

فصل اول: روش‌های فاصله‌سنجی ویدیویی

در این فصل به معرفی اجمالی روش‌های فاصله‌سنجی بر اساس پردازش تصاویر ویدیویی می‌پردازیم. سعی گردیده است که ایده اصلی و اصول پایه‌ای در هر روش شرح داده شود.

۱-۱- فاصله‌سنجی در بینایی ماشین

فاصله‌سنجی یا عمق‌سنجی مقوله‌ای پراهمیت و پرکاربرد در بینایی ماشین می‌باشد؛ تا آنجا که هرگونه اندازه‌گیری با استفاده از دوربین مستلزم بکارگیری روش‌های پردازش تصویر دقیق برای تخمین فاصله (یا عمق) نقاط شاخص می‌باشد. از طرفی، عمق‌سنجی و سرعت‌سنجی دو مقوله بسیار نزدیک به هم می‌باشند تا آنجا که در بسیاری از موارد، ارائه پاسخ برای یکی منجر به حل مساله دیگر می‌شود. به عنوان مثال، برای اندازه‌گیری سرعت یک متحرک (V) کافی است فاصله آن تا دوربین (Z) در هر لحظه اندازه‌گیری شود:

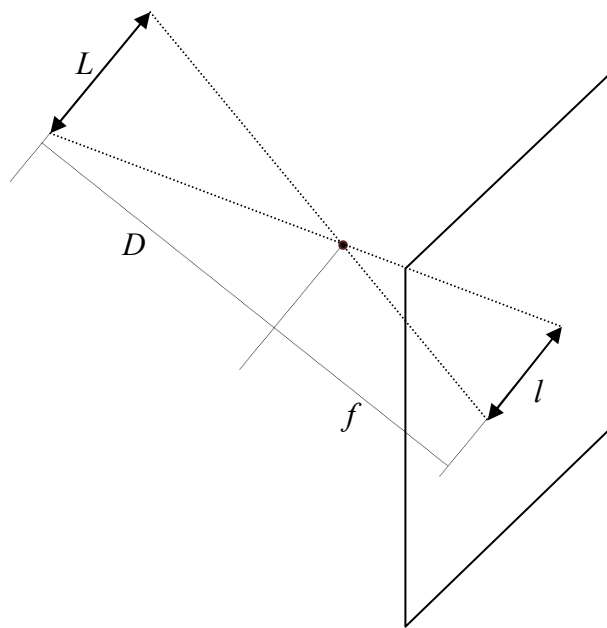
$$V = -\frac{dZ}{dt} \quad (1-1)$$

بعلاوه، در بسیاری موارد، دوربین بر روی یک متحرک قرار گرفته و تصویربرداری در حین حرکت انجام می‌شود. با فرض داشتن سرعت متوسط متحرک (\bar{V})، جابجایی دوربین در مدت زمان Δt عبارت است از:

$$\Delta X = -\bar{V} \Delta t \quad (2-1)$$

از طرف دیگر، اولین و مهمترین نیاز در یک سیستم عمق‌سنجی/سرعت‌سنجی برپایه بینایی ماشین، داشتن اطلاعات کالیبراسیون دوربین می‌باشد. در حقیقت، مهمترین شاخص برای اندازه‌گیری فاصله/سرعت داشتن پارامترهای ذاتی دوربین شامل فاصله کانونی، ابعاد CCD، فاصله دوربین‌ها از یکدیگر (تنها در بینایی دو یا چند دوربینی) و مدل اعوجاجات لنز می‌باشد. تاکنون روش‌های متعددی برای محاسبه پارامترهای فوق و کالیبراسیون دوربین ارائه شده است. باید توجه داشت که اگر لنز از نوع زوم (مثلا با فاصله کانونی بالاتر از ۵۰ میلیمتر) باشد، می‌توان لنز را ایده‌آل و بدون اعوجاج فرض نمود. زیرا در این حالت، تنها از ناحیه مرکزی لنز برای تشکیل تصویر استفاده می‌شود.

با وجود این، اطلاعات کالیبراسیون به تنهایی برای عمق‌سنجی/سرعت‌سنجی کافی نمی‌باشد. زیرا، با داشتن ابعاد CCD می‌توان اندازه واقعی تصویر (l) بر روی سنسور را محاسبه نمود. بعلاوه، فاصله کانونی بیانگر فاصله تصویر از کانون لنز می‌باشد، لذا مطابق شکل ۱-۱، با فرض موازی بودن هدف با سنسور تصویربرداری، می‌توان یک رابطه تشابه بین مثلث‌های تصویر و هدف نوشت:



شکل ۱-۱: عمق سنجی تنها با داشتن پارامترهای ذاتی دوربین امکان پذیر نیست.

$$\frac{L}{D} = \frac{l}{f} \quad (۳-۱)$$

به این ترتیب، یک معادله با دو مجهول L و D (به ترتیب، طول هدف و فاصله مرکز اپتیکی دوربین تا آن) بدست می آید. حل معادله فوق، نیازمند استفاده از اطلاعات کمکی، جدای از سیستم تصویربرداری می باشد. شاید بتوان کلیه روش هایی که تاکنون محققین برای حل معادله (۳) ارائه داده اند را به چهار دسته کلی تقسیم نمود: (۱) بینایی دو/چند دوربینی، (۲) نگاشت صفحه به تصویر، (۳) استفاده از شاخص های محیطی و (۴) روش رشد ناحیه که در ادامه هر یک به اجمال معرفی می شود.

۱-۲- بینایی دو/چند دوربینی

در روش بینایی دو/چند دوربینی^۱ با استفاده از تطبیق شی در دو/چند دوربین، فاصله شی تا دوربین محاسبه می شود. به عنوان مثال، فرض کنید مطابق شکل ۱-۲، تصویر نقطه w با مختصات جهانی (X_w, Y_w, Z_w) در دو دوربین متفاوت، منطبق بر نقاط با مختصات به ترتیب (x_1, y_1) و (x_2, y_2) باشد. اگر فاصله کانونی دوربین های مذکور یکسان، محورهای اپتیکی (نوری) آنها موازی، و محور مختصات x آنها نیز برهم منطبق باشد (این سه

^۱ Stereo/multi camera vision

شرط پس از کالیبراسیون دوربین‌ها، به صورت مجازی و با تصحیح تصاویر دریافتی از آنها قابل تحقق است)،
 آنگاه می‌توان نوشت:

$$\begin{cases} X_1 = \frac{x_1}{f}(f - Z_1) \\ X_2 = \frac{x_2}{f}(f - Z_2) \end{cases} \quad (4-1)$$

که (X_1, Y_1, Z_1) و (X_2, Y_2, Z_2) مختصات نقطه w در دستگاه مختصات به ترتیب دوربین‌های اول و دوم می‌باشد.
 اگر دستگاه مختصات جهانی منطبق بر دستگاه مختصات دوربین اول باشد، می‌توان نوشت:

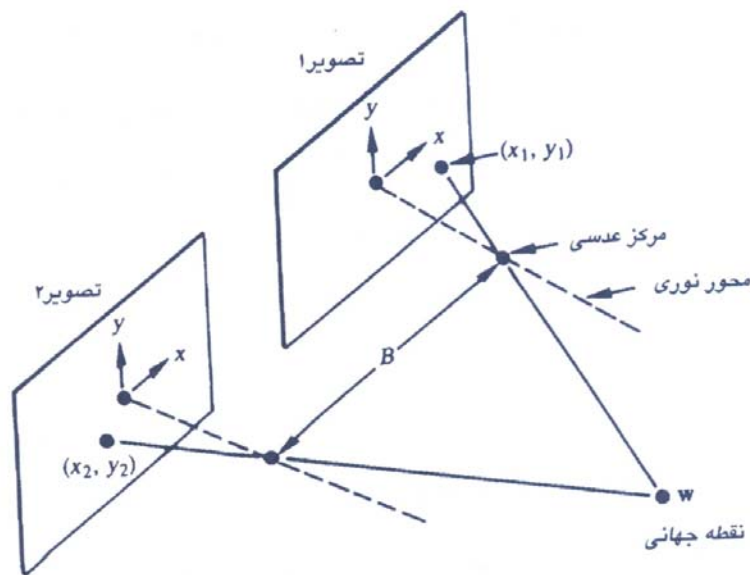
$$\begin{cases} Z_1 = Z_2 = Z_w \\ Y_1 = Y_w \\ X_1 = X_w \\ X_2 = X_1 + B \end{cases} \quad (5-1)$$

از حل دستگاه معادلات (4-1) و (5-1) داریم:

$$Z = f - \frac{X_w B}{x_2 - x_1} \quad (6-1)$$

بنابراین با داشتن فاصله کانونی (f) و فاصله دوربین‌ها (B) می‌توان عمق نقطه w را محاسبه نمود. مهمترین مزیت بینایی دو/چند دوربینی محاسبه عمق، به صورت ذاتی و مستقل از عوامل مزاحم بیرونی است.
 با وجود این، عمق‌یابی با بینایی دو/چند دوربینی خالی از دشواری نیز نیست. اول، باید موقعیت هدف در دو تصویر با دقت بالا بدست آید. این خواسته مستلزم استفاده از روش‌های دقیق پردازشی برای تطبیق تصاویر دو دوربین می‌باشد. دوم، اگر هدف و دوربین نسبت به هم متحرک باشند، تصاویر دو دوربین باید کاملاً به صورت همزمان اخذ گردند تا x_1 و x_2 مختصات یک شی در دو دوربین باشد. سوم، مطابق معادله (6-1)، با افزایش عمق، پارامتر توازن $\Delta x = x_2 - x_1$ کاهش می‌یابد، لذا با توجه به محدود بودن رزولوشن تصویر و اثر نویز (ناشی از اختلاف زمانی در تصویربرداری یا خطا در تطبیق تصاویر)، خطای عمق‌یابی به سرعت رشد می‌کند. در این حالت، بجز افزایش فاصله دوربین‌ها (B)، راه‌کار مؤثر دیگری برای مقابله با نویز و خطاهای پردازشی وجود ندارد. به طور کلی، با توجه به محدودیت در افزایش فاصله بین دو دوربین، عمق‌سنجی دقیق برای فواصل خیلی دور با دو/چند دوربین بسیار دشوار خواهد بود.

¹ Parity



شکل ۱-۲: روش محاسبه عمق در بینایی دو دوربینی

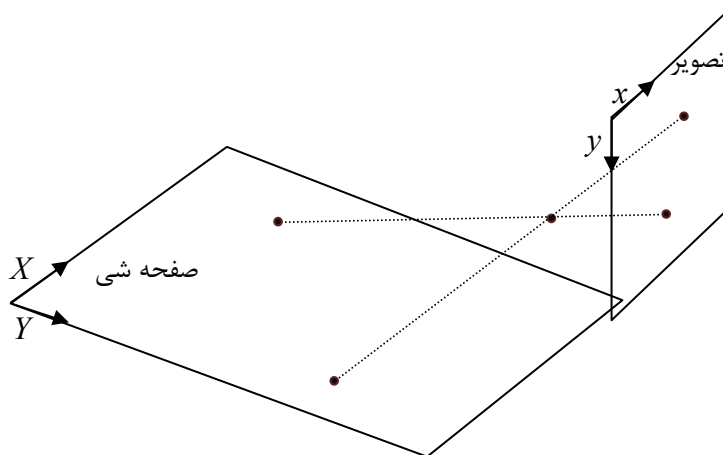
۱-۳- نگاشت صفحه به تصویر

مطابق شکل ۱-۳، مشکل اصلی عمق‌سنجی با یک دوربین در آن است که تمام نقاط واقع بر یک خط نوری عبوری از مرکز کانونی لنز، بر یک نقطه واحد در تصویر منطبق می‌گردند. یعنی تصویربرداری یک نگاشت چند به یک است. حال اگر بدانیم که شی ما (مثلا یک خودرو) حتما در یک صفحه مشخص (مثلا سطح جاده) قرار گرفته است، آنگاه مشکل فوق به آسانی حل خواهد گردید. زیرا، هر نقطه از تصویر متناظر با محل تلاقی شعاع نوری مربوطه (عبوری از آن نقطه) با صفحه شی خواهد بود. در نتیجه بین صفحه تصویر و صفحه شی یک نگاشت یک به یک ایجاد می‌گردد. اگر فرض کنیم مختصات یک نقطه در صفحه شی (X, Y) و مختصات نقطه متناظر آن در تصویر (x, y) باشد، می‌توان نشان داد که بین این دو مختصات رابطه زیر برقرار است:

$$\begin{cases} x = \frac{aX + bY + c}{gX + hY + 1} \\ y = \frac{dX + eY + f}{gX + hY + 1} \end{cases} \quad (Y-1)$$

که a تا h هشت پارامتر ثابت می‌باشند. بدیهی است که با داشتن مختصات حداقل چهار جفت نقطه متناظر می‌توان پارامترهای مجهول در معادله (۳) را محاسبه نمود.

اگرچه روش نگاشت صفحه به تصویر بسیار سراسر است بوده و اندازه‌گیری‌های دقیقی ارائه می‌دهد، اما در این روش، دوربین باید ثابت باشد تا نگاشت بین دو صفحه معتبر باقی بماند. لذا استفاده از آن برای پروژه پیشنهادی غیرعملی خواهد بود.



شکل ۱-۳: نگاشت صفحه شی به صفحه تصویر

۴-۱- استفاده از شاخص‌های محیطی

در این روش، از اجسامی که اندازه آنها از قبل مشخص می‌باشد، به عنوان شاخصی برای عمق‌سنجی استفاده می‌شود. به این ترتیب که ابتدا با توجه به اندازه اشیا مشخص موجود، صحنه تحلیل شده و به صورت سه بعدی مدل می‌گردد. سپس، از مقایسه موقعیت شی در تصویر با مدل سه بعدی، فاصله آن تا دوربین محاسبه می‌شود. با توجه به اینکه پس‌زمینه تصاویر مورد نظر در پروژه پیشنهادی یکنواخت (دریا) می‌باشد لذا، استفاده از این روش منتفی خواهد بود.

۵-۱- روش رشد ناحیه

روش رشد ناحیه در حقیقت بیشتر برای سرعت‌سنجی مطرح می‌باشد. فرض کنید اندازه هدف (L)، مشخص بوده و در راستای عمود بر محور دید دوربین قرار گرفته باشد، در این حالت، با توجه به شکل ۱-۱، فاصله هدف تا دوربین برابر است با:

$$D = \frac{f}{l} L \quad (۸-۱)$$

اکنون برای محاسبه سرعت متحرک کافی است از معادله فوق برحسب زمان مشتق بگیریم:

$$\frac{dD}{dt} = -\frac{fL}{l^2} \frac{dl}{dt} \quad (۹-۱)$$

همانطور که مشاهده می‌شود تغییرات طول شی در تصویر، متناسب با عکس تغییرات فاصله آن تا دوربین است. یعنی با نزدیک شدن هدف به دوربین تصویر آن بزرگتر می‌شود و برعکس.

به عنوان مثال، در رابطه با پروژه پیشنهادی، چنانچه هیکل هدف دریایی در تصویر بخوبی مشاهده شده و ابعاد آن از قبل مشخص باشد، می توان از این روش برای عمق سنجی استفاده نمود. اما، بدیهی است که فرآهم آوردن اطلاعات فوق در همه شرایط، غیرممکن است. زیرا اولاً در تصاویر مادون قرمز معمولاً کل هیکل هدف قابل مشاهده نمی باشد و ثانیاً ابعاد شناورها نیز بسیار متنوع می باشد.