

سورة التين



وزارت علوم، تحقیقات و فناوری

دانشگاه تفرش

دانشکده مهندسی نقشه برداری

پایان نامه کارشناسی ارشد

روشی جهت تنظیم و کالیبراسیون یک سیستم تصویربرداری استریو متشکل از دوربین های

آماتور به منظور تولید ویدیوهای سه بعدی

نگارش:

عطیه گنجعلی

استاد راهنما:

دکتر علیرضا صفدری نژاد

تاریخ: ۱۴۰۲/۰۴/۲۹
شماره: ۴۹۳۵۹



صور تجلسه دفاعیه پایان نامه کارشناسی ارشد

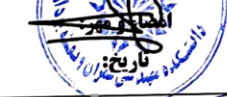
نام و نام خانوادگی: عطیه گنجلی
شماره دانشجویی: ۹۷۳۱۹۱۰۰۶
رشته تحصیلی/گرایش: مهندسی نقشه برداری-نقشه برداری-فتوگرامتری
عنوان پروژه: روشی جهت تنظیم و کالیبراسیون یک سیستم تصویربرداری استریو متشکل از دوربین های آماتور بمنظور تولید ویدئوهای سه بعدی
تعداد واحد: ۶
تاریخ تصویب: 1399/02/08
تاریخ دفاع: ۱۴۰۰/۰۴/۲۹

درجه:

عالی (۱۹-۲۰) خیلی خوب (۱۸-۱۸/۹۹) خوب (۱۷-۱۷/۹۹) متوسط (۱۴-۱۵/۹۹) مردود (کمتر از ۱۴)

نام و نام خانوادگی	سمت	رتبه	محل اشتغال	محل امضاء
دکتر علیرضا صفدری نژاد	استاد راهنمای اول	استادیار	دانشگاه تفرش	
-	استاد راهنمای دوم	-	-	-
-	استاد مشاور	-	-	-
دکتر حدیثه سادات حسینی	داور داخلی	استادیار	دانشگاه تفرش	
دکتر محمدجواد عبداللهی فرد	داور خارجی	دانشیار	دانشگاه تفرش	
دکتر حدیثه سادات حسینی	نماینده تحصیلات تکمیلی	استادیار	دانشگاه تفرش	

مدیر گروه: دکتر نژاد به بیان



معاون آموزشی و تحصیلات تکمیلی دانشگاه: اسماعیل نظری گیاشی

امضاء و مهر:

تاریخ:





وزارت علوم، تحقیقات و فناوری

دانشگاه نفرش

بنام خدا

تعهدنامه اصالت اثر

تاریخ:

۱۴۰۰/۰۶/۰۳

اینجانب عطیه گنجعلی به شماره دانشجویی ۹۷۳۱۹۱۰۰۶ دانش آموخته مقطع کارشناسی ارشد در رشته نقشه برداری گرایش فتوگرامتری که در تاریخ ۱۴۰۰/۴/۲۹ از پایان نامه خود تحت عنوان روشی جهت تنظیم و کالیبراسیون یک سیستم تصویر برداری استریو دفاع کرده ام متعهد می شوم: متشکل از دوربین های آماتور به منظور تولید ویدیوهای سه بعدی

که مطالب مندرج در این پایان نامه حاصل کار پژوهشی اینجانب تحت نظارت و راهنمایی اساتید دانشگاه تفرش بوده و به دستاوردهای دیگران که در این پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان نامه قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارائه نگردیده است.

در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان نامه متعلق به دانشگاه تفرش بوده و هرگونه استفاده از نتایج علمی و عملی واگذاری اطلاعات به دیگران یا چاپ و تکثیر، نسخه برداری، ترجمه و اقتباس از این پایان نامه با کسب اجازه از استاد (استادان) راهنما باشد.

نقل مطالب با ذکر مآخذ بلامانع است.

عطیه گنجعلی

نام و نام خانوادگی:

امضا و تاریخ:

۱۴۰۰/۰۶/۰۳

با احترام به پدر و مادر عزیزم

تقدیم به همسر و دختر مهربانم

که در سایه همیاری و همدلی آنها به این منظور نائل شدم.

پروردگار را سپاسگزارم که بار دیگر به من فرصت آموختن داد.

اکنون که با یاری خداوند این دوره را به اتمام رسانیده ام، بر خود واجب می دانم از استاد گرانقدر و فرزانه جناب آقای دکتر علیرضا صفدری نژاد که همواره راهنما و پیشتیبان اینجانب بوده اند و هرگز از هیچ کمک و همکاری دریغ ننموده اند و با مساعدت خویش مسیر انجام تحقیق را هموار ساختند، کمال تشکر و قدردانی را به جای آورم.

چکیده

در این تحقیق راهکاری خودکار به منظور تولید ویدیوهای سه بعدی 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 نسبی و همچنین محدودیت‌های مربوط به تنظیم سخت‌افزاری ارتباط نسبی دوربین‌ها، چالش‌های این راهکار قلمداد می‌شود. در راهکار پیشنهادی، ابتدا هم‌زمانی ویدیوها از طریق تناظریابی شاخص‌های زمانی تامین شده و در ادامه مقاطع زمانی توأم با سکون در طول ویدیوها شناسایی می‌شوند. سپس، مجموعه‌ای از نقاط متناظر در دو ویدیو 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 استفاده می‌شوند. در آخر، ویدیوهای سه بعدی نرمال شده از طریق بازنمونه‌برداری مبتنی بر هندسه‌ی اپی‌پلار تولید می‌گردند. این روش در مورد چندین ویدیوی سه بعدی از چهار جنبه‌ی مختلف کمی و کیفی مورد ارزیابی قرار گرفته است. دقت هندسی مطلوب در تولید ویدیوهای نرمال شده، هم‌زمانی دقیق ویدیوهای سه بعدی، تعمیم‌پذیری مطلوب روش پیشنهادی در تولید ویدیوهای سه بعدی در شرایط محیطی مختلف و رضایت تماشاگران ویدیوهای سه بعدی از منظر درک بصری عمق، از ویژگی‌های نتایج این روش محسوب می‌شوند.

واژگان کلیدی: فیلم سه بعدی، انتروپی، شار نوری، هم‌زمان سازی خودکار، تصاویر نرمال، توجیه

نسبی

فهرست مطالب

فصل اول (مقدمه)	۹
۱-۱- مقدمه	۱۰
۲-۱- طرح مسأله، ضرورت و جایگاه تحقیق	۱۰
۳-۱- اهداف تحقیق	۱۲
۴-۱- ساختار پایان نامه	۱۳
فصل دوم (مبانی نظری و پیشینه تحقیق)	۱۴
۱-۲- مقدمه	۱۵
۲-۲- فیلم‌های سه‌بعدی (تئوری و روش‌ها)	۱۹
۳-۲- ابزارهای مشاهده‌ی ویدیوهای سه‌بعدی	۱۹
۲-۳-۱- فیلم آناگلیف	۱۹
۲-۳-۲- فیلم پنجره‌های مجزا	۲۰
۲-۳-۳- فیلم پولاریزه	۲۰
۴-۲- بازنمونه برداری اپی پولار	۲۱
۵-۲- تناظریابی عارضه مبنای Sift	۲۲
۶-۲- ارزیابی کمی	۲۷
۸-۲- پیشینه‌ی تحقیقات مرتبط	۲۸
۹-۲- جمع بندی	۳۱
فصل سوم (روش تحقیق و پیاده سازی)	۳۲
۱-۳- مقدمه	۳۳
۲-۳- روش پیشنهادی	۳۴
۱-۲-۳- گام اول (آماده سازی دوربین‌ها و اخذ داده)	۳۵
۲-۲-۳- گام دوم (همزمان سازی ویدیوها)	۳۷
۳-۲-۳- گام سوم (انتخاب جفت فریمهای متناظر به منظور برآورد ارتباط نسبی ویدیوها)	۴۳

۴۹ (کالیبراسیون داخلی، نسبی و بازنمونه برداری اپی پلار)
۵۵ (تولید ویدیوهای سه بعدی)
۵۸ فصل چهارم (بررسی نتایج حاصل از پیاده سازی)
۵۹ ۱-۴- مقدمه
۵۹ ۲-۴- داده های مورد استفاده
۵۹ ۳-۴- ارزیابی نتایج و بحث
۵۹ ۱-۳-۴- رویکرد نخست (صحت سنجی فرایند هم زمان سازی ویدیوها)
 ۲-۳-۴- رویکرد دوم (انحراف در برآورد پارامترهای داخلی و نسبی دوربین های استریو)
۶۱
۶۴ ۳-۳-۴- رویکرد سوم (ثبات در راستای شکل گیری پارالاکس عوارض در ویدیوهای سه بعدی)
۶۷ ۴-۳-۴- رویکرد چهارم (ارزیابی کیفی حاصل از تماشای ویدیوهای سه بعدی)
۶۸ فصل پنجم (نتیجه گیری و پیشنهادات)
۷۲ منابع

فهرست اشکال

- شکل ۱-۲- شمای کلی از روند تولید فیلم سه بعدی در استراتژی تک دوربین ۱۷
- شکل ۲-۲- نمونه ای از دوربین استریو..... ۱۸
- شکل ۳-۲- نمونه ای از یک فریم فیلم آنالگلیف و عینک آنالگلیف ۲۰
- شکل ۴-۲- نمونه ای از یک فریم فیلم پنجره های مجزا و عینک هوک ۲۰
- شکل ۵-۲- نمونه ای از یک فریم فیلم پلاریزه و عینک پلاریزه ۲۱
- شکل ۶-۲- نمای شماتیک از بازنمونه برداری در دوربین های استریو ۲۶
- شکل ۱-۳- فلوچارتی از روند کلی اجرای الگوریتم پیشنهادی ۳۵
- شکل ۲-۳- نمایی از دستگاه ساخته شده به منظور تامین ثبات نسبی دوربین ها..... ۳۶
- شکل ۳-۳- نمایی شماتیک از تمایز بین fps و اختلاف زمانی آغاز ویدیوها ۳۸
- شکل ۴-۳- نمایی شماتیک از روند تغییرات زمانی معیار C برای یک ویدیو ۴۶
- شکل ۵-۳- ارتباط هندسی فضای دوبعدی عکس و فضای سه بعدی شئی..... ۵۰
- شکل ۶-۳- شمایی از محور X سیستم مختصات شئی موازی با خط واصل مراکز تصویر دوربین های استریو
- ۵۲
- شکل ۱-۴- نمایش نقاط متناظر در یک فریم نرمال شده در یک فریم از ویدیو سه بعدی ۶۵

فهرست جداول

جدول ۳-۱ - شاخص‌های آماری پیشنهادی به‌منظور تولید سری زمانی در روند تناظریابی زمانی خودکار	۴۱
جدول ۴-۱ - گزارش از آماره‌های دقت شاخص‌های زمانی در روند تناظریابی زمانی	۶۱
جدول ۴-۲ - اختلاف برآورد پارامترهای کالیبراسیون داخلی دستی و روش پیشنهادی در مجموعه داده‌های نمونه	۶۳
جدول ۴-۳ - اختلاف برآورد پارامترهای ارتباط نسبی دستی و روش پیشنهادی در مجموعه داده‌های نمونه	۶۴
جدول ۴-۴ - گزارش آماری انحراف در راستای پارالاکس نقاط متناظر در مجموعه داده نمونه	۶۷

فصل اول

مقدمه

امروزه ظهور فن‌آوری‌های نوین در حوزه‌ی تولید فیلم‌های سه‌بعدی سطح انتظار مخاطبان به این صنعت را بیش از پیش افزایش داده است. در دنیای امروز، فیلم‌های سینمایی، بازی‌های رایانه‌ای، مستندنگاری آثار تاریخی و ساخت تورهای مجازی همگی به سمت ایجاد شرایط دید سه‌بعدی سوق داده شده‌اند [۱].

شاید شناخته‌ترین کاربرد فیلم‌های سه‌بعدی، فیلم‌ها و انیمیشن‌های سه‌بعدی باشد. اما فیلم‌های سه‌بعدی تنها جنبه سرگرمی ندارند. به عنوان مثال با ایجاد یک فیلم سه‌بعدی از یک محیط شهری، زمینه‌ی تکمیل و بروزرسانی نقشه‌های موجود ایجاد می‌گردد. تهیه اطلاعات هندسی قابل اعتماد از سطح یک پدیده از جمله اهداف تولید فیلم سه‌بعدی بشمار می‌رود. بنابراین تولید فیلم‌های سه‌بعدی در طیف وسیعی از کاربردها نظیر حفظ و نگهداری میراث فرهنگی، زمین شناسی، طرح‌های عمرانی، پزشکی، سیستم‌های حمل و نقل هوشمند، خودروهای خودران و ... مورد استفاده قرار می‌گیرند [۱].

یک ویدیوی سه‌بعدی از تلفیق دو فیلم اخذ شده توسط دوربین‌های توأم با اختلاف منظر ساخته می‌شود. سازوکار تماشای این ویدیوها مبتنی بر بکارگیری تکنیک‌های برجسته‌بینی بوده که به شکل‌های مختلفی هر چشم را به سمت مشاهده‌ی یکی از این فیلم‌ها هدایت می‌کنند [۲].

۱-۲- طرح مسأله، ضرورت و جایگاه تحقیق

تولید فیلم سه‌بعدی را میتوان به دو دسته روش‌های تک دوربین [۳] و روش‌های دوربین‌های استریو [۴] تقسیم‌بندی می‌شوند. در روش‌های تک دوربین، برای تولید فیلم سه‌بعدی، جفت تصویر متناظر برای هر فریم بصورت مجازی و با استفاده از مدل سه‌بعدی تولید می‌شود. بدیهی است که تمامی عدم قطعیت‌های مرتبط با تولید مدل سه‌بعدی و به تبع آن تصاویر عمق در کیفیت تولید تصاویر

^۱Monocular Camera

^۲Stereo Camera

مجازی اثرگذار خواهد بود. بدین ترتیب، نتایج روش‌های تک دوربین بطور مستقیم به کیفیت بازسازی سه‌بعدی سطح به تبع آن نقشه‌ی عمق وابسته است. نقشه عمق یک تصویر خاکستری دو بعدی بوده که در آن شدت کم، فاصله دور از دوربین و شدت زیاد فاصله نزدیک به دوربین را نشان می‌دهد و می‌تواند درکی از بعد سوم برای بیننده ایجاد نماید [۵].

اما در سیستم‌های تصویربرداری استریو، نیازی به در دسترس بودن مدل سه‌بعدی نبوده و براحتی می‌توان با معلوم بودن ارتباط نسبی میان دوربین‌های استریو شرایط لازم برای تولید فیلم‌های سه‌بعدی را فراهم آورد. هزینه‌ی بالای تولید دوربین‌های استریو یکی از محدودیت‌هایی است که منجر به توسعه‌ی هرچه بیشتر روش‌های تک دوربین شده است. (در بخش‌های بعد روش‌های تک دوربین و دوربین‌های استریو بطور جامع تشریح و معرفی می‌گردد). این تحقیق تلاش دارد با در نظر گرفتن مزایای هر دو روش یک سیستم تصویربرداری استریوسکوپیک و آماتور طراحی نماید، تا در نهایت با استفاده از سازوکاری خودکار، بعد از ثابت قرار گرفتن دو دوربین آماتور، شرایط لازم به منظور تولید فیلم سه‌بعدی را فراهم آورد.

ساختار یک سیستم تصویربرداری استریوسکوپیک و آماتور برای تولید فیلم سه‌بعدی نیازمند فراهم آوردن شرایطی است که دو دوربین آماتور بدون تغییر موقعیت و وضعیت نسبی در طول زمان فیلم‌برداری ثابت باشند. در سامانه‌های حرفه‌ای تصویربرداری استریو، موقعیت دو دوربین با کمک ابزارهای دقیق مکانیکی تنظیم می‌شود تا کمترین اعوجاج بعد از نمونه‌برداری اپی‌پلار به فیلم‌های تولیدی تحمیل گردد. همچنین در ساختار تصویربرداری این سیستم‌ها، نرخ تصویربرداری بصورت یکسان و همزمان در نظر گرفته می‌شود. اما بکارگیری راهکارهای آماتور به منظور تولید فیلم‌های سه‌بعدی با چالش‌هایی همچون:

^۱Depth Map

^۲Epipolar Resampling

۱- عدم امکان تنظیم دقیق زمان آغاز تصویربرداری در دو دوربین آماتور ۲- عدم امکان تنظیم دقیق موقعیت و وضعیت دو دوربین به منظور تولید فیلم با محورهای تصویربرداری کاملاً موازی و ۳- یکسان نبودن و عدم پایداری پارامترهای کالیبراسیون داخلی دوربین‌های بکارگرفته شده روبرو است. به نظر می‌رسد در صورت استقرار دوربین‌ها در یک قاب ثابت و آگاهی از کالیبراسیون داخلی و نسبی، کماکان چالش تولید فیلم سه‌بعدی محدود به یافتن فریم‌های متناظر و به بیان بهتر اختلاف زمانی آغاز فیلم‌برداری خواهد بود. بدیهی است که این چالش در هر بار شروع فیلم‌برداری تکرار می‌گردد. با آنکه امکان کالیبراسیون داخلی و نسبی دو دوربین به کمک تکنیک‌های مرسوم فتوگرامتری وجود دارد؛ اما در این تحقیق راهکاری خودکار به منظور برآورد پارامترهای مذکور مطرح می‌شود تا روند تولید ویدیو سه‌بعدی برای کاربران ساده‌تر و در ضمن بتوانند بدون هیچگونه اطلاعاتی از پارامترهای داخلی و نسبی دو دوربین فیلم سه‌بعدی با کیفیتی تولید نمایند.

۱-۳- اهداف تحقیق

در این پژوهش ضمن بررسی چالش‌های موجود در بکارگیری دوربین‌های آماتور، بدنبال پاسخ این سوال است، که آیا صرفاً با استفاده از دوربین‌های آماتور و روش‌های فتوگرامتری می‌توان راهکارهایی خودکار به منظور کالیبراسیون سیستم‌های استریوی آماتور با هدف تولید فیلم سه‌بعدی ارائه داد؟ بدیهی است که در صورت وجود پارامترهای دورانی نامشابه بین دوربین‌ها، امکان برجسته‌بینی سراسری فیلم‌های اخذ شده صرفاً با اعمال یک بردار انتقال میان تصاویر وجود نخواهد داشت. در چنین شرایطی تولید فیلم‌های نرمال شده به‌عنوان بخشی از فرایند تولید فیلم‌های سه‌بعدی مطرح بوده که نیازمند دانش به پارامترهای ارتباط نسبی دو دوربین است. از سوی دیگر، در صورت عدم اطلاع از اختلاف زمانی دقیق میان فریم‌های متناظر در دو دوربین، حتی در زمان آگاهی از ارتباط نسبی دو دوربین نمی‌توان انتظار داشت که پارامترهای نسبی فریم‌های متناظر انتخاب شده بصورت نادرست، در طول زمان فیلم‌برداری ثابت بماند. این خطا در زمان حرکت نسبی صحنه یا دوربین تشدید

شده و روند تولید فیلم‌های سه‌بعدی را پیچیده می‌سازد. سومین چالش موجود در روند بکارگیری دوربین‌های آماتور، عدم مشابهت و حتی شناخت نسبت به پارامترهای کالیبراسیون داخلی دوربین‌ها خواهد بود.

سازوکار روش توسعه‌یافته در این تحقیق 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 ویدیوها اصلاح می‌گردد. در گام دوم، مجموعه‌ای از فریم‌های مناسب در ویدیوهای همزمان‌سازی شده به‌منظور برآورد پارامترهای کالیبراسیون داخلی و ارتباط نسبی دوربین‌ها شناسایی می‌گردد. بعد از برآورد پارامترهای ارتباط نسبی دوربین‌ها، در گام سوم نگاهی به‌منظور نرمال‌سازی ویدیوها از طریق بازنمونه‌برداری مبتنی بر هندسه‌ی اپی‌پلار انجام می‌شود. آخرین گام نیز به آماده‌سازی ویدیوها برای اهداف برجسته‌بینی اختصاص دارد. در این روش با بکارگیری اقدامات پردازشی سخت‌گیرانه تلاش شده تا دستیابی به ویدیوهای سه‌بعدی بصورت خودکار و با قابلیت اعتماد مطلوب به اجرا رسد.

۱-۴- ساختار پایان نامه

پایان نامه حاضر در پنج فصل تدوین شده است. در فصل اول، پس از طرح مقدمه، ضرورت و اهداف تحقیق بیان شده است. در فصل دوم مبانی نظری تولید تصاویر نرمال و شاخص‌های آماری و مطالعات صورت گرفته در این زمینه ارائه شده است. در فصل سوم به تشریح روش پیشنهادی پرداخته شده است. در فصل چهارم نتایج کسب شده از روش پیشنهادی ارائه شده و مورد بحث قرار گرفته و آخرین فصل از پایان نامه نیز به نتیجه‌گیری و ارائه پیشنهادات برای تحقیقات آتی اختصاص دارد.

فصل دوم

مبانی نظری و پیشینه تحقیق

۲-۱- مقدمه

فیلم سه‌بعدی تولید فیلم یا تصویر متحرک براساس دید استریوسکوپي است که به ما امکان میدهد با استفاده از ابزارهای خاص برجسته‌بینی بتوانیم واقعیت‌های نمایش داده شده در آن را بصورت سه‌بعدی مشاهده نماییم [۴]. سازوکار درک تغییرات عمق عوارض در فیلم‌های سه‌بعدی منطبق بر سازوکار بینایی انسان است. اخذ تصویر از یک پدیده از دو منظر متفاوت شرایط لازم جهت تشخیص عمق را برای انسان فراهم می‌آورد [۳].

در این پژوهش ابتدا پیش‌شرط‌های هندسی لازم به‌منظور تامین شرایط برجسته‌بینی تشریح شده، سپس ابزارها و روش‌های مشاهده ویدیوهای سه‌بعدی معرفی و ارائه می‌گردد. سپس صفحات نرمال معرفی و مراحل تولید آن تشریح شده، در انتها روش تناظریابی مورد استفاده در این پژوهش مورد بررسی قرار می‌گیرد.

۲-۲- فیلم‌های سه‌بعدی (تئوری و روش‌ها)

از نقطه نظر هندسی و محاسباتی، جابجایی مکانی ثبت عوارض در تصاویر پرسپکتیو اخذ شده از منظرهای متفاوت، محتوای لازم به‌منظور تشخیص عمق را فراهم می‌آورد [۶]. 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 ویدیوهای سه‌بعدی با سازوکار مشابه سامانه‌ی بینایی انسان طراحی می‌شوند [۸]. عدم وجود اختلاف مقیاس محسوس، توازی محورهای نوری دوربین‌های عکسبرداری، اختلاف منظر مناسب و همزمانی در نمایش فیلم‌های اخذ شده از دوربین‌های استریو، شروط لازم به‌منظور درک بصری مطلوب از یک ویدیوی سه‌بعدی می‌باشند [۴]. تاکنون راهکارهای مختلفی به‌منظور تولید ویدیوهای سه‌بعدی توسعه یافته است. در یک دسته‌بندی

¹stereo camera

کلی می‌توان این راهکارها را به دو حوزه‌ی روش‌های نیازمند به بازسازی سه‌بعدی صحنه و روش‌های مستقل از آن دسته‌بندی نمود [۳]. در دسته‌بندی نخست، صحنه‌ی فیلم‌برداری به کمک ابزارها و تکنیک‌های متفاوتی به‌صورت سه‌بعدی بازسازی شده و در ادامه، با طراحی و حرکت دوربین‌های استریو به‌صورت 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 به‌منظور اخذ فیلم‌های مناسب برای اهداف برجسته‌بینی، هندسه‌ی نسبی دوربین‌های مجازی را جهت تولید ویدیوهای پرسپکتیو استریو تنظیم می‌نماید [۹].

تعمیم این راهکار به صحنه‌های واقعی با چالش‌هایی همچون کیفیت و کامل بودن بازسازی سه‌بعدی و وجود اشیاء متحرک مواجه است. به همین دلیل در مورد صحنه‌های واقعی حداقل یک دوربین به‌صورت فیزیکی جهت اخذ ویدیو مبنا مورد استفاده قرار داشته و فیلم مکمل با هدف برجسته‌بینی، به‌صورت مجازی و در خلال پردازش ویدیوی مبنا و به کمک مدل سه‌بعدی سطح تولید می‌شود [۱۰].

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 به‌منظور تامین درک بصری عمق، بازنمونه‌برداری می‌گردد [۱۱].

دسترسی به مدل سه‌بعدی سطح و همچنین معلوم بودن پارامترهای لحظه‌ای موقعیت و وضعیت دوربین در ویدیوی مبنا، پیش‌شرط اجرای این رویکرد قلمداد شده که بطور معمول از طریق اجرای تکنیک‌های مثلث‌بندی دسته‌اشعه و تناظریابی‌های متراکم به‌صورت خودکار تامین می‌گردد [۱۲] و [۱۳].

پیش‌شرط‌های ذکر شده می‌تواند به‌صورت منفرد توسط ویدیوی مبنا و یا با مشارکت همزمان سنجنده‌های فعال و غیرفعال دیگر تامین گردد [۱۴].^۳ به‌کاربردن سنجنده‌های فعال همچون

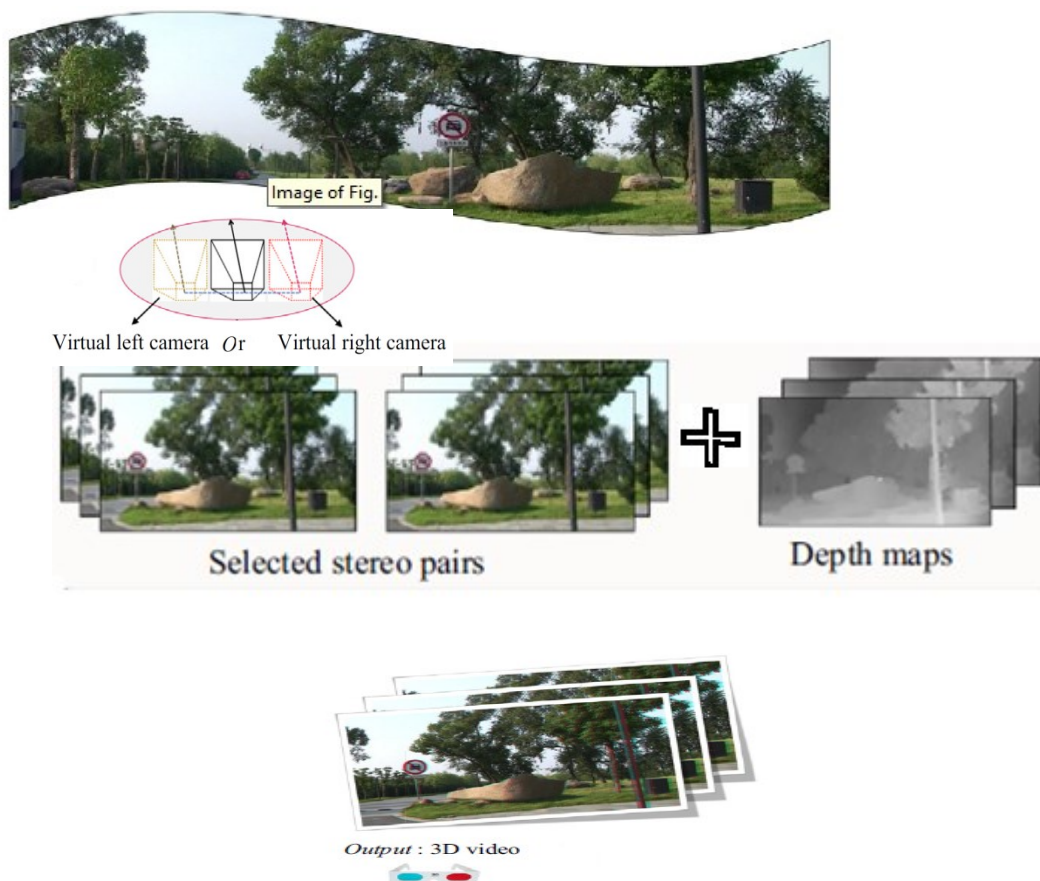
^۱3D animation

2 Active

3 Passive

پوشگرهای لیزری و همچنین استفاده از تکنیک‌های SfM در سنجنده‌های غیرفعال در این حوزه استفاده شده است [۱۵]. در شکل (۱-۲) شمای کلی از روند تولید فیلم سه بعدی در استراتژی تک دوربین را نشان می‌دهد.

در رویکرد دوم، از دوربین‌های استریو به منظور اخذ ویدیوهای همزمان توأم با اختلاف منظر استفاده می‌شود [۱۶]. در این روند نیازی به مدل رقومی سطح نبوده و صرفاً تأمین هم‌زمانی، پارامترهای داخلی و ارتباط نسبی مناسب بین دوربین‌های استریو شرایط کافی به منظور تولید ویدیوهای سه‌بعدی را فراهم می‌آورد [۱۷].



شکل (۱-۲) - شمای کلی از روند تولید فیلم سه بعدی در استراتژی تک دوربین

در دوربین‌های استریو تجاری سازی شده، پیش‌نیازهایی همچون تامین همزمانی، نرخ زمانی مشابه در فیلم‌برداری و ارتباط نسبی دوربین‌ها بصورت سخت‌افزاری تنظیم شده به نحوی که فیلم‌های اخذ شده با کمترین سطح نیاز به پیش‌پردازش، قابلیت بکارگیری در سامانه‌های برجسته‌بینی را خواهند داشت [۱۸].



شکل (۲-۲) - نمونه‌ای از دوربین استریو

اما در صورت استفاده از ویدیوهای اخذ شده توسط دو دوربین آماتور با هدف تولید ویدیوهای 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 ویدیوهای مناسب به منظور درک بصری عمق به‌عنوان چالش مطرح خواهد بود [۱۹]. شکل (۲-۲) نمونه‌ای از دوربین استریو را نشان می‌دهد. در این تحقیق تلاش شده تا با مبنا قرار دادن ویدیوهای اخذ شده از دو دوربین گوشی همراه، روندی خودکار به‌منظور آماده‌سازی ویدیوهای اخذ شده توسط هر دوربین برای اهداف برجسته‌بینی توسعه یابد.

۲-۳- ابزارهای مشاهده ویدیوهای سه‌بعدی

فیلم‌های سه‌بعدی را به روش‌های متفاوتی مشاهده کرد. هدف برجسته بینی هدایت فیلم‌های اخذ شده از دوزبین‌های چپ و راست به سمت چشم‌ها می‌باشد. در اصل ابزارهای برجسته بینی با سازوکار متفاوتی این هدف را دنبال می‌کنند. تاکنون به طرق مختلفی روش‌های برجسته‌بینی به اجرا رسیده است. که در این ذیل برخی از این راهکارها به اختصار ارائه می‌گردد.

۲-۳-۱ فیلم آناگلیف

۱

یکی از ساده‌ترین و پرکاربردترین این روش‌ها نمایش فیلم به صورت آناگلیف بوده در این روش پس از تولید تصاویر اصلاح شده با اتصال تصاویر چپ و راست از طریق نقاط متناظر و نمایش تصاویر نرمال در دو باند آبی و قرمز، دید سه‌بعدی ایجاد می‌گردد. فیلم‌های آناگلیف با استفاده از عینک آناگلیف قابل مشاهده هستند. عینک آناگلیف براساس قانون عبور انتخابی (بیر-لامبرت) ^۲ در فیزیک اپتیک که می‌گوید: یک شیء شفاف نورهای هم‌رنگ خود را عبور داده و نورهایی از رنگ دیگر را جذب می‌کند، ساخته شده است. عینک‌های آناگلیف شامل شیشه‌هایی از دو رنگ مکمل هستند. عینکی با این ویژگی توانایی جداسازی دو تصویر مستقل که با یکدیگر ترکیب شده را دارا است. اگر با این عینک به تصویر نمادین شکل (۲-۳) نگاه کنیم، چشم از پشت شیشه‌ی سبز رنگ سطوح سرخ و سیاه را به رنگ سیاه و سطوح سبز و سفید را به رنگ سبز می‌بیند [۲۰].

^۱Anaglyph

^۲ Beer-Lambert law

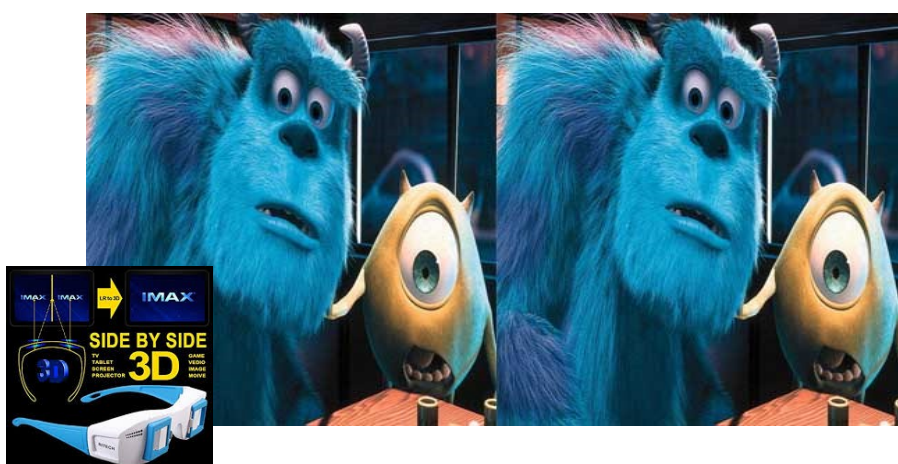


شکل (۲-۳) - نمونه ای از یک فریم فیلم آناگلیف و عینک آناگلیف

۲-۳-۲ فیلم پنجره های مجزا

۱

در این روش صفحه‌ی نمایش به دو بخش مجزا تقسیم و فیلم نرمال چپ و راست در هر یک از بخش‌ها به نمایش در می‌آید. این نوع از فیلم سه‌بعدی با استفاده از عینک هوک قابل مشاهده است [۲۱].



شکل (۲-۴) - نمونه ای از یک فریم فیلم پنجره های مجزا و عینک هوک

۲-۳-۳ فیلم پولاریزه (قطبیده)

۲

از دیگر روش‌های مشاهده فیلم سه‌بعدی میتوان به فیلم‌های پولاریزه اشاره کرد. در فیلم‌های پولاریزه از خاصیت قطبیدگی نور استفاده شده است. پرتوی نور عمود بر جهت تابش خود در تمام جهات حرکت

^۱Split Screen

^۲Polarizer

لرزشی دارد. پلاریزاسیون بدین معنی است که تمام جهات این حرکت غیر از یک جهت خاص حذف شود. این کار توسط فیلترهای پلاریزه کننده صورت می گیرد. در سینمای سه بعدی (که پرده نمایش آنها برای بازتابش نور پلاریزه باید براق بوده و معمولاً از جنس آلومینیوم است) دو تصویر پلاریزه مجزا با دو محور پلاریزه عمود بر هم نمایش داده می شود. عینک پلاریزه نیز از دو محور پلاریزه عمود بر هم تشکیل شده است. نتیجه اینکه چشم راست شما فقط تصویر مربوط به خودش را می بیند و فیلتر این چشم به تصویر اختصاصی چشم چپ اجازه عبور نمی دهد و برعکس. بنابراین هر چشم تصویری مجزا می بیند که درک سه بعدی ماحصل کار است [۲۲].



شکل (۲-۵)- نمونه ای از یک فریم فیلم پلاریزه و عینک پلاریزه

۲-۴- بازنمونه برداری اپی پلار

نتیجه نهایی تولید فیلم سه بعدی بر روی صفحه تلویزیون یا پرده سینما به نمایش در خواهد آمد، که صفحه ای کاملاً تخت بوده و برخلاف دستگاه های تبدیل مکانیکی اعمال انتقال و دوران در برخی از راستای محورهای مختصات در هر شرایطی در آن میسر نیست. به منظور ایجاد دید سه بعدی در این شرایط از مفهوم تصاویر نرمال استفاده می شود. در تصاویر نرمال هدف تولید تصاویر تصحیح شده ای است که بدون اعمال هیچ دورانی و تنها با یک انتقال ساده امکان دید سه بعدی در سراسر تصویر ایجاد

نماید. در این شرایط هر دو تصویر بر روی صفحه جدیدی بازنمونه برداری شده به طوری که این صفحه موازی با خط واصل مراکز تصاویر بوده و پارالاکس قائم نقاط متناظر حذف خواهد شد. با مشخص بودن موقعیت و وضعیت دو دوربین نسبت به یکدیگر در لحظه تصویر برداری، پارامترهای هندسی لازم جهت بازنمونه برداری دو تصویر اخذ شده بر روی صفحه‌ی نرمال قابل برآورد خواهد بود [۲۳]. بنابراین پس از مشخص شدن اختلاف زمانی آغاز فیلم برداری و برآورد پارامترهای نسبی بین دو دوربین، مقادیر لازم جهت تبدیل فریم‌های متناظر به فریم‌های نرمال بدست می‌آید. با در نظر گرفتن این نکته که فیلم مجموعه‌ای از تصاویر متوالی بوده، که با سرعت و پشت سرهم نمایش داده می‌شوند؛ با کنار هم قرار دادن فریم‌های نرمال، فیلم سه بعدی تولید خواهد شد. در ادامه مراحل تولید تصاویر نرمال از دو تصویر استریو، تشریح می‌گردد.

۲-۵- تناظریابی عارضه مبنای SIFT

راهکار پیشنهادی خودکار بوده و برآورد پارامترهای کالیبراسیون داخلی و نسبی بین دوربین‌ها نیازمند نقاط متناظر است. تاکنون الگوریتم‌های متنوعی به منظور استخراج عوارض و تناظریابی توسعه یافته، از جمله آن به الگوریتم‌هایی همچون SURF [۲۴]، Hariss [۲۵] و SIFT [۲۶] میتوان اشاره کرد. در این پروژه باهدف یافتن عوارض کاندید تناظریابی، از الگوریتم SIFT استفاده شده است. بطور کلی روشهای تناظریابی شامل سه مرحله است:

(۱) استخراج عوارض کاندید تناظریابی، (۲) تولید بردار ویژگی برای هر یک از عوارض کاندید تناظریابی، (۳) انتخاب عوارض متناظر از طریق مقایسه بردارهای ویژگی، اولین مرحله از الگوریتم SIFT، شناسایی موقعیتهایی از تصویر است که مستقل از تغییر مقیاس هستند. برای این منظور، ویژگی‌های پایدار تصویر در مقیاس مختلف آن با استفاده از فضای مقیاس استخراج می‌شوند. فضای^۱ مقیاس، نمایش ساختارهای تصویر در مقیاس‌های مختلف بوده و متشکل از مجموع‌های از تصاویر

^۱Scale Space

گوسی و تفاضل گوسی در ابعاد مختلف، که در لایه‌های مختلفی تحت عنوان اکتاو مرتب شده‌اند. برای تصویر $I(x,y)$ ، تصاویر گوسی فضای مقیاس با استفاده از تابع کرنل گوسی از رابطه (۱) محاسبه می‌گردد.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

رابطه‌ی (۱)

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y, \sigma)$$

که در این رابطه پارامتر σ ، بیانگر مقیاس هر تصویر بوده و مقدار اولیه آن $1/6$ است. با استفاده از یک پارامتر ثابت در یک روند افزایشی، در سطوح مختلف هر اکتاو افزایش یافته.

بعد از ایجاد هر اکتاو، تصویر گوسی (مقیاس آن دو برابر مقیاس اولیه است) انتخاب، که ابعاد آن نصف شده و به‌عنوان تصویر ابتدایی اکتاو بعدی در نظر گرفته می‌شود. هدف از ایجاد فضای مقیاس، استخراج ویژگی‌هایی است که مستقل از مقیاس هستند. بنابراین جهت استخراج موقعیت‌های پایدار میانی تصاویر DOG، در هر اکتاو هر پیکسل با هشت پیکسل همسایه خود و نه پیکسل در همسایگی تصویر DOG بالایی و پایینی مقایسه و در صورتی که اکستریم (بیشینه یا کمینه) باشد به‌عنوان یک ویژگی کاندید ذخیره می‌شود. هدف از ایجاد اکتاو و تولید DOG یافتن عوارض کاندید تناظریابی فارغ از مقیاس است، اگر تصاویر مورد نظر اختلاف مقیاس فاحشی نداشته باشند یک اکتاو هم کفایت می‌کند. در پروسه تولید فیلم سه بعدی اندازه دو تصویر یکسان در نظر گرفته شده و دو دوربین به‌نحوی در کنار یکدیگر قرار گرفته‌اند که محور تصویربرداری آنها با یکدیگر موازی هستند، بنابراین اختلاف مقیاس زیادی میان فریم‌ها وجود نداشته و با در نظر گرفتن تعداد کم اکتاو و لایه‌های DOG،

^۱Difference of Gaussian(DOG)

^۲Octave

عوارض بیشتری در تصویر جهت کاندیدای تناظریابی یافت می‌شود؛ چرا که با زیاد شدن لایه های اکتاو بخشی از عوارض در لایه های DOG حذف می‌شود. در این پروژه تعداد اکتاو ۵ و تعداد لایه های DOG، ۴ در نظر گرفته شده است. که این مقدار با توجه به منطقه فیلمبرداری در پروژه های متفاوت قابل تنظیم است.

بعد از استخراج نقاط کلیدی، الگوریتم SIFT توصیفگرهایی به منظور تطبیق میان تصاویر ایجاد می‌گردد. این توصیفگرها نسبت به تغییر مقیاس و چرخش و سایر تغییرات نظیر تغییرات روشنایی و تصویربرداری تغییرات ناشی از زاویه دید تا حد ممکن مقاوم هستند. برای ایجاد توصیفگر در الگوریتم استاندارد SIFT، در ابتدا ناحیه ای به صورت یک شبکه 4×4 در اطراف هر ویژگی در تصویر گوسی مربوط به آن، در نظر گرفته می‌شود. سپس برای سلول در شبکه، یک هیستوگرام از جهت گرادیان‌ها تخصیص داده می‌شود. در نهایت توصیفگر SIFT به صورت یک بردار با ۱۲۸ مولفه ایجاد می‌شود.

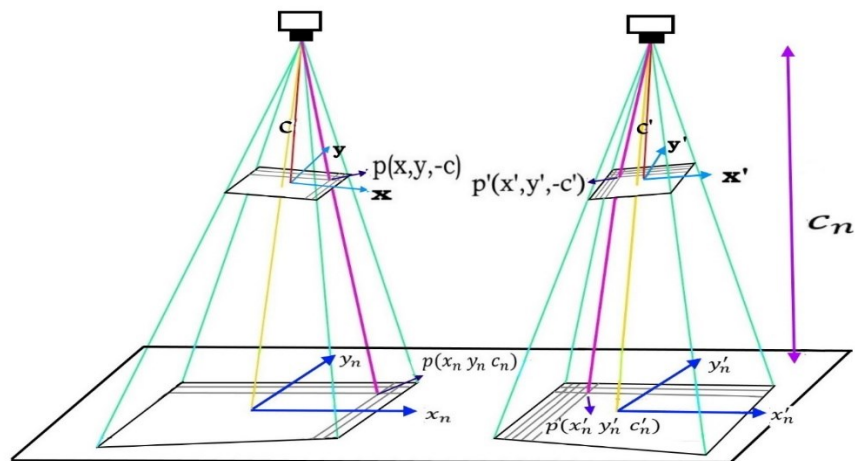
پس از ایجاد یک توصیفگر برای هر ویژگی، میبایست این توصیفگرها در دو تصویر با هم تطبیق داده شوند. معمولاً از فاصله اقلیدسی برای این هدف استفاده می‌شود. اگر فاصله بین توصیفگرهای دو نقطه کلیدی کمتر از مقدار حد آستانه مورد نظر باشد، دو نقطه کلیدی تطبیق می‌یابند [۲۶]. مهمترین نقص این روش این است که یک توصیفگر ممکن است دارای چندین تطبیق باشد. به منظور رفع این مشکل دومین فاصله اقلیدسی بدست آمده نیز در نظر گرفته و اگر دومین فاصله اختلاف زیادی با اولین فاصله داشت این دو عارضه متمایز و در غیر این صورت دو عارضه متشابه هستند. تناظریابی که در آن عوارض متمایز از یکدیگر هستند قابل اعتمادتر از تناظریابی است که عوارض متشابه باشند.

عواملی مانند تکرار عوارض مشابه و بازتاب‌ها موجب اشتباه در تناظریابی می‌شوند، از جمله روش‌های رایج در تشخیص اشتباهات استفاده از الگوریتم RANSAC در ترکیب با توابع مختلف است [۲۷]. RANSAC الگوریتمی مبتنی بر تکرار است که طی آن تعدادی از داده‌ها به طور تصادفی بعنوان نمونه

انتخاب می‌شود. با توجه به نمونه‌ی انتخابی و حدآستانه‌ی t_e برای خطاها، داده‌ها به دو دسته و درست‌ها (inliers) و اشتباهات (outliers) تقسیم‌بندی می‌شوند. تکرار تا جایی ادامه پیدا می‌کند که احتمال (P) پیدا کردن یک نمونه با خطای کمتر از t_e از بین داده‌های درست، برابر یک حدآستانه باشد. حداکثر تعداد تکرارهای مورد نیاز از رابطه (۲) بدست می‌آید.

$$N = \frac{\log(1-P)}{\log(1-f^s)} \quad \text{رابطه‌ی (۲)}$$

ویدیوهای اخذ شده در سازوکار پیشنهادی بواسطه‌ی عدم امکان تنظیم دقیق دوربین‌ها بصورت نسبی، برای استفاده مستقیم در روش‌های برجسته‌بینی مناسب نیستند. نرمال‌سازی یک نگاشت هندسی با هدف تغییر در منظر ویدیوها است. به عبارت بهتر، در این روند تلاش می‌شود ویدیوهای جدیدی با محورهای موازی تصویربرداری و مقیاس‌های مشابه بازنمونه‌برداری گردند. با این اتفاق جهت‌گیری پارالاکس عوارض در سراسر تصاویر نرمال شده همسو شده و خطوط اپی‌پلار موازی می‌شوند. برای این هدف، ویدیوهای اخذ شده از دوربین‌های استریو به یک صفحه (صفحه‌ی نرمال) با موقعیت و جهت‌گیری مشخص بصورت محاسباتی تابانده شده و در قالب شبکه‌های منظم رستری بازنمونه‌برداری می‌شوند [۲۳]. شکل (۲-۶) این فرایند را بصورت شماتیک نشان می‌دهد.



شکل (۲-۶)- نمای شماتیک از بازنمونه برداری در دوربین های استریو

علیرغم اهمیت موازی بودن خطوط اپی پلار با راستای سطرهای تصاویر بازنمونه برداری شده در کاربردهایی همچون تناظریابی های متراکم ، تامین این شرایط در تولید ویدیوهای سه بعدی الزامی نبوده و صرفاً توازی خطوط اپی پلار شرایط کافی را به منظور برجسته بینی ویدیوهای نرمال شده فراهم می آورد [۲۸]. بر این اساس، با مبنا قرار دادن ω ، φ و κ به عنوان زوایای دورانی یکی از دوربین های استریو (بدست آمده از فرایند برآورد ارتباط نسبی)، نگاشت بین مختصات تصویری فریم های اولیه و نرمال شده بصورت رابطه ی (۳) خواهد بود.

$$\begin{bmatrix} x_n \\ y_n \\ -c_n \end{bmatrix} = \lambda R_3(\kappa_n) R_1(\omega_n) R_3(-\kappa) R_2(-\varphi) R_1(-\omega) \begin{bmatrix} x - x_0 \\ y - y_0 \\ -c \end{bmatrix} \quad \text{رابطه ی (۳)}$$

$$R_\omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix}$$

$$R_\varphi = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix}$$

$$R_k = \begin{bmatrix} \cos k & \sin k & 0 \\ -\sin k & \cos k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\omega_n = \frac{\omega_l + \omega_r}{2}$$

$$k_n = \frac{k_l + k_r}{2}$$

$$\varphi_n = -\tan^{-1} \frac{Z_l - Z_r}{X_l - X_r}$$

در رابطه‌ی (۳)، $[x, y]$ مختصات متریک پیکسل‌ها در فضای ویدیو اولیه، $[x_n, y_n]$ مختصات متریک متناظر در فضای تصویر نرمال شده، $[x_o, y_o, c]$ پارامترهای کالیبراسیون داخلی ویدیوی اولیه، λ ضریب مقیاس، R_1, R_2 و R_3 به ترتیب ماتریس‌های دوران حول محورهای اول، دوم و سوم مختصاتی و c_n فاصله‌ی کانونی ویدیوی نرمال شده می‌باشند. مقادیر ω_n, k_n و c_n به عنوان پارامترهای دورانی و فاصله کانونی ویدیوی نرمال شده، معادل میانگین مقادیر ω, k و c بدست آمده برای دوربین‌های استریو در فرایند برآورد پارامترهای کالیبراسیون نسبی و داخلی انتخاب می‌شوند. از آنجاکه توجیه نسبی دوطرفه در نظر گرفته شده است، مقدار φ_n مساوی صفر است.

رابطه‌ی (۳) بعد از حذف ضریب مقیاس بصورت یک تبدیل پروجکتیو دوبعدی در دو حالت مستقیم و معکوس برای بازنمونه برداری فریم‌های هر یک از ویدیو‌ها مورد استفاده قرار می‌گیرد.

۲-۶ - ارزیابی کمی

شاخص‌های کمی متعددی به منظور ارزیابی کیفیت و دقت توسعه یافته که هر یک به نحوی میزان مشابهت فیلم مرجع و فیلم تولیدی با استفاده از روش پیشنهادی را می‌سنجند. در این تحقیق از شاخص‌های آماری همچون انحراف معیار و RMSE جهت ارزیابی استفاده شده است.

$RMSE^1$ در اندازه‌گیری اختلاف میان کمیت‌های محاسباتی و واقعی شناخته می‌شود. در تحقیق حاضر این شاخص میزان اختلافات مقادیر پارامترهای مختلف فیلم سه‌بعدی تولیدی با راهکار پیشنهادی نسبت به فیلم سه‌بعدی مرجع را محاسبه می‌نماید [۲۹].

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{رابطه‌ی (۴)}$$

۷-۲ - پیشینه تحقیقات مرتبط

در این بخش مروری بر پژوهش‌های صورت گرفته در تهیه فیلم‌های سه‌بعدی با استفاده از روش‌های مختلف و همچنین پژوهش‌هایی در حوزه همگام‌سازی ویدیوهای توأم با همپوشانی (زمانی و مکانی) ارائه شده است.

دوران تصاویر متحرک استریوسکوپی از سال ۱۸۹۰ میلادی توسط ویلیام فرایس گرین [۳۰] آغاز شد در اختراع او دو فیلم در کنار هم بر روی پرده سینما پخش میشد و بوسیله استریوسکوپ قابل دیدن بود. این روش بدلیل مکانیک مزاحم مورد قبول قرار نگرفت. تاکنون از هر دو رویکرد سخت‌افزاری و پردازشی به‌منظور تنظیم نسبی ویدیوهای استریو با هدف تولید تصاویر و ویدیوهای سه‌بعدی استفاده شده است.

هی‌بو و همکارانش با مبنا قرار دادن فریم‌های متوالی اخذ شده در یک فیلم، فیلم سه‌بعدی تهیه کردند. روند کار در این تحقیق به این صورت است که در ابتدا فریم $i+1$ به‌عنوان فریم مجازی فریم i در نظر گرفته می‌شود. سپس با تناظریابی بین فریم‌های متوالی با استفاده از الگوریتم SIFT و الگوریتم RANSAC، پارامترهای ماتریس ضروری برآورد می‌گردد، سپس با استخراج پارامترهای

¹Root Mean Square Error

^۲Essential Matrix

توجه نسبی از ماتریس ضروری و تولید صفحات نرمال فیلم سه بعدی تهیه می گردد. سادگی، هزینه کم اجرایی و عدم نیاز به نقشه عمق در روش تک دوربین از مزایای این تحقیق بشمار می رود [۳۱].

از آنجاکه در تولید فیلم سه بعدی با استفاده از دوربین های استریو با کیفیت پایین زمان آغاز تصویربرداری در دوربین ها تا حدودی با یکدیگر متفاوت است و این اختلاف زمانی در عوارض متحرک مشهود تر می باشند؛ الکساندر پلوکشن و همکاران با استفاده از بردارهای شار نوری و اختلاف مختصات و با فرض ثابت بودن اعوجاج هندسی در تمام صحنه ها، مشکل همگام سازی نماهای استریوسکوپي برطرف نمودند. در این تحقیق جهت کاهش تاثیر نویز بردارهای شار نوری تصاویر بلوک بندی می شود. بدیهی است این امر موجب از بین رفتن برخی اطلاعات و کاهش دقت می گردد. شارنوری نیز بواسطه ی:

۱- دقت پایین مقادیر تخمین زده شده در لبه ها، ۲- اثرپذیری از تغییرات روشنایی و ۳- وجود مناطق بدون بافت به عنوان یک راهکار قدرتمند برای رسیدن به دقت مناسب در سراسر فیلم در نظر گرفته نمی شود؛ که این خود یکی دیگر از محدودیت های روش ذکر شده قلمداد می گردد [۳۲].

سیون دو آن و همکارانش سعی بر این داشته اند با استفاده از استخراج اطلاعات عمق، فیلم دوبعدی را به فیلم سه بعدی تبدیل کنند. این روش شامل چهار مرحله است:

۱- تبدیل فیلم به تصاویر متوالی، ۲- اختصاص عمق به تصاویر ایجاد شده، ۳- تولید تصاویر مجازی به منظور ایجاد دید سه بعدی و در نهایت ۴- با استفاده از فیلم اصلی، نقشه عمق و فیلم مجازی، فیلم سه بعدی تولید می شود.

در این روش جهت پیدا کردن نقشه عمق هر تصویر از شبکه عصبی کانولوشنی استفاده شده است.^۲ داده آموزشی در شبکه عصبی توسط دوربین کینکت جمع آوری شده و به شبکه عصبی معرفی

^۱Disparity map

^۲Convolutional Neural Network (CNN)

^۳Microsoft Kinect cameras

می‌گردد. در گام بعد با توجه به پارامترهای نسبی دوربین برای هر یک از فریم‌ها یک فریم مجازی به‌منظور ایجاد دید سه‌بعدی تولید می‌شود. استفاده از شبکه‌های عصبی کانولوشن را می‌توان از ابتکارات این روش قلمداد کرد، چرا که شبکه‌های عصبی کانولوشن نسبت به سایر شبکه‌های عصبی انعطاف بیشتری داشته و عموماً در زمینه پردازش تصویر و یا پردازش ویدیو استفاده می‌شوند. تعداد لایه‌های پنهان و داده آموزشی انتخابی در شبکه‌های عصبی در دقت نتیجه نهایی اثرگذار بوده که این امر موجب کم شدن سرعت پردازش در این روش می‌شود [۳۳].

جانگ سوک لو و همکارانش در ابتدا تأثیر نهم زمانی بر کیفیت فیلم استریوسکوپي را تحلیل کرده، و سپس راهکاری به‌منظور همزمان‌سازی ویدیوهای اخذ شده از دو دوربین مبتنی بر تناظریابی بین مشتق تصاویر مطرح می‌گردد. در این روند ابتدا به ازای هر فریم هر دو ویدیو مشتق گرفته می‌شود و با مقایسه میانگین مشتق تصاویر، فریم‌هایی که بیشینه‌ی معیار همبستگی را دارند، به‌عنوان فریم متناظر انتخاب می‌شوند. در این روش شیفت زمانی با حداکثر خطای یک فریم یافت می‌شود [۳۴].

تاکنون از هر دو رویکرد سخت‌افزاری و پردازشی به‌منظور تنظیم نسبی ویدیوهای استریو با هدف تولید تصاویر و ویدیوهای سه‌بعدی استفاده شده است [۳۵]. 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 [۳۶]. روش‌های پردازشی نیز با اجرای فرایند بازنمونه‌برداری اپی‌پلار، توازی محورهای نوری و یکسان‌سازی مقیاس ویدیوهای استریو را فراهم می‌آورند [۳۷].

در حوزه‌ی هم‌زمان‌سازی ویدیوهای توأم با هم‌پوشانی (زمانی و مکانی) نیز راهکارهای مختلفی توسعه یافته است [۳۸]. تکنیک‌های هندسی شاخه‌ای از راهکارهای هم‌زمان‌سازی ویدیوها بوده که در آنها از طریق به‌کارگیری تارگت‌های متحرک با هندسه‌ی حرکتی معلوم و یا هنگام معلوم بودن ارتباط نسبی دوربین‌های فیلم‌برداری استفاده می‌شود. این روش‌ها از طریق تناظریابی میان فریم‌های هم‌زمان

محتمل، هندسه‌ی حرکتی تارگت متحرک و یا موضع نسبی دوربین‌ها برآورد می‌شود. مقایسه‌ی نتایج برآورد شده یا مقادیر معلوم معیار تشخیص هم‌زمانی در این راهکارها تلقی می‌شود [۳۹]. 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 [۴۰].

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 به‌منظور یافتن بیشینه‌ی همبستگی میان ویدئوها استفاده می‌شود [۴۱] در این رویکرد نیز عدم وجود حرکت نسبی میان صحنه و دوربین عامل ایجاد ابهام در هم‌زمانی محسوب می‌شود [۴۲].

۲-۸- جمع بندی

این فصل به مرور انواع شاخص‌های آماری و مروری بر روش تناظریابی و روش حذف تناظرهای اشتباه اختصاص یافت. در این تحقیق از میان شاخص‌های آماری ۷ شاخص به‌منظور تولید سری‌های زمانی استفاده شده است. از این بین شش شاخص از آماره‌های مستخرج از هیستوگرام تصویر بدست آمده و یک شاخص مربوط به بردارهای جریان تصویری است.

فصل سوم

روش تحقیق و پیاده سازی

مغز و سیستم عصبی انسان به طور تکوینی آموزش یافته تا بتواند محیط سه بعدی را به کمک تلفیق محتوای تصویری چشم‌ها درک نماید. پرتوهای رسیده از عوارض سه بعدی به سامانه‌ی بینایی انسان توأم 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 نقطه نظر هندسی، هر چشم به عنوان 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 شبکیه 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 سه بعدی از عمق را رقم می‌زنند [۴۳]. 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.

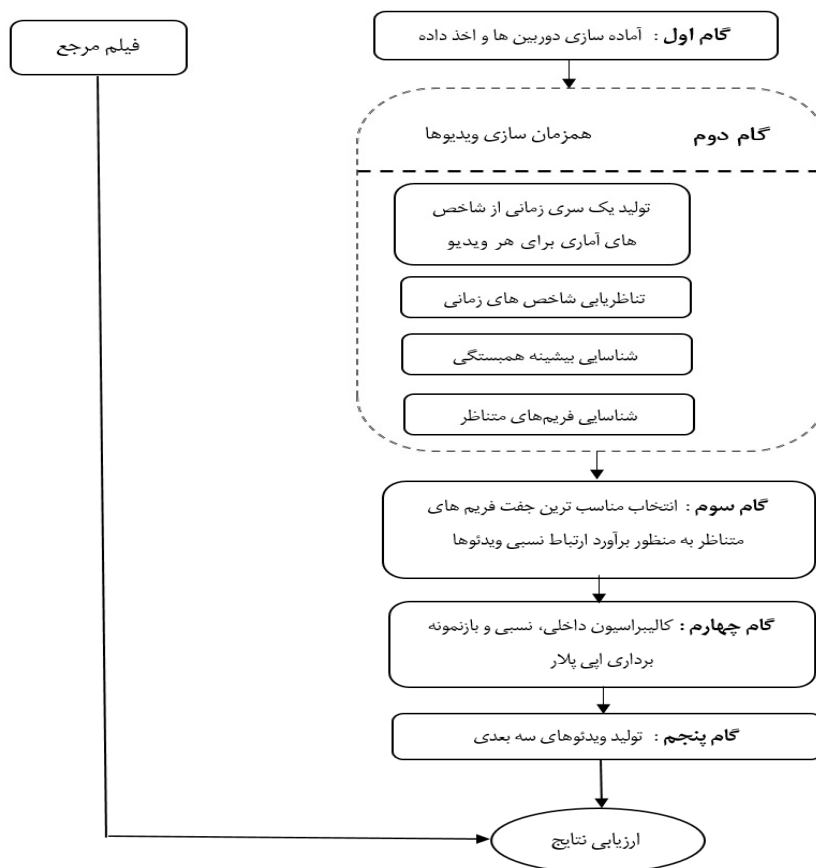
از منظر هندسی، ۶ درجه‌ی آزادی در ارتباط نسبی دو دوربین وجود دارد. با این حال، ارتباط نسبی چشم‌ها درجه‌ی آزادی کمتری را بواسطه‌ی محدودیت‌های حرکتی، تقارن در چرخش و فاصله‌ی ثابت چشم‌ها برخوردارند. پر تکرارترین تنظیم نسبی، تامین توازی محورهای اصلی چشم‌ها و اخذ تصاویر استریو با اختلاف نامحسوس در مقیاس بوده که کمترین فشار را به سامانه‌ی بینایی انسان تحمیل می‌سازند [۴۴]. در صورتیکه ویدیوهای اخذ شده توسط دوربین‌های استریو شرایط مشابه با حالات نسبی قابل درک در سامانه‌ی بینایی انسان را نداشته باشند؛ ارسال مستقل محتوای هر ویدیو به چشم‌ها به کمک ابزارهای برجسته‌بینی منجر به دوبینی و یا برجسته‌بینی موضعی می‌گردد. به همین جهت، ساخت ویدیوهای سه بعدی 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، ویدیوهای مناسب به‌منظور اهداف برجسته‌بینی اخذ نماید. در راهکارهای نرم‌افزاری نیز، ناهمگونی در پارامترهای داخلی و ارتباط نسبی دوربین‌ها به‌کمک نگاشت‌های هندسی، اصلاح و ویدیوهای بازنمونه‌برداری شده تولید می‌گردند. در این روند، ویدیوهای پرسپکتیو با محورهای موازی تصویربرداری و مقیاس یکسان به‌کمک ویدیوهای خام و با برآورد پارامترهای داخلی و ارتباط نسبی دوربین‌ها تولید می‌شوند [۴۵].

۳-۲ - روش پیشنهادی

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

- ۱- عدم امکان تنظیم نسبی دقیق دوربین‌ها به‌صورت مکانیکی به‌منظور اخذ ویدیوهای 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 ویدیوهای اخذ شده توسط دوربین‌های استریو)، ۳- عدم امکان تامین هم‌زمانی دقیق در آغاز فیلم‌برداری برای دوربین‌های استریو، ۴- نرخ نمونه‌برداری زمانی (fps) متمایز ویدیوهای اخذ شده توسط دوربین‌ها به‌عنوان چالش‌های این مسیر قلمداد می‌شوند. انتظار می‌رود که راهکار پیشنهاد شده بتواند به‌صورت خودکار ویدیوهای اصلاح شده‌ی هندسی و زمانی را به‌منظور به‌کارگیری در تکنیک‌های برجسته‌بینی تولید نماید.
- در شکل (۱-۳) فلوچارتی از روند اجرای الگوریتم پیشنهادی ارائه شده است.



شکل (۳-۱) - فلوچارتی از روند کلی اجرای الگوریتم پیشنهادی

آماده سازی دوربین ها و اخذ داده، همزمان سازی ویدیوها، انتخاب جفت فریم های متناظر و مناسب به منظور تخمین ارتباط نسبی و داخلی دوربین ها، تولید ویدیوهای نرمال شده و تولید ویدیوهای سه بعدی به ترتیب گام های روش پیشنهادی را تشکیل داده که در ادامه تشریح شده اند.

۳-۲-۱ - گام اول (آماده سازی دوربین ها و اخذ داده)

در راهکار پیشنهادی از کنار هم قرار دادن دو دوربین آماتور به صورت ثابت به منظور تولید ویدیوهای سه بعدی استفاده شده است. عامل انسانی مفروض به منظور آماده سازی دوربین ها و اخذ ویدیو غیر متخصص تصور شده، که امکانات تنظیم دقیق نسبی دوربین ها 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 به‌منظور اخذ داده اشاره شده که اجرای آنها برای عامل آماتور میسر بوده و همچنین کیفیت نتایج ویدیوهای 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 به‌منظور آماده‌سازی دوربین‌ها محسوب می‌شوند. برای تامین ثبات نسبی دوربین‌ها، پایه‌ای چوبی به صورت شکل (۲-۳) طراحی شده و مورد استفاده قرار گرفت.



شکل (۲-۳) - نمایی از دستگاه ساخته شده به‌منظور تامین ثبات نسبی دوربین‌ها

عدم وجود تقارب و یا واگرایی محور دوربین‌ها نیز صرفاً به‌صورت بصری از طریق هم‌راستایی گوشه‌های 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 به دنبال داشته و

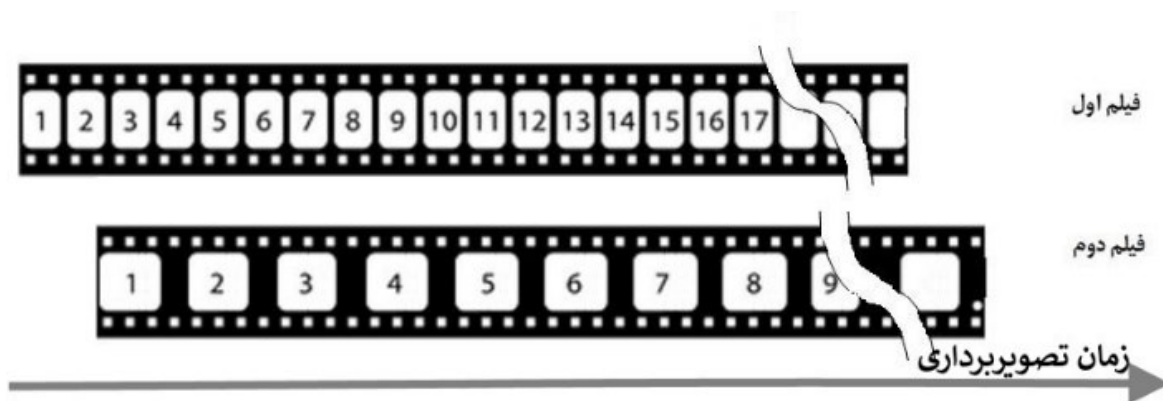
پارالاکس‌های بزرگ منجر به دوبینی می‌شوند. از طرفی، یکسان بودن فاصله‌ی کانونی اختلاف مقیاس میان دو ویدیو را کاهش داده و تصاویر نرمال شده کمتر دچار تغییرات مقیاس خواهند شد. همانطور که اشاره شد از منظر هندسی، ۶ درجه‌ی آزادی در ارتباط نسبی دو دوربین در حالیکه محورهای دو دوربین موازی یکدیگرند وجود دارد، پس از تست‌های انجام شده از بین این ۶ حالت، موردی که شامل کمترین میزان دوبینی و اعوجاج بوده انتخاب شد. بنابراین در طی فرایند تولید فیلم سه بعدی هر دو گوشی بصورت افقی و دوربین‌ها با کمترین فاصله نسبت به یکدیگر قرار گرفتند. از نقطه نظر نرم‌افزاری، پایداری فوکوس و عدم فعال بودن امکاناتی مانند فوکوس خودکار مهمترین موارد لازم به‌منظور تنظیم دوربین‌ها هنگام فیلم‌برداری خواهد بود. اعمال قیودی مانند تنظیم دستی fps دوربین‌ها به‌منظور اخذ ویدیو 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 به‌منظور افزایش کیفیت ویدیوهای سه‌بعدی محسوب می‌شوند.

۳-۲-۲ - گام دوم (همزمان‌سازی ویدیوها)

در روش پیشنهادی دوربین‌های استریو نسبت به یکدیگر ثابت هستند. در صورت استفاده از ویدیوهای همزمان‌سازی نشده، نمی‌توان انتظار داشت که ارتباط نسبی پایداری بین ویدیوهای استریو وجود داشته باشد. به‌عبارت بهتر، پایداری ارتباط نسبی میان ویدیوهای استریو هنگام همزمان بودن آنها قابل تأمین خواهد بود. در شرایط استفاده از ویدیوهای همزمان‌سازی نشده، به ازای هر جفت فریم متناظر از دو ویدیو، پارامترهای نسبی متمایزی وجود خواهد داشت و روند تغییرات زمانی آنها تابعی از حرکت نسبی دوربین و صحنه خواهد بود. بدیهی است که در صورت همزمان نبودن ویدیوها، به ازای هر جفت فریم متناظر، برآورد پارامترهای نسبی و به تبع آن نگاشت 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 ویدیوهای 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 ویدیوهای 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. علاوه، 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 ویدیوی سه‌بعدی گردد. شکل (۳-۳) مساله‌ی اختلاف زمانی میان ویدیوها و همچنین تمایز در fps را بصورت شماتیک در یک نمودار یک‌بعدی نشان می‌دهد.



شکل (۳-۳) - نمایی شماتیک از تمایز بین fps و اختلاف زمانی آغاز ویدیوها

با توجه به شکل (۳-۳)، بدیهی است که در صورت وجود تمایز در fpsها، یافتن فریم‌های متناظر برای ویدیوی 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 تلاش شده تا به ازای تمامی فریم‌های ویدیو با fps کمتر، فریمی متناظر از ویدیو با fps بیشتر شناسایی و به آن الصاق گردد.

راهکار توسعه یافته برای همزمان‌سازی ویدیوهای اخذ شده از دو دوربین، مبتنی بر تناظریابی شاخص‌های زمانی است. در این روند ابتدا به ازای هر فریم از هر یک از ویدیوها یک شاخص آماری مستخرج از محتوای تصویری آن فریم تولید می‌گردد. بر این اساس، به ازای هر ویدیو 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 ویدیو خواهد بود. بر این اساس، به دلیل عدم تطابق احتمالی fps ویدیوها، در یک روند درونیابی خطی توان تفکیک زمانی سری زمانی تولید شده برای ویدیو با fps کمتر با دوربین با fps بیشتر یکسان‌سازی می‌شود. این اقدام، توان تفکیک زمانی شاخص‌های آماری دو ویدیو را یکسان و برابر با ویدیوی با fps بیشتر می‌نماید. لازم به ذکر است که در این روند یکی از شاخص‌های زمانی دچار فرایند درونیابی خطی شده و سری زمانی مستخرج از ویدیوی 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 ویدیو، تمامی مجموعه‌های $\{Intersection_i\}$ متشکل از زوج مرتب‌های رابطه‌ی (۵) ایجاد می‌گردند.

$$Intersection_i = \{(k_m, q_i), (k_{m-1}, q_{i-1}), (k_{m-2}, q_{i-2}), \dots, (k_1, q_{i-m}) \mid k \in K \text{ and } q \in Q\}$$

$$i = 0, 1, 2, \dots, (m + n) \quad \text{رابطه‌ی (۵)}$$

در رابطه‌ی (۵)، $K = \{k_1, k_2, \dots, k_m\}$ به‌عنوان مجموعه‌ی یکی از شاخص‌های آماری استخراج شده برای ویدیو با fps کمتر متشکل از m عضو و $Q = \{q_1, q_2, \dots, q_n\}$ مجموعه‌ی مشابه برای ویدیوی 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 K و Q به کمک رابطه‌ی (۶) به زمان متناظر تبدیل می‌شوند.

$$t_k = \frac{j-1}{fps_{max}} \quad \text{رابطه‌ی (۶)}$$

در رابطه‌ی (۶)، j اندیس اعضای مجموعه‌ی K و Q و همچنین fps_{max} معادل با fps ویدیوی سریعتر می‌باشند. اندیس‌های مذکور برای مجموعه‌های P و Q به ترتیب در بازه‌ی $[1 \sim m]$ و $[1 \sim n]$ معنادار بوده و زوج‌مرتب‌های 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 خواهد بود که در ادامه با k_i شناخته می‌شوند. در ادامه، به ازای زوج مرتب‌های هر $\{Intersection_i\}$ معیار همبستگی پیرسون بصورت رابطه‌ی (۷) محاسبه می‌شود [۴۶].

$$\rho_i = \frac{\sum_{j=1}^{k_i} (k_j - \bar{k})(q_j - \bar{q})}{\sqrt{\sum_{j=1}^{k_i} (k_j - \bar{k})^2 \sum_{j=1}^{k_i} (q_j - \bar{q})^2}} \quad \text{رابطه‌ی (۷)}$$

در صورتیکه Δ به عنوان آرگومان بیشینه‌کننده‌ی معیار همبستگی تعریف گردد (رابطه‌ی ۸)؛ اندیس فریم‌های متناظر بین دو ویدیو از طریق اندیس زوج مرتب‌های مجموعه‌ی $\{Intersection_{\Delta}\}$ شناسایی می‌شوند.

$$\Delta = \underset{i}{\operatorname{argmax}}(\rho_i) \quad \text{رابطه‌ی (۸)}$$

به واسطه‌ی درونیابی خطی بکاررفته در تولید شاخص‌های زمانی، توان تفکیک زمانی مربوط به اندیس زوج مرتب‌های مجموعه‌ی $\{Intersection_{\Delta}\}$ معادل با ویدیوی با fps بیشتر خواهد بود. بر این اساس، در ادامه با معلوم بودن زمان اخذ هر فریم از ویدیوی با fps کمتر (t_K)، زمان متناظر در ویدیوی با fps بیشتر (t_Q) به کمک رابطه‌ی (۹) محاسبه می‌گردد.

$$t_Q = t_K + \frac{m - \Delta}{fps_{max}} \quad \text{رابطه‌ی (۹)}$$

در این تحقیق از شاخص‌های آماری مختلفی در روند تناظریابی استفاده شده که در جدول (۱-۳) معرفی شده‌اند. هدف از این اقدام یافتن شاخص‌های مطلوب و همچنین حساسیت‌سنجی نتایج به انتخاب نوع شاخص بوده است.

جدول ۱-۳ - 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	$\frac{\sum_{j=1}^n \sqrt{v_{xj}^2 + v_{yj}^2}}{n}$
انترپی	$H = - \sum_{i=0}^{L-1} P_i \log P_i$
مساحت تابع توزیع تجمعی هیستوگرام	$A = \sum_{i=0}^{L-1} \sum_{j=0}^i P_j$
<p>P_i: احتمال رخداد درجه خاکستری در تصویر</p> <p>L: تعداد درجات خاکستری</p> <p>v_x: مولفه بردار جریان تصویری در جهت x</p>	<p>v_y: مولفه بردار جریان تصویری در جهت y</p> <p>n: تعداد بخش‌های تصویر به منظور تولید بردارهای نوری</p>

انترپی شانون خاکستری یکی از روش‌های تصمیم‌گیری چند شاخصه می باشد که در آن از اعداد خاکستری^۲ استفاده می‌گردد. عدد خاکستری عددی است که مقدار واقعی آن مشخص نیست اما بازه‌ای که عدد در آن قرار می‌گیرد مشخص است. در واقع انترپی اندازه‌گیری توزیع شدت و معیاری برای اندازه‌گیری عدم اطمینان بیان شده بوسیله توزیع احتمال گسسته می باشد. در نظریه انترپی اطلاعات بصورت کمی اندازه‌گیری می‌گردد، این اطلاعات مجموعه‌ای از نمادها و رمزها و احتمال وقوع رویداد است و با اندازه‌گیری بسامد نمادها و رمزها میتوان درجه احتمال رویداد را سنجید [۴۷].

در انترپی براساس هیستوگرام تصویر، هر چقدر انترپی بزرگتر باشد تصویر حاوی اطلاعات بیشتری است و بالعکس هر چقدر محتوای یک تصویر کمتر باشد مقدار انترپی تصویر عدد کوچکتری خواهد بود، بدین معنی که تصویر محتوای اطلاعاتی پایینی دارد.

$$H = - \sum_{i=0}^{L-1} P_i \log P_i \quad \text{رابطه‌ی (۱۰)}$$

شار نوری الگوی آشکار اشیاء، سطوح و لبه‌ها در یک صحنه بصری بوده، که ناشی از حرکت نسبی بین یک ناظر و یک صحنه است. شار نوری سعی بر محاسبه‌ی حرکت، بین دو فریم از تصاویر در زمان‌های t و $t+\Delta t$ دارد.

بردارهای جریان نوری با فرض ثبات در محتوای تصویری فریم‌های متوالی، میزان جابجایی هر ناحیه از تصویر را به کمک رابطه‌ی (۱۱) تخمین می‌زنند.

$$I_x \times v_x + I_y \times v_y = -I_t \quad \text{رابطه‌ی (۱۱)}$$

^۱Entropy

^۲Grey number

در رابطه‌ی (۱۱)، I_x ، I_y و I_t بترتیب مشتق در جهات x ، y و زمان بوده و همچنین $[v_x, v_y]$ مولفه‌های بردار جریان تصویری هستند. این بردارها را می‌توان به صورت یک شبکه‌ی منظم در فضای تصویر متناسب با محتوای رادیومتریکی هر سلول از شبکه محاسبه نمود [۴۸].

ضریب چولگی بیانگر میزان عدم تقارن توزیع احتمال داده‌ها حول عدد میانگین است. اگر داده‌ها نسبت به میانگین متقارن باشند، چولگی برابر صفر است. مقدار چولگی می‌تواند منفی یا مثبت باشد. برای یک توزیع نامتقارن با کشیدگی به سمت مقادیر بالاتر چولگی مثبت و برای توزیع نامتقارن با کشیدگی به سمت مقادیر کوچکتر مقدار چولگی منفی است [۴۹].

$$\mu_3 = \frac{\sum_{i=0}^{L-1} P_i \times (i-\mu)^3}{\sigma^3 \times \sum_{i=0}^{L-1} P_i} \quad \text{رابطه‌ی (۱۲)}$$

تابع توزیع تجمعی یک تابع صعودی (و نه لزوماً صعودی اکید) است که احتمال اینکه یک متغیر تصادفی کوچکتر یا مساوی مقدار معینی باشد را ارائه می‌دهد. از این رو به این تابع تجمعی گفته می‌شود که بیان‌کننده تجمع (جمع) احتمالات رخداد کلیه مقادیر کمتر یا مساوی یک عدد است.

$$A = \sum_{i=0}^{L-1} \sum_{j=0}^i P_j \quad \text{رابطه‌ی (۱۳)}$$

۳-۲-۳- گام سوم (انتخاب جفت فریم‌های متناظر به منظور برآورد ارتباط نسبی ویدیوها)

از منظر سخت‌افزاری ارتباط نسبی دوربین‌های استریو در طول زمان پایدار است. با این حال در صورت همزمان نبودن فریم‌های نمایش یافته در ویدیوی سه‌بعدی، ارتباط نسبی ناپایدار در خلال پخش ویدیو 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 لازمی بازنمونه‌برداری اپی‌پلار به منظور موازی‌سازی راستاهای پارالاکس و هم‌مقیاس نمودن ویدیوهای اخذ شده است. هم‌زمان‌سازی انجام

شده در بخش (۲-۲-۳) برای ویدیوها، شرایط پایداری ارتباط نسبی را برای فریم‌های متناظر فراهم می‌آورد. با این حال انتخاب فریم‌های متناظر در شرایط سکون قابلیت اعتماد بیشتری را در برآورد ارتباط نسبی تامین خواهد ساخت. به عبارت بهتر، انتخاب فریم‌های متناظر در زمان وجود حرکت بواسطه‌ی عواملی همچون:

۱- دقت تناظریابی زمانی، ۲- بازنمونه‌برداری زمانی ویدیوها به منظور یکسان‌سازی fps و ۳- اثر شاتر جاروب‌شونده [۵۰] می‌تواند صحت تخمین ارتباط نسبی را مخدوش سازد. در این بخش تلاش شده تا با تمهیداتی، به صورت خودکار فریم‌های متناظر مناسب، در زمان‌های سکون به منظور برآورد ارتباط نسبی دوربین‌های استریو شناسایی گردند.

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 (متقارب نبودن) و ۴- محتوای رادیومتریکی نامعلوم ویدیوها، مساله‌ی یافتن فریم‌های مناسب و تناظریابی میان آنها را با چالش‌های بیشتری روبرو می‌سازد. یافتن زمان‌های سکون در طول ویدیوهای همزمان‌سازی شده، انتخاب توزیع شده‌ی نقاط متناظر متناسب با محتوای تصویری زمان‌های سکون و حصول اطمینان از توزیع مکانی مطلوب نقاط متناظر، مواردی بشمار رفته که در روند برآورد پارامترهای داخلی و نسبی دوربین‌ها لحاظ شده‌اند.

زمان‌های رخداد سکون در ویدیوها را می‌توان به کمک شاخص‌های مختلفی اندازه‌گیری کرد. بردارهای جریان نوری را می‌توان به عنوان یکی از معیارهای حرکت نسبی صحنه و دوربین قلمداد نمود. علی‌رغم مزایای بکارگیری بردارهای جریان نوری، جابجایی‌های بزرگ و فقدان محتوای رادیومتریکی مناسب از عوامل عدم قطعیت در برآورد صحیح بردارهای جریان نوری محسوب می‌شوند.

¹Rolling Shutter

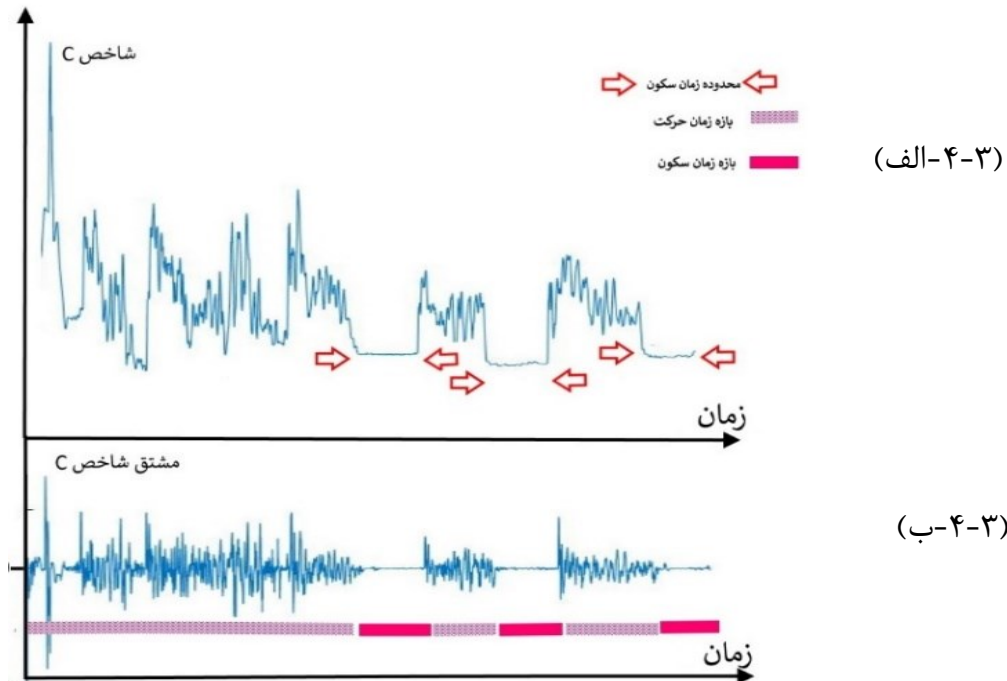
از سوی دیگر، رخداد سکون در ویدیو، ثبات در محتوای تصویری فریم‌های متوالی را به‌همراه داشته که این موضوع می‌تواند پایداری زمانی را در برخی شاخص‌های آماری مستخرج از فریم‌های تصویر ایجاد سازد. شاخص‌هایی مثل انترویی و آماره‌های مستخرج از هیستوگرام تصویر در زمره‌ی این شاخص‌های آماری قرار دارند. با این حال ممکن است بواسطه‌ی شرایط محیطی ویدیوها، علیرغم حرکت نسبی دوربین و صحنه، تغییرات محسوسی در این شاخص‌ها رخ ندهد. محیط‌های بدون بافت و فضا‌های متخلخل با توزیع یکنواخت در روشنایی مصداقی از این صحنه‌ها بشمار می‌روند. در این تحقیق بطور همزمان از پتانسیل سری زمانی شاخص‌های مبتنی بر جریان نوری و همچنین شاخص‌های آماری محتوای تصویری به‌منظور کاهش عدم قطعیت‌های هر یک در شناسایی مقاطع زمانی رخداد سکون در ویدیوها استفاده شده است.

با توجه به موفقیت‌آمیز بودن شاخص انترویی نسبت به سایر شاخص‌های آماری در آزمون تناظرایی زمانی، در این بخش نیز این شاخص برای شناسایی مناطق سکون انتخاب شده است. برای این منظور به ازای تمامی فریم‌ها در ویدیوی با fps بیشتر، انترویی نرمال‌شده‌ی هیستوگرام ($e_i, i = 1, 2, \dots, f\Delta$) و همچنین میانگین طول بردارهای جریان تصویری ($o_i, i = 1, 2, \dots, f\Delta$) محاسبه می‌گردد. انتظار می‌رود در شرایط سکون، میانگین طول بردارهای جریان نوری مقداری نزدیک به صفر و شاخص انترویی نیز مقدار عددی ثابتی را در 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 بصورت رابطه‌ی (۱۴) برای شناسایی فریم‌های توام با سکون استفاده شده است.

$$C_i = \frac{o_i}{e_i} \quad i = 1, 2, \dots, f\Delta \quad \text{رابطه‌ی (۱۴)}$$

شکل (۳-۴-الف) نمایی شماتیک از روند تغییرات زمانی معیار C را برای یک ویدیو نمایش می‌دهد. همانطور که در این شکل مشخص است، مناطق توام با سکون بصورت 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 ۳-۴- (الف) نمایشی شماتیک از روند تغییرات زمانی معیار C برای یک ویدیو

شکل (ب) نمایشی 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 بصورت باینری قطع‌بندی می‌گردد. قطعات بهم‌پیوسته، بازه‌های زمانی سکون را نشان داده که برای هر بازه می‌توان اعتباری متناسب با متوسط بزرگی C اختصاص داد (شکل ۳-۴-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 به کمک روش تناظریابی عارضه مبنای SIFT با اعمال قیود سخت‌گیرانه‌ی شباهت و تمایز توصیف‌گرها و همچنین اعتبار سنجی هندسی نقاط متناظر

از طریق اعمال روش RANSAC 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 به‌منظور رفع مشکل محتمل در توزیع یا تعداد نامناسب نقاط متناظر به‌کار گرفته شود. بر این اساس، به‌منظور حصول اطمینان از تأمین 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 شباهت‌توصیفگرهای SIFT و مقدار C مربوط به زوج فریم بکارگرفته شده، اختصاص می‌یابد. هر یک از این مقادیر به‌منظور همسان‌سازی مقیاسی قبل از جمع شدن در بازه صفر تا یک نرمال‌سازی می‌گردند.

بعد از اجرای روند تناظریابی مطابق با ضوابط تشریح شده برای تمامی جفت‌فریم‌های متناظر، لیستی تجمیع شده از تمامی نقاط تناظریابی شده تولید می‌گردد. بدیهی است که بواسطه 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 بشمار 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 بصورت $[X_I, Y_I, X_I - X_T]$ برای تمامی نقاط متناظر تولید می‌گردد. فضای سه‌بعدی گسترش‌یافته توسط مختصات‌های مورد بحث به $7 \times 11 \times 1$ سلول شبکه‌بندی می‌شود. سپس میانگین تعداد نقاط واقع در سلول‌های ناتهی (μ) محاسبه شده وضریبی از آن ($\alpha\mu$) 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 $\alpha\mu$ 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 $\alpha\mu$ از تعداد 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 بکارگرفته شده در این تحقیق، مقدار عددی این ضریب معادل یک انتخاب شده است. شبه‌کد ارائه شده در الگوریتم (۱) روند اجرای این مرحله را نشان می‌دهد.

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

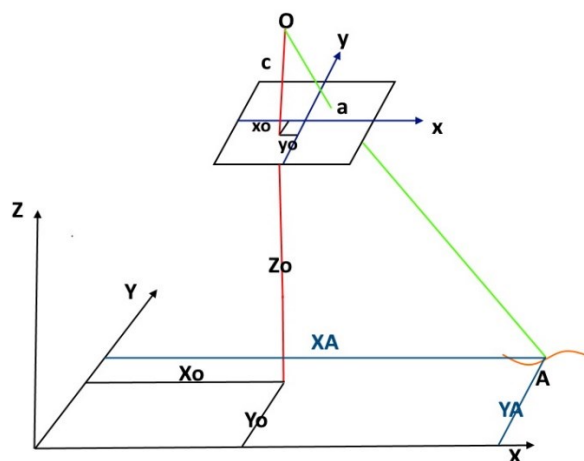
ورودی‌ها:	
<p>نمودار زمانی $C=[c_1, c_2, \dots, c_{t\Delta}]$</p> <p>حد آستانه مشتق C به‌منظور جداسازی قطعات زمانی سکون (T_s)</p> <p>ابعاد شبکه‌ی اندازه‌گیری توزیع نقاط متناظر ($t \times u \times v$)</p> <p>ویدیوهای هم‌زمان‌سازی شده (بخش ۲-۲-۳)</p> <p>ضریب میانگین (α) به‌منظور کاهش تراکم نقاط متناظر</p>	
خروجی‌ها:	
نقاط متناظر به‌منظور برآورد پارامترهای ارتباط نسبی	
۱	محاسبه‌ی مشتق زمان‌بردار C بصورت $\left(\frac{dc}{dt_i} = c_{i+1} - c_i\right)$
۲	اعمال حد آستانه T_s به‌منظور باینری‌سازی بردار $\frac{dc}{dt}$ بصورت $R(t) = \begin{cases} \left \frac{dc}{dt_i}\right < T_s \rightarrow 1 \\ else \rightarrow 0 \end{cases}$
۳	قطعه‌بندی نواحی بهم‌پیوسته‌ی دارای مقدار ۱ 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 عارضه مبنای 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 هزینه‌ی هر نقطه‌ی متناظر
۷	نگاشت نقاط متناظر شناسایی شده به فضای سه‌بعدی $[x_l, y_l, x_r - x_l]$ 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

۳-۲-۴ - گام چهارم (کالیبراسیون داخلی، نسبی و بازنمونه‌برداری اپی‌پلار)

در این فرایند با فرض معلوم بودن پارامترهای کالیبراسیون داخلی در دوربین‌ها (فاصله‌ی اصلی، موقعیت مرکز تصویر، ابعاد پیکسل‌ها و در صورت لزوم پارامترهای مرتبط با خطاهای سیستماتیک ناشی از اعوجاجات سیستم اپتیکی) ارتباط نسبی دوربین‌ها از طریق اندازه‌گیری حداقل ۵ نقطه‌ی متناظر (گسترده شده در فضای همپوشانی تصاویر استریو) قابل برآورد است. بر این اساس، اجرای خودکار این فرایند وابسته به یافتن خودکار نقاط متناظر در تصاویر استریو بوده و باقی فرایند در قالب حل یک دستگاه معادلات غیرخطی به اجرا می‌رسد.

لازم به ذکر است که در صورت افزوده شدن پارامترهای کالیبراسیون داخلی به مجهولات ارتباط نسبی، این روند به نوعی یک free network self-calibration تلقی شده، با این تمایز که بجای بکارگیری تعداد زیاد تصاویر متقارب، صرفاً از یک زوج تصویر استریو استفاده شده است.

بطور معمول در هر دو حالت (توجیه نسبی یا خودکالیبراسیون) از معادله‌ی شرط هم‌خطی (تبیین‌کننده‌ی هندسه‌ی هم‌خطی موقعیت عکسی، مرکز تصویر و موقعیت متناظر زمینی برای هر نقطه‌ی ثبت شده در تصویر در لحظه‌ی اخذ آن) به‌عنوان معادله‌ی ارتباط هندسی فضای دوبعدی عکس و فضای سه‌بعدی شئی استفاده می‌شود (شکل ۳-۵). در رابطه‌ی (۱۵) نیز معادله‌ی شرط هم‌خطی ارائه شده است.



شکل (۳-۵) - ارتباط هندسی فضای دوبعدی عکس و فضای سه‌بعدی شئی

$$x - x_0 = -c \frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

رابطه‌ی (۱۵)

$$y - y_0 = -c \frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}$$

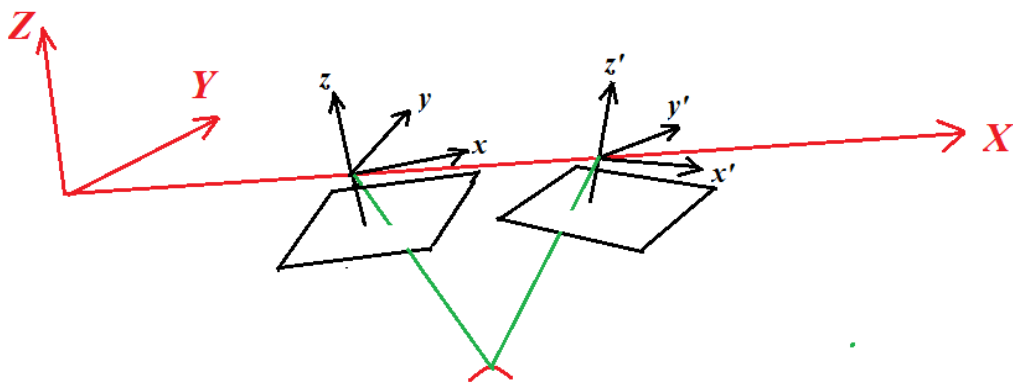
در رابطه‌ی فوق، مختصات متریک یک نقطه در فضای تصویر، $[x_0, y_0]$ مختصات متریک پای قائم مرکز تصویر نسبت به مرکز ماتریس تصویر، r_{ij} عناصر ماتریس دوران متشکل از سه دوران مستقل حول محورهای اول (ω) ، دوم (φ) و سوم (κ) مختصاتی، $[X_0, Y_0, Z_0]$ مختصات مرکز تصویر در سیستم مختصات شئی، $[X, Y, Z]$ مختصات زمینی متناظر و c فاصله اصلی دوربین هستند. متناسب با معادله‌ی شرط هم‌خطی ارتباط هندسی یک نقطه‌ی سه‌بعدی در فضای شئی با متناظر آن در فضای تصویر در زمان معلوم بودن پارامترهای داخلی (x_0, y_0, c) و خارجی $(\omega, \varphi, \kappa, X_0, Y_0, Z_0)$ میسر است. در روند کالیبراسیون داخلی و نسبی دوربین‌های استریو مجموعاً ۶ پارامتر داخلی و ۱۲ پارامتر

خارجی برای دو دوربین تعیین می‌گردد. بدیهی است که به دلیل بکارگیری نقاط متناظر در روند کالیبراسیون داخلی و نسبی، مجهول تصور کردن تمامی ۱۲ پارامتر خارجی منجر به نقص در تعریف سیستم مختصات فضای شئی شده که این امر به طرق مختلف و با اعمال قید در پارامترهای خارجی در روند کالیبراسیون نسبی مرتفع می‌گردد [۵۱].

در روند کالیبراسیون نسبی از نقاط متناظر استفاده می‌شود. نامعلوم بودن مختصات شئی نقاط متناظر باعث شده که به ازای هر نقطه‌ی متناظر، مختصات سه‌بعدی آن در فضای شئی نیز به‌عنوان مجهول به مجموعه‌ی مجهولات افزوده گردد. با این حال هر نقطه‌ی متناظر با تامین ۴ معادله (۲ معادله‌ی شرط هم‌خطی برای هر تصویر) امکان برآورد همزمان پارامترهای خارجی و مختصات سه‌بعدی نقاط متناظر را فراهم می‌آورد. بر این اساس در صورت مجهول بودن پارامترهای خارجی دوربین‌های استریو (۱۲ پارامتر)، به حداقل ۱۲ نقطه‌ی متناظر نیاز خواهد بود. اما به دلیل نقص در تعریف سیستم مختصات فضای شئی، این دستگاه معادلات دچار تکینگی شده و کمبود رتبه‌ی ماتریس ضرایب مجهولات در این شرایط معادل ۷ خواهد بود. ۷ نقص مربوط به تعریف سیستم مختصات شئی شامل ۳ پارامتر انتقال، ۳ پارامتر دوران و ۱ پارامتر مقیاس است.

در روند برآورد پارامترهای کالیبراسیون نسبی به شکل‌های مختلفی نقص در تعریف سیستم مختصات در فضای شئی جبران می‌شود. یکی از ساده‌ترین راهکارهای انجام این مهم، ثابت در نظر گرفتن برخی از پارامترهای خارجی دوربین‌های استریو است. در روال‌های اجرایی فتوگرامتری به دو شکل این موضوع به اجرا می‌رسد. در حالت نخست، تمامی پارامترهای خارجی یکی از دوربین‌ها ثابت تصور شده و پارامترهای خارجی دوربین دوم بطور نسبی برآورد می‌شود (توجه نسبی یکطرفه). در این شرایط ۶ پارامتر به‌عنوان پارامترهای ارتباط نسبی مطرح بوده که کماکان یک نقص را بدنبال خواهد داشت. ارتباط نسبی دو دوربین در فواصل متفاوتی نسبت به یکدیگر قابل استحصال است. بر این اساس، ثابت در نظر گرفتن فاصله‌ی دو دوربین به‌عنوان قید دیگری در روند برآورد پارامترهای

کالیبراسیون نسبی لحاظ می‌شود. این قید از طریق ثابت لحاظ شدن یکی از سه مولفه‌ی انتقالی مرکز تصویر (X_0, Y_0, Z_0) دوربین دوم به اجرا می‌رسد. با این اوصاف مجموع پارامترهای مجهول کالیبراسیون نسبی شامل ۵ پارامتر (۲ پارامتر انتقالی و ۳ پارامتر دورانی) برای دوربین دوم خواهد بود. در حالت دوم اجرای کالیبراسیون نسبی (روش بکارگرفته شده در این تحقیق)، با یک فرض ابتدایی، محور X سیستم مختصات شئی موازی با خط واصل مراکز تصویر دوربین‌های استریو تصور می‌شود.



شکل (۳-۶) - شمایی از محور X سیستم مختصات شئی موازی با خط واصل مراکز تصویر دوربین‌های استریو با این فرض، تمامی پارامترهای انتقالی (X_0, Y_0, Z_0) دو دوربین ثابت تصور می‌شوند. در این رویکرد پارامترهای دورانی دو دوربین به‌عنوان پارامترهای ارتباط نسبی شناخته می‌شوند. با این حال دوران محور X سیستم مختصات‌های تصویری نسبت به محور X سیستم مختصات شئی، کماکان به‌عنوان پارامتر افزونه و توأم با ابهام تلقی می‌شود. همبستگی بین زاویه‌ی چرخش دوربین‌ها حول محور X سیستم مختصات شئی نیز باعث ابهام و نقص در تعریف سیستم مختصات شئی شده که با ثابت لحاظ شدن این زاویه در یکی از دوربین‌ها مرتفع می‌گردد. بر این اساس پارامترهای ارتباط نسبی دوربین‌ها به ۵ زاویه‌ی دوران برای دو دوربین محدود می‌شود.

به عبارت بهتر، در روش بکارگرفته شده صرفاً المان‌های دورانی به عنوان پارامترهای کالیبراسیون نسبی استفاده شده و پنج پارامتر به ترتیب دوران دوران Φ و K برای دوربین اول و Φ ، ω و K برای دوربین دوم خواهد بود. دلیل استفاده از این رویکرد همبستگی کمتر پارامترهای دورانی با پارامترهای کالیبراسیون داخلی دوربین‌ها (x_0, y_0, c) است [۵۰].

با توجه به ماهیت غیرحرفه‌ای دوربین‌های به کارگرفته شده، بدیهی است که علاوه بر پارامترهای ارتباط نسبی دوربین‌ها، مقادیر فاصله کانونی و همچنین موقعیت پای قائم مرکز تصویر در آنها نیز بطور دقیق مشخص نباشد. بر این اساس در این مرحله به کمک نقاط متناظر شناسایی شده در بخش (۳-۲-۳)، 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 [۵۲].

در مثلث‌بندی دسته اشعه، با مبنا قرار دادن معادله‌ی شرط هم‌خطی، 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 [۵۲]. در شبکه‌ی آزاد به دلیل عدم وجود نقاط کنترلی (نقاط با مختصات معلوم در فضای شی) صرفاً از پتانسیل نقاط متناظر در برآورد پارامترهای مجهول استفاده می‌شود.

راهکار به کارگرفته شده در بخش (۳-۲-۳) 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 مشاهدات به کاررفته در مثلث‌بندی دسته‌اشعه تامین سازد.

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 ۳-۲). 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 به عنوان شرایط بدوضعی شناخته می‌شود [۵۳]. همچنین به دلیل وجود صرفاً دو موضع ثابت در اخذ ویدیو و عدم تقارب کافی در دوربین‌های استریو، برآورد همزمان پارامترهای کالیبراسیون داخلی و نسبی مستعد بروز شرایط بدوضعی خواهد بود.

در مورد چگونگی تعریف سیستم مختصات شیء در فرآیند برآورد پارامترهای مجهول، راهکار تثبیت برخی از پارامترهای خارجی دوربین‌ها استفاده شده است. در این راهکار که به‌عنوان فرآیند توجیه نسبی دوطرفه شناخته 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 تامین نقاط متناظر با توزیع مناسب در پارالاکس و فضای تصویر می‌تواند شرایط بدوضعی دستگاه معادلات متأثر از هندسه‌ی نامستحکم دوربین‌ها را کاهش دهد.

در روند برآورد همزمان پارامترهای کالیبراسیون داخلی و نسبی از رویکردهای رایج مواجهه با بدوضعی دستگاه معادلات غیرخطی که در مراجع شناخته‌شده‌ی مسائل ژئودتیکی توسعه یافته‌اند، استفاده شده است. این موارد شامل اقداماتی همچون افزوده شدن تدریجی پارامترهای مجهول با هدف تدقیق مقادیر اولیه قبل از حل همزمان تمامی پارامترهای مجهول و بکارگیری پارامترهای تدقیق شده به‌عنوان شبه [۵۴] مشاهدات وزندار با هدف جلوگیری از تشدید همبستگی می‌باشند.

در هنگام خرید گوشی همراه برخی اطلاعات داخلی دوربین گوشی مثل کیفیت دوربین، مقدار تقریبی فاصله کانونی، اندازه پیکسل سائز، f-stop و ... در داخل کاتالوگ آن نوشته شده است.

¹Ill posed or ill condition

همچنین تلاش سازندگان دوربینها بر این بوده تا نقطه اصلی بر مرکز تصویر منطبق باشد. بنابراین در طی پروسه کالیبراسیون داخلی و نسبی مقدار تقریبی x_0 و y_0 و فاصله کانونی به عنوان شبه مشاهده وارد معادلات می شود، بنابراین در این تحقیق مقدار x_0 و y_0 مساوی صفر و مقدار تقریبی فاصله کانونی $1/5$ میلی متر در نظر گرفته شد. همانطور که در بالا اشاره شد در توجیه نسبی یکطرفه یکی از دوربینها ثابت فرض میشود بنابراین وابستگی کمتری بین مراکز تصاویر وجود دارد، بنابراین با جایگذاری مقادیر اولیه و همچون افزوده شدن تدریجی پارامترهای مجهول در معادلات توجیه نسبی یکطرفه، مقادیر تقریبی مراکز تصاویر و زوایای دوران یکی از دوربینها و پارامترهای داخلی دو دوربین برآورد می گردد، سپس این مقادیر تقریبی در معادلات شرط هم خطی دوطرفه قرار داده تا مقادیر دقیق پارامترها جهت تولید تصاویر نرمال بدست آید.

۳-۲-۵ - گام پنجم (تولید ویدیوی سه بعدی)

ویدیوهای اخذ شده در سازوکار پیشنهادی بواسطه عدم امکان تنظیم دقیق دوربینها بصورت نسبی، برای استفاده مستقیم در روشهای برجسته بینی مناسب نیستند. نرمال سازی یک نگاشت هندسی با هدف تغییر در منظر ویدیوها است. به عبارت بهتر، در این روند تلاش می شود ویدیوهای جدیدی با محورهای موازی تصویربرداری و مقیاسهای مشابه بازنمونه برداری گردند. با این اتفاق جهت گیری پارالاکس عوارض در سراسر تصاویر نرمال شده همسو شده و خطوط اپی پلار موازی می شوند. برای این هدف، ویدیوهای اخذ شده از دوربینهای استریو به یک صفحه (صفحه ی نرمال) با موقعیت و جهت گیری مشخص بصورت محاسباتی تابانده شده و در قالب شبکه های منظم رستری بازنمونه برداری می شوند.

بعد از تولید ویدیوهای نرمال شده، تمامی اقدامات هندسی لازم به منظور آماده سازی ویدیوها جهت نمایش در سامانه های پخش ویدیوهای سه بعدی به اجرا رسیده است. تنها تنظیم لازم به منظور نمایش ویدیوهای سه بعدی، جابجایی دوبعدی آنها نسبت به یکدیگر جهت حذف دوبینی و تشکیل عمق

خواهد بود. نقاط متناظر شناسایی شده در بخش (۳-۲-۳) امکان این تنظیم را به صورت خودکار فراهم می‌آورند.

یک ویدیوی سه‌بعدی از طریق ارسال مجزای دو ویدیوی نرمال شده و همزمان‌سازی شده به هر چشم، نمایش می‌یابد. نمایش آنالگلیف، نور قطبیده و یا تکنیک‌های مکانیکی راهکارهای توسعه‌یافته به منظور ارسال مجزای ویدیوهای استریو به سامانه بینایی انسان محسوب می‌شوند. در این تحقیق از روش نمایش آنالگلیف به واسطه‌ی سازوکار ساده و اجرای کم‌هزینه‌ی آن جهت نمایش ویدیوهای سه‌بعدی برای تماشاگران استفاده شده است. به همین دلیل، در ادامه ضمن تشریح مختصر این روش، الزامات نمایش باکیفیت ویدیوی سه‌بعدی در این روش تشریح شده است. لازم به ذکر است که محدودیتی در زمینه به‌کارگیری سایر راهکارهای ذکر شده جهت تماشای ویدیوهای وجود ندارد.

در روش نمایش آنالگلیف از فیلترهای رنگی برای ارسال مجزای ویدیوهای استریو استفاده می‌شود. شیشه‌های عینک آنالگلیف از دو رنگ کاملاً متضاد (زاویه‌ی ۹۰ درجه‌ای در مکعب رنگی RGB) ساخته شده که به طور معمول این دو رنگ آبی و قرمز است. هر یک از ویدیوهای استریو به منظور نمایش آنالگلیف با رنگی متناظر با یکی از شیشه‌های عینک نمایش می‌یابد. بدین ترتیب در زمان همپوشانی نمایش دو ویدیو، نمایش ویدیوی غیر مرتبط برای هر چشم از طریق فیلترهای رنگی عینک محدود می‌گردد [۲۰].

در صورت به‌کارگیری فیلترهای رنگی غیر متضاد، احتمال کاهش کیفیت نمایش ویدیوی سه‌بعدی از طریق تداخل بصری ویدیوها وجود دارد. خلوص و تضاد رنگی فیلترهای عینک را می‌توان از طریق روی هم قرار دادن شیشه‌های رنگی تشخیص داد. عدم عبور نور در این شرایط نشان از تضاد کافی در فیلترهای رنگی است. ویدیوهای نمایش یافته نیز می‌بایست از تطابق رنگی با فیلترهای عینک برخوردار باشند. اینکار از طریق شناسایی فام هر یک از فیلترهای رنگی^۱ و تخصیص آن به مولفه‌ی

^۱Hue

رنگ در ویدیوی مرتبط به اجرا می‌رسد. به عبارت بهتر، بعد از تبدیل هر یک از ویدیوها از فضای رنگی RGB به HSI، رنگ مربوط به فیلتر جایگزین مقادیر H شده و سپس ویدیوهای اصلاح شده به فضای RGB باز می‌گردند [۵۵].

فصل چهارم

بررسی نتایج حاصل از پیاده سازی

۴-۱- مقدمه

در این فصل به منظور اعتبار سنجی عملکرد الگوریتم پیشنهادی به منظور تولید ویدیو سه بعدی نسبت به فیلم مرجع مورد مقایسه و ارزیابی قرار گرفته است. در فیلم مرجع فرض بر این است که پارامترهای توجیه داخلی و نسبی دوربین‌ها مفروض بوده و طی یک عملیات میدانی برآورد شده‌اند.

۴-۲- داده‌های مورد استفاده

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

۴-۳- ارزیابی نتایج و بحث

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

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

۴-۳-۱ - رویکرد نخست (صحت‌سنجی فرایند هم‌زمان‌سازی ویدیوها)

پارامترهای داخلی دوربین‌های بکاررفته در هر ویدیو در یک فرایند خودکالیبراسیون و از طریق اخذ تصاویر متعدد از یک میدان آزمون تارگت‌گذاری شده برآورد شده‌اند. برای این منظور از نرم‌افزار

فتوگرامتری بردکوتاه Australis ver 6.02 استفاده شده است. به عبارت دیگر، برای هر دوربین پارامترهای کالیبراسیون داخلی برآورد می‌شود.

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

مطابق با روند متوالی تشریح شد، اولین اقدام از روش پیشنهادی به هم‌زمان‌سازی ویدیوها اختصاص دارد. در روند هم‌زمان‌سازی، پس از برآورد اختلاف زمانی بین ویدیوها از طریق تناظریابی شاخص‌های زمانی، فریم‌های متناظر و به تبع آن ویدیوهای هم‌زمان شده تولید می‌شود. مقادیر عددی اختلاف زمانی ویدیوها معلوم بوده و در اولین ارزیابی، با مقادیر به دست آمده توسط روش پیشنهادی مقایسه شده‌اند. با توجه به انتخاب چند شاخص زمانی متفاوت در روند تناظریابی زمانی (جدول ۳-۱)، در جدول (۴-۱) آماره‌های دقت، شامل دامنه و RMSE بین روش دستی و خودکار گزارش شده است.

با بررسی نتایج ارائه شده در جدول (۴-۱)، موفقیت شاخص آماری انتروپی نسبت به سایر شاخص‌های آماری بکارگرفته 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. شازنوری بدلیل وابستگی به مقیاس تصویر، دقت پایین مقادیر تخمین زده شده در لبه‌ها، اثرپذیری از تغییرات روشنایی و همچنین فقدان محتوای رادیومتریکی مناسب در مرتبه‌ی دوم برآورد تناظریابی زمانی قرار گرفت.

جدول ۴-۱- گزارش از آماره های دقت شاخص های 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

۴-۳-۲- رویکرد دوم (انحراف در برآورد پارامترهای داخلی و نسبی دوربین های استریو)

از منظر تئوری، مقایسه‌ی پارامترهای کالیبراسیون داخلی و نسبی دوربین‌ها در دو حالت برآورد خودکار و دستی می‌تواند معیاری برای اندازه‌گیری تطابق روند تولید ویدیوهای نرمال شده تلقی گردد. بر این اساس، اختلاف پارامترهای کالیبراسیون داخلی و نسبی بدست آمده به روش دستی با مقادیر بدست آمده از روش پیشنهادی به‌عنوان دومین رویکرد ارزیابی نتایج انتخاب شده است. جدول (۲-۴) آماره‌های مربوط به اختلاف در برآورد پارامترهای کالیبراسیون داخلی (x_0, y_0, c) را برای ۲۰ ویدیوی آزمون (۱۰ ویدیوی سه‌بعدی) نشان می‌دهد. اختلافات اندک گزارش شده در جدول (۲-۴) حاکی از موفقیت روش پیشنهادی در کالیبراسیون هم‌زمان داخلی و نسبی ویدیوهای سه‌بعدی است. به‌نظر می‌رسد که تلاش صورت گرفته به‌منظور تامین نقاط متناظر با تنوع در عمق و توزیع مناسب در فضای تصویر و همچنین بکارگیری 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 ویدیوهای 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 (۳-۴) گزارشی آماری از میزان انحراف در برآورد پارامترهای ارتباط نسبی برای ویدیوهای آزمون ارائه شده است. در این جدول علاوه بر انحراف زوایای دورانی برآورد شده برای دوربین‌ها نسبت به مقادیر کالیبره شده بصورت 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 شئی در هنگام تمایز بین پارامترهای خارجی شکل می‌گیرد. لازم به ذکر است که ستون سوم ماتریس دوران سه‌بعدی $R_3(-K) \times R_2(-\varphi) \times R_1(-\omega)$ معادل با جهت‌گیری بردار یکه محور سوم سیستم مختصات تصویری در فضای شئی است. نتایج گزارش شده در جدول (۳-۴) نیز حاکی از اختلافات زاویه‌ای اندک میان پارامترهای ارتباط نسبی در دو حالت دستی و روش پیشنهادی است.

همانطور که در مورد نتایج جدول (۲-۴) 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 (۲-۴) و (۳-۴) ناشی از وجود همبستگی میان پارامترهای مجهول در روند کالیبراسیون داخلی و نسبی است. البته این موضوع مادامیکه اثری در ویژگی‌های مورد انتظار در ویدیوهای 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 ویدیوهای 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 ویدیوهای مختلف، متمایز ساخته است. به همین دلیل، بخش دیگری از انحرافات ایجاد شده در گزارشات آماری جدول (۲-۴) و (۳-۴) را می‌توان متأثر از این شرایط قلمداد نمود. به نظر می‌رسد، با توجه به تنوع مناسب در محتوای ویدیوهای آزمون، می‌توان گزارش‌های آماری ارائه شده در جداول (۲-۴) و (۳-۴) را به‌عنوان محدوده‌ی مورد انتظار از دقت و صحت فرایند کالیبراسیون داخلی و نسبی در روش پیشنهادی در نظر گرفت. لازم به ذکر است که ابعاد متریک پیکسل‌های دوربین‌های بکار گرفته شده در این تحقیق معادل یک میکرون است.

جدول ۲-۴- اختلاف برآورد پارامترهای کالیبراسیون داخلی دستی 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	۰٫۰۰۲۱
c	۰٫۰۲۷۸	۰٫۰۰۳۳	۰٫۰۱۱۳

جدول ۴-۳- اختلاف برآورد پارامترهای ارتباط نسبی دستی و روش پیشنهادی در مجموعه داده های 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

۴-۳-۳ - رویکرد سوم (ثبات در راستای شکل‌گیری پارالاکس عوارض در ویدیوهای سه‌بعدی)

همانطور که در بخش (۳-۲-۴) اشاره شد، یک ویدیوی سه‌بعدی از دو ویدیو 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 متراکم، در تولید ویدیوهای سه‌بعدی، اقدامی به‌منظور هم‌راستا نمودن خطوط اپی‌پلار با جهت‌گیری سطرهای تصویر رقومی در ویدیوهای نرمال‌شده انجام نمی‌شود. علیرغم اهمیت موازی بودن خطوط اپی‌پلار با راستای سطرهای تصاویر بازنمونه‌برداری شده در کاربردهایی همچون تناظریابی‌های متراکم، تامین این شرایط در تولید ویدیوهای سه‌بعدی الزامی نبوده و صرفاً توازی خطوط اپی‌پلار شرایط کافی را به‌منظور برجسته‌بینی

ویدیوهای نرمال شده فراهم می‌آورد. به عبارت بهتر، ترازوی 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 ویدیوهای سه‌بعدی، در روش پیشنهادی از همراستایی خطوط اپی‌پلار با راستای سطرهای تصویر رقومی صرف‌نظر شده است. با این مقدمه، در مرحله سوم از ارزیابی نتایج روش پیشنهادی، کیفیت ویدیوی‌های سه‌بعدی از منظر ترازوی خطوط اپی‌پلار بررسی شده است. برای انجام این کار، به کمک تناظریابی خودکار نقاط تناظری بین فریم‌های چپ و راست ویدیوی سه‌بعدی شناسایی می‌شود. انتظار می‌رود در صورت ترازوی راستای رخداد پارالاکس در ویدیوهای سه‌بعدی، همگی بردارهای اختلاف موقعیت مکانی نقاط متناظر در فریم‌های متناظر چپ و راست موازی بوده و از زاویه‌ی شیب ثابتی برخوردار باشند.

شکل (۴-۱) نمونه‌ای از بردارهای اختلاف مکانی برای نقاط متناظر در یک فریم از ویدیوی سه‌بعدی را به شکل آنالیز نشان می‌دهد.



شکل (۴-۱) - نمایش نقاط متناظر در یک فریم نرمال شده در یک فریم از ویدیوی سه‌بعدی

مطابق با شکل (۴-۱)، اختلاف میان شیب این بردارها با مقدار میانگین در هر ویدیوی سه‌بعدی، کیفیت تولید ویدیوهای نرمال شده را نشان می‌دهد. نقاط متناظر در این مرحله برای هر ویدیوی 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 ویدیو انجام شده است. جدول (۴-۴) گزارشی آماری از میزان انحراف راستای پارالاکس برای نقاط متناظر نسبت به مقدار میانگین را در تمامی ویدیوهای سه‌بعدی نشان می‌دهد.

نتایج گزارش شده در جدول (۴-۴) بدون نیاز به ویدیوهای سه‌بعدی مرجع تولید می‌شوند. به عبارت بهتر، این ارزیابی مستقل از مراجع خارجی تولید شده و به همین دلیل می‌تواند به ازای هر محصول تولیدی از روش پیشنهادی برآورد گردد. ناهمسو بودن راستای پارالاکس‌های در ویدیوهای سه‌بعدی می‌تواند متأثر از بروز خطا و اختلال در هر یک از گام‌های تولید ویدیوی سه‌بعدی باشد. بر این اساس، این ارزیابی محصول محور می‌تواند به عنوان 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 ویدیوهای 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.

جدول ۴-۴- گزارش آماری انحراف در راستای پارالاکس نقاط متناظر در مجموعه داده نمونه

شماره ویدیوی سه‌بعدی	دامنه (درجه)	انحراف معیار (درجه)
ویدیوی ۱	۱٫۳۴۷۰	۰٫۴۵۲
ویدیوی ۲	۱٫۴۲۷	۰٫۵۱۵
ویدیوی ۳	۱٫۹۳۱	۰٫۳۶۱
ویدیوی ۴	۱٫۲۴۱	۰٫۵۲۷
ویدیوی ۵	۱٫۴۷۳	۰٫۵۰۴۹
ویدیوی ۶	۱٫۴۳۸	۰٫۴۸۱۵
ویدیوی ۷	۱٫۹۶۳	۰٫۴۵۳
ویدیوی ۸	۱٫۹۳۱	۰٫۴۸۱۵
ویدیوی ۹	۲٫۱۴۹۷	۰٫۵۹۶۲
ویدیوی ۱۰	۱٫۳۴۷۰	۰٫۴۲۹

۴-۳-۴ - رویکرد چهارم (ارزیابی کیفی حاصل از تماشای ویدیوهای سه‌بعدی)

در مرحله‌ی چهارم ارزیابی از چند مخاطب متخصص و غیرمتخصص درخواست شد تا تجربیات خود را از تماشای ویدیوهای سه‌بعدی بصورت آنالیز گزارش کنند. هیچ‌یک از بینندگان گزارشی مبنی بر عدم توانایی تشخیص عمق را برای کل ویدیو گزارش نکردند. بطور عمومی درک عمق برای همه‌ی ویدیوها میسر بوده اما در لحظاتی و در برخی از ویدیوها تشخیص عمق مختل شده و دوبینی اتفاق افتاده است. در واکاوی این رخداد بررسی‌های مختلفی از روند اجرای الگوریتم پیشنهادی و ویدیوهای اولیه صورت گرفت. عوامل مختلفی در بروز اختلال در روند تولید ویدیوهای سه‌بعدی وجود داشته که عمدتاً ناشی از محدودیت‌های فنی و سخت‌افزاری است.

فصل پنجم

نتیجه گیری و پیشنهادات

امروزه ظهور فن‌آوری‌های نوین در حوزه‌ی تولید فیلم‌های سه‌بعدی سطح انتظار مخاطبان به این صنعت را بیش از پیش افزایش داده است. در دنیای امروز، فیلم‌های سینمایی، بازی‌های رایانه‌ای، مستندنگاری آثار تاریخی و ساخت تورهای مجازی همگی به سمت ایجاد شرایط دید سه‌بعدی سوق داده شده‌اند.

در این تحقیق، توسعه‌ی یک سامانه‌ی آماتور فیلم‌برداری استریو با هدف تولید فیلم‌های سه‌بعدی در دستورکار قرار گرفت که روند آماده‌سازی و خودکارسازی تولید فیلم‌سه‌بعدی محور روش پیشنهاد شده در آن محسوب می‌شود. با توسعه‌ی روش پیشنهادی بخش‌هایی از پیش‌پردازش‌های لازم به‌منظور تولید فیلم‌های سه‌بعدی بصورت خودکار به انجام رسید. اینکار سهولت استفاده از چنین سامانه‌هایی برای کاربران عادی افزایش داده است.

روش پیشنهادی از چهار گام متوالی:

۱- تولید ویدیوهای هم‌زمان‌سازی شده، ۲- شناسایی نقاط متناظر مناسب به‌منظور کالیبراسیون داخلی و نسبی، ۳- تخمین پارامترهای کالیبراسیون داخلی و نسبی دوربین‌های استریو و ۴- تولید ویدیوهای نرمال شده و سه‌بعدی، تشکیل شده است.

تعمیم‌پذیری این روش در مورد چندین ویدیو ارزیابی شده که نتایج حاکی از موفقیت روش پیشنهادی در تمامی ویدیوها بود.

دوربین‌های معمولی و غیرحرفه‌ای از سیستم‌های تصویربرداری مبتنی بر شاتر جاروب‌شونده استفاده می‌کنند. در این سیستم تصویربرداری بر خلاف سیستم‌های تصویربرداری با شاتر سراسری، اخذ هر فریم از بصورت 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

¹Global Shutter

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 رخداد پارالاکس عوارض می‌گردد. توسعه‌ی راهکارهای پردازشی به‌منظور حذف اثر شاتر جاروب‌شونده در فریم‌های توأم با حرکت و یا بکارگیری دوربین‌های با شاتر سراسری می‌تواند به‌عنوان روش‌های رفع این چالش شناخته شوند. در شرایط فعلی توصیه می‌شود که در زمان اخذ ویدیوهای استریو، سرعت حرکت نسبی دوربین و صحنه تا حد ممکن پایین باشد. این موضوع می‌تواند بروز اعوجاجات هندسی ناشی از شاتر جاروب‌شونده تا حد زیادی کاهش دهد.

از سوی دیگر، محدودیت‌های سخت‌افزاری گاهی منجر به بروز انقطاع در روند متوالی اخذ فریم‌های یک ویدیو می‌گردد. در این شرایط بیننده شاهد وقوع سکت در نمایش ویدیو 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 بصورت مقطعی ویدیوهای سه‌بعدی را تخریب نماید. با این حال نتایج تولید ویدیوهای سه‌بعدی در این تحقیق حاکی از آن است که علیرغم بروز چنین خطاهایی در روند اخذ ویدیوهای اولیه، این مشکلات نتوانسته شکله‌ی اصلی ویدیوهای سه‌بعدی را دچار اختلال نماید. به‌عبارت 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.

با این حال چالش‌هایی همچون بروز خطاهای سیستمی ناشی از نقصان‌های سخت‌افزاری دوربین‌ها و اثرات شاتر جاروب‌شونده هنگام بروز سرعت نسبی بالا بین دوربین و صحنه می‌تواند نتایج تولید شده در این روش را بصورت مقطعی مخدوش سازد. توسعه‌ی راهکارهای پردازشی به‌منظور جبران اثرات

شاتر جاروب شونده و همچنین تنظیم خودکار مقیاس ویدیوی سه‌بعدی متناسب با عمق لحظه‌ای میدان از مواردی است که در این حوزه می‌تواند به‌عنوان افق تحقیقاتی پیش‌رو مطرح باشد.

منابع

[1] Bernard Mendiburu, 2012, 3D Movie Making Stereoscopic Digital Cinema from Script to Screen

- [2] Kapp, K.M. and T. O'Driscoll. Learning in 3D: Adding a new dimension to enterprise learning and collaboration. John Wiley & Sons,2009.
- [3] Lambooi, Marc, Wijnand A. IJsselsteijn, and Ingrid Heynderickx. "Visual discomfort of 3D TV: Assessment methods and modeling." *Displays* 32.4 (2011).
- [4] Zone, R., Stereoscopic cinema and the origins of 3-D film, University Press of Kentucky, 2014.
- [5] Shweta Patil, Prof. Priya Charles.2015. Review on 2D to 3D Image and Video Conversion Methods, International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue 11.
- [6] Sergi Pujades Rocamora. Camera Models and algorithms for 3D video content creation,Ph.D. Thesis-Universityof GRENOBLE,2016.
- [7] Zhang, Licong, et al. "Real-time human motion tracking using multiple depth cameras." 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2012.
- [8] Dou, Pengfei, and Ioannis A. Kakadiaris. "Multi-view 3D face reconstruction with deep recurrent neural networks." *Image and Vision Computing* 80 (2018)
- [9] Emoto, Michiko, Shinya Miyamoto, and Kazuyoshi Yamamoto. "Navigation apparatuses, methods, and programs for generation of a 3D movie." U.S. Patent No. 7,974,781. 5 Jul. 2011.
- [10] Roberts, Steve. Character animation fundamentals: developing skills for 2D and 3D character animation. CRC Press, 2012.
- [11] Garrido, Pablo, et al. "Reconstructing detailed dynamic face geometry from monocular video." *ACM Trans. Graph.* 32.6 (2013).
- [12] 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.
- [13] Passmore, Charles. "Director-style based 2d to 3d movie conversion system and method." U.S. Patent 8,730,232, issued May 20, 2014.
- [14] Lakis Christodoulou,3D Stereo Vision Camera-sensors, Advancements, and Technologies, Electrical & Computer Engineering & Computer Science, Cyprus University of Technology,2013.
- [15] 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.
- [16] Vijayanarasimhan, Sudheendra, et al. "Sfm-net: Learning of structure and motion from video." arXiv preprint arXiv:1704.07804 (2017).
- [17] Heinzle, Simon, et al. "Computational stereo camera system with programmable control loop." *ACM Transactions on Graphics (TOG)* 30.4 (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 (2010): 946-950.
- [20] Kim, Jonghyun, et al. "A full-color anaglyph three-dimensional display system using active color filter glasses." *Journal of Information Display* 12.1 (2011): 37-41.
- [21] May, Paul. "A survey of 3-D display technologies." *Information Display* 32 (2005): 28-33.
- [22] Yang, Lin, et al. "See in 3D: state of the art of 3D display technologies." *Multimedia Tools and Applications* 75.24 (2016): 17121-17155.
- [23] Michel Morgan, *Epipolar Resampling of Linear Array Scanner Scenes*, Ph.D. Dissertation, UNIVERSITY OF CALGARY, Canada, 2004.
- [24] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." *European conference on computer vision*. Springer, Berlin, Heidelberg, 2006.
- [25] Zhao, Wan-jin, et al. "Adaptive Harris corner detection algorithm." *Computer Engineering* 10.5 (2008): 212-215.
- [26] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60.2 (2004): 91-110.
- [27] Wang, X., Zhang, H. and Liu, SH. 2013. Reliable RANSAC Using a Novel Preprocessing Model. Xiaoyan Wang et al. An open access article distributed under the Creative Commons Attribution License.
- [28] Banks, Martin S., et al. "Stereoscopy and the human visual system." *SMPTE motion imaging journal* 121.4 (2012): 24-43.
- [29] Chai, Tianfeng, and Roland R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)-Arguments against avoiding RMSE in the literature." *Geoscientific model development* 7.3 (2014): 1247-1250.
- [30] Coe, Brian. "William Friese Greene and the origins of cinematography." *Screen* 10.2 (1969): 25-41.
- [31] Zhihan LU, Shafiq ur Rehman, Muhammad Sikandar Lal Khan, Haibo Li; 2013, Anaglyph 3D Stereoscopic Visualization of 2D Video based on Fundamental Matrix, International Conference on Virtual Reality and Visualization.
- [32] Aleksandr Ploshkin and Dmitriy Vatolin. 2018. Accurate Method of Temporal-Shift Estimation for 3D Video, IEEE Conference on 3D-TV.
- [33] 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.

- [34] Goldmann, Lutz, Jong-Seok Lee, and Touradj Ebrahimi. "Temporal synchronization in stereoscopic video: Influence on quality of experience and automatic asynchrony detection." *2010 IEEE International Conference on Image Processing*. IEEE, 2010.
- [35] Merkle, Philipp, Karsten Müller, and Thomas Wiegand. "3D video: acquisition, coding, and display." *IEEE Transactions on Consumer Electronics* 56.2, pp. 946-950, 2010.
- [36] 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.
- [37] Negahdaripour, Shahriar. "Analyzing epipolar geometry of 2-D forward-scan sonar stereo for matching and 3-D reconstruction." *OCEANS MTS/IEEE Charleston*. IEEE, 2018.
- [38] Meyer, Benjamin, et al. "Subframe Temporal Alignment of Non-Stationary Cameras." *BMVC*. 2008.
- [39] 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.
- [40] 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.
- [41] 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.
- [42] 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.
- [43] Zone, R. *Stereoscopic cinema and the origins of 3-D film, 1838-1952*: University Press of Kentucky, 2014.
- [44] Ng CJ, Blake R, Banks MS, Tadin D, Yoon G. Seeing the World Like Never Before: Human stereovision through perfect optics. *bioRxiv*. 2021.
- [45] Holliman, Nick. "Cosmic origins: experiences making a stereoscopic 3D movie." *Stereoscopic Displays and Applications XXI*. Vol. 7524. International Society for Optics and Photonics, 2010.
- [46] Nahler, Gerhard. "Pearson correlation coefficient." *Dictionary of Pharmaceutical Medicine*. Springer, Vienna, 2009.
- [47] Wen, K. L., Chang, T. C., & You, M. L. (1998). The grey entropy and its application in weighting analysis. *IEEE International Conference on Systems, Man, and Cybernetics*, 2, 1842-1844.

- [48] Horn, Berthold KP, and Brian G. Schunck. "Determining optical flow." *Artificial intelligence* 17.1-3 (1981): 185-203.
- [49] Seglen, Per O. "The skewness of science." *Journal of the American society for information science* 43.9 (1992): 628-638.
- [50] Baker, Simon, et al. "Removing rolling shutter wobble." *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010.
- [51] Wolf PR. *Elements of photogrammetry; with air photointerpretation and remote sensing*. 1983.
- [52] Brown, D.C. *The Bundle Adjustment - progress and prospects*, IAPRS, Vol.XIII, No. 3(1979).
- [53] Mikhail, Edward M., James S. Bethel, and J. Chris McGlone. "Introduction to modern photogrammetry." *New York* 19 (2001).
- [54] Yuna, Wang Zhizhong Guo Xingcui Zhou. "Quasi Observation Approach of Nonlinear Semi-parametric Model." *Mathematical Theory and Applications* 2 (2007).
- [55] Govindasamy, Saravanan, and Yamuna Govindarajan. "Real-time implementation of various colour space models." *International Journal of Circuits and Architecture Design* 2.3-4 (2016): 258-271.

A method for automatic temporal and relative calibration of the amateur cameras to produce 3D videos

Abstract

In this paper, a novel method has been proposed to automatically produce 3D videos through amateur digital cameras that have been fixed with each other. Asynchrony of the videos (frames per second rates and start times), unavailability of the exact camera internal parameters, and the technical limitation of precise relative adjusting the stereo cameras could be known as the main challenges of producing 3D videos via the amateur cameras. In the proposed method, the videos acquired by the stereo camera are synchronized through the automatic matching between temporal indices. Then, the calm periods of videos (the times with zero relative velocity between camera and scenes) are detected to be used for selecting proper matched image frames. Proper matched frames are then applied for finding matched points via a geometrical constrained feature-based matching method. The matched points are used for self-calibration as well as relative parameters estimation of the stereo cameras. In the last step, epipolar resampling procedure has been used to generate normalized videos. Automatic and precise synchronization of the stereo videos as well as the proper generalization of the proposed approach in the different sample datasets has been seen in the evaluation processes. Spectator satisfaction in the depth perception of the 3D videos is another quality achievement of the proposed method.

Keywords: 3D Video, Entropy, Optical Flow, Automatic Synchronization, Normalized Images, Relative Orientation



Ministry of Science, Research and Technology Tafresh University

Tafresh University

Department of Geodesy and Surveying Engineering

MSc Thesis

Title

**A method for automatic temporal and relative calibration of the amateur cameras to
produce 3D videos**

By Atiyeh Ganjali

Supervisors:

Dr. Alireza Safdarinezhad