

بررسی روش بینایی استریو به عنوان روشی برای داده‌برداری دوبعدی و سه‌بعدی از اهداف مورد نظر

بابک هوشمند^{۱*}، حسین امیرآبادی^۲، خلیل خلیلی^۳

^۱ کارشناسی ارشد، مهندسی مکانیک، دانشگاه بیرجند، بیرجند

^۲ دانشیار، مهندسی مکانیک، دانشگاه بیرجند، بیرجند

^۳ استاد، مهندسی مکانیک، دانشگاه بیرجند، بیرجند

*مسئول مکاتبات: bhoushmand1986@yahoo.com

چکیده

واژگان کلیدی

ماشین بینایی
بینایی استریو
کالیبراسیون
انطباق
هندسه اپیپولار

تاریخچه مقاله

تاریخ دریافت: ۱۳۹۸/۰۸/۰۱

تاریخ پذیرش: ۱۳۹۹/۰۲/۲۳

از جمله فناوری‌هایی که امروزه به طور گسترده مورد توجه قرار گرفته است، روش‌های منطبق بر بینایی است. این روش‌ها، در بخش‌های مختلفی از جمله خط تولید، کنترل و بازرسی و حتی کنترل ایمنی افراد به کار گرفته شده است. تصویر اصلی اساسی ماشین بینایی است و در واقع بازنمایی است که از طریق دستگاه‌های تصویربرداری (دوربین‌ها) به دست می‌آیند. در این بحث، روش بینایی استریو به طور گسترده‌ای برای اندازه‌گیری و داده‌برداری دوبعدی و سه‌بعدی مورد استفاده قرار می‌گیرد. در این روش، از دو دوربین در دو زاویه مختلف برای برداشت تصویر از یک منظره استفاده می‌شود. این مقاله ماهیت روش بینایی استریو و چالش‌های موجود در استفاده از این روش شامل کالیبراسیون دوربین‌ها، انطباق بین دو تصویر گرفته‌شده توسط دو دوربین در دو زاویه مختلف و در نظر گرفتن اصل هندسه اپیپولار با هدف تسهیل در تطبیق بین دو تصویر را مورد بررسی قرار داده است.

۱ مقدمه

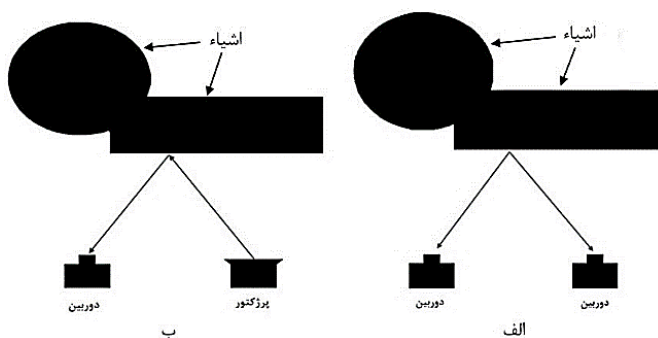
منجر به درک عمق می‌شود [۳]. استفاده از روش‌های فعال با توجه به ارزیابی و قابلیت‌هایشان، در کاربرد بینایی کامپیوتر افزایش داشته است، ولی دارای نقاط ضعفی نیز هستند. آن‌ها مستعد خطاهای سیستماتیک همچون نوسان‌های ناخواسته در سیگنال تصویر^۱ و ابهام در یافتن پیکسل صحیح هستند که مربوط به خواص حسگر^۲ مورد استفاده است [۴].

ماشین بینایی عبارت است از تجزیه و تحلیل خودکار تصاویر و ویدیوها توسط کامپیوتر به نحوی که مقدار درکی از دنیای اطراف خود به دست آورد [۱]. این روش در واقع از جمله روش‌های غیرتماسی در بحث اندازه‌گیری و داده‌برداری محسوب می‌شود. روش‌های غیرتماسی برای پیدا کردن عمق به دو دسته فعال و غیرفعال تقسیم می‌شود. روش‌های فعال شامل روش‌های مبتنی بر ارسال انواع پرتوهای انرژی به سمت شیء و یا هر عامل کمکی دیگر که در امر اندازه‌گیری از روی تصاویر هدف کمک‌کننده باشد. برای مثال از یک دوربین و یک نورافکن به عنوان عامل کمکی استفاده می‌شود. دسته دوم، روش‌های غیرفعال است که مبتنی بر تصویر بوده و در آن‌ها از دو دوربین استفاده می‌شود [۲].

در شکل ۱، پیدا کردن عمق با دو روش فعال و غیرفعال نشان داده شده است.

بینایی استریو با توجه به استفاده از دو دوربین، جزء روش‌های غیرفعال محسوب می‌شود. استخراج عمق در این روش براساس تحلیل دو تصویر گرفته‌شده یک صحنه از زوایای مختلف است.

این روش از ساختار بینایی انسان الهام گرفته شده است. در ساختار بینایی انسان، وجود دو چشم و همچنین انحراف و عدم تطابق اشیا در نمای سه‌بعدی دوچشمی در مقایسه با دو نمای یک چشمی است که



شکل ۱: پیدا کردن عمق با دو روش (الف) غیرفعال و (ب) فعال [۲].

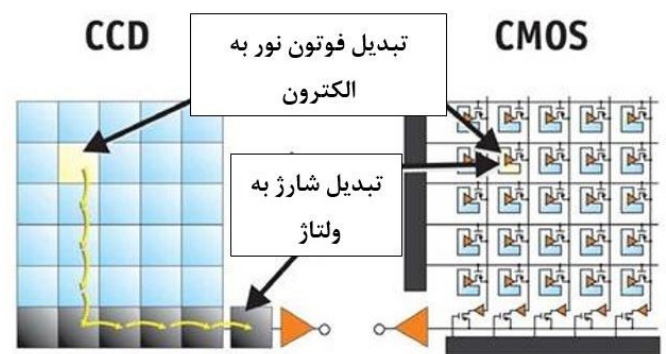
طبق تحلیل‌های صورت گرفته توسط فویکس^۳ و همکاران [۵]، روش زمان پرواز به عنوان یکی از روش‌های فعال، در فواصل بیش از ۵ تا ۷ متر به طور رضایت بخش عمل کرده و به شرایط محیطی به خصوص مناطق روشن بسیار حساس هستند. به دلیل همین محدودیت، بینایی استریو قابل اطمینان‌تر است. چرا که این روش قادر به ارائه تمایز^۴ با تفکیک‌پذیری بالا در محیط‌های داخلی و بیرونی است [۶].

¹noise ²sensor ³Foix ⁴disparity

۲ دوربین و هندسه توصیف‌کننده آن

اساس عملکرد ماشین بینایی و بینایی استریو، تصویر است. یک تصویر در واقع یک سیگنال است که می‌توان آن را با یک آرایه دوبعدی نمایش داد. دوربین ایجادکننده تصویر متشکل از دو المان اصلی حسگر (CMOS و CCD) و لنز است [۱].

حسگرها: حسگرهای CCD از سیلیکون حساس به نور ساخته شده‌اند. در این حسگرها، الکترون‌ها به ازای هر فوتون نور برخوردی با سطح سیلیکون تولید می‌شوند. شارژ ایجادشده که شامل الکترون‌های جمع شده است، در داخل محفظه نازک مجازی نگه‌داری می‌شود و پیکسل‌ها که اجزای منحصربه‌فرد تصویر هستند را ایجاد می‌کنند. حسگرهای CMOS، بر اساس پدیده فتولتاییک^۵ عمل می‌کنند؛ یعنی با برخورد نور با نیمه‌هادی اکسید فلز، بین دو الکتروود مثبت و منفی، اختلاف پتانسیل رخ داده و جریان جاری می‌گردد. در واقع در این نوع حسگر، شارژ ایجادشده در هر پیکسل به ولتاژ تبدیل می‌شود. قابلیت یکپارچه‌سازی و مصرف انرژی کمتر از جمله مزایای حسگر CMOS است. در این نوع حسگر، امکان داشتن خروجی از چندین کانال که منجر به نرخ بالای انتقال اطلاعات تا بیش از ۲ گیگا بایت بر ثانیه و نرخ برداشت تصویر تا ۱۰۰۰ تصویر بر ثانیه می‌شود؛ وجود دارد که این امر برای تصویربرداری سرعت بالا مناسب است. در شکل ۲، تفاوت نحوه عملکرد حسگرهای CMOS و CCD قابل مشاهده است. در حسگر CCD، فوتون‌های دریافتی توسط پیکسل‌ها در یک خروجی به ولتاژ تبدیل می‌شوند؛ در حالی که در CMOS، تبدیل فوتون به ولتاژ در داخل هر پیکسل رخ می‌دهد. با توجه به شکل ۲، در CMOS، طراحی و چیدمان پیکسل‌ها با توجه به الزام استفاده از ترانزیستور حساس به نور برای هر پیکسل به منظور دستیابی به کارایی و انعطاف‌پذیری مطلوب، کار پیچیده‌ای است. با وجود این، در مقایسه با نوع CCD، پیشرفت‌های اخیر آن‌ها سریع‌تر و در عین حال آشفته‌تر بوده است.



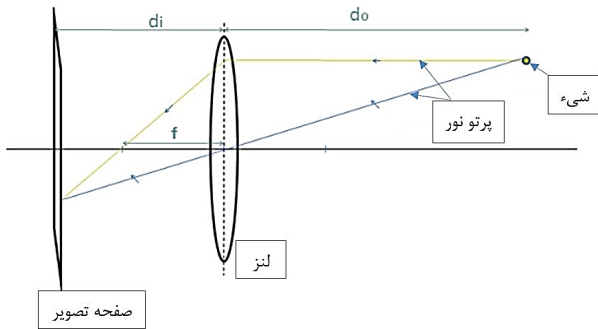
شکل ۲: مقایسه ساختار فیزیکی حسگرهای CMOS و CCD [۱].

لنز: وظیفه لنز، تمرکز نور در یک منطقه کوچک (نقطه) است. نور ساطع شده از یک منظره، توسط دوربین و از طریق دریچه جلوی دوربین^۶ دریافت می‌شود. پرتوهای نور عبوری از لنز، بر صفحه تصویر قرارگرفته در پشت لنز قرار می‌گیرند. این شرایط در شکل ۳، قابل مشاهده است. صفحه تصویر در واقع همان سطح حسگر است که پرتوهای نور از فضای سه‌بعدی

بر روی آن تابیده می‌شود.

در شکل ۳، d_o ، فاصله لنز تا هدف مشاهده‌شده، d_i ، فاصله لنز تا صفحه تصویر و f ، فاصله کانونی لنز است. برای محاسبه فاصله کانونی از رابطه (۱) استفاده می‌شود.

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (1)$$



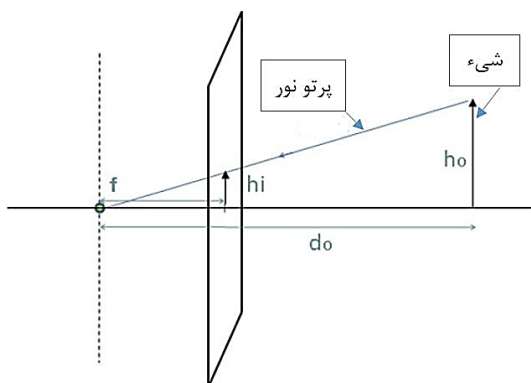
شکل ۳: نحوه تشکیل تصویر در دوربین [۱].

در بحث ماشین بینایی، مدل دوربین می‌تواند به روش‌های مختلفی ساده‌سازی شود. در ابتدا می‌توان با فرض بینهایت بودن دریچه دوربین، از اثر لنز صرف نظر کرد. از نظر تئوری این امر، ظاهر تصویر را تغییر نمی‌دهد. هر چند با این فرض، اثر تمرکز به وسیله ایجاد تصویر با عمق بینهایت نادیده گرفته می‌شود. در این شرایط، تنها پرتو مرکزی در نظر گرفته می‌شود. از طرفی چون $d_o \gg d_i$ است (در اغلب اوقات)، می‌توان فرض کرد که صفحه تصویر در فاصله کانونی واقع شده است. مدل ساده شده با عنوان مدل دوربین حفره کوچک^۷ نام‌گذاری شده است که در شکل ۴، قابل مشاهده است. بر طبق تشابه مثلثاتی، رابطه پرتوافکنی که یک شیء را با تصویر آن مرتبط می‌کند، طبق رابطه (۲) است.

$$h_i = f \frac{h_o}{d_o} \quad (2)$$

در رابطه (۲)، h_o ، ارتفاع واقعی شیء، h_i ، ارتفاع شیء در تصویر و d_o ، فاصله شیء تا مرکز کانونی است.

مقدار h_i در تصویر، به طور معکوس با فاصله شیء از دوربین متناسب است. بنابراین رابطه (۲)، توصیف‌کننده محل پرتوافکنی یک نقطه در فضای سه‌بعدی بر روی تصویر با توجه به هندسه دوربین است.



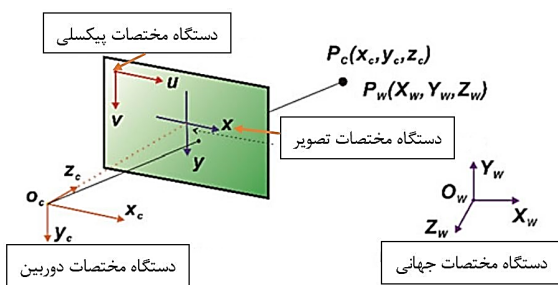
شکل ۴: مدل حفره کوچک دوربین [۱].

^۵photovoltaic ^۶diaphragm ^۷pin hole

بدیهی است که می‌توان از مشخصات ارائه شده توسط سازنده دوربین استفاده کرد، اما برای بعضی از کارها مانند بازسازی سه‌بعدی، این مشخصات به اندازه کافی دقیق نیست. کالیبراسیون دوربین، با ارائه الگوهای شناخته شده به دوربین و تجزیه و تحلیل تصاویر به دست آمده ادامه می‌یابد. سپس تخمین حالت (تعیین جهت و موقعیت دوربین) که کاربرد آن تشخیص اشیا مبتنی بر مدل دوبعدی به سه‌بعدی است را می‌توان به صورت یک مسئله بهینه‌سازی فرموله کرد تا خطا بین مختصات تصویر و مدل حداقل شود [۷]. لازم به ذکر است که در تشخیص اشیا مبتنی بر مدل دوبعدی به سه‌بعدی، یک مدل سه‌بعدی از جسم و تصویر دوبعدی‌اش از یک نقطه مناسب دوربین گرفته می‌شود. هدف تعیین سه چرخش و سه انتقال نسبت به یک سیستم مختصات پایه است. این فرآیند پیچیده توسط نرم‌افزاری مانند متلب به راحتی انجام می‌شود. به منظور کالیبره کردن، ایده، نمایش یک مجموعه نقاط است که مختصات سه‌بعدی آنها شناخته شده است و سپس بایستی مشخص شود که این نقاط در کجای تصویر نمایش داده شده‌اند. با داشتن تعداد کافی نقاط سه‌بعدی و نقاط دوبعدی تصویر مرتبط، می‌توان مقادیر دقیق پارامترهای دوربین را از معادلات مربوطه استخراج کرد. لذا هر چه تعداد نقاط شناخته شده سه‌بعدی بیشتر باشد صحت مقادیر اندازه‌گیری شده (نزدیکی مقادیر به دست آمده با مقادیر از قبل تعیین شده) افزایش می‌یابد [۸].

۱.۳ مدل خطی کالیبراسیون دوربین

هدف از کالیبراسیون دوربین، تعیین رابطه نگاشت نقطه بین دستگاه مختصات جهانی و دستگاه مختصات پیکسلی است [۸]. همان‌طور که در شکل ۶ نمایش داده شده است، نقطه P دارای مختصات (X_W, Y_W, Z_W) نسبت به دستگاه مختصات جهانی، (X_C, Y_C, Z_C) نسبت به دستگاه مختصات دوربین، (x, y) نسبت به دستگاه مختصات پیکسلی است.



شکل ۶: دستگاه‌های مختصات در مدل خطی دوربین [۸].

رابطه نگاشت بین دستگاه‌های مذکور به شکل زیر ارائه شده است [۹]:

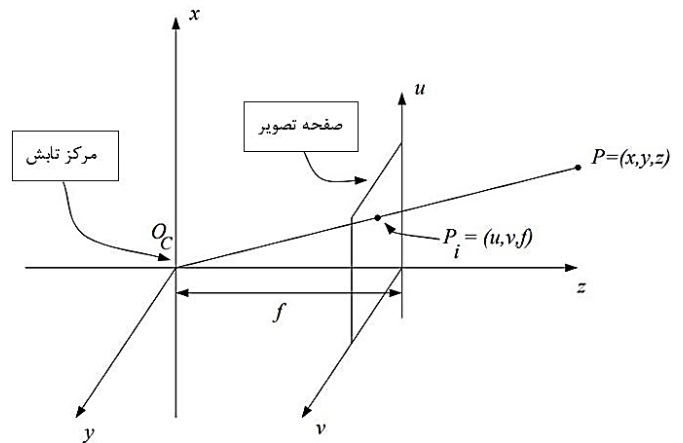
$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} R & t \\ O & I \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (۴)$$

$$s = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} f_u & \alpha & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (۵)$$

در شکل ۴، هر نقطه از هدف، نقطه تصویر شده متناظر آن و حفره، یک خط را تشکیل می‌دهند. این نوع تصویربرداری از فضای سه‌بعدی بر یک صفحه دوبعدی، تصویربرداری پرسپکتیو^۸ نام دارد. در این نوع تصویربرداری، صفحه نمایش، یک جنبه از نمایش دوبعدی تصویر از دید ناظر است. با فرض اینکه ناظر به صفحه تصویر نگاه می‌کند (با فرض مختصات ناظر در فضای سه‌بعدی $(0, 0, 0)$ باشد) آن چیزی که از یک نقطه در صفحه تصویر می‌بیند، علاوه بر دارا بودن مختصات (u, v) بر حسب دستگاه مختصات دوبعدی صفحه تصویر، یک فاصله از ناظر نیز دارد. لذا می‌توان آن را یک پرتو در نظر گرفت. در شکل ۵، مدل دوربین حفره کوچک و تبدیل پرسپکتیو نمایش داده شده است. محل ناظر O_C با مختصات $(0, 0, 0)$ در شکل مذکور در محل مرکز پرتوافکنی است. P ، نقطه‌ای در فضای سه‌بعدی و P_i ، نقطه تابیده شده روی صفحه تصویر است. بر اساس تشابه مثلثاتی داریم:

$$v = f \frac{x}{z}, \quad u = f \frac{y}{z} \quad (۳)$$

در روابط (۳)، f فاصله کانونی، (x, y, z) مختصات نقطه نسبت به دستگاه مختصات مرکز پرتوافکنی (دوربین) و (u, v) مختصات نقطه تابیده شده بر صفحه تصویر نسبت به دستگاه مختصات در صفحه تصویر هستند.



شکل ۵: مدل حفره کوچک دوربین [۱].

۳ کالیبراسیون

بر طبق مدل حفره کوچک، پارامترهای ضروری عبارتند از فاصله کانونی و ابعاد صفحه تصویر که میدان دید^۹ را تعیین می‌کند [۷]. از طرفی چون تصویر از نوع دیجیتال است، تعداد پیکسل‌های صفحه تصویر^{۱۰} ویژگی مهم دیگری است که بایستی مدنظر باشد. در نهایت، به منظور محاسبه موقعیت یک نقطه بر روی تصویر بر حسب مختصات پیکسلی، تعریف یک مبنا الزامی است. اگر فرض شود خطی که از نقطه کانونی عبور می‌کند بر صفحه تصویر عمود باشد، بایستی مشخص شود.

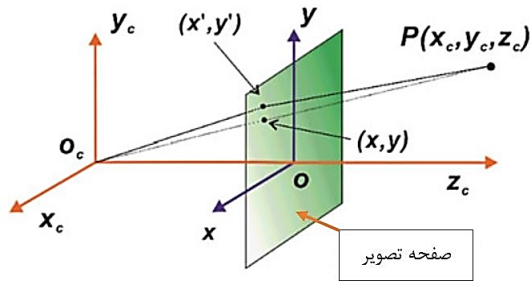
که در کدام پیکسل این خط از صفحه تصویر عبور می‌کند. این نقطه، نقطه اصلی^{۱۱} نامیده می‌شود. در عمل ممکن است نقطه اصلی به اندازه چند پیکسل از مرکز فاصله داشته باشد. این امر به دقت ساخت دوربین بستگی دارد. برای تعیین پارامترهای مذکور، دوربین را کالیبره می‌کنند.

⁸perspective ⁹field of view ¹⁰resolution ¹¹principal point

است.

$$\begin{cases} x' = x + \delta_x(x, y) \\ y' = y + \delta_y(x, y) \end{cases} \quad (۶)$$

در رابطه (۶) مختصات واقعی نقطه P تاییده شده بر روی صفحه تصویر است [۸].

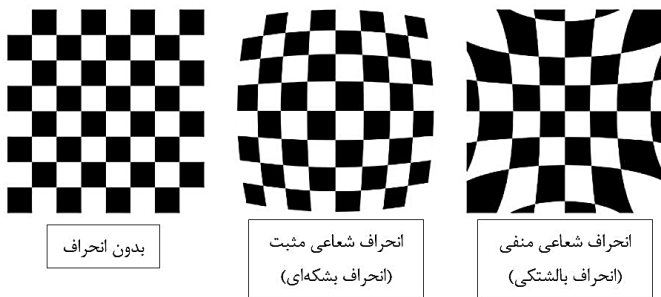


شکل ۸: مدل غیرخطی دوربین [۸].

در رابطه (۶)، (x, y) مختصات نقطه تاییده شده با فرض شرایط خطی است. δ_x و δ_y مقدار اختلاف به وجود آمده حاصل از انحرافات ساختی و مونتاژی در امتداد x و y هستند. این انحرافات شامل انحرافات شعاعی^{۱۴} و انحرافات مماسی^{۱۵} هستند. انحرافات شعاعی حاصل از انحراف شعاعی در ساخت لنز است [۹]. این انحراف، باعث می شود تا نقطه در تصویر در امتداد شعاعی به بیرون یا به داخل جابه جا شود. به طوری که هر چه از مرکز دورتر یا نزدیک تر شود تغییر شکل بیشتر می شود. δ_{yr} و δ_{xr} انحرافات شعاعی در امتداد محورهای x و y هستند. انحرافات مماسی حاصل ایراد در مونتاژ لنز است؛ به طوری که لنز با صفحه تصویر موازی نباشد. δ_{yt} و δ_{xt} انحرافات مماسی در امتداد x و y هستند. در شکل های ۹ و ۱۰ این انحرافات و نحوه ایجاد آن ها نمایش داده شده است.

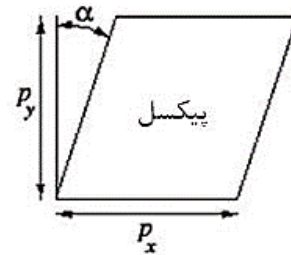
$$\begin{cases} r^2 = x^2 + y^2 \\ \delta_{xr} = x(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \delta_{yr} = y(k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \delta_{xt} = p_1(r^2 + 2x^2) + 2p_2 xy \\ \delta_{yt} = p_1(r^2 + 2y^2) + 2p_2 xy \end{cases} \quad (۷)$$

در روابط (۷) تا k_3, k_2, k_1 و p_2, p_1 ضرایب انحراف شعاعی و p_2 و p_1 ضرایب انحراف مماسی هستند.



شکل ۹: اثر انحرافات شعاعی بر صفحه شطرنجی [۱۰].

در رابطه (۴)، ماتریس A یک ماتریس همگن است که حاوی پارامترهای داخلی دوربین است. α ، فاکتور انحراف^{۱۲} است. این فاکتور، تعیین کننده میزان انحراف از چهارضلعی با اضلاع عمود بر هم (مستطیلی بودن) هر پیکسل در صفحه تصویر است. در دوربین های امروزی این فاکتور تقریباً برابر صفر است. در شکل ۷، α ، میزان انحراف از عمود بودن ضلع مستطیل نسبت به امتداد y در مختصات پیکسلی را نمایش می دهد.



شکل ۷: نمایش هندسی فاکتور انحراف.

ماتریس دوران t شامل پارامترهای خارجی دوربین است؛ به طوری که $\begin{bmatrix} R & t \\ O & I \end{bmatrix}$ بین دو یا تعداد بیشتری دوربین را نسبت به یکدیگر تعیین کنند. برای تبدیل مقیاس از واحد متریک به واحد پیکسل، فاکتوری با عنوان s که به وسیله ابعاد تصویر تعریف می شود، نیاز است. s_x, s_y, s_z تغییر مقیاس در راستاهای X, Y, Z هستند. u و v در ماتریس (۵)، محل نقطه اصلی را به ترتیب در امتداد محور x و محور y تصویر و بر حسب پیکسل مشخص می کنند. مقادیر به دست آمده برای محل نقطه اصلی الزاماً در مرکز تصویر نیست؛ چرا که حسگرها ممکن است دقیقاً موقعیت دهی نشده باشند و یا ممکن است به دلیل فرآیند بهینه سازی که در طی آن همه پارامترها طوری تعیین شده اند که خطای پرتوافکنی حداقل شود، محل نقطه اصلی را تحت تأثیر قرار دهد [۱]. در نهایت f_u و f_v در ماتریس (۵)، فاصله کانونی در دو امتداد x و y را مشخص می کنند. علت ارائه فاصله کانونی با دو مقدار این است که پیکسل ها الزاماً مربعی نیستند.

۲.۳ مدل غیرخطی کالیبراسیون دوربین

در عمل، فرآیند ساخت لنزهای به کار رفته در دوربین ها بدون عیب نیست و حتی طی مرحله مونتاژ ممکن است کمی خطا در نصب آن ها وجود داشته باشد. این عوامل، منجر به خطاهای کوچکی می شود که انحراف تصویر را به دنبال دارد. مدل خطی که برای دوربین در نظر گرفته می شود، یک مدل ایده آل بوده که انحرافات حاصل از مشکلات مذکور را در بر نمی گیرد. ولی با توجه به سادگی، به طور گسترده مورد استفاده قرار می گیرد. با وجود این، مدل خطی با خطا همراه است.

انحرافات غیرخطی معمولاً با عنوان انحرافات هندسی^{۱۳} شناخته می شوند. این انحراف باعث می شود بین نقطه تصویربرداری شده با موقعیت واقعی آن می شود. لذا با تصحیح این انحرافات تا حد امکان می توان این فاصله را کاهش داد. در شکل ۸، مدل غیرخطی دوربین نمایش داده شده

¹²skew factor ¹³geometric distortion ¹⁴radial distortion ¹⁵tangential distortion

مقدار d ، اختلاف مکانی^{۱۸} (اختلاف محل تشکیل تصویر یک نقطه در صفحه تصاویر چپ و راست) نامیده می‌شود. بدین ترتیب عمق هر نقطه با توجه به اختلاف مکانی آن قابل محاسبه است:

$$Z = \frac{f}{d} b \quad (11)$$

از رابطه (۱۱) می‌توان به معکوس بودن رابطه بین عمق نقطه با اختلاف مکانی پی برد.

در کالیبراسیون استریو برای هر نقطه مانند P در فضای سه‌بعدی، هر یک از دوربین‌ها به طور جداگانه کالیبره می‌شوند. در روابط (۱۲) و (۱۳)، ارتباط بین محل تابش نقطه P روی صفحه تصویر هر یک از دوربین‌های چپ (P_l) و راست (P_r) با ماتریس دوران و انتقال برای هر دوربین آورده شده است.

$$P_l = R_l P + T_l, \quad P_r = R_r P + T_r \quad (12)$$

$$R^T (P_r - T) = P_l \quad (13)$$

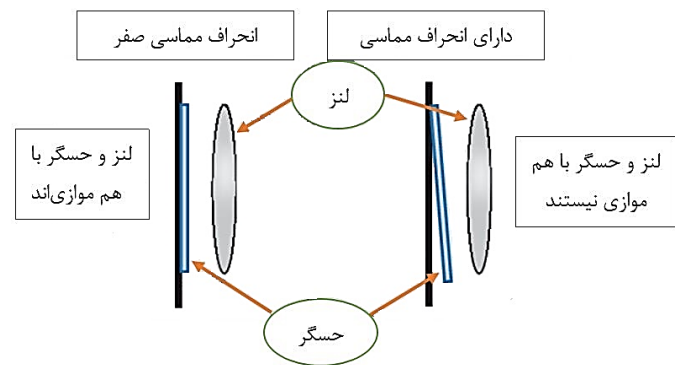
با کالیبراسیون هر دوربین به طور جداگانه ماتریس‌های R_l, T_l, R_r و R_r, T_r که شامل ماتریس‌های انتقال و دوران برای هر یک از دوربین‌های چپ (l) و راست (r) هستند و سپس ماتریس‌های انتقال و دوران R و T بین دو دوربین به دست می‌آیند. در روابط (۱۴) و (۱۵) نحوه محاسبه ماتریس‌های انتقال و دوران بین دو دوربین با توجه به ماتریس‌های انتقال و دوران به دست آمده برای هر یک از دوربین‌ها آورده شده است.

$$R = R_r (R_l)^T \quad (14)$$

$$T = T_r - R T_l \quad (15)$$

۵ انطباق در بینایی استریو

انطباق تصویر، فرآیند مقایسه و یافتن ویژگی‌های هندسی مشترک بین حداقل دو تصویر از یک منظره است که در شرایط متفاوت تصاویر برداری مانند زمان‌های متفاوت، زوایای متفاوت یا حتی حسگرهای متفاوت گرفته شده‌اند. این فرآیند از نظر هندسی، دو تصویر را که یکی از آن دو به عنوان مرجع^{۱۹} در نظر گرفته می‌شود را هم‌تراز می‌کند [۱۲]. در واقع شرایط متفاوت در تصویربرداری منجر به ایجاد اختلافات قابل توجه بین تصاویر می‌شود. این اختلافات به چهار دسته مشکلات هندسی، مشکلات رادیومتری^{۲۰}، مشکلات بافت و تغییر مناظر تقسیم می‌شوند [۱۳]. اختلافات در مقیاس تصاویر، چرخش تصاویر و تغییر شکل‌های ناشی از تغییر موقعیت در هنگام برداشت تصویر از یک منظره به عنوان مشکلات هندسی شناخته می‌شوند. نوسان‌های ناخواسته، تفاوت شدت روشنایی^{۲۱}، موقعیت منابع روشنایی در صحنه و اخذ تصویر در حسگرها و باندهای طیفی متفاوت، باعث ایجاد تغییر در شدت روشنایی تصویر می‌شود که با عنوان مشکلات رادیومتری شناخته می‌شود. تصاویر ممکن است دارای سطوح با بافت ضعیف (مانند سطح آب)، بدون بافت و بافت‌های تکراری (مانند خودروهای مشابه در یک ردیف)



شکل ۱۰: نمایش انحراف مماسی [۱۰].

انحراف شعاعی و مماسی در روابط (۶) می‌توانند به صورت معادله (۸)، با هم ترکیب شوند.

$$\delta_x = x(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1 (r^2 + 2x^2) + 2p_2 xy \quad (8)$$

$$\delta_y = y(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_2 (r^2 + 2y^2) + 2p_1 xy$$

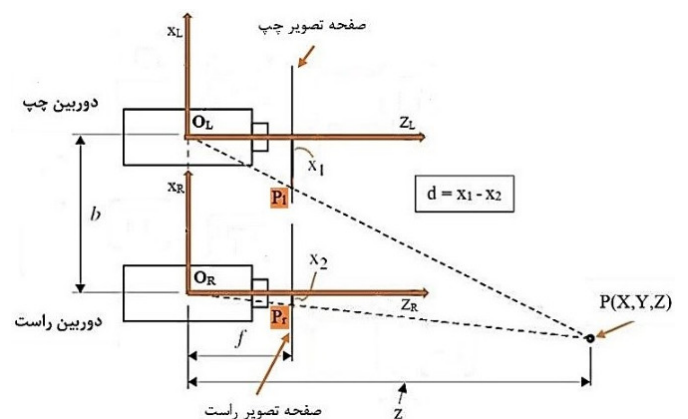
۴ هندسه روش استریو

در تصویربرداری پرسپکتیو، یکی از ابعاد هر نقطه از فضا هنگام پرتوافکنی بر صفحه تصویر حذف می‌شود. برای بازیابی این بعد از دست رفته، روش استریو به عنوان روشی برای بازیابی بعد سوم از تصاویر دوبعدی مطرح می‌شود. در این روش توسط دو دوربین از یک صحنه معین دو تصویر گرفته می‌شود و با استفاده از پارامترهای دوربین‌ها و روابط حاکم بر مدل هندسی تصاویر استریو، هر سه بعد نقاط مورد نظر قابل محاسبه می‌شوند [۱۱]. در شکل ۱۱، O_R و O_L موقعیت دوربین‌های چپ و راست در صحنه، b فاصله بین آن‌ها (خط مبنا^{۱۶}) و P_l و P_r پرتوافکنی^{۱۷} با مختصات (X, Y, Z) نسبت به دستگاه مختصات جهانی در فضای سه‌بعدی روی تصاویر تشکیل شده در دوربین‌های چپ و راست هستند.

از شکل می‌توان نتیجه گرفت:

$$x_{pl} = \frac{f}{Z} X, \quad x_{pr} = \frac{f}{Z} (X - b) \quad (9)$$

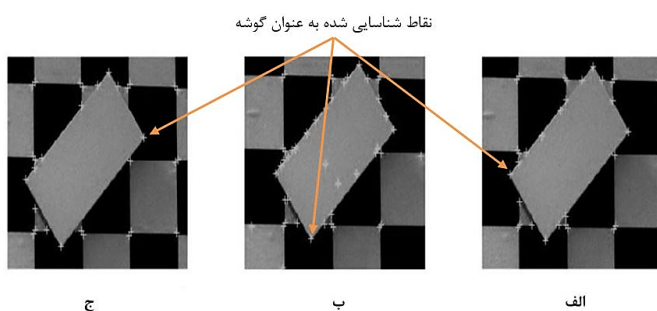
$$d = x_{pl} - x_{pr} = \frac{f}{Z} b \quad (10)$$



شکل ۱۱: هندسه روش استریو.

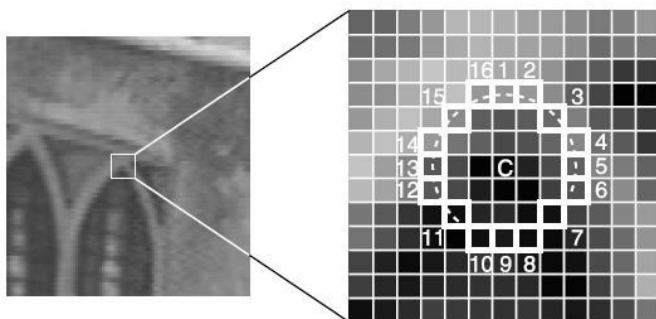
^{۱۹} تصویر مرجع: تصویری که به عنوان مبنا برای انطباق تصویر استفاده می‌شود.

به نوع تصویر و محتوای ویژگی‌های تصویر شده دارد. با توجه به سه نوع ویژگی مختلف در روش‌های مبتنی بر ویژگی، می‌توان نتیجه گرفت که ویژگی‌های نقطه‌ای از دیگر ویژگی‌ها پایدارتر هستند؛ یعنی شناساگرهای ویژگی نقاط، معمولاً در برابر انحراف‌های هندسی پیچیده بین تصاویر، پایدارتر از شناساگرهای نواحی است [۱۵]. توصیف‌گرهای ویژگی نقطه‌ای فست^{۲۶} [۱۶]، هریس^{۲۷} [۱۷] و شای-توماسی^{۲۸} [۱۸] از جمله روش‌های استخراج ویژگی‌های نقطه‌ای می‌باشند که در این پژوهش مورد بررسی قرار گرفتند. شکل ۱۲، نحوه تشخیص نقاط گوشه توسط الگوریتم‌های رایج تشخیص گوشه فست، هریس و شای-توماسی را نشان می‌دهد. هر سه الگوریتم نتایج تقریباً یکسانی در یافتن نقاط گوشه خواهند داشت، و اختلاف بین آنها در تعداد نقاط ناخواسته‌ای است که به عنوان گوشه تشخیص داده شده است.



شکل ۱۲: تشخیص نقاط توسط سه الگوریتم (الف) هریس، (ب) شای توماسی و (ج) فست.

در بین توصیف‌گرهای مذکور، فست به عنوان روش تشخیص گوشه با بازده بالای محاسباتی شناخته می‌شود. در این روش، اساس عملکرد این است که برای هر پیکسل، یک دایره آزمون شامل ۱۶ پیکسل که آن پیکسل را احاطه کرده‌اند در نظر گرفته و پیکسلی به عنوان گوشه انتخاب می‌شود که شدت روشنایی حداقل ۱۲ پیکسل پیوسته در دایره مذکور از پیکسل مورد نظر بیشتر یا کمتر باشد. معیار ارزیابی شدت، مقایسه با یک مقدار آستانه است. این روش در شکل ۱۳ نشان داده شده است.



شکل ۱۳: مربع‌ها با اضلاع روشن‌تر، پیکسل‌های استفاده شده برای تشخیص گوشه را نمایش می‌دهند. پیکسل C، مرکز گوشه شناسایی شده است. خط چین از ۱۲ پیکسل پیوسته که روشن‌تر از پیکسل مرکزی هستند، عبور کرده است [۱۶].

توصیف‌گر هریس، برای شناسایی گوشه از یک پنجره دایره‌ای هموار برای مثال پنجره گوسی استفاده می‌کند و سپس با استفاده از بسط تیلور، پنجره را

باشد که مشکلات بافت نامیده می‌شود. ویژگی‌های متحرک (مانند حرکت میز ماشین) و تغییراتی که در اثر گذر زمان در تصاویر وجود دارد، باعث ایجاد تفاوت‌هایی در مناظر می‌شود. با توجه به موارد مطرح شده، به منظور افزایش دقت در مرحله پردازش تصویر، انجام فرآیند انطباق ضروری است. با توجه به وجود تنوع در مشکلات ذکر شده، الگوریتم‌های ارائه شده برای انطباق تنها نوع مشخصی از تصاویر ایجاد شده‌اند. هدف از انطباق دو یا تعداد بیشتری تصویر، استخراج اطلاعات سه‌بعدی تصویر (عمق) و کسب اطلاعات کامل‌تری از تصاویر مورد نظر است. به طور کلی دوروش تعاملی^{۲۹} و خودکار برای انجام عمل انطباق وجود دارد [۱۴]. در روش تعاملی، یک مجموعه از نقاط کنترلی در تصاویر به صورت دستی انتخاب می‌شوند و سپس از این نقاط برای برآورد تابع تبدیل میان دو تصویر و نمونه‌برداری مجدد استفاده می‌کنند. در حالی که در روش‌های خودکار همین‌طور که از نام آن مشخص است، به صورت خودکار اقدام به شناسایی ویژگی‌ها و تطبیق بین آن‌ها می‌کند.

۱.۵ شناسایی ویژگی‌ها

به طور کلی روش‌های انطباق تصویر در انطباق خودکار، بر اساس ویژگی‌ها به دو دسته روش‌های مبتنی بر ناحیه و روش‌های مبتنی بر ویژگی تقسیم می‌شوند.

۱.۱.۵ روش‌های مبتنی بر ناحیه

روش‌های مبتنی بر ناحیه زمانی مطرح می‌شود که تصاویر جزئیات مهم زیادی را نداشته باشند. تمایز در اطلاعات از طریق تفاوت در شدت روشنایی تعیین می‌شود. در این دسته از روش‌ها، هیچ ویژگی از تصویر شناسایی نمی‌شود؛ لذا در این روش‌ها بیشتر بر انطباق تصویر تأکید می‌شود. روش‌های مبتنی بر ناحیه نیاز به فضای جستجو و مقدار اولیه مناسب داشته و در مناطق با بافت یکنواخت ضعف دارند.

۲.۱.۵ روش‌های مبتنی بر ویژگی

در این روش‌ها، ویژگی‌های تصاویر شناسایی می‌شوند و سپس تطبیق میان آن‌ها انجام می‌شود. از این دسته از روش‌ها، معمولاً زمانی استفاده می‌شود که اطلاعات ساختار محلی، مهم‌تر از اطلاعات شدت روشنایی باشد. این روش نسبت به انحراف‌های بین تصاویر پایدار است. یک ویژگی مناسب برای فرآیند تطبیق این است که باید نسبت به همسایگی‌های خود متمایز بوده و در میان دیگر ویژگی‌ها، منحصربه‌فرد باشد و مستقل از اعوجاج هندسی و رادیومتری و پایدار در برابر نوسان‌های ناخواسته در سیگنال تصویر باشد. نوع ویژگی‌هایی که در تصاویر انتخاب می‌شود، بستگی به نوع تصاویر دارد. به طور کلی، در تصاویر سه نوع ویژگی وجود دارد که عبارتند از: ویژگی‌های نقطه‌ای^{۳۰}، ویژگی‌های خطی^{۳۱} و ویژگی‌های ناحیه‌ای^{۳۲} [۱۲].

روش‌های مبتنی بر ویژگی در مقابل روش‌های مبتنی بر ناحیه از قابلیت اطمینان بیشتری برخوردارند و در برابر اعوجاج‌های هندسی و رادیومتری عملکرد بهتری دارند [۱۲]. نوع ویژگی‌های انتخاب شده در تصویر بستگی

²²interactive ²³point features ²⁴line features ²⁵region features

²⁸Shi-Thomasi

²⁶FAST (Features from Accelerated Segment Test) algorithm ²⁷Harris

تغییرات در تمام جهات صورت می‌گیرد که منجر به تغییرات مشتق در تمام جهات می‌شود.

۶ تطبیق

پس از شناسایی ویژگی‌ها در دو تصویر، لازم است نقاط متناظر بین آنها تعیین شود. ویژگی‌های شناسایی شده در بین تصاویر می‌تواند با استفاده از روش‌های مقادیر شدت روشنایی، روابط فضایی و توصیف‌گرها تطبیق یابند [۱۲].

در تطبیق بر اساس مقادیر روشنایی از توزیع سطوح خاکستری تصاویر، درون پنجره‌هایی با ابعاد یکسان استفاده می‌کنند. برای انطباق با این روش، با استفاده از معیارهای مختلف تشابه، به طور آماری شدت روشنایی دو تصویر مقایسه و موقعیت بیشینه شباهت را به عنوان موقعیت‌های مشابه تعیین می‌کنند. در تطبیق مبتنی بر روابط فضایی، ابتدا یک مدل هندسی میان ویژگی‌های استخراج شده در دو تصویر چپ و راست تعریف می‌شود و سپس پارامترهای مدل هندسی برآورد می‌شود. در این روش، از اطلاعات فاصله بین نقاط کنترلی یا همان نقاط ویژگی و اطلاعات فضایی توزیع شده بین آنها استفاده می‌شود.

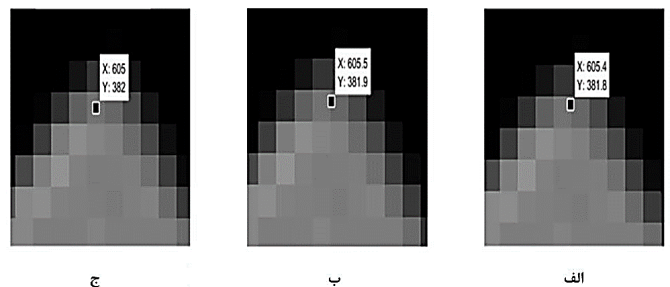
در تطبیق بر اساس توصیف‌گرها، بعد از استخراج ویژگی‌ها، یک توصیف‌گر با استفاده از خصوصاتی که در ناحیه‌ای اطراف هر ویژگی وجود دارد ایجاد شده و سپس فرآیند تطبیق با استفاده از یک معیار اختلاف (مانند فاصله اقلیدسی) میان بردارهای مشخصه، انجام می‌گیرد.

به طور کلی، روش‌های مبتنی بر اطلاعات مشترک از میان دیگر روش‌های تطبیق مبتنی بر مقادیر روشنایی مقاوم‌تر است؛ چون این روش تنها از رابطه آماری بین تصاویر استفاده می‌کند [۲۵]. روش‌های مبتنی بر مقادیر شدت روشنایی توانایی تطبیق با دقت بسیار بالایی دارند [۲۱]، اما این روش‌ها ضعف‌هایی نیز دارند. یکی از محدودیت‌های اصلی این روش‌ها این است که دو تصویر چپ و راست باید تا حد زیادی دارای توابع شدت روشنایی مشابه یا یکسانی باشند یا حداقل از نظر آماری وابسته باشند [۱۲]. این روش‌ها، تنها برای تصاویر با اعوجاج کم (اختلاف مقیاس، چرخش، تغییر زاویه کم میان تصاویر) مناسب هستند. نوسان‌های ناخواسته در سیگنال تصویر، تغییرات شدت روشنایی و استفاده از حسگرهای متفاوت در برداشت تصاویر باعث تغییراتی در میزان شدت روشنایی تصویر می‌شود که در این نوع تصاویر، روش‌های مبتنی بر مقدار شدت روشنایی مناسب نیست، چون این روش‌ها حساسیت زیادی به تغییرات روشنایی دارند. در حالی که روش‌های مبتنی بر روابط فضایی، زمانی استفاده می‌شوند که ویژگی‌های شناسایی شده مبهم باشند یا همسایه‌های آنها از نظر محلی دارای اعوجاج باشند [۱۲]. روش‌های تطبیق مبتنی بر توصیف‌گرها به آسانی پیاده سازی می‌شود، اما زمانی که ویژگی‌ها و توصیف‌گرها، شکل‌های مشابهی داشته باشند، ممکن است با شکست مواجه شوند. عملکرد توصیف‌گرها به نوع تصویر هم بستگی دارد [۱۲].

در تمام جهات حرکت می‌دهد و مقدار شدت روشنایی را در تمام جهات بررسی می‌کند. این شناساگر، نسبت به مقیاس تغییرپذیر است؛ یعنی ممکن است یک پیکسل در یک مقیاس تصویر به عنوان گوشه در نظر گرفته شود، اما همان پیکسل در مقیاس دیگر همان تصویر به عنوان گوشه در نظر گرفته نشود.

توصیف‌گر شای-توماسی در واقع نمونه بهبود یافته توصیف‌گر هریس است. در نوع شای-توماسی، تغییر جزئی در معیار انتخاب اعمال شده که منجر به عملکرد بسیار بهتر آن نسبت به هریس شده است. در مواردی که شناساگر هریس موفق عمل نمی‌کند، این شناساگر موفق‌تر می‌باشد.

نکته قابل توجه در بین شناساگرهای ویژگی نقطه‌ای مذکور این است که توجه به مطرح بودن بحث تناظر یابی بین دو تصویر چپ و راست، از آنجایی که الگوریتم فست، پیکسل به پیکسل عمل جستجو را انجام می‌دهد، نسبت به دو الگوریتم دیگر که به صورت آماری محل گوشه را تشخیص می‌دهند مناسب‌تر است. این امر در شکل ۱۴ مشهود است. در روش فست، پیکسل گوشه مشخص شده است در حالی که در دو روش دیگر بین دو پیکسل به عنوان گوشه تعیین شده است. مزیت دیگر استفاده از شناساگر فست امکان تعیین حد آستانه برای تعیین مقدار اختلاف شدت روشنایی برای جستجو گوشه است.



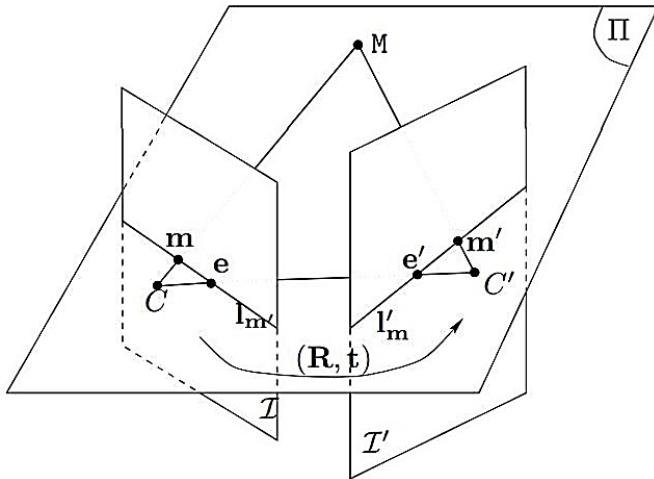
شکل ۱۴: پیکسل تشخیص داده شده به عنوان گوشه با استفاده از سه الگوریتم (الف) هریس، (ب) شای-توماسی و (ج) فست.

در کار گش ۲۹ و همکاران [۱۹] در سال ۲۰۱۵، صحت روش فست مورد بررسی قرار گرفت. نتایج کار آن‌ها نشان داد که صحت نتایج در استفاده از این شناساگر در مقایسه با الگوریتم‌هایی مانند هریس و سرف^{۳۰} به مراتب بهتر است. همچنین آنها اذعان کردند که در استفاده از توصیف‌گرهای یابنده محتمل‌ترین گوشه در تصویر مانند فریک^{۳۱} که مبتنی بر این شناساگر هستند، بهترین صحت به دست آمده است.

لازم به ذکر است که روش‌های معمول تشخیص لبه مانند آشکارساز کنی^{۳۲}، آشکارساز لاپلاس گوسی^{۳۳} آشکارساز پرویت^{۳۴} و سبل^{۳۵} [۱۲]، به عنوان فیلتر بالا گذر (حساسیت آن به تغییرات شدید شدت روشنایی است؛ به طوری که مناطق یکنواخت را تیره کرده و در محل‌های تغییرات شدید، شدت روشنایی را حداکثر می‌کند) در بحث تناظر یابی مناسب نیستند. چرا که به دلیل یکنواختی در برخی لبه‌ها، حتی حرکت در امتداد آن‌ها قابل تشخیص نیست. با توجه به اینکه در لبه، مشتق اول تنها در یک جهت تغییر می‌کند (عمود بر لبه)، لذا لبه‌ها ویژگی‌های مفیدی نیستند. در حالی که در گوشه،

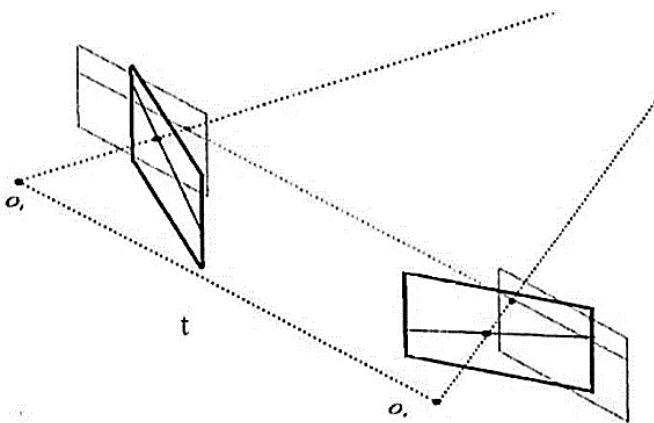
²⁹Ghosh ³⁰SURF (Speeded-up Robust Feature) ³¹FREAK (Fast Retina Key Point) ³²Canny ³³Laplacian of Gaussian ³⁴Prewitt ³⁵Sobel

۷ هندسه اپیپولار در استریو



شکل ۱۵: هندسه اپیپولار [۲۲].

در شکل ۱۶ نمای کلی تبدیل به وضعیت موازی نمایش داده شده است. برای این امر لازم است تا ابتدا دوربین چپ تحت یک ماتریس چرخش، دوران یابد. ترتیب دوران دستگاه مختصات دوربین عبارت‌اند از [۲۳]:



شکل ۱۶: نمایش تبدیل به وضعیت موازی.

۱. ابتدا محور x ، باید با بردار انتقال (t) هم راستا شود.

$$R_{\text{rec}} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_t = \frac{\mathbf{t}}{\|\mathbf{t}\|} = \mathbf{e}_1 \quad (19)$$

۲. سپس جهت محور y دوربین باید قائم باشد، یعنی بر بردار انتقال و محور نوری عمود باشد.

$$\begin{aligned} R_{\text{rec}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \text{Normalized} \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{t_x^2 + t_y^2}} \begin{bmatrix} -t_y \\ t_x \\ 0 \end{bmatrix} = \mathbf{e}_2 \quad (20) \end{aligned}$$

۳. در انتها محور z ، بر دو محور قبلی عمود خواهد شد.

$$\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2 \quad (21)$$

به منظور کاهش فضای جستجو و افزایش سرعت در فرآیند تطبیق، از محدودیت خط اپیپولار^{۳۶} استفاده می‌شود [۲۲]. بدین صورت که با یکسوسازی تصاویر چپ و راست می‌توان جستجو در تصویر سمت راستی را از دو بعد به یک بعد کاهش داد و در واقع برای یافتن متناظر هر نقطه ویژگی، عمل جستجو صرفاً در خط اپیپولار انجام پذیرد.

هندسه اپیپولار برای هر نوع روش بینایی در کاربرد استریو حاکم است. در شکل ۱۵، وضعیت دوربین‌ها را نشان می‌دهد که در آن I و I' ، صفحه تصویر دوربین‌های چپ و راست و C و C' ، مراکز نوری آن‌ها هستند. همچنین خط CC' ، خط مبناست که مرکز دوربین‌ها را به هم وصل می‌کند. محل تقاطع خط CC' با صفحات I و I' ، نقاط e و e' است که به ترتیب اپیپول^{۳۷} چپ و راست نامیده می‌شوند. برای هر نقطه دلخواه M در فضا، یک صفحه اپیپولار متناظر (II) تعریف می‌گردد که از آن نقطه و مرکز دوربین‌ها عبور می‌کند. تصویر M بر صفحه تصویر دوربین‌ها با m و m' نشان داده شده است. همچنین محل تقاطع صفحه اپیپولار با I و I' ، خطوط اپیپولار متناظر با M ، I'_m و I_m هستند که از نقاط e و e' عبور می‌کنند [۲۲].

موقعیت دوربین راست نسبت به دوربین چپ توسط ماتریس R و t تعیین می‌شود. بنابراین هر نقطه (X, Y, Z) در سیستم مختصات دوربین چپ با نقطه (X', Y', Z') در سیستم مختصات دوربین راست متناسب است به طوری که:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + t, \quad (16)$$

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}, \quad (17)$$

$$\begin{cases} R_{11} = \cos \phi_y \cos \phi_z, \\ R_{12} = \cos \phi_x \sin \phi_z + \sin \phi_x \sin \phi_y \cos \phi_z, \\ R_{13} = \sin \phi_x \sin \phi_z - \cos \phi_x \sin \phi_y \cos \phi_z, \\ R_{21} = -\cos \phi_y \sin \phi_z, \\ R_{22} = \cos \phi_x \cos \phi_z + \sin \phi_x \sin \phi_y \sin \phi_z, \\ R_{23} = \sin \phi_x \cos \phi_z + \cos \phi_x \sin \phi_y \sin \phi_z, \\ R_{31} = \sin \phi_y, \\ R_{32} = -\sin \phi_x \cos \phi_y, \\ R_{33} = \cos \phi_x \cos \phi_y. \end{cases} \quad (18)$$

$$t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}.$$

زوایای ϕ_x ، ϕ_y ، ϕ_z چرخش اوپلر حول محورهای X ، Y و Z هستند. طبق روابط مطرح شده در بحث هندسه اپیپولار بین دو صفحه تصویر چپ و راست، به منظور پیدا کردن نقطه متناظر m در صفحه دوربین C' ، جستجو بر روی خط اپیپولار کافی خواهد بود.

³⁶epipolar line ³⁷epipole

مراجع

- [1] Hornberg, A. *Handbook of Machine Vision*. Weinheim WILEY-VCH Verlag GmbH & Co KGaA, 2017.
- [2] Balafar. *Machine Vision principle and Image processing*. ilar, 2015.
- [3] Bhatti, A. Current advancements in stereo vision. *InTechy*, pp. 11–22, 2012.
- [4] Hansard, M., Lee, S., Choi, O., and Horaud, R. P. Time-of-flight cameras principles, methods and applications. *Springer*, 2012.
- [5] Foix, S., Alenya, G., and Torras, C. Lock-in time-of-flight (tof) cameras: a survey. *IEEE Sensors Journal*, 11:1917–1926, 2011.
- [6] Kazmi, W., Foix, S., Alenya, G., and Andersen, H. J. Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: analysis and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88:128–146, 2014.
- [7] Corso, J. *Geometric Camera Calibration*. Electrical Engineering & Computer Science University of Michigan, 2014.
- [8] Yangao, L., Wang, B., R. Zhang, H. Zhou, and Wang, R. Analysis on location accuracy for the binocular stereo vision system. *IEEE Photonics Journal*, 10, 2018.
- [9] Li, X., Yao, J., and Zhang, J. Comparison and selection of camera models with lens distortion. *J. Comput.-Aided Des. Comput. Graph*, 27:824–831, 2015.
- [10] Bouguet, J.-Y. www.mathworks.com.
- [11] Hamzah, R. A., Kadmin, A. F., Hamid, M. Saad, and Ghani, F. A. Improvement of stereo matching algorithm for 3d surface reconstruction. *Signal Processing: Image Communication*, 65:165–172, 2018.
- [12] J.Flusser and B.Zitova. Image registration methods: a survey. *Image and vision computing*, 21:977–1000, 2003.
- [13] Radke, R. J., Andra, S., Al-kofahi, O., and Roysam, B. Image change detection algorithms: a systematic survey. *Image Processing, IEEE Transactions*, 14:294–307, 2005.
- [14] Lukashovich, P., Zalesky, B., and Ablameyko, S. Medical image registration based on surf detector. *Pattern Recognition and Image Analysis*, 21:519–521, 2011.
- [15] Zhang, Q., Wang, Y., and Wang, L. Registration of images with affine geometric distortion based on maximally stable extremal regions and phase congruency. *Image and Vision Computing*, 36:23–39, 2015.
- [16] Rosten, E. and Drummond, T. Fusing points and lines for high performance tracking. vol. 2, pp. 1508–1511, 2005.
- [17] Harris, C. and Stephens, M. *Handbook of Machine Vision*. 1988.
- [18] Shi, J. and Tomasi, C. Good features to track. pp. 593–600, 1994.
- [19] Ghosh, P., Pandey, A., and Pati, U. C. Comparison of different feature detection techniques for image mosaicing, 2015.

در نهایت ماتریس چرخش برای یکسوسازی تصویر چپ به دست می‌آید:

$$R_{rec} \begin{bmatrix} e_1^T \\ e_2^T \\ e_3^T \end{bmatrix} \quad (22)$$

ماتریس به دست آمده برای چرخش دوربین چپ است. برای یافتن چرخش نقاط از ماتریس R_{rec}^{-1} استفاده می‌شود. با توجه به متعامد بودن ماتریس R_{rec} ، رابطه $R_{rec}^{-1} = R_{rec}^T$ برقرار است. ماتریس دوران برای دوربین راست به صورت رابطه (۲۳)، قابل محاسبه خواهد بود.

$$R_r = RR_{rec} \quad (23)$$

R ماتریس دوران بین دوربین‌های چپ و راست است.

۸ خلاصه و جمع‌بندی

درک خصوصیات سه‌بعدی یک منظره واقعی و یافتن ویژگی‌های اجسام موجود در آن از مهم‌ترین زمینه‌های تحقیقاتی در بینایی ماشین و کامپیوتر است که توسط روش بینایی استریو انجام می‌شود. اصل اساسی ماشین بینایی تصویر است. یک تصویر در واقع یک سیگنال است که می‌توان آن را با یک آرایه دوبعدی نمایش داد. این تصویر توسط وسیله‌ای به نام دوربین قابل تولید است. مدل ساده شده برای توصیف دوربین با عنوان مدل دوربین حفره کوچک نام‌گذاری شده است. با استفاده از این مدل روابط اولیه حاکم در داخل دوربین به راحتی قابل توصیف است. در تصویربرداری پرسپکتیو، یکی از ابعاد هر نقطه از فضا هنگام پرتوافکنی بر صفحه تصویر حذف می‌شود. برای بازیابی این بعد از دست‌رفته، روش‌های بازیابی بعد سوم از تصاویر دوبعدی مطرح می‌شوند که یکی از پر کاربردترین آن‌ها روش استریو است. به منظور تعیین رابطه نگاهت نقطه بین دستگاه مختصات جهانی و دستگاه مختصات پیکسلی در صفحه تصویر، فرآیند کالیبراسیون برای هر دوربین استفاده شده در روش استریو مطرح می‌شود. برای کالیبره کردن دقیق دوربین‌ها علاوه بر مدل خطی بایستی مدل غیرخطی دوربین را نیز در نظر گرفت، به طوری که انحرافات حاصل از خطاهای ساخت دوربین و مونتاژ آن نیز در نظر گرفته شود. به منظور استخراج مختصات سه‌بعدی نقاط ویژگی بحث انطباق و پیدا کردن نقاط متناظر بین تصاویر چپ و راست مطرح می‌شود. به طور کلی روش‌های انطباق تصویر بر اساس ویژگی‌ها به دو دسته روش‌های مبتنی بر ناحیه و روش‌های مبتنی بر ویژگی در انطباق خودکار تقسیم می‌شوند. روش‌های مبتنی بر اطلاعات مشترک از میان دیگر روش‌های تطبیق مبتنی بر مقادیر روشنایی مقاوم‌تر است؛ چون این روش تنها از رابطه آماری بین تصاویر استفاده می‌کند. روش‌های مبتنی بر مقادیر شدت روشنایی توانایی تطبیق با دقت بسیار بالایی دارند، اما این روش‌ها ضعف‌هایی نیز دارند. به منظور کاهش فضای جستجو و افزایش سرعت در فرآیند تطبیق، از محدودیت خط اپیپولار استفاده می‌شود. بدین صورت که با یکسوسازی تصاویر چپ و راست می‌توان جستجو در تصویر سمت راستی را از دو بعد به یک بعد کاهش داد و در واقع برای یافتن متناظر هر نقطه ویژگی، عمل جستجو تنها در خط اپیپولار انجام پذیرد.

- [20] Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., and Gopinath, K. Brain functional localization: a survey of image registration techniques, 2007.
- [21] Remondino, F., El-Hakim, S., Gruen, A., and Zhang, L. Turning images into 3-d models, 2008.
- [22] Xu, Z. Z. Gang. Epipolar geometry in stereo, motion and object recognition a unified approach, 1996.
- [23] Monasse, P., Morel, J. M., and Tang, Z. Three-step image rectification. Aberystwyth, United Kingdom, 2010.