

روشی خودکار به منظور کالیبراسیون نسبی و زمانی دوربین‌های غیرحرفه‌ای با هدف تولید ویدئوهای سه‌بعدی

عطیه گنجعلی^۱ و علیرضا صفدری نژاد^{۲*}

چکیده

در این مقاله راهکاری خودکار به منظور تولید ویدئوهای سه‌بعدی از طریق کنارهم قراردادن دو دوربین غیرحرفه‌ای پیشنهاد شده است. عدم امکان تامین هم‌زمانی دوربین‌ها در شروع فیلم‌برداری، نرخ نامشابه نمونه‌برداری فریم‌ها، معلوم نبودن پارامترهای کالیبراسیون داخلی و همچنین محدودیت‌های مربوط به تنظیم سخت‌افزاری ارتباط نسبی دوربین‌ها، چالش‌های این راهکار قلمداد می‌شوند. در راهکار پیشنهادی، ابتدا هم‌زمانی ویدئوها از طریق تناظریابی شاخص‌های زمانی تامین شده و در ادامه مقاطع زمانی توأم با سکون در طول ویدئوها شناسایی می‌شوند. در ادامه، مجموعه‌ای از نقاط متناظر در دو ویدئو به کمک اجرای تناظریابی خودکار شناسایی شده و در روندی اصلاحی مورد پالایش قرار می‌گیرند. نقاط متناظر پالایش شده در برآورد هم‌زمان پارامترهای کالیبراسیون داخلی و نسبی دوربین‌های استریو به کار گرفته شده و در آخر، ویدئوهای سه‌بعدی نرمال شده از طریق بازنمونه‌برداری مبتنی بر هندسه‌ی اپی‌پلار تولید می‌گردند. این روش در مورد چندین ویدئوی سه‌بعدی از چهار جنبه‌ی مختلف کمی و کیفی مورد ارزیابی قرار گرفته است. دقت هندسی مطلوب در تولید ویدئوهای نرمال شده، هم‌زمانی دقیق ویدئوهای سه‌بعدی، تعمیم‌پذیری مطلوب روش پیشنهادی در تولید ویدئوهای سه‌بعدی در شرایط محیطی مختلف و رضایت تماشاگران ویدئوهای سه‌بعدی از منظر درک بصری عمق، از ویژگی‌های نتایج این روش محسوب می‌شوند.

کلید واژه‌ها

انترپوی، توجیه نسبی، تصاویر نرمال، فیلم سه‌بعدی، شار نوری، هم‌زمان سازی خودکار

۱ مقدمه

امروزه فن‌آوری تولید ویدئوهای سه‌بعدی در کاربردهای مختلفی مانند ساخت فیلم‌های سینمایی، بازی‌های رایانه‌ای، مستندنگاری آثار باستانی و هدایت ربات‌ها استفاده می‌شود [۱]. یک ویدئوی سه‌بعدی از تلفیق دو فیلم اخذ شده توسط

این مقاله در بهمن‌ماه سال ۱۳۹۹ دریافت، در فروردین‌ماه سال ۱۴۰۰ بازنگری و در اردیبهشت همان سال پذیرفته شد.

^۱ دانشجوی کارشناسی ارشد فتوگرامتری، گروه ژئودزی و مهندسی نقشه‌برداری، دانشگاه تفرش، تفرش، ایران

رایانامه: geo97.ganjali@tafreshu.ac.ir

^۲ گروه ژئودزی و مهندسی نقشه‌برداری، دانشگاه تفرش، تفرش، ایران

رایانامه: safdarinezhad@tafreshu.ac.ir

نویسنده مسئول: علیرضا صفدری نژاد

دوربین‌های توأم با اختلاف منظر ساخته می‌شود. سازوکار تماشای این ویدئوها مبتنی بر به‌کارگیری تکنیک‌های برجسته‌بینی بوده که به شکل‌های مختلفی هر چشم را به سمت مشاهده‌ی یکی از این فیلم‌ها هدایت می‌کنند [۲].

از نقطه نظر هندسی و محاسباتی، جابجایی مکانی ثبت عوارض در تصاویر پرسپکتیو اخذ شده از منظرهای متفاوت، محتوای لازم به منظور تشخیص عمق را فراهم می‌آورد [۳]. با این حال، صرفنظر از امکان‌پذیر بودن تخمین عمق در تصاویر استریو، درک بصری عمق نیازمند تامین شرایطی خاص در این تصاویر خواهد بود [۴]. این شرایط با هدف تطابق روند اخذ ویدئوهای سه‌بعدی با سازوکار مشابه سامانه‌ی بینایی انسان طراحی می‌شوند [۵]. عدم وجود اختلاف مقیاس محسوس، توازی محورهای نوری دوربین‌های عکسبرداری، اختلاف منظر مناسب

می‌آورد [۱۶]. در دوربین‌های استریو تجاری‌سازی شده، پیش‌نیازهایی همچون تامین هم‌زمانی، نرخ زمانی^۴ مشابه در فیلم‌برداری و ارتباط نسبی دوربین‌ها به صورت سخت‌افزاری تنظیم شده به نحوی که فیلم‌های اخذ شده با کمترین سطح نیاز به پیش‌پردازش، قابلیت به‌کارگیری در سامانه‌های برجسته‌بینی را خواهند داشت [۱۷]. اما در صورت استفاده از ویدئوهای اخذ شده توسط دو دوربین آماتور با هدف تولید ویدئوهای سه‌بعدی، عواملی مانند عدم تشابه پارامترهای داخلی و نرخ نمونه‌برداری زمانی دوربین‌ها، عدم تامین هم‌زمانی در شروع فیلم‌برداری و پیچیدگی تنظیم سخت‌افزاری دوربین‌ها با هدف اخذ ویدئوهای مناسب به منظور درک بصری عمق، به‌عنوان چالش مطرح خواهد بود [۱۸].

تاکنون از هر دو رویکرد سخت‌افزاری و پردازشی به منظور تنظیم نسبی ویدئوهای استریو با هدف تولید تصاویر و ویدئوهای سه‌بعدی استفاده شده است [۱۹]. تنظیم نسبی دوربین‌های استریو به‌کمک تارگت‌های مثلثی با هدف موازی‌سازی خطوط اپی‌پلار، نمونه‌ای از اقدامات مکانیکی در حوزه تنظیم نسبی دوربین‌ها محسوب شده که علیرغم مبانی تئوری ساده، اجرای آن نیازمند تجهیزات و تجربه‌ی تخصصی است [۲۰]. روش‌های پردازشی نیز با اجرای فرایند بازنمونه‌برداری اپی‌پلار، ترازوی محورهای نوری و یکسان‌سازی مقیاس ویدئوهای استریو را فراهم می‌آورند [۲۱].

در حوزه‌ی هم‌زمان‌سازی ویدئوهای توأم با هم‌پوشانی (زمانی و مکانی) نیز راهکارهای مختلفی توسعه یافته است [۲۲]. تکنیک‌های هندسی شاخه‌ای از راهکارهای هم‌زمان‌سازی ویدئوها بوده که در آنها از تارگت‌های متحرک با هندسه‌ی حرکتی معلوم و یا از معلوم بودن ارتباط نسبی دوربین‌های فیلم‌برداری جهت هم‌زمان‌سازی استفاده می‌شود. در این روش‌ها از طریق تناظریابی میان فریم‌های هم‌زمان محتمل، هندسه‌ی حرکتی تارگت متحرک و یا موضع نسبی دوربین‌ها برآورد می‌شود. مقایسه‌ی نتایج برآورد شده با مقادیر معلوم معیار تشخیص هم‌زمانی در این راهکارها تلقی می‌شود [۲۳]. طراحی مسیر حرکتی نامناسب برای تارگت متحرک و یا عدم وجود پویایی حرکتی در ویدئوهای اخذ شده توسط دوربین‌های با موضع نسبی معلوم، چالش‌های این حوزه برشمرده می‌شوند [۲۴].

در مقابل و در زمان فقدان اطلاعات هندسی از صحنه و ارتباط نسبی دوربین‌های فیلم‌برداری، متناسب با میزان هم‌پوشانی مکانی میان ویدئوها از روش‌های مبتنی بر مقایسه‌ی رفتار زمانی محتوای تصویری به منظور یافتن بیشینه‌ی همبستگی میان ویدئوها استفاده می‌شود [۲۵] در این رویکرد نیز عدم وجود

و هم‌زمانی در نمایش فیلم‌های اخذ شده از دوربین‌های استریو، شروط لازم به منظور درک بصری مطلوب از یک ویدئوی سه‌بعدی می‌باشند [۶].

تاکنون راهکارهای مختلفی به منظور تولید ویدئوهای سه‌بعدی توسعه یافته است. در یک دسته‌بندی کلی می‌توان این راهکارها را به دو حوزه‌ی روش‌های نیازمند به بازسازی سه‌بعدی صحنه و روش‌های مستقل از آن دسته‌بندی نمود [۷]. در دسته‌بندی نخست، صحنه‌ی فیلم‌برداری به‌کمک ابزارها و تکنیک‌های متفاوتی به صورت سه‌بعدی بازسازی شده و در ادامه، با طراحی و حرکت دوربین‌های استریو به صورت مجازی از صحنه‌ی سه‌بعدی فیلم‌برداری می‌شود. ساخت پویانمایی‌های سه‌بعدی از این قانده تبعیت کرده به نحوی که پویانما پس از ساخت سه‌بعدی صحنه در محیط‌های طراحی سه‌بعدی و با درک پیش‌شرط‌های لازم به منظور اخذ فیلم‌های مناسب برای اهداف برجسته‌بینی، هندسه‌ی نسبی دوربین‌های مجازی را جهت تولید ویدئوهای پرسپکتیو استریو تنظیم می‌نماید [۸]. تعمیم این راهکار به صحنه‌های واقعی با چالش‌هایی همچون کیفیت و کامل بودن بازسازی سه‌بعدی و وجود اشیاء متحرک مواجه است. به‌همین دلیل در مورد صحنه‌های واقعی حداقل یک دوربین به صورت فیزیکی جهت اخذ ویدئو مینا مورد استفاده قرار داشته و فیلم مکمل با هدف برجسته‌بینی، به صورت مجازی و در خلال پردازش ویدئوی مینا و به‌کمک مدل سه‌بعدی سطح تولید می‌شود [۹]. در این روند، به ازای هر فریم از فیلم اخذ شده توسط دوربین، یک فریم مجازی با در نظر گرفتن پارامترهای نسبی و داخلی مناسب به منظور تامین درک بصری عمق، بازنمونه‌برداری می‌گردد [۱۰]. دسترسی به مدل سه‌بعدی سطح و همچنین معلوم بودن پارامترهای لحظه‌ای موقعیت و وضعیت دوربین در ویدئوی مینا، پیش‌شرط اجرای این رویکرد در قلمداد شده که بطور معمول از طریق اجرای تکنیک‌های مثلث‌بندی دسته‌اشعه و تناظریابی‌های متراکم به صورت خودکار تامین می‌گردد [۱۱ و ۱۲]. پیش‌شرط‌های ذکر شده می‌تواند به صورت منفرد توسط ویدئوی مینا و یا با مشارکت هم‌زمان سنجنده‌های فعال^۱ و غیرفعال^۲ دیگر تامین گردد [۱۳]. به‌کاربردن سنجنده‌های فعال همچون پوششگرهای لیزری و همچنین استفاده از تکنیک‌های SfM^۳ در سنجنده‌های غیرفعال در این حوزه استفاده شده است [۱۴].

در رویکرد دوم، از دوربین‌های استریو به منظور اخذ ویدئوهای هم‌زمان توأم با اختلاف منظر استفاده می‌شود [۱۵]. در این روند نیازی به مدل رقمی سطح نبوده و صرفاً تامین هم‌زمانی، پارامترهای داخلی و ارتباط نسبی مناسب بین دوربین‌های استریو شرایط کافی به منظور تولید ویدئوهای سه‌بعدی را فراهم

- 1 Active
- 2 Passive
- 3 Structure from Motion

4 Frame Per Second Rate (FPS Rate)

سامانه‌ی بینایی انسان تحمیل می‌سازند [۲۸]. در صورتیکه ویدئوهای اخذ شده توسط دوربین‌های استریو شرایط مشابه با حالات نسبی قابل درک در سامانه‌ی بینایی انسان را نداشته باشند؛ ارسال مستقل محتوای هر ویدئو به چشم‌ها به کمک ابزارهای برجسته‌بینی منجر به دوبینی و یا برجسته‌بینی موضعی می‌گردد. به همین جهت، ساخت ویدئوهای سه‌بعدی نیازمند تنظیم شرایط نسبی دوربین‌ها بوده که به شکل‌های سخت‌افزاری و یا نرم‌افزاری به اجرا می‌رسند. در شکل سخت‌افزاری، تولیدکننده‌ی دوربین استریو تلاش می‌کند که با انتخاب نوع دوربین‌ها و تنظیم ارتباط نسبی آنها، ویدئوهای مناسب به منظور اهداف برجسته‌بینی اخذ نماید. در راهکارهای نرم‌افزاری نیز، ناهمگونی در پارامترهای داخلی و ارتباط نسبی دوربین‌ها به کمک نگاشت‌های هندسی، اصلاح و ویدئوهای بازنمونه‌برداری شده تولید می‌گردند. در این روند، ویدئوهای پرسپکتیو با محورهای موازی تصویربرداری و مقیاس یکسان به کمک ویدئوهای خام و با معلوم بودن پارامترهای داخلی و ارتباط نسبی دوربین‌ها به صورت محاسباتی تولید می‌شوند [۲۹].

در این پژوهش تلاش شده تا با کنارهم قراردادن دو دوربین گوشی تلفن همراه به صورت ثابت نسبت به یکدیگر و پردازش ویدئوهای اخذ شده از آنها، ویدئوهای سه‌بعدی به صورت خودکار تولید گردد. عواملی همچون:

- ۱- عدم امکان تنظیم نسبی دقیق دوربین‌ها به صورت مکانیکی به منظور اخذ ویدئوهای استریو متناسب با شرایط قابل درک در سامانه‌ی بینایی انسان
- ۲- یکسان نبودن پارامترهای داخلی دوربین‌ها (فاصله اصلی، موقعیت پای قائم مرکز تصویر و ابعاد ماتریس ویدئوهای اخذ شده توسط دوربین‌های استریو)
- ۳- عدم امکان تامین هم‌زمانی دقیق در آغاز فیلم‌برداری برای دوربین‌های استریو
- ۴- نرخ نمونه‌برداری زمانی (fps) متمایز ویدئوهای اخذ شده توسط دوربین‌ها

به‌عنوان چالش‌های این مسیر قلمداد می‌شوند. انتظار می‌رود که راهکار پیشنهاد شده بتواند به صورت خودکار ویدئوهای اصلاح شده‌ی هندسی و زمانی را به منظور به‌کارگیری در تکنیک‌های برجسته‌بینی تولید نماید.

آماده‌سازی دوربین‌ها و اخذ داده، هم‌زمان‌سازی ویدئوها، انتخاب جفت فریم‌های متناظر و مناسب به منظور تخمین ارتباط نسبی و داخلی دوربین‌ها، تولید ویدئوهای نرمال شده و تولید ویدئوهای سه‌بعدی به ترتیب گام‌های روش پیشنهادی را تشکیل داده که در ادامه تشریح شده‌اند.

۱-۲- آماده‌سازی دوربین‌ها و اخذ داده

در راهکار پیشنهادی از کنارهم قرار دادن دو دوربین آماتور به صورت ثابت به منظور تولید ویدئوهای سه‌بعدی استفاده شده

حرکت نسبی میان صحنه و دوربین عامل ایجاد ابهام در هم‌زمانی محسوب می‌شود [۲۶].

در این مقاله تلاش شده تا با مینا قرار دادن ویدئوهای اخذ شده از دو دوربین گوشی همراه توام با هم‌پوشانی مکانی بالا، روندی خودکار به منظور آماده‌سازی ویدئوهای اخذ شده توسط هر دوربین برای اهداف برجسته‌بینی توسعه یابد.

سازوکار روش توسعه‌یافته مبتنی بر چهار گام اصلی است. در گام نخست، عدم تطابق زمانی ناشی از تفاوت در آغاز و همچنین تمایز در نرخ نمونه‌برداری زمانی فیلم‌های اخذ شده از دوربین‌ها، از طریق تناظریابی شاخص‌های زمانی مستخرج از ویدئوها اصلاح می‌گردد. در گام دوم، مجموعه‌ای از فریم‌های مناسب در ویدئوهای هم‌زمان‌سازی شده به منظور برآورد پارامترهای کالیبراسیون داخلی و ارتباط نسبی دوربین‌ها شناسایی می‌گردد. بعد از برآورد پارامترهای ارتباط نسبی دوربین‌ها، نگاشتی به منظور نرمال‌سازی ویدئوها از طریق بازنمونه‌برداری مبتنی بر هندسه‌ی اپی‌پلار انجام می‌شود. آخرین گام نیز به آماده‌سازی ویدئوها برای اهداف برجسته‌بینی اختصاص دارد. در این روش با به‌کارگیری اقدامات پردازشی سخت‌گیرانه تلاش شده تا دستیابی به ویدئوهای سه‌بعدی به صورت خودکار و با قابلیت اعتماد مطلوب به اجرا رسد.

۲ روش شناسایی

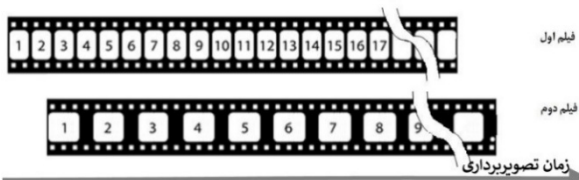
مغز و سیستم عصبی انسان بطور تکوینی آموزش یافته تا بتواند محیط سه‌بعدی را به کمک تلفیق محتوای تصویری چشم‌ها درک نماید. پرتوهای رسیده از عوارض سه‌بعدی به سامانه‌ی بینایی انسان توام با پارالاکس ثبت می‌شوند. جابجایی عوارض در ثبت تصاویر پرسپکتیو متأثر از اختلاف منظر، پارالاکس نامیده می‌شود. درک بصری عمق در انسان از طریق تفسیر ذهنی پارالاکس اتفاق می‌افتد. از نقطه‌نظر هندسی، هر چشم به‌عنوان یک سامانه‌ی تصویربرداری پرسپکتیو از دو درجه‌ی آزادی شامل امکان تغییر در قطر دیافراگم و فاصله کانونی برخوردار است. تغییر این دو پارامتر با هدف تشکیل تصویر واضح در شبکه‌ی صورت می‌پذیرد. ارتباط نسبی چشم‌ها نیز به کمک عضلات بینایی تنظیم می‌شوند. حالات متعددی در ارتباط نسبی چشم‌ها از طریق عضلات بینایی تامین شده که تصاویر اخذ شده در آنها درک سه‌بعدی از عمق را رقم می‌زنند [۲۷]. بر این اساس، تمام حالاتی که ارتباط نسبی چشم‌ها بدون اختلال منجر به درک عمق در سامانه‌ی بینایی انسان شود را می‌توان در طراحی و تنظیم نسبی دوربین‌های استریو مورد استفاده قرار داد.

از منظر هندسی، ۶ درجه‌ی آزادی در ارتباط نسبی دو دوربین وجود دارد. با این حال، ارتباط نسبی چشم‌ها درجه‌ی آزادی کمتری را به واسطه‌ی محدودیت‌های حرکتی، تقارن در چرخش و فاصله‌ی ثابت چشم‌ها برخوردارند. پرتکرارترین تنظیم نسبی، تامین توازی محورهای اصلی چشم‌ها و اخذ تصاویر استریو با اختلاف نامحسوس در مقیاس بوده که کمترین فشار را به

۲-۲- هم‌زمان‌سازی ویدئوها

در روش پیشنهادی دوربین‌های استریو نسبت به یکدیگر ثابتند. در صورت استفاده از ویدئوهای هم‌زمان‌سازی نشده، نمی‌توان انتظار داشت که ارتباط نسبی پایداری بین ویدئوهای استریو وجود داشته باشد. به عبارت بهتر، پایداری ارتباط نسبی میان ویدئوهای استریو هنگام هم‌زمان بودن آنها قابل تأمین خواهد بود. در شرایط استفاده از ویدئوهای هم‌زمان‌سازی نشده، به ازای هر جفت فریم متناظر از دو ویدئو، پارامترهای نسبی متمایزی وجود داشته که روند تغییرات زمانی آنها تابعی از حرکت نسبی دوربین و صحنه خواهد بود. بدیهی است که در صورت هم‌زمان نبودن ویدئوها، به ازای هر جفت فریم متناظر، برآورد پارامترهای نسبی و به تبع آن نگاهت به فضای تصاویر نرمال ضروری خواهد بود. صرف‌نظر از احتمال بروز مشکلاتی همچون فقدان محتوای رادیومتریکی مناسب به منظور شناسایی نقاط متناظر کافی در برآورد پارامترهای نسبی برای تمامی فریم‌های متناظر، این اقدام می‌تواند منجر به نوسانات ناخواسته در درک عمق عوارض و یا اختلال در برجسته‌بینی هنگام وجود ارتباط نسبی نامناسب میان فریم‌های متناظر گردد. بر این اساس، هم‌زمان‌سازی فریم‌های متناظر می‌تواند علاوه بر کاهش محسوس پیچیدگی در تولید ویدئوهای سه‌بعدی، ثبات پارامترهای ارتباط نسبی فریم‌های متناظر را در سراسر فیلم سه‌بعدی فراهم آورد.

نرخ نمونه‌برداری زمانی (fps) ناهمگون و همچنین نامعلوم بودن اختلاف زمانی میان ویدئوهای اخذ شده از دوربین‌ها، دو عامل اصلی بروز اختلالات زمانی قلمداد می‌شوند. پیامد تمایز شدید در fpsها می‌تواند فقدان نمونه‌برداری فریم هم‌زمان را برای دوربین با fps بالاتر بدنبال داشته باشد. بعلاوه، هم‌زمان نبودن فریم‌های متناظر می‌تواند ضمن ایجاد تغییر در پارامترهای نسبی، موجب عدم تطابق در محتوای نمایش یافته برای فریم‌های متناظر هنگام افزایش سرعت نسبی دوربین و صحنه در ویدئوی سه‌بعدی گردد. شکل (۲) مساله‌ی اختلاف زمانی میان ویدئوها و همچنین تمایز در fps را به صورت شماتیک در یک نمودار یک‌بعدی نشان می‌دهد.



شکل ۲- نمایشی شماتیک از تمایز بین fps و اختلاف زمانی آغاز ویدئوها

با توجه به شکل (۲)، بدیهی است که در صورت وجود تمایز در fpsها، یافتن فریم‌های متناظر برای ویدئوی با fps کمتر، احتمال کمتری را از منظر مواجهه با شرایط فقدان فریم متناظر مناسب به همراه خواهد داشت. بر این اساس، در راهکار اتخاذ شده

است. عامل انسانی مفروض به منظور آماده‌سازی دوربین‌ها و اخذ ویدئو غیر متخصص تصور شده که امکانات تنظیم دقیق نسبی دوربین‌ها را نداشته و صرفاً توانایی اجرای اقدامات محدود سخت‌افزاری و نرم‌افزاری را برخوردار است. به همین دلیل، در این بخش صرفاً به مواردی از روند آماده‌سازی دوربین‌ها به منظور اخذ داده اشاره شده که اجرای آنها برای عامل آماتور میسر بوده و همچنین کیفیت نتایج ویدئوهای سه‌بعدی را ارتقاء می‌بخشد.

ثابت قرار دادن دوربین‌ها نسبت به یکدیگر، جلوگیری از تقارب و یا واگرا قرار گرفتن نسبی دوربین‌ها، عدم استفاده از دوربین‌هایی با اختلاف فاحش در فاصله کانونی، فاصله اندک بین دوربین‌ها خصوصاً در مورد دوربین‌های با میدان دید بزرگ (فاصله کانونی کوچک)، مهمترین اقدامات سخت‌افزاری به منظور آماده‌سازی دوربین‌ها محسوب می‌شوند. نویسندگان این مقاله برای تأمین ثبات نسبی دوربین‌ها، پایه‌ای چوبی به صورت شکل (۱) را طراحی و مورد استفاده قرار داده‌اند.



شکل ۱- نمایشی از دستگاه ساخته شده به منظور تأمین ثبات نسبی دوربین‌ها

عدم وجود تقارب و یا واگرایی محور دوربین‌ها نیز صرفاً به صورت بصری از طریق هم‌راستایی گوشی‌های تلفن همراه ایجاد می‌گردد. باید توجه داشت که وجود زوایای تقارب یا واگرایی بزرگ، اعوجاج شدید تصاویر نرمال شده را به دنبال داشته و جدایی بیش از حد دوربین‌ها، پارالاکس‌های بسیار بزرگ را به وجود می‌آورد. اعوجاج بیش از حد، تفسیر نادرست از عمق را بدنبال داشته و پارالاکس‌های بزرگ منجر به دوبینی می‌شوند. از طرفی، یکسان بودن فاصله‌ی کانونی اختلاف مقیاس میان دو ویدئو را کاهش داده و تصاویر نرمال شده کمتر دچار تغییرات مقیاس خواهند شد.

از نقطه نظر نرم‌افزاری، پایداری فوکوس و عدم فعال بودن امکاناتی مانند فوکوس خودکار مهمترین موارد لازم به منظور تنظیم دوربین‌ها هنگام فیلم‌برداری خواهد بود. اعمال قیودی مانند تنظیم دستی fps دوربین‌ها به منظور اخذ ویدئو با نرخ زمانی مشابه نیز می‌تواند کیفیت بصری فیلم‌های تولیدی را ارتقا بخشد. هرچند که روش پیشنهاد شده صرفاً با فرض ثبات پارامترهای نسبی دوربین‌ها بنا شده و تمامی موارد ذکر شده در این بخش توصیه‌هایی به منظور افزایش کیفیت ویدئوهای سه‌بعدی محسوب می‌شوند.

$\{Intersection_i\}$ معیار همبستگی پیروسون به صورت رابطه‌ی (۳) محاسبه می‌شود [۳۰].

$$\rho_i = \frac{\sum_{j=1}^{k_i} (p_j - \bar{p})(q_j - \bar{q})}{\sqrt{\sum_{j=1}^{k_i} (p_j - \bar{p})^2 \sum_{j=1}^{k_i} (q_j - \bar{q})^2}} \quad (3)$$

در صورتیکه Δ به عنوان آرگمان بیشینه‌کننده‌ی معیار همبستگی تعریف گردد (رابطه‌ی ۴): اندیس فریم‌های متناظر بین دو ویدئو از طریق اندیس زوج مرتب‌های مجموعه‌ی $\{Intersection_\Delta\}$ شناسایی می‌شوند.

$$\Delta = \arg \max_i (\rho_i) \quad (4)$$

به واسطه‌ی درونیایی خطی به‌کاررفته در تولید شاخص‌های زمانی، توان تفکیک زمانی مربوط به اندیس زوج مرتب‌های مجموعه‌ی $\{Intersection_\Delta\}$ معادل با ویدئوی با fps بیشتر خواهد بود. بر این اساس، در ادامه با معلوم بودن زمان اخذ هر فریم از ویدئوی با fps کمتر (t_p)، زمان متناظر در ویدئوی با fps بیشتر (t_q) به کمک رابطه‌ی (۵) محاسبه می‌گردد.

$$t_q = t_p + \frac{m - \Delta}{fps_{max}} \quad (5)$$

در این تحقیق از شاخص‌های آماری مختلفی در روند تناظریابی استفاده شده که در جدول (۱) معرفی شده‌اند. هدف از این اقدام یافتن شاخص‌های مطلوب و همچنین حساسیت‌سنجی نتایج به انتخاب نوع شاخص بوده است.

جدول ۱ - شاخص‌های آماری پیشنهادی به منظور تولید سری زمانی در روند تناظریابی زمانی خودکار

شاخص	فرمول
میانگین	$\mu = \frac{\sum_{i=0}^{L-1} i \times P_i}{\sum_{i=0}^{L-1} P_i}$
انحراف معیار	$\sigma = \sqrt{\frac{\sum_{i=0}^{L-1} i^2 P_i}{\sum_{i=0}^{L-1} P_i} - \mu^2}$
نسبت انحراف معیار به میانگین	$S = \frac{\sigma}{\mu}$
ضریب چولگی	$\mu_3 = \frac{\sum_{i=0}^{L-1} P_i \times (i - \mu)^3}{\sigma^3 \times \sum_{i=0}^{L-1} P_i}$
میانگین دامنه‌ی بردارهای شار نوری	$\frac{\sum_{j=1}^n \sqrt{v_{x_j}^2 + v_{y_j}^2}}{n}$
انتروپی	$H = -\sum_{i=0}^{L-1} P_i \log P_i$
مساحت تابع توزیع تجمعی هیستوگرام	$A = \sum_{i=0}^{L-1} \sum_{j=0}^i P_i$
P_i : احتمال رخداد درجه خاکستری در تصویر L: تعداد درجات خاکستری V_x : مولفه بردار جریان تصویری در جهت X	V_y : مولفه بردار جریان تصویری در جهت Y n: تعداد بخش‌های تصویر به منظور تولید بردارهای نوری

تلاش شده تا به ازای تمامی فریم‌های ویدئو با fps کمتر، فریمی متناظر از ویدئو با fps بیشتر شناسایی و به آن الصاق گردد. راهکار توسعه یافته برای هم‌زمان‌سازی ویدئوهای اخذ شده از دو دوربین، مبتنی بر تناظریابی شاخص‌های زمانی است. در این روند ابتدا به ازای هر فریم از هر یک از ویدئوها یک شاخص آماری مستخرج از محتوای تصویری آن فریم تولید می‌گردد. بر این اساس، به ازای هر ویدئو و با معلوم بودن زمان اخذ هر فریم، یک سری زمانی از شاخص‌های آماری تولید می‌گردد. واضح است که فواصل زمانی بین هر شاخص آماری متناسب با fps هر ویدئو خواهد بود. بر این اساس، به دلیل عدم تطابق احتمالی fps ویدئوها، در یک روند درونیایی خطی توان تفکیک زمانی سری زمانی تولید شده برای ویدئو با fps کمتر با دوربین با fps بیشتر یکسان‌سازی می‌شود. این اقدام، توان تفکیک زمانی شاخص‌های آماری دو ویدئو را یکسان و برابر با ویدئوی با fps بیشتر می‌نماید. لازم به ذکر است که در این روند یکی از شاخص‌های زمانی دچار فرایند درونیایی خطی شده و سری زمانی مستخرج از ویدئوی دیگر بدون تغییر به مرحله تناظریابی زمانی منتقل می‌شود.

در ادامه و در اولین گام از روند مشابهت‌سنجی با هدف یافتن اختلاف زمانی بین دو ویدئو، تمامی مجموعه‌های $\{Intersection_i\}$ متشکل از زوج مرتب‌های رابطه‌ی (۱) ایجاد می‌گردند.

$$Intersection = \{(p_m, q_i), (p_{m-1}, q_{i-1}), (p_{m-2}, q_{i-2}), \dots, (p_1, q_{i-m}) \mid p \in P \text{ and } q \in Q\} \\ i = 1, 2, \dots, (m+n) \quad (1)$$

در رابطه‌ی (۱)، $P = \{p_1, p_2, \dots, p_m\}$ به عنوان مجموعه‌ی یکی از شاخص‌های آماری استخراج شده برای ویدئو با fps کمتر متشکل از m عضو و $Q = \{q_1, q_2, \dots, q_n\}$ مجموعه‌ی مشابه برای ویدئوی با fps بیشتر با n عضو است. بدیهی است که در صورت تساوی fps دوربین‌ها نامگذاری P و Q اختیاری خواهد بود. اندیس اعضای مجموعه‌ی P و Q به کمک رابطه‌ی (۲) به زمان متناظر تبدیل می‌شوند.

$$t_k = \frac{j-1}{fps_{max}} \quad (2)$$

در رابطه‌ی (۲)، j اندیس اعضای مجموعه‌ی P و Q و همچنین fps_{max} معادل با fps ویدئوی سریعتر می‌باشند. اندیس‌های مذکور برای مجموعه‌های P و Q به ترتیب در بازه‌ی $[1 \sim m]$ و $[1 \sim n]$ معنادار بوده و زوج مرتب‌های هر $\{Intersection_i\}$ صرفاً به ازای اعضای معنادار تولید می‌گردند. به همین دلیل، تعداد اعضای مجموعه‌های $\{Intersection_i\}$ متناسب با میزان اشتراک ایجاد شده در موقعیت i ام متفاوت بوده که در ادامه با k_i شناخته می‌شوند. در ادامه، به ازای زوج مرتب‌های هر

با توجه به:

- ۱- خودکار بودن فرایند برآورد ارتباط نسبی
- ۲- عدم قطعیت‌های تناظریابی خودکار
- ۳- نامعلوم بودن پارامترهای دقیق داخلی دوربین‌ها (مختصات مرکز تصویر و فاصله اصلی)
- ۴- هندسه‌ی ضعیف دوربین‌های استریو (مقاربت نبودن)

مساله‌ی یافتن فریم‌های مناسب و تناظریابی میان آنها را با چالش‌های بیشتری روبرو می‌سازد. یافتن زمان‌های سکون در طول ویدئوهای هم‌زمان‌سازی شده، انتخاب توزیع شده‌ی نقاط متناظر متناسب با محتوای تصویری زمان‌های سکون و حصول اطمینان از توزیع مکانی مطلوب نقاط متناظر، مواردی به‌شمار رفته که در روند برآورد پارامترهای داخلی و نسبی دوربین‌ها لحاظ شده‌اند.

زمان‌های رخداد سکون در ویدئوها را می‌توان به‌کمک شاخص‌های مختلفی اندازه‌گیری کرد. بردارهای جریان نوری را می‌توان به‌عنوان یکی از معیارهای حرکت نسبی صحنه و دوربین قلمداد نمود [۳۲]. علیرغم مزایای به‌کارگیری بردارهای جریان نوری، جابجایی‌های بزرگ و فقدان محتوای رادیومتریکی مناسب از عوامل عدم قطعیت در برآورد صحیح بردارهای جریان نوری محسوب می‌شوند.

در صورت بی‌حرکت ماندن صحنه و دوربین در زمان اخذ ویدئو، محتوای تصویری ثبت شده در ویدئوها در بازه‌ی زمانی سکون، ثابت می‌ماند. به‌همین دلیل، پارامترهای آماری استخراج شده از فریم‌های بدون حرکت در این محدوده‌های زمانی، ثابت خواهد بود. شاخص‌هایی مثل انترپوی و آماره‌های مستخرج از هیستوگرام تصویر در زمره‌ی این شاخص‌های آماری قرار دارند. اما صحنه‌های توأم با بافت تکراری می‌تواند به‌عنوان یک چالش در این زمینه تلقی شده که علیرغم وجود حرکت در ویدئوها، شاخص‌های آماری پایدار باقی بمانند.

در این تحقیق بطور هم‌زمان از پتانسیل سری زمانی شاخص‌های مبتنی بر جریان نوری و همچنین شاخص‌های آماری محتوای تصویری استفاده شده است. اینکار با هدف هم‌افزایی و پوشش نقصان‌های هر یک، در روند شناسایی مقاطع زمانی سکون در ویدئوها استفاده شده است.

با توجه به موفقیت‌آمیز بودن شاخص انترپوی نسبت به سایر شاخص‌های آماری در آزمون تناظریابی زمانی، در این بخش نیز این شاخص برای شناسایی مناطق سکون انتخاب شده است. برای این منظور به ازای تمامی فریم‌ها در ویدئوی با fps بیشتر، انترپوی نرمال‌شده‌ی هیستوگرام $(e_i, i = 1, 2, \dots, k_A)$ و همچنین میانگین طول بردارهای جریان تصویری $(o_i, i = 1, 2, \dots, k_A)$ محاسبه می‌گردد. انتظار می‌رود در شرایط سکون، میانگین طول بردارهای جریان نوری مقداری نزدیک به صفر و شاخص انترپوی نیز مقدار عددی ثابتی را در

مطابق با مفاد جدول (۱)، از ۷ شاخص مختلف به‌منظور تولید سری‌های زمانی استفاده شده است. از این بین شش شاخص از آماره‌های مستخرج از هیستوگرام تصویر به‌دست آمده و یک شاخص مربوط به بردارهای جریان تصویری است.

بردارهای جریان نوری با فرض ثبات در محتوای تصویری فریم‌های متوالی، میزان جابجایی هر ناحیه از تصویر را به‌کمک رابطه‌ی (۶) تخمین می‌زنند.

$$I_x \times v_x + I_y \times v_y = -I_t \quad (6)$$

در رابطه‌ی (۶)، I_x ، I_y و I_t برترتیب مشتق در جهات x ، y و زمان بوده و همچنین $[v_x, v_y]$ مولفه‌های بردار جریان تصویری هستند. این بردارها را می‌توان به‌صورت یک شبکه‌ی منظم در فضای تصویر متناسب محتوای رادیومتریکی هر سلول از شبکه محاسبه نمود. تاثیر انتخاب نوع شاخص‌های زمانی به‌کارگرفته شده و نتایج کسب شده از این مرحله در بخش نتایج ارائه و مورد بحث قرار گرفته است.

۳-۲- انتخاب جفت فریم‌های متناظر به‌منظور برآورد ارتباط نسبی ویدئوها

از منظر سخت‌افزاری ارتباط نسبی دوربین‌های استریو در طول زمان پایدار است. با این‌حال در صورت هم‌زمان نبودن فریم‌های نمایش یافته در ویدئوی سه‌بعدی، ارتباط نسبی ناپایدار در خلال پخش ویدئو مشاهده خواهد شد. این ناپایداری در شرایط افزایش سرعت نسبی دوربین و صحنه تشدید شده و اساساً در زمان سکون، ناپایداری رخ نمی‌دهد. ارتباط نسبی دوربین‌ها از طریق نقاط متناظر بین دو تصویر برآورد می‌شود. معلوم بودن پارامترهای ارتباط نسبی لازمه‌ی بازنمونه‌برداری اپی‌پلار به‌منظور موازی‌سازی راستاهای پارالاکس و هم‌مقیاس نمودن ویدئوهای اخذ شده است. هم‌زمان‌سازی انجام شده در بخش (۲-۲) برای ویدئوها، شرایط پایداری ارتباط نسبی را برای فریم‌های متناظر فراهم می‌آورد. با این‌حال انتخاب فریم‌های متناظر در شرایط سکون قابلیت اعتماد بیشتری را در برآورد ارتباط نسبی تأمین خواهد ساخت. به‌عبارت بهتر، انتخاب فریم‌های متناظر در زمان وجود حرکت به واسطه‌ی عواملی همچون:

- ۱- دقت تناظریابی زمانی
- ۲- بازنمونه‌برداری زمانی ویدئوها به‌منظور یکسان‌سازی fps
- ۳- اثر شاتر جاروب‌شونده^۱ [۳۱] می‌تواند صحت تخمین ارتباط نسبی را مخدوش سازد. در این بخش تلاش شده تا با تمهیداتی، به‌صورت خودکار فریم‌های متناظر مناسب، در زمان‌های سکون به‌منظور برآورد ارتباط نسبی دوربین‌های استریو شناسایی گردند.

¹Rolling Shutter

متناظر از طریق اعمال روش RANSAC [۳۴] با معیار شایستگی فاصله از خطوط اپی پلار [۳۵]، نقاط متناظر شناسایی می‌شوند. اما با توجه به محتوای تصویری غیرقابل پیش‌بینی در فریم‌های توأم با سکون، تلاش شده تا از مشارکت جفت‌فریم‌های متناظر بیشتری به منظور رفع مشکل محتمل در توزیع یا تعداد نامناسب نقاط متناظر به‌کار گرفته شود. بر این اساس، به‌منظور حصول اطمینان از تامین افزونگی کافی از نظر تعداد و توزیع مکانی برای نقاط متناظر، از پتانسیل تمامی مقاطع زمانی سکون استفاده شده است. به‌عبارت بهتر، به ازای هر مقطع زمانی شناسایی شده به‌عنوان سکون، جفت فریمی که کمترین مقدار C را برخوردار باشد برای تناظریابی انتخاب می‌شود. به هر نقطه‌ی متناظر، هزینه‌ای معادل با مجموع شباهت توصیف‌گرهای SIFT و مقدار C مربوط به زوج فریم به‌کارگرفته شده، اختصاص می‌یابد. هر یک از این مقادیر به منظور همسان‌سازی مقیاسی تقبل از جمع شدن در بازه صفر تا یک نرمال‌سازی می‌گردند.

بعد از اجرای روند تناظریابی مطابق با ضوابط تشریح شده برای تمامی جفت فریم‌های متناظر، لیستی تجمیع شده از تمامی نقاط تناظریابی شده تولید می‌گردد. بدیهی است که به واسطه تمایز در محتوای تصویری مناطق سکون، تراکم و توزیع نقاط متناظر در هر جفت‌فریم متناظر متفاوت خواهد بود. بعد از تجمیع تمامی تناظرهای شناسایی شده، تراکم و توزیع نقاط متناظر در فضای تصویر اصلاح می‌گردد. این روند یک فرایند حذفی به‌شمار رفته که با هدف انتخاب نقاط متناظر، توأم با تنوع کافی در عمق و توزیع شده در سراسر فضای تصویر طراحی شده است.

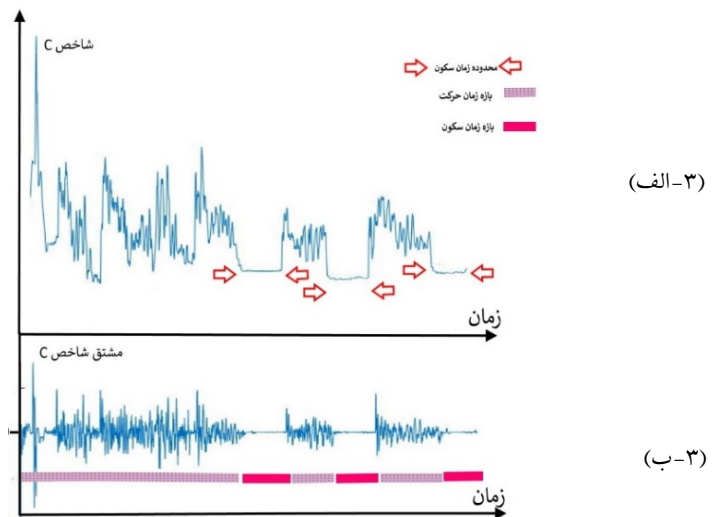
برای این منظور، با فرض اینکه هر نقطه‌ی متناظر در فریم‌های متناظر از دو مختصات تصویری $[X_I, Y_I]$ و $[X_T, Y_T]$ تشکیل شده باشد؛ لیستی از مختصات سه‌بعدی به‌صورت $[X_I, Y_I, X_I - X_T]$ برای تمامی نقاط متناظر تولید می‌گردد. فضای سه‌بعدی گسترش یافته توسط مختصات‌های مورد بحث به $7 \times 6 \times 6$ سلول شبکه‌بندی می‌شود. سپس میانگین تعداد نقاط واقع در سلول‌های ناتهی (μ) محاسبه شده و ضریبی از آن ($\mu\alpha$) به‌عنوان حداکثر مجاز تعداد نقاط واقع در هر سلول از شبکه انتخاب می‌شود. هر سلول از شبکه که تعداد نقاط واقع در آن بیش از آستانه‌ی $\mu\alpha$ باشد؛ وارد فرایند حذفی شده و نقاط واقع در آن سلول متناسب با هزینه هر نقطه‌ی متناظر حذف می‌گردند. تعداد نقاط حذف شده معادل با اختلاف بین $\mu\alpha$ از تعداد نقاط نقاط واقع در هر سلول خواهد بود. ضریب α به‌عنوان پارامترهای تنظیمی این روش محسوب شده که در شرایط ازدیاد نقاط متناظر کمتر از یک و در شرایط کم بودن نقاط متناظر بیشتر از یک در نظر گرفته می‌شود. در تمامی آزمون‌های به‌کارگرفته شده در این تحقیق، مقدار عددی این ضریب معادل

بازه‌ی زمانی سکون داشته باشند. از سوی دیگر، مقدار عددی انترویی میزان یکنواختی توزیع درجات خاکستری در تصویر را کمی‌سازی می‌نماید که ارتباط معناداری را با محتوای رادیومتریکی تصویر برخوردار است. نسبت دو شاخص ذکر شده به‌صورت رابطه‌ی (۷) برای شناسایی فریم‌های توأم با سکون استفاده شده است.

$$C_i = \frac{O_i}{e_i} \quad \text{رابطه‌ی (۷)}$$

$$i = 1, 2, \dots, k_{\Delta}$$

شکل (۳-الف) نمایی شماتیک از روند تغییرات زمانی معیار C را برای یک ویدئو نمایش می‌دهد. همان‌طور که در این شکل مشخص است، مناطق توأم با سکون به‌صورت پاره‌خط‌هایی موازی با محور زمان (شیب صفر) دیده می‌شوند. طول این نواحی بازه‌ی زمانی سکون را نشان داده و بزرگی C با محتوای تصویری آن بازه‌ی زمانی ارتباط معکوس دارد.



شکل ۳- (الف) نمایی شماتیک از روند تغییرات زمانی معیار C برای یک ویدئو

(ب) نمایی شماتیک از مشتق نمودار زمانی C

شناسایی خودکار بازه‌های زمانی سکون از نمودار C با روندی ساده مبتنی بر مشتق‌گیری از نمودار زمانی C به اجرا می‌رسد. بعد از تولید نمودار مشتق، با تعریف حدآستانه‌ای در اطراف صفر، نمودار مشتق به‌صورت باینری قطعه‌بندی می‌گردد. قطعات بهم‌پیوسته، بازه‌های زمانی سکون را نشان داده که برای هر بازه می‌توان اعتباری متناسب با متوسط بزرگی C اختصاص داد (شکل ۳-ب). به نظر می‌رسد که انتخاب یک جفت فریم متناظر در هر قطعه‌ی شناسایی شده به‌عنوان بازه‌های زمانی سکون، بتواند به‌عنوان نماینده‌ی مناسبی جهت اجرای فرایند تناظریابی خودکار و برآورد پارامترهای ارتباط نسبی دوربین‌ها باشد. در هر زوج فریم متناظر به‌کمک روش تناظریابی عارضه مبنای SIFT [۳۳] با اعمال قیود سخت‌گیرانه‌ی شباهت و تمایز توصیف‌گرها و همچنین اعتبار سنجی هندسی نقاط

در مثلث‌بندی دسته‌اشعه، با مبنا قرار دادن معادله‌ی مشهور شرط هم‌خطی [۳۷]، پارامترهای خارجی دوربین‌ها مشتمل بر موقعیت و وضعیت هر دوربین و همچنین پارامترهای کالیبراسیون داخلی (در صورت مجهول بودن) بطور هم‌زمان و در قالب یک دستگاه معادلات غیرخطی برآورد می‌شود [۳۶]. در شبکه‌ی آزاد به دلیل عدم وجود نقاط کنترلی (نقاط با مختصات معلوم در فضای شیء) صرفاً از پتانسیل نقاط متناظر در برآورد پارامترهای مجهول استفاده می‌شود. نقص در تعریف سیستم مختصات شیء یکی از چالش‌های مثلث‌بندی در شبکه‌های آزاد تلقی شده که برای رفع آن راهکارهای مختلفی اعم از اعمال قیود داخلی و یا معلوم فرض نمودن برخی از پارامترهای خارجی ایستگاه‌های عکسبرداری، استفاده می‌شود [۳۸]. این اقدامات منجر به برآورد نسبی پارامترهای خارجی دوربین‌ها در شبکه‌های آزاد می‌شود.

در این تحقیق، برآورد پارامترهای ارتباط نسبی و کالیبراسیون داخلی دو دوربین با اختلاف منظر اندک و راستای نسبتاً موازی در دستورکار قرار دارد (شکل ۱). این شرایط از منظر اصول طراحی شبکه‌های فتوگرامتری، نامستحکم قلمداد می‌شود. در شبکه‌ی نامستحکم، خطاهای مشاهداتی معمول، عدم قطعیت بیشتری را در برآورد پارامترهای مجهول به دنبال دارد. این شرایط در حل دستگاه معادلات به‌عنوان شرایط بدو وضعی^۱ شناخته می‌شود [۳۸].

راهکار به‌کارگرفته شده در بخش (۲-۳) به منظور شناسایی نقاط متناظر، به واسطه‌ی پتانسیل بهره‌گیری از تمامی مقاطع زمانی سکون می‌تواند تنوع بیشتری در عمق نقاط متناظر و به تبع آن تامین دامنه‌ی گسترده‌تری از پارالاکس را برای مشاهده‌های به‌کاررفته در مثلث‌بندی دسته‌اشعه تامین سازد. با این حال به دلیل وجود صرفاً دو موضع ثابت در اخذ ویدئو و همچنین عدم تقارب کافی در دوربین‌های استریو، برآورد هم‌زمان پارامترهای کالیبراسیون داخلی و نسبی مستعد بروز شرایط بدو وضعی خواهد بود. به همین دلیل اجرای تدابیری همچون برآورد تدریجی مقادیر پارامترهای کالیبراسیون داخلی به منظور تدقیق مقادیر اولیه و به‌کارگیری مقادیر تدقیق شده‌ی کالیبراسیون داخلی به صورت شبه‌مشاهده، جهت جلوگیری از بدو وضعی دستگاه معادلات به اجرا رسیده است.

در مورد چگونگی تعریف سیستم مختصات شیء در فرایند برآورد پارامترهای مجهول، راهکار تثبیت برخی از پارامترهای خارجی دوربین‌ها استفاده شده است. در این راهکار که به‌عنوان فرایند توجیه نسبی دوطرفه شناخته می‌شود؛ راستای محور X در سیستم مختصات شیء در راستای خط واصل مراکز تصویر دوربین‌های استریو در نظر گرفته می‌شود. با این فرض و با ثابت لحاظ کردن فاصله‌ی بین مراکز تصویر، شش پارامتر دورانی

یک انتخاب شده است. شبکه‌کد ارائه شده در الگوریتم (۱) روند اجرای این مرحله را نشان می‌دهد.

الگوریتم (۱): شناسایی نقاط متناظر به‌منظور برآورد پارامترهای ارتباط نسبی دوربین استریو

ورودی‌ها:	
نمودار زمانی $C=[c_1, c_2, \dots, c_{kA}]$	
حدآستانه مشتق C به‌منظور جداسازی قطعات زمانی سکون (T_s)	
ابعاد شبکه‌ی اندازه‌گیری توزیع نقاط متناظر $(t \times u \times v)$	
ویدئوهای هم‌زمان‌سازی شده (بخش ۲-۲)	
ضریب میانگین (α) به‌منظور کاهش تراکم نقاط متناظر	
خروجی‌ها:	
نقاط متناظر به‌منظور برآورد پارامترهای ارتباط نسبی	
۱	محاسبه‌ی مشتق زمانی بردار C به‌صورت $(\frac{dC}{dt_i} = C_{i+1} - C_i)$
۲	اعمال حد آستانه T_s به‌منظور باینری‌سازی بردار $\frac{dC}{dt}$ به‌صورت
	$R(t) = \begin{cases} \left \frac{dC}{dt_i} \right < T_s \rightarrow 1 \\ else \rightarrow 0 \end{cases}$
۳	قطعه‌بندی نواحی بهم‌پیوسته‌ی دارای مقدار 1 در $R(t)$ و حذف قطعات با طول کمتر از fps
۴	انتخاب یک جفت فریم متناظر به ازای کمترین مقدار C در هر قطعه (c_{min})
۵	تناظریابی عارضه‌مبنای RANSAC+SIFT با معیار فاصله از خطوط اپی‌پلار برای هر جفت فریم متناظر
۶	اختصاص مجموع فاصله‌ی توصیفگر SIFT و c_{min} جفت فریم متناظر، به‌عنوان هزینه هر نقطه‌ی متناظر
۷	نگاشت نقاط متناظر شناسایی شده به فضای سه‌بعدی $[X_I, Y_I, X_T - X_T]$ و شبکه‌بندی فضا به $t \times u \times v$ سلول
۸	محاسبه‌ی میانگین آماری تعداد نقاط واقع شده در سلول‌های ناتهی (μ)
۹	حذف نقاط مازاد بر $(\alpha\mu)$ در سلول‌های شبکه بر اساس هزینه نقاط متناظر

۲-۴ - کالیبراسیون داخلی، نسبی و بازنمونه‌برداری

اپی‌پلار

با توجه به ماهیت غیرحرفه‌ای دوربین‌های به‌کارگرفته شده، بدیهی است که علاوه بر پارامترهای ارتباط نسبی دوربین‌ها، مقادیر فاصله‌کانونی و همچنین موقعیت پای قائم مرکز تصویر در آنها نیز بطور دقیق مشخص نباشد. بر این اساس در این مرحله به‌کمک نقاط متناظر شناسایی شده در بخش (۲-۳)، روند برآورد پارامترهای ارتباط نسبی و کالیبراسیون داخلی دوربین‌ها به اجرا می‌رسد. این موضوع از طریق اجرای یک مثلث‌بندی دسته‌اشعه در یک شبکه‌ی آزاد متشکل از دو ایستگاه تصویربرداری (دوربین استریو) به اجرا می‌رسد [۳۶].

¹ Ill posed or ill condition

کافی را به منظور برجسته‌بینی ویدئوهای نرمال شده فراهم می‌آورد [۳۹]. بر این اساس، با مبنا قرار دادن (ω, φ) و κ به عنوان زوایای دورانی یکی از دوربین‌های استریو (به دست آمده از فرایند برآورد ارتباط نسبی)، نگاشت بین مختصات تصویری فریم‌های اولیه و نرمال شده به صورت رابطه‌ی (۸) خواهد بود.

$$\begin{bmatrix} x_n \\ y_n \\ -c_n \end{bmatrix} = \lambda R_3(\kappa_n) R_1(\omega_n) R_3(-\kappa) R_2(-\varphi) R_1(-\omega) \begin{bmatrix} x-x_o \\ y-y_o \\ -c \end{bmatrix} \quad (8)$$

در رابطه‌ی (۸)، $[x, y]$ مختصات متریک پیکسل‌ها در فضای ویدئو اولیه، $[x_n, y_n]$ مختصات متریک متناظر در فضای تصویر نرمال شده، $[x_o, y_o, c]$ پارامترهای کالیبراسیون داخلی ویدئوی اولیه، λ ضریب مقیاس، R_1, R_2, R_3 به ترتیب ماتریس‌های دوران حول محورهای اول، دوم و سوم مختصات c_n فاصله‌ی کانونی ویدئوی نرمال شده می‌باشند. مقادیر κ_n, ω_n و c_n به عنوان پارامترهای دورانی و فاصله کانونی ویدئوی نرمال شده، معادل میانگین مقادیر κ, ω و c به دست آمده برای دوربین‌های استریو در فرایند برآورد پارامترهای کالیبراسیون نسبی و داخلی انتخاب می‌شوند.

رابطه‌ی (۸) بعد از حذف ضریب مقیاس به صورت یک تبدیل پروجکتیو دوبعدی در دو حالت مستقیم و معکوس برای بازنمونه برداری فریم‌های هر یک از ویدئوها مورد استفاده قرار می‌گیرد. لازم به ذکر است که حذف $R_3(\kappa_n)$ از رابطه‌ی (۸) منجر به جهت‌گیری راستای پارالاکس‌ها همسو با سطرهای تصویر نرمال شده می‌گردد.

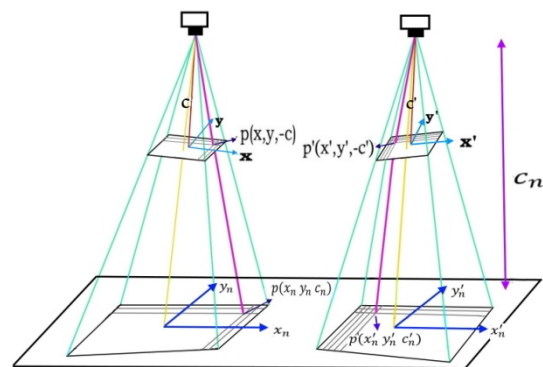
۵-۲- نمایش ویدئوی سه‌بعدی

بعد از تولید ویدئوهای نرمال شده، تمامی اقدامات هندسی لازم به منظور آماده‌سازی ویدئوها جهت نمایش در سامانه‌های پخش ویدئوهای سه‌بعدی به اجرا رسیده است. تنها تنظیم لازم به منظور نمایش ویدئوهای سه‌بعدی، جابجایی دوبعدی آنها نسبت به یکدیگر جهت حذف دوبینی و تشکیل عمق خواهد بود. نقاط متناظر شناسایی شده در بخش (۲-۳) امکان این تنظیم را به صورت خودکار فراهم می‌آورند.

یک ویدئوی سه‌بعدی از طریق ارسال مجزای دو ویدئوی نرمال شده و هم‌زمان‌سازی شده به هر چشم، نمایش می‌یابد. نمایش آنالگلیف، نور قطبیده و یا تکنیک‌های مکانیکی راهکارهای توسعه‌یافته به منظور ارسال مجزای ویدئوهای استریو به سامانه بینایی انسان محسوب می‌شوند [۴۰]. در این تحقیق از روش نمایش آنالگلیف به واسطه‌ی سازوکار ساده و اجرای کم‌هزینه‌ی آن جهت نمایش ویدئوهای سه‌بعدی برای تماشاگران استفاده شده است. به همین دلیل، در ادامه ضمن تشریح مختصر این روش، الزامات نمایش باکیفیت ویدئوی سه‌بعدی

(سه پارامتر برای هر دوربین) به عنوان پارامترهای نسبی انتخاب می‌شوند. همبستگی بین زاویه‌ی چرخش دوربین‌ها حول محور X سیستم مختصات شیء نیز باعث ابهام و نقص در تعریف سیستم مختصات شیء شده که با ثابت لحاظ شدن این زاویه در یکی از دوربین‌ها مرتفع می‌گردد. بر این اساس پارامترهای ارتباط نسبی دوربین‌ها به ω زاویه‌ی دوران برای دو دوربین محدود می‌شود. به عبارت بهتر، تمامی زوایای دوران یکی از دوربین‌ها و همچنین دو زاویه‌ی دوران در دوربین دیگر، پارامترهای ارتباط نسبی تلقی شده که بطور هم‌زمان با سایر مجهولات کالیبراسیون داخلی برآورد می‌شوند. بطور تجربی انتخاب این پارامترهای نسبی، همبستگی کمتری را با پارامترهای کالیبراسیون داخلی داشته و احتمال بروز بدوضعی دستگاه معادلات را می‌کاهد.

ویدئوهای اخذ شده در سازوکار پیشنهادی به واسطه‌ی عدم امکان تنظیم دقیق دوربین‌ها به صورت نسبی، برای استفاده مستقیم در روش‌های برجسته‌بینی مناسب نیستند. نرمال‌سازی یک نگاشت هندسی با هدف تغییر در منظر ویدئوها است. به عبارت بهتر، در این روند تلاش می‌شود ویدئوهای جدیدی با محورهای موازی تصویربرداری و مقیاس‌های مشابه بازنمونه برداری گردند. با این اتفاق جهت‌گیری پارالاکس عوارض در سراسر تصاویر نرمال شده همسو شده و خطوط اپی‌پلار موازی می‌شوند. برای این هدف، ویدئوهای اخذ شده از دوربین‌های استریو به یک صفحه (صفحه‌ی نرمال) با موقعیت و جهت‌گیری مشخص به صورت محاسباتی تابانده شده و در قالب شبکه‌های منظم رستری بازنمونه برداری می‌شوند. شکل (۴) این فرایند را به صورت شماتیک نشان می‌دهد.



شکل (۴) - نمای شماتیک از بازنمونه برداری در دوربین‌های استریو

علیرغم اهمیت موازی بودن خطوط اپی‌پلار با راستای سطرهای تصاویر بازنمونه برداری شده در کاربردهایی همچون تناظریابی‌های مترکم^۱، تامین این شرایط در تولید ویدئوهای سه‌بعدی الزامی نبوده و صرفاً توازی خطوط اپی‌پلار شرایط

¹Dense Matching

پارامترهای داخلی دوربین به‌کاررفته در هر ویدئو در یک فرایند خودکالیبراسیون و از طریق اخذ تصاویر متعدد از یک میدان آزمون تارگت‌گذاری شده برآورد شده‌اند. برای این منظور از نرم‌افزار فتوگرامتری بردکوتا Australis ver 6.02 استفاده شده است. به‌عبارت بهتر، برای هر دوربین پارامترهای کالیبراسیون داخلی برآورد می‌شود.

پس از نصب دوربین‌ها در کنار هم و قبل از اخذ ویدئوی اصلی، از یک میدان آزمون شطرنجی در حالت سکون فیلم‌برداری شده است. فریم‌های اخذ شده از میدان آزمون امکان برآورد پارامترهای ارتباط نسبی را به‌صورت دستی فراهم می‌آورند. اختلاف زمانی ویدئوها نیز بعد از یکسان‌سازی fps آنها، به‌کمک تفسیر بصری فریم‌های متناظر با دقتی معادل با یک فریم به‌صورت دستی اندازه‌گیری شده‌اند. مقادیر اندازه‌گیری شده و یا برآورد شده به‌صورت دستی برای مقایسه‌ی با نتایج کسب شده در روش پیشنهادی استفاده شده‌اند.

مطابق با روند متوالی تشریح شده در بخش (۲)، اولین اقدام از روش پیشنهادی به هم‌زمان‌سازی ویدئوها اختصاص دارد. در روند هم‌زمان‌سازی، پس از برآورد اختلاف زمانی بین ویدئوها از طریق تناظریابی شاخص‌های زمانی، فریم‌های متناظر و به تبع آن ویدئوهای هم‌زمان شده تولید می‌شود. مقادیر عددی اختلاف زمانی ویدئوها معلوم بوده و در اولین ارزیابی، با مقادیر به‌دست‌آمده توسط روش پیشنهادی مقایسه شده‌اند. با توجه به انتخاب چند شاخص زمانی متفاوت در روند تناظریابی زمانی (جدول ۱)، در جدول (۲) آماره‌های دقت، شامل دامنه و RMSE بین روش دستی و خودکار گزارش شده است.

جدول ۲- گزارش از آماره‌های دقت شاخص‌های زمانی در روند

تناظریابی زمانی

شاخص زمانی	دامنه (ثانیه)	RMSE (ثانیه)
شار نوری	۰,۲۴۸۹	۰,۱۶۶۱
انترویی	۰,۰۹۶۷	۰,۰۵۳۷
مساحت تابع توزیع تجمعی هیستوگرام	۰,۶۱۱۴	۰,۱۹۶۳
انحراف معیار	۱,۶۳۳	۰,۶۰۳۳
نسبت انحراف معیار به میانگین	۱,۹۳۳	۰,۶۵۷۲
میانگین	۱,۵۸۵۳	۰,۶۳۶۷
ضریب چولگی	۳,۰۵۷۳	۱,۳۴۰۷

با بررسی نتایج ارائه شده در جدول (۲)، موفقیت شاخص آماری انترویی نسبت به سایر شاخص‌های آماری به‌کارگرفته شده مشهود است. حساسیت بالای این شاخص به تغییر در محتوای فریم‌های متوالی، یکتایی بیشتری را برای سری‌های زمانی فراهم آورده که این موضوع دقت تناظریابی زمانی را ارتقاء بخشیده است.

از منظر تئوری، مقایسه‌ی پارامترهای کالیبراسیون داخلی و نسبی دوربین‌ها در دو حالت برآورد خودکار و دستی می‌تواند معیاری برای اندازه‌گیری تطابق روند تولید ویدئوهای نرمال شده تلقی

در این روش تشریح شده است. لازم به ذکر است که محدودیتی در زمینه به‌کارگیری سایر راهکارهای ذکر شده جهت تماشای ویدئوهای وجود ندارد.

در روش نمایش آنالگیف از فیلترهای رنگی برای ارسال مجزای ویدئوهای استریو استفاده می‌شود. شیشه‌های عینک آنالگیف از دو رنگ کاملاً متضاد (زاویه‌ی ۹۰ درجه‌ای در مکعب رنگی RGB) ساخته شده که به‌طور معمول این دو رنگ آبی و قرمز است. هر یک از ویدئوهای استریو به‌منظور نمایش آنالگیف با رنگی متناظر با یکی از شیشه‌های عینک نمایش می‌یابد. بدین ترتیب در زمان هم‌پوشانی نمایش دو ویدئو، نمایش ویدئوی غیر مرتبط برای هر چشم از طریق فیلترهای رنگی عینک محدود می‌گردد [۴۱].

در صورت به‌کارگیری فیلترهای رنگی غیر متضاد، احتمال کاهش کیفیت نمایش ویدئوی سه‌بعدی از طریق تداخل بصری ویدئوها وجود دارد. خلوص و تضاد رنگی فیلترهای عینک را می‌توان از طریق روی‌هم‌قرار دادن شیشه‌های رنگی تشخیص داد. عدم عبور نور در این شرایط نشان از تضاد کافی در فیلترهای رنگی است. ویدئوهای نمایش یافته نیز می‌بایست از تطابق رنگی با فیلترهای عینک برخوردار باشند. اینکار از طریق شناسایی فام^۱ هر یک از فیلترهای رنگی و تخصیص آن به مولفه‌ی رنگ در ویدئوی مرتبط به اجرا می‌رسد. به‌عبارت بهتر، بعد از تبدیل هر یک از ویدئوها از فضای رنگی RGB به HSI، رنگ مربوط به فیلتر جایگزین مقادیر H شده و سپس ویدئوهای اصلاح شده به فضای RGB باز می‌گردند [۴۲].

۳ نتایج و بحث

نتایج و عملکرد روش پیشنهادی در چهار مرحله‌ی مجزا ارائه و مورد ارزیابی قرار گرفته است.

- ۱- صحت‌سنجی فرایند هم‌زمان‌سازی ویدئوها
- ۲- ارزیابی میزان انحراف در برآورد پارامترهای داخلی و نسبی دوربین‌های استریو
- ۳- بررسی میزان ثبات در راستای شکل‌گیری پارالاکس عوارض در ویدئوهای سه‌بعدی
- ۴- ارزیابی کیفی حاصل از تماشای ویدئوهای سه‌بعدی، جنبه‌های چهارگانه از روند ارزیابی روش پیشنهادی را شکل می‌دهند.

برای این منظور از ۱۰ ویدئو با محتوای متفاوت استفاده شده است. در اخذ هر ویدئو، ابتدا دوربین‌های استریو نسبت به یکدیگر ثابت شده و سپس فرایند فیلم‌برداری آغاز گردیده است. متغیرهای مختلفی در اخذ ویدئوها لحاظ شده‌اند. فیلم‌برداری از محیط‌های داخلی و فضای باز، همچنین تنوع در محتوای تصویری، طول ویدئو و تمایز در دفعات و تداوم سکون، متغیرهای لحاظ شده در ویدئوهای اخذ شده می‌باشند.

¹ Hue

جدول ۴- اختلاف برآورد پارامترهای ارتباط نسبی دستی و روش پیشنهادی در مجموعه داده‌های نمونه

دوربین	پارامتر	دامنه (درجه)	میانگین (درجه)	RMSE (درجه)
چپ	دوران حول محور دوم	۳,۵۹۴۳	۰,۱۱۴۶	۲,۴۸۷۹
	دوران حول محور سوم	۳,۶۹۱۷	۰,۰۷۴۵	۱,۷۱۹۷
	انحراف محور اصلی	۳,۰۷۸۳	۰,۰۰۸۹	۱,۶۷۳۹
راست	دوران حول محور اول	۲,۳۶۱۸	۰,۳۱۵۳	۱,۵۲۲
	دوران حول محور دوم	۳,۷۲۰۴	۰,۲۷۵۲	۲,۵۰۵۱
	دوران حول محور سوم	۴,۷۹۲۴	۰,۴۸۱۵	۲,۹۹۲۴
	انحراف محور اصلی	۲,۰۰۶۴	۰,۷۴۵۲	۱,۷۷۱۳

نتایج گزارش شده در جدول (۴) نیز حاکی از اختلافات زاویه‌ای اندک میان پارامترهای ارتباط نسبی در دو حالت دستی و روش پیشنهادی است. همان‌طور که در مورد نتایج جدول (۳) نیز اشاره شد؛ علیرغم تلاش به‌منظور تامین توزیع متوازن در عمق نقاط متناظر، برآورد هم‌زمان پارامترهای کالیبراسیون داخلی و نسبی برای دوربین‌های استریو به دلیل عدم تامین زاویه‌ی تقارب کافی بین آنها، توأم با بروز همبستگی بین پارامترهای مجهول خواهد بود. بر این اساس، بخشی از انحرافات گزارش شده در جداول (۳) و (۴) ناشی از وجود همبستگی میان پارامترهای مجهول در روند کالیبراسیون داخلی و نسبی است. البته این موضوع مادامیکه اثری در ویژگی‌های مورد انتظار در ویدئوهای سه‌بعدی نداشته باشد قابل اغماض خواهد بود. بررسی این موضوع در دستورکار مرحله سوم ارزیابی نتایج روش پیشنهادی است. با این حال، شرایط مشابهی در ویدئوهای آزمون از نظر:

۱- تعداد و دقت مکانی نقاط متناظر

۲- توزیع مناسب نقاط متناظر در فضای تصویر

۳- وجود تناظرهای نادرست (با احتمال اندک)

و ۴- وجود نقاط متناظر در عمق‌های مختلف، وجود نداشته که این عوامل کیفیت فرایند کالیبراسیون نسبی و داخلی را برای ویدئوهای مختلف، متمایز ساخته است. به‌همین دلیل، بخش دیگری از انحرافات ایجاد شده در گزارشات آماری جدول (۳) و (۴) را می‌توان متأثر از این شرایط قلمداد نمود. به نظر می‌رسد، با توجه به تنوع مناسب در محتوای ویدئوهای آزمون، می‌توان گزارش‌های آماری ارائه شده در جداول (۳) و (۴) را به‌عنوان محدوده‌ی مورد انتظار از دقت و صحت فرایند کالیبراسیون داخلی و نسبی در روش پیشنهادی در نظر گرفت.

همان‌طور که در بخش (۲-۴) اشاره شد، یک ویدئوی سه‌بعدی از دو ویدئو با مقیاس مشابه و محورهای اصلی موازی تشکیل شده است. این دو عامل منجر به توازی خطوط اپی‌پلار و به تبع آن همراستایی در ثبت پارالاکس برای تمامی پدیده‌های حاضر در صحنه می‌گردد. بر خلاف روند تولید تصاویر نرمال شده با هدف اجرای تناظریابی‌های مترکم، در تولید ویدئوهای سه‌بعدی، اقدامی به‌منظور هم‌راستا نمودن خطوط اپی‌پلار با

گردد. بر این اساس، اختلاف پارامترهای کالیبراسیون داخلی و نسبی به‌دست آمده به روش دستی با مقادیر به‌دست آمده از روش پیشنهادی به‌عنوان دومین رویکرد ارزیابی نتایج انتخاب شده است. جدول (۳) آماره‌های مربوط به اختلاف در برآورد پارامترهای کالیبراسیون داخلی (x_0, y_0, c) را برای ۲۰ ویدئوی آزمون (۱۰ ویدئوی سه‌بعدی) نشان می‌دهد. لازم به ذکر است که ابعاد متریک پیکسل‌های دوربین‌های بکار گرفته شده در این تحقیق معادل یک میکرون است.

جدول ۳- اختلاف برآورد پارامترهای کالیبراسیون داخلی و روش پیشنهادی در مجموعه داده‌های نمونه

پارامتر	دامنه (میلیمتر)	میانگین (میلیمتر)	RMSE (میلیمتر)
x_0	۰,۰۰۷۳	-۰,۰۰۰۰۰۵	۰,۰۰۱۳
y_0	۰,۰۰۶۱	۰,۰۰۴۹	۰,۰۲۱
c	۰,۰۲۷۸	۰,۰۰۳۳	۰,۰۱۱۳

اختلافات اندک گزارش شده در جدول (۳) حاکی از موفقیت روش پیشنهادی در کالیبراسیون هم‌زمان داخلی و نسبی ویدئوهای سه‌بعدی است. به‌نظر می‌رسد که تلاش صورت گرفته به‌منظور تامین نقاط متناظر با تنوع در عمق و توزیع مناسب در فضای تصویر و همچنین به‌کارگیری تکنیک‌های حل دستگاه معادلات غیرخطی در شرایط وجود همبستگی بین پارامترهای مجهول منجر به این اتفاق شده است. هرچند که بخشی از اختلافات گزارش شده در این جدول کماکان متأثر از وجود همبستگی بین پارامترهای کالیبراسیون داخلی و نسبی بوده که انتظار می‌رود این موضوع نتایج نهایی تولید ویدئوهای سه‌بعدی را مخدوش نسازد.

دو زاویه‌ی دورانی حول محورهای دوم و سوم برای یک دوربین و هر سه زاویه دورانی در دوربین دیگر پارامترهای ارتباط نسبی محسوب شده که در روند کالیبراسیون نسبی دوربین‌ها برآورد می‌شوند. در جدول (۴) گزارشی آماری از میزان انحراف در برآورد پارامترهای ارتباط نسبی برای ویدئوهای آزمون ارائه شده است. در این جدول علاوه بر انحراف زوایای دورانی برآورد شده برای دوربین‌ها نسبت به مقادیر کالیبره شده به‌صورت دستی، آماره‌های مربوط به انحراف زاویه‌ای محور اصلی دوربین‌ها نیز گزارش شده است. این زاویه بین محور سوم سیستم مختصات تصویری از منظر سیستم مختصات فضای شیء در هنگام تمایز بین پارامترهای خارجی شکل می‌گیرد. لازم به ذکر است که ستون سوم ماتریس دوران سه‌بعدی $R_3(-\kappa) \times R_2(-\varphi) \times R_1(-\omega)$ معادل با جهت‌گیری بردار یکه محور سوم سیستم مختصات تصویری در فضای شیء است.

جدول (۵) - گزارش آماری انحراف در راستای پارالاکس نقاط متناظر در مجموعه داده نمونه

شماره ویدئوی سه‌بعدی	دامنه (درجه)	انحراف معیار (درجه)
ویدئوی ۱	۱,۳۴۷۰	۰,۴۵۲
ویدئوی ۲	۱,۴۲۷	۰,۰۵۱۵
ویدئوی ۳	۱,۹۳۱	۰,۳۶۱
ویدئوی ۴	۱,۲۴۱	۰,۵۲۷
ویدئوی ۵	۱,۴۷۳	۰,۵۰۴۹
ویدئوی ۶	۱,۴۳۸	۰,۴۸۱۵
ویدئوی ۷	۱,۹۶۳	۰,۴۵۳
ویدئوی ۸	۱,۹۳۱	۰,۴۸۱۵
ویدئوی ۹	۲,۱۴۹۷	۰,۵۹۶۲
ویدئوی ۱۰	۱,۳۴۷۰	۰,۴۲۹

نتایج گزارش شده در جدول (۵) بدون نیاز به ویدئوهای سه‌بعدی مرجع تولید می‌شوند. به عبارت بهتر، این ارزیابی مستقل از مراجع خارجی تولید شده و به همین دلیل می‌تواند به‌ازای هر محصول تولیدی از روش پیشنهادی برآورد گردد. ناهمسو بودن راستای پارالاکس‌های در ویدئوهای سه‌بعدی می‌تواند متاثر از بروز خطا و اختلال در هر یک از گام‌های تولید ویدئوی سه‌بعدی باشد. بر این اساس، این ارزیابی محصول محور می‌تواند به‌عنوان یک ارزیابی نهایی بر تایید یا رد تمامی اقدامات اجرایی در روش پیشنهادی تلقی گردد. نتایج گزارش شده حاکی از آن است که انحرافات بوجود آمده در راستای پارالاکس عوارض متناظر در ویدئوهای سه‌بعدی اندک و قابل اغماض است. عدم قطعیت‌های مکانی در شناسایی نقاط متناظر و همچنین گسسته بودن فضای تصویر در بوجود آمدن انحرافات گزارش شده اثر داشته که این موضوع جزئی لاینفک این ارزیابی محسوب می‌شود.

در مرحله‌ی چهارم ارزیابی از چند مخاطب متخصص و غیرمتخصص درخواست شد تا تجربیات خود را از تماشای ویدئوهای سه‌بعدی به‌صورت آناگلیف گزارش کنند. هیچ یک از بینندگان گزارشی مبنی بر عدم توانایی تشخیص عمق را برای کل ویدئو گزارش نکردند. به عبارت بهتر، بطور عمومی درک عمق برای همه‌ی ویدئوها میسر بوده اما در لحظاتی و در برخی از ویدئوها تشخیص عمق مختل شده و دویینی اتفاق افتاده است. در واکاوی این رخداد بررسی‌های مختلفی از روند اجرای الگوریتم پیشنهادی و ویدئوهای اولیه صورت گرفت. عوامل مختلفی در بروز اختلال در روند تولید ویدئوهای سه‌بعدی وجود داشته که عمدتاً ناشی از محدودیت‌های فنی و سخت‌افزاری است. به‌عنوان مثال، دوربین‌های معمولی و غیرحرفه‌ای از سیستم‌های تصویربرداری مبتنی بر شاتر جاروب شونده استفاده می‌کنند. در این سیستم تصویربرداری بر خلاف

جهت‌گیری سطرهای تصویر رقومی در ویدئوهای نرمال شده انجام نمی‌شود. به عبارت بهتر، توازی خطوط اپی‌پلار برای اهداف برجسته‌بینی کافی بوده و همراستا ساختن خطوط اپی‌پلار با راستای سطرها در تصاویر رقومی عموماً در اجرای الگوریتم‌های تناظریابی متراکم اهمیت می‌یابد. به همین دلیل و با هدف جلوگیری از چرخش‌های اضافی در ویدئوهای سه‌بعدی، در روش پیشنهادی از همراستایی خطوط اپی‌پلار با راستای سطرهای تصویر رقومی صرف‌نظر شده است. با این مقدمه، در مرحله‌ی سوم از ارزیابی نتایج روش پیشنهادی، کیفیت ویدئوی‌های سه‌بعدی از منظر توازی خطوط اپی‌پلار بررسی شده است. برای انجام این کار، به کمک تناظریابی خودکار نقاط متناظری بین فریم‌های چپ و راست ویدئوی سه‌بعدی شناسایی می‌شود. انتظار می‌رود در صورت توازی راستای رخداد پارالاکس در ویدئوهای سه‌بعدی، همگی بردارهای اختلاف موقعیت مکانی نقاط متناظر در فریم‌های متناظر چپ و راست موازی بوده و از زاویه‌ی شیب ثابتی برخوردار باشند. شکل (۵) نمونه‌ای از بردارهای اختلاف مکانی برای نقاط متناظر در یک فریم از ویدئوی سه‌بعدی را به شکل آناگلیف نشان می‌دهد.



شکل (۵) نمایش نقاط متناظر در یک فریم نرمال شده در یک فریم از ویدئو سه‌بعدی

مطابق با شکل (۵)، اختلاف میان شیب این بردارها با مقدار میانگین در هر ویدئوی سه‌بعدی، کیفیت تولید ویدئوهای نرمال شده را نشان می‌دهد. نقاط متناظر در این مرحله برای هر ویدئوی سه‌بعدی، به ازای انتخاب چند فریم در مقاطع زمانی مختلف شناسایی می‌شوند. این موضوع با هدف ارزیابی ثبات در جهت‌گیری پارالاکس‌ها در بازه‌ی زمانی هر ویدئو انجام شده است. جدول (۵) گزارشی آماری از میزان انحراف راستای پارالاکس برای نقاط متناظر نسبت به مقدار میانگین را در تمامی ویدئوی‌های سه‌بعدی نشان می‌دهد.

۳- تخمین پارامترهای کالیبراسیون داخلی و نسبی دوربین‌های استریو
 ۴- تولید ویدئوهای نرمال شده و سه‌بعدی، تشکیل شده است.
 تعمیم‌پذیری این روش در مورد چندین ویدئو ارزیابی شده که نتایج حاکی از موفقیت روش پیشنهادی در تمامی ویدئوها بود. با این حال چالش‌هایی همچون بروز خطاهای سیستمی ناشی از نقصان‌های سخت‌افزاری دوربین‌ها و اثرات شاتر جاروب شونده هنگام بروز سرعت نسبی بالا بین دوربین و صحنه می‌تواند نتایج تولید شده در این روش را به صورت مقطعی مخدوش سازد. توسعه‌ی راهکارهای پردازشی به منظور جبران اثرات شاتر جاروب شونده و همچنین تنظیم خودکار مقیاس ویدئوی سه‌بعدی متناسب با عمق لحظه‌ای میدان از مواردی است که در این حوزه می‌تواند به عنوان افق تحقیقاتی پیش‌رو مطرح باشد.

مراجع

- [1] Kapp, K.M. and T. O'Driscoll. Learning in 3D: Adding a new dimension to enterprise learning and collaboration. John Wiley & Sons, 2009.
- [2] Sergi Pujades Rocamora. Camera Models and algorithms for 3D video content creation, Ph.D. Thesis-University of GRENOBLE, 2016.
- [3] Zhang, Licong, et al. "Real-time human motion tracking using multiple depth cameras." International Conference on Intelligent Robots and Systems. IEEE, 2012.
- [4] Dou, Pengfei, and Ioannis A. Kakadiaris. "Multi-view 3D face reconstruction with deep recurrent neural networks." Image and Vision Computing 80 2018.
- [5] Zone, R., Stereoscopic cinema and the origins of 3-D film, University Press of Kentucky, 2014.
- [6] Lambooi, Marc, Wijnand A. Ijsselstein, and Ingrid Heynderickx. "Visual discomfort of 3D TV: Assessment methods and modeling." Displays 32.4 2011.
- [7] Emoto, Michiko, Shinya Miyamoto, and Kazuyoshi Yamamoto. "Navigation apparatuses, methods, and programs for generation of a 3D movie." U.S. Patent No. 7, pp.974,781. 5 Jul. 2011.
- [8] Roberts, Steve. Character animation fundamentals: developing skills for 2D and 3D character animation. CRC Press, 2012.
- [9] Garrido, Pablo, et al. "Reconstructing detailed dynamic face geometry from monocular video." ACM Trans. Graph. 2013.
- [10] Smolic, Aljoscha, et al. "An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution." 2009 Picture Coding Symposium. IEEE, 2009.

سیستم‌های تصویربرداری با شاتر سراسری ۱، اخذ هر فریم از به صورت تدریجی و در خلال جاروب آرایه‌های حساس به نور اتفاق می‌افتد. علیرغم سرعت بالای جاروب، کماکان حرکت دوربین در بازه‌ی زمانی جاروب هر فریم می‌تواند اعوجاجات هندسی را در ثبت تصویر بوجود آورد. الگو و شدت این اثر وابسته به سرعت نسبی صحنه و دوربین و همچنین الگوی هندسی جاروب آرایه‌های حساس به نور است. اعوجاجات ناشی از این پدیده هندسه‌ی فریم‌های نرمال شده را برای این مقاطع زمانی مخدوش ساخته که منجر به بروز پیچیدگی در الگوی رخدادهای پارالاکس عوارض می‌گردد. توسعه‌ی راهکارهای پردازشی به منظور حذف اثر شاتر جاروب شونده در فریم‌های توأم با حرکت و یا به‌کارگیری دوربین‌های با شاتر سراسری می‌تواند به‌عنوان روش‌های رفع این چالش شناخته شوند. در شرایط فعلی توصیه می‌شود که در زمان اخذ ویدئوهای استریو، سرعت حرکت نسبی دوربین و صحنه تا حد ممکن پایین باشد. این موضوع می‌تواند بروز اعوجاجات هندسی ناشی از شاتر جاروب شونده تا حد زیادی کاهش دهد.

از سوی دیگر، محدودیت‌های سخت‌افزاری گاهی منجر به بروز انقطاع در روند متوالی اخذ فریم‌های یک ویدئو می‌گردد. در این شرایط بیننده شاهد وقوع سکت در نمایش ویدئو می‌گردد. این اتفاق یک خطای سیستمی محسوب شده که می‌تواند انطباق محتوای فریم‌های متناظر را برای لحظاتی دچار اختلال نماید. بدیهی است که چنین خطاهایی ناشی از عوامل مختلف سخت‌افزاری و نرم‌افزاری در ساختار دوربین‌های غیر حرفه‌ای بوده که به صورت مقطعی ویدئوهای سه‌بعدی را تخریب نماید. با این حال نتایج تولید ویدئوهای سه‌بعدی در این تحقیق حاکی از آن است که علیرغم بروز چنین خطاهایی در روند اخذ ویدئوهای اولیه، این مشکلات نتوانسته شاکله‌ی اصلی ویدئوهای سه‌بعدی را دچار اختلال نماید. به عبارت بهتر، بروز خطا در مقاطع زمانی محدود منجر به مشکلات تشخیص عمق شده اما این موضوع اثراتی در زمان صحت عملکرد دوربین‌های استریو نداشته و در سایر مقاطع زمانی برجسته‌بینی و تشخیص عمق بدون اختلال ممکن است.

۴ نتیجه‌گیری

در این مقاله راهکاری خودکار به منظور تولید ویدئوهای سه‌بعدی از طریق ویدئوهای اخذ شده توسط دوربین‌های آماتور ارائه شد. روش پیشنهادی از چهار گام متوالی:

- ۱- تولید ویدئوهای هم‌زمان‌سازی شده
- ۲- شناسایی نقاط متناظر مناسب به منظور کالیبراسیون داخلی و نسبی

¹Global Shutter

- [26] Elhayek, Ahmed, et al. "Feature-based multi-video synchronization with subframe accuracy." Joint DAGM (German Association for Pattern Recognition) and OAGM Symposium. Springer, Berlin, Heidelberg, 2012.
- [27] Zone, R. Stereoscopic cinema and the origins of 3-D film, 1838-1952: University Press of Kentucky, 2014.
- [28] Ng CJ, Blake R, Banks MS, Tadin D, Yoon G. Seeing the World Like Never Before: Human stereovision through perfect optics. bioRxiv, 2021.
- [29] Holliman, Nick. "Cosmic origins: experiences making a stereoscopic 3D movie." Stereoscopic Displays and Applications XXI. Vol. 7524. International Society for Optics and Photonics, 2010.
- [30] Nahler, Gerhard. "Pearson correlation coefficient." Dictionary of Pharmaceutical Medicine. Springer, Vienna, 2009.
- [31] Baker, Simon, et al. "Removing rolling shutter wobble." IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010.
- [32] Horn, B.K.P. & Schunk, Determining optical Flow, AI Memo 572, AI Lab, MIT, B.G.1980.
- [33] David G. Lowe, Distinctive Image Features from Scale-Invariant Key points, publication in the International Journal of Computer Vision, 2004.
- [34] Wang, X., Zhang, H. and Liu, SH. Reliable RANSAC Using a Novel Preprocessing Model. Xiaoyan Wang et al. An open access article distributed under the Creative Commons Attribution License, 2013.
- [35] Michel Morgan, Epipolar Resampling of Linear Array Scanner Scenes, Ph.D. Dissertation, UNIVERSITY OF CALGARY, Canada, 2004.
- [36] Brown, D.C. *The Bundle Adjustment - progress and prospects*, IAPRS, Vol. XIII, No. 3, 1979.
- [37] Wolf PR. Elements of photogrammetry; with air photointerpretation and remote sensing. 1983.
- [38] Mikhail, Edward M., James S. Bethel, and J. Chris McGlone. "Introduction to modern photogrammetry." New York 19, 2001.
- [39] Banks, Martin S., et al. "Stereoscopy and the human visual system." SMPTE motion imaging journal 121.4, pp. 24-43, 2012.
- [40] Yang, Lin, et al. "See in 3D: state of the art of 3D display technologies." Multimedia Tools and Applications 75.24, pp. 17121-17155, 2016.
- [41] Kim, Jonghyun, et al. "A full-color anaglyph three-dimensional display system using active color filter glasses." Journal of Information Display 12.1, pp. 37-41, 2011.
- [42] Govindasamy, Saravanan, and Yamuna Govindarajan. "Real-time implementation of various colour space models." *International Journal of Circuits and Architecture Design* 2.3-4, pp. 258-271, 2016.
- [11] Passmore, Charles. "Director-style based 2d to 3d movie conversion system and method." U.S. Patent 8,730,232, issued May 20, 2014.
- [12] Lakis Christodoulou, 3D Stereo Vision Camera-sensors, Advancements, and Technologies, Electrical & Computer Engineering & Computer Science, Cyprus University of Technology, 2013.
- [13] Zhang, Yingchun, Jianbo Huang, and Siwen Duan. "3D video conversion system based on depth information extraction." MATEC Web of Conferences. Vol. 232. EDP Sciences, 2018.
- [14] Vijayanarasimhan, Sudheendra, et al. "Sfm-net: Learning of structure and motion from video." arXiv preprint arXiv, 2017.
- [15] Heinzle, Simon, et al. "Computational stereo camera system with programmable control loop." ACM Transactions on Graphics (TOG) 30.4, 2011.
- [16] Lang, Manuel, et al. "Nonlinear disparity mapping for stereoscopic 3D." ACM Transactions on Graphics (TOG), 2010.
- [17] Tam, Wa James, et al. "Stereoscopic 3D-TV: visual comfort." IEEE transactions on broadcasting 57.2, 2011.
- [18] Akbar, Muhammad Ali, et al. "Making short movie using 3D professional camera." 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC). IEEE, 2018.
- [19] Merkle, Philipp, Karsten Müller, and Thomas Wiegand. "3D video: acquisition, coding, and display." IEEE Transactions on Consumer Electronics 56.2, pp. 946-950, 2010.
- [20] J. P. Ramirez-Paredes, D. J. Lary, and N. R. Gans, "Low-altitude Terrestrial Spectroscopy from a Pushbroom Sensor," Journal of Field Robotics, vol. 33, no. 6, pp. 837-852, 2016.
- [21] Negahdaripour, Shahriar. "Analyzing epipolar geometry of 2-D forward-scan sonar stereo for matching and 3-D reconstruction." OCEANS MTS/IEEE Charleston. IEEE, 2018.
- [22] Meyer, Benjamin, et al. "Subframe Temporal Alignment of Non-Stationary Cameras." BMVC. 2008.
- [23] Wang, Xue, et al. "Motion-based temporal alignment of independently moving cameras." IEEE Transactions on Circuits and Systems for Video Technology 27.11 pp. 2344-2354, 2016.
- [24] Lu, Cheng, and Mrinal Mandal. "A robust technique for motion-based video sequences temporal alignment." IEEE Transactions on Multimedia 15.1, pp. 70-82, 2012.
- [25] Evangelidis, Georgios D., and Christian Bauckhage. "Efficient subframe video alignment using short descriptors." IEEE transactions on pattern analysis and machine intelligence 35.10, pp. 2371-2386, 2013.



عطیه گنجعلی دانشجوی مقطع کارشناسی ارشد رشته مهندسی نقشه‌برداری گرایش فتوگرامتری در دانشگاه تفرش است. ایشان کارشناس آزمایشگاه تخصصی فتوگرامتری دانشگاه تفرش بوده و علاقمند به موضوعات تحقیقاتی مرتبط با فتوگرامتری رقومی و ویدئوگرامتری است.



علیرضا صفدری نژاد فارغ‌التحصیل مقطع دکترای مهندسی نقشه‌برداری در گرایش سنجش از دور از دانشگاه صنعتی خواجه نصیرالدین طوسی است. ایشان در حال حاضر استادیار گروه ژئودزی و مهندسی نقشه‌برداری دانشگاه تفرش بوده و علاقمندی‌های پژوهشی وی فتوگرامتری، ویدئوگرامتری، تناظریابی خودکار، تشخیص الگو در فضاها با ابعاد بالا و سری‌های زمانی است.