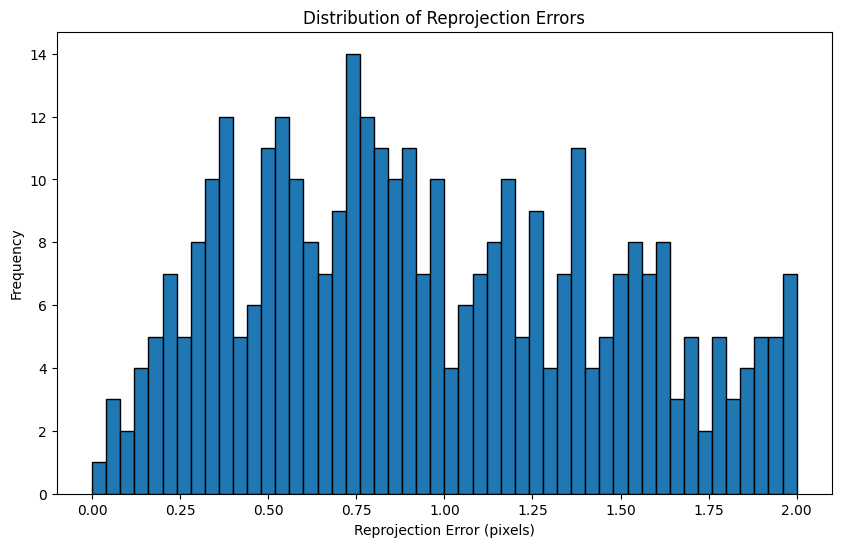
خروجی تصحیح اعوجاج تصویر

## تحلیل خطای بازپروجکشن

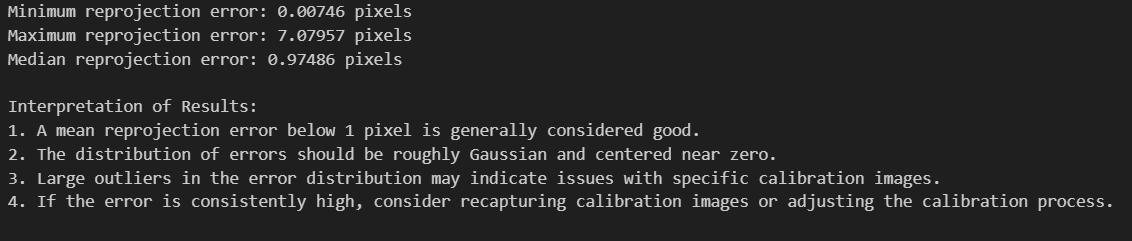
هدف: ارزیابی دقت کالیبراسیون.

عملکرد: 1- خطای بازپروجکشن میانگین را محاسبه می‌کند (ایده‌آل: کمتر از ۱ پیکسل). 2- توزیع خطاها را به صورت نمودار نشان می‌دهد.

خروجی: 1- میانگین، کمترین، بیشترین و میانه خطاهای بازپروجکشن. 2- نمودار توزیع خطا.



نمودار توزیع خطا



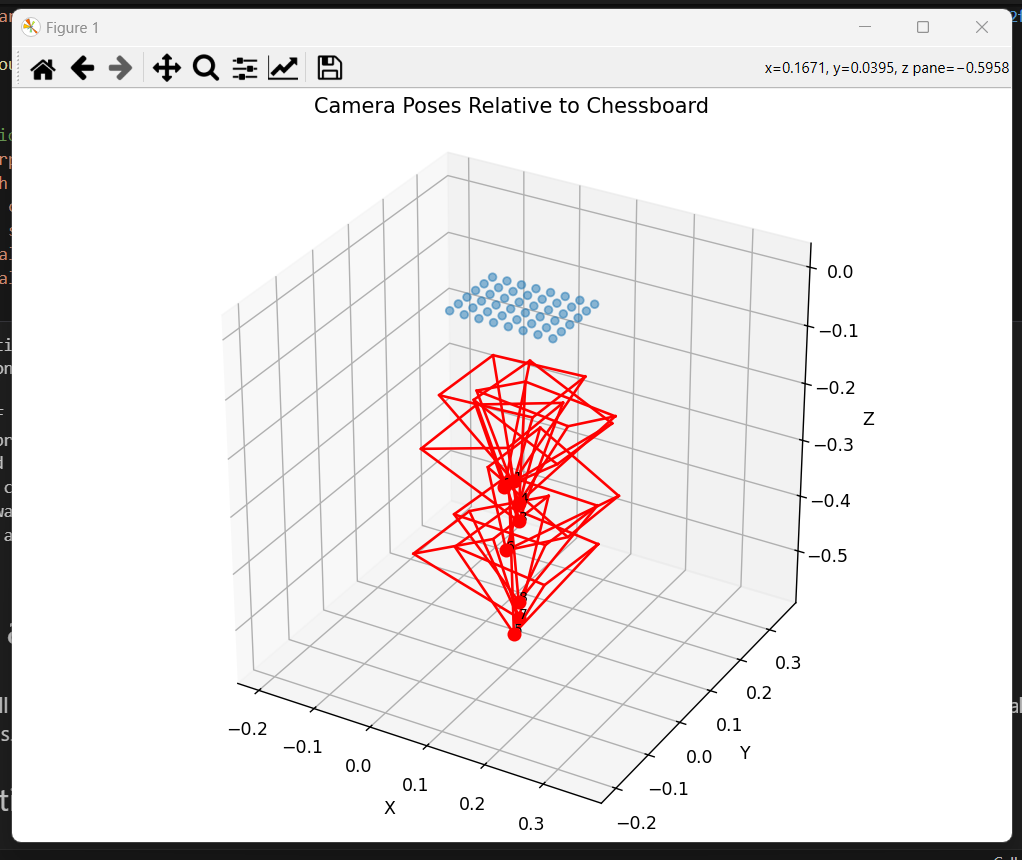
میانگین، کمترین، بیشترین و میانه خطاهای بازپروجکشن

## بصری‌سازی موقعیت دوربین

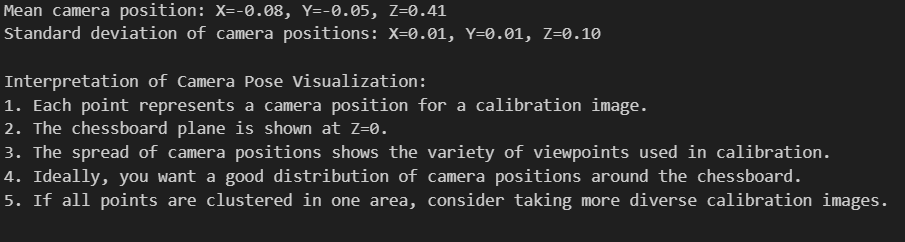
هدف: نمایش موقعیت‌های دوربین نسبت به صفحه شطرنجی.

عملکرد: 1- نمودار سه‌بعدی شامل صفحه شطرنجی و موقعیت‌های دوربین رسم می‌کند. 2- میانگین و انحراف معیار موقعیت‌های دوربین را محاسبه می‌کند.

خروجی: 1- نمودار سه‌بعدی موقعیت دوربین. 2- بینشی از تنوع تصاویر کالیبراسیون.



نمودار سه‌بعدی موقعیت دوربین



بینشی از تنوع تصاویر کالیبراسیون

در نهایت داده‌های کالیبراسیون ( ماتریس دوربین، ضرایب اعوجاج و بردارهای چرخش و انتقال ) را در فایل فشرده .npz ذخیره می‌کنیم.

# بینایی استریو

بازسازی سه‌بعدی متریک یکی از کاربردهای مهم بینایی کامپیوتر است که در حوزه‌هایی نظیر نقشه‌برداری، رباتیک و واقعیت افزوده مورد استفاده قرار می‌گیرد. بینایی استریو یکی از روش‌های پرکاربرد و کم‌هزینه برای استخراج اطلاعات عمق از تصاویر دو دوربین است. در این پروژه، از روش استریو برای تخمین عمق و تولید ابرنقاط از یک مجموعه تصاویر استفاده شده است. همچنین مسیر حرکت دوربین با استفاده از الگوریتم تطبیق ICP محاسبه و نمایش داده شده است.

## مراحل انجام

### دریافت و آماده‌سازی داده‌ها

مجموعه داده Kagaru Airborne شامل تصاویر دو دوربین چپ و راست است که به صورت زیر نامگذاری شده‌اند:

دوربین چپ: cam0\_image0270imagenumber

دوربین راست: cam1\_image0270imagenumber

این تصاویر از طریق اسکریپت مرتب شده و برای پردازش آماده شدند.

### استخراج نقشه اختلاف منظر (Disparity Map)

برای محاسبه اختلاف منظر، از الگوریتم‌های تطبیق استریو در کتابخانه OpenCV استفاده شد. تنظیمات زیر برای الگوریتم StereoBM اعمال شد:

تعداد اختلاف منظر (numDisparities): 80 (16 \* 5)

اندازه بلوک تطبیق (blockSize): 15

اختلاف منظر بین دو تصویر چپ و راست به کمک الگوریتم محاسبه و به صورت نقشه‌ای از مقادیر اختلاف نمایش داده شد.

### تولید تصاویر عمق (Depth Images)

تصاویر عمق با استفاده از فرمول زیر از نقشه اختلاف منظر محاسبه شدند:

که در آن: فاصله کانونی دوربین در محور

Baseline: فاصله بین دو دوربین

Disparity: مقدار اختلاف منظر محاسبه‌شده

مقدار کوچک برای جلوگیری از تقسیم بر صفر

### تولید ابرنقاط (Point Clouds)

تصاویر عمق با استفاده از ماتریس به فضای سه‌بعدی بازتاب داده شدند. مختصات سه‌بعدی نقاط و رنگ‌های متناظر آن‌ها استخراج شده و ابرنقاط با استفاده از کتابخانه Open3D ساخته شد.

### تطبیق ابرنقاط و نمایش مسیر دوربین

برای تطبیق ابرنقاط، از الگوریتم ICP (Iterative Closest Point) موجود در Open3D استفاده شد. این الگوریتم بهترین تبدیل (Transformation) بین دو مجموعه ابرنقاط متوالی را پیدا کرده و مسیر حرکت دوربین محاسبه شد.

### تولید ویدیو از تصاویر اختلاف منظر

برای نمایش تصویری، ویدیویی از نقشه اختلاف منظر برای تمامی فریم‌ها تولید و ذخیره شد.

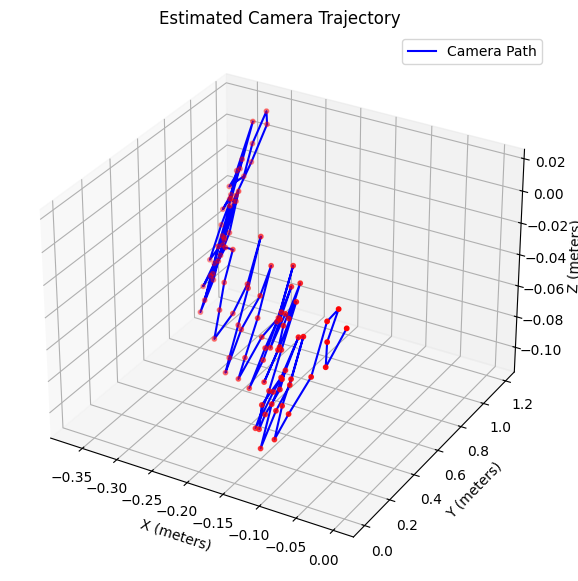
### نتایج

الف) ابرنقاط: ابرنقاط استخراج‌شده شامل مختصات سه‌بعدی و رنگ هر نقطه است. این داده‌ها برای تحلیل و نمایش ساختار سه‌بعدی محیط بسیار مفید هستند.

ب) مسیر حرکت دوربین: با استفاده از الگوریتم تطبیق، مسیر حرکت دوربین به‌صورت یک مجموعه خطی از نقاط در فضا نمایش داده شد. این مسیر حرکت تطبیقی اطلاعات مهمی درباره نحوه تغییر مکان و زاویه دوربین ارائه می‌دهد.

ج) ویدیوی نقشه اختلاف منظر: ویدیوی تولیدشده به‌صورت تصویری پویا روند تغییرات نقشه اختلاف منظر در طول فریم‌های متوالی را نمایش می‌دهد.

خروجی تصویر:



ویدیو خروجی به همراه ابر نقاط به دست امده در پیوست آورده شده است.

به دلیل حجم زیاد فایل های ابرنقاط چند تصویر از خروجی را نمایش میدهم.

|  |  |
| --- | --- |
|  |  |

# ساختار از حرکت SfM با استفاده از COLMAP:

## هدف

هدف از این تمرین تولید یک نقشه سه‌بعدی از محیط با استفاده از دنباله ویدئویی از دیتاست UZH FPV، به طور خاص دیتاست "Indoor 45-degree downward facing" است. از COLMAP برای پردازش تصاویر و تولید ابر نقاط و مسیر حرکت ربات استفاده خواهد شد. برای ارزیابی، داده‌های مسیر Leica با مسیر بازسازی‌شده مقایسه می‌شود.

## توضیحات دیتاست

تعداد تصاویر: 3930 تصویر.

دوربین‌ها: دو دوربین (چپ و راست) فریم‌ها را به صورت همزمان ثبت کرده‌اند.

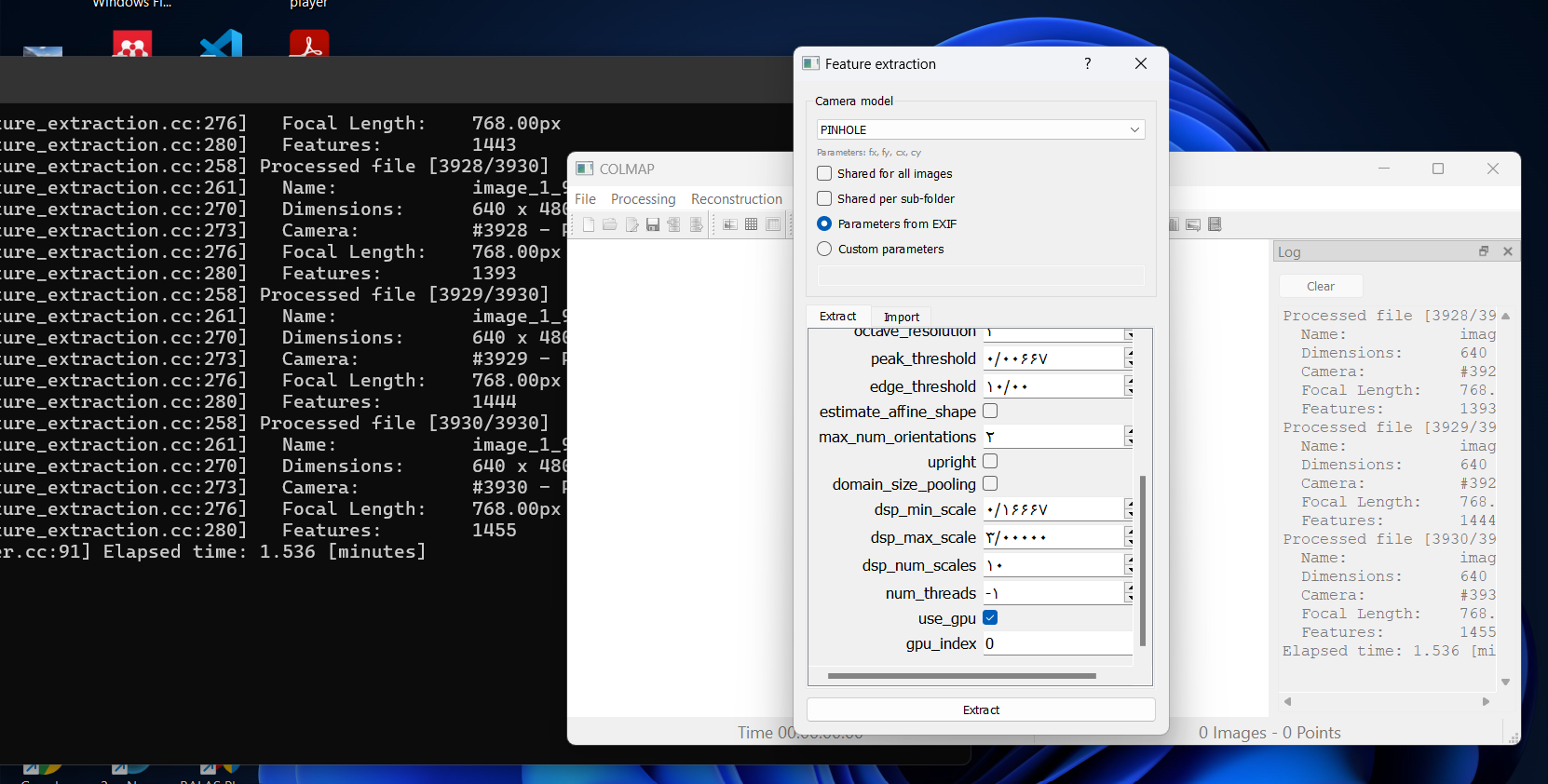
دوربین چپ: تصاویر با نام‌گذاری image\_0\_imagenumber.

دوربین راست: تصاویر با نام‌گذاری image\_1\_imagenumber.

**البته برای این که سرعت انجام این قسمت بالا برود و همچنین به دلیل محدودیت منابع تصاویر را به نصف کاهش دادم.**

## استخراج ویژگی‌ها

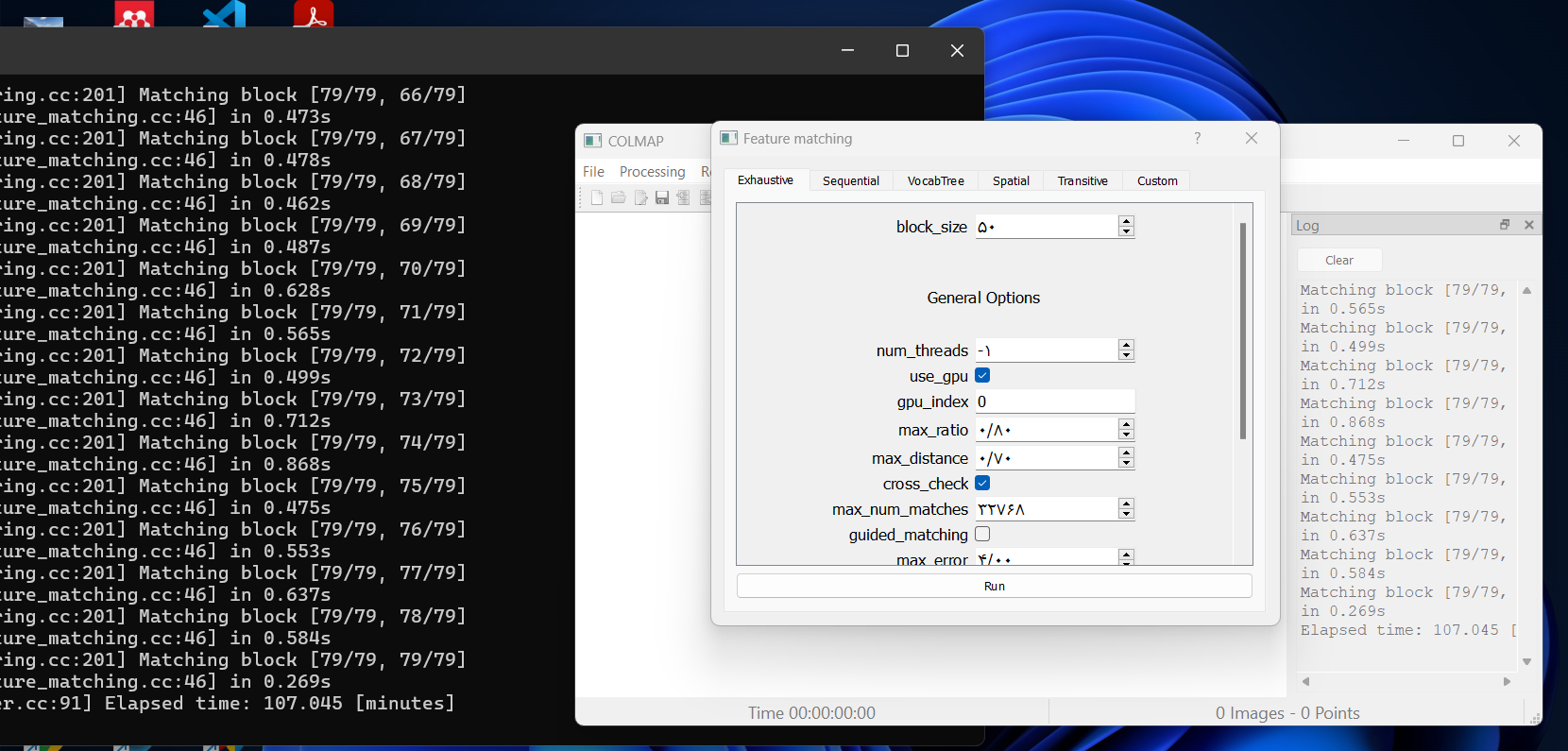
استخراج ویژگی‌ها با ابزار FeatureExtractor برای تصاویر انجام شد در تصویر پایین نتیجه کار را مشاهده می‌کنید:



استخراج ویژگی تصاویر

## انطباق ویژگی‌ها

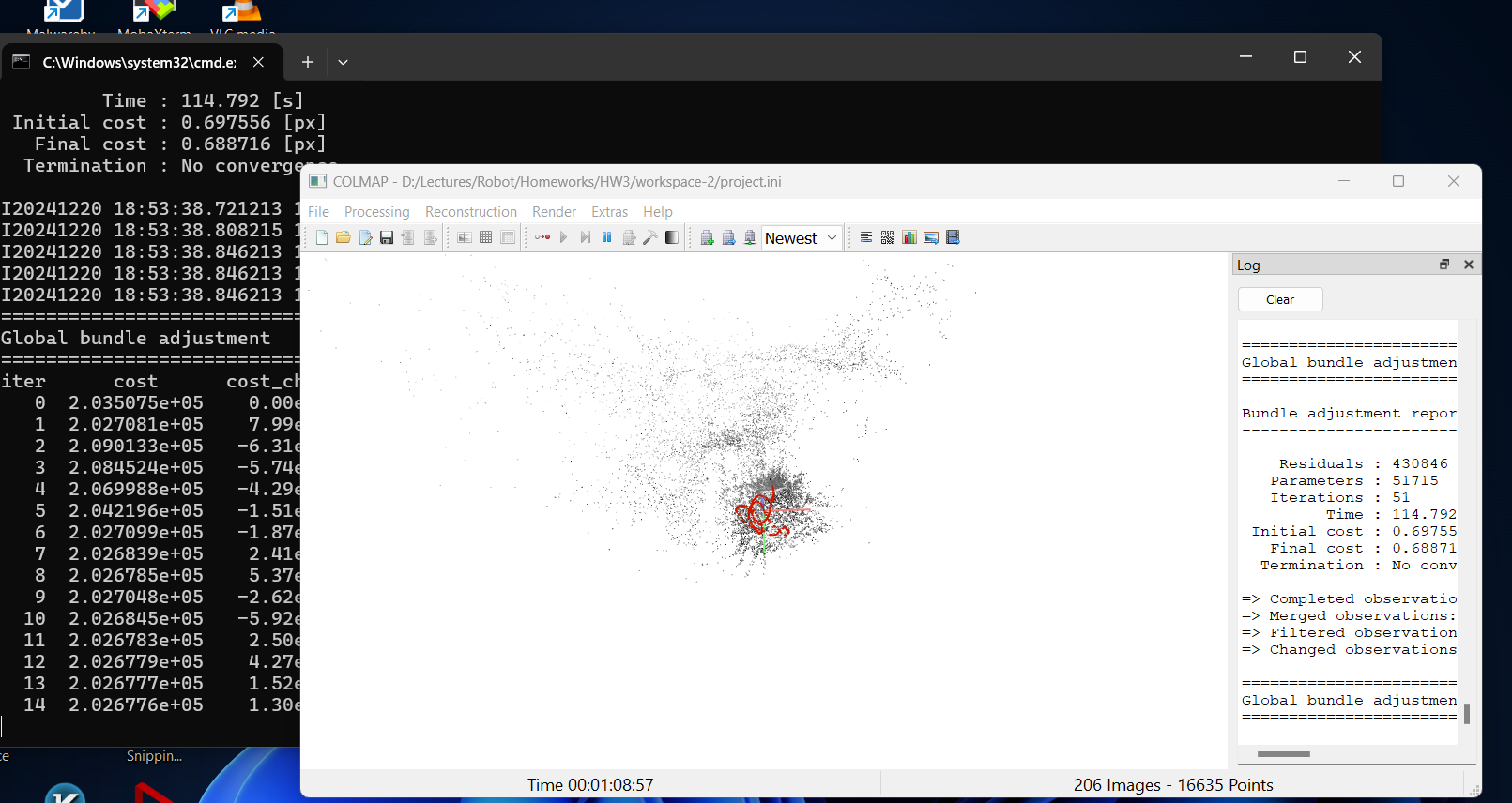
برای یافتن تطابق‌های میان تصاویر، ابزار ExhaustiveMatcher اجرا شد:



مرحله تطبیق ویژگی های تصاویر

## بازسازی پراکنده و متراکم با اسفاده از ابزار reconstruction

|  |  |
| --- | --- |
|  |  |
|  | |



تصاویر مربوط به مرحله بازسازی

## مقایسه با مسیر Leica

داده‌های مسیر Leica برای مقایسه با مسیر بازسازی شده استفاده شد. هر دو مسیر با استفاده از پایتون بارگذاری و ترسیم شدند.

باتوجه به دسترسی به داده های Leica نتایج مقایسه ای با این داده و داده های خروجی موقعیت ابزار colmap را نمایش می‌دهیم:

|  |  |
| --- | --- |
|  |  |

باتوجه به این که تعداد تصاویر کاهش یافته دقت مسیر بازسازی شده کمی دچار خطا می‌باشد.

# هندسه تک نمایی

**محاسبه ارتفاع دوربین از نقطه ناپدید شدن با استفاده از ارتفاع یک شیء شناخته‌شده**

در این گزارش، روشی برای محاسبه ارتفاع دوربین نسبت به سطح زمین با استفاده از تصویر، نقطه ناپدید شدن و ارتفاع یک جسم شناخته‌شده ارائه می‌شود. این تکنیک بر پایه اصول هندسه پرسپکتیو و مثلث‌های مشابه بنا شده است و می‌تواند در حوزه‌های مختلفی مانند بینایی ماشین، نقشه‌برداری و عکاسی کاربرد داشته باشد.

## نقطه ناپدید شدن (Vanishing Point)

نقطه ناپدید شدن نقطه‌ای در تصویر است که خطوط موازی در یک فضای سه‌بعدی در آنجا به همگرا می‌شوند. این مفهوم یک اصل اساسی در رسم پرسپکتیو و درک بصری فضا توسط انسان‌ها و سیستم‌های بینایی ماشین است. شناسایی نقطه ناپدید شدن به تعیین جهت‌های افقی و عمودی در تصویر کمک می‌کند و نقش مهمی در محاسبات هندسی ایفا می‌کند.

## اطلاعات مورد نیاز

برای انجام محاسبات مذکور، به اطلاعات زیر نیاز داریم:

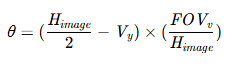
1. مختصات نقطه ناپدید شدن (Vx, Vy): مختصات محور x و y نقطه ناپدید شدن در تصویر.
2. ارتفاع شیء شناخته‌شده (H): ارتفاع واقعی شیء‌ای که در تصویر مشخص شده است (مثلاً ارتفاع یک ساختمان یا یک فرد).
3. ارتفاع شیء در تصویر (h): ارتفاع شیء به صورت پیکسل در تصویر.
4. ارتفاع تصویر (Himage): ارتفاع کل تصویر به صورت پیکسل.
5. میدان دید عمودی دوربین (FOVv): میدان دید عمودی دوربین به درجه.

## مراحل محاسبه

پرداختن به محاسبات مرحله به مرحله به درک بهتر فرآیند کمک می‌کند:

### محاسبه زاویه افقی (θ)

* تعیین زاویه افقی نسبت به افق: با استفاده از موقعیت عمودی نقطه ناپدید شدن (Vy) در تصویر، می‌توان زاویه θ را نسبت به افق محاسبه کرد. فرض بر این است که دوربین به صورت افقی تراز است.
* فرمول محاسبه زاویه:



### محاسبه فاصله تا شیء شناخته‌شده (D)

* استفاده از مثلثات: با داشتن زاویه θ و ارتفاع واقعی شیء (H)، می‌توان فاصله افقی (D) تا پایه شیء را با استفاده از تابع مماس (tan) محاسبه کرد.
* فرمول محاسبه فاصله:



این فرمول نشان می‌دهد که هرچه زاویه θ بزرگ‌تر باشد، فاصله D کمتر خواهد بود و بالعکس.

### محاسبه ارتفاع دوربین (Camera Height)

* استفاده از مثلث‌های مشابه: با استفاده از مثلث‌های مشابه که از موقعیت دوربین و شیء تشکیل شده‌اند، می‌توان ارتفاع دوربین را محاسبه کرد.
* فرمول تنظیم نسبتی:



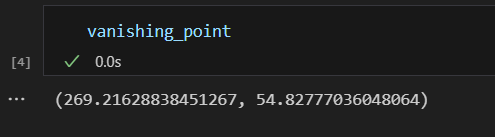




تصویری ورودی برای محاسبه vanishing point



خروجی نمایش vanishing point بر روی تصویر



مختصات نقطه