

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ



دانشگاه صنعتی شاہروود

دانشکده برق و رباتیک

گروه رباتیک

عنوان :

طراحی و پیاده‌سازی الگوریتم بلاذرنگ تطابق استریو در سطح
ویژگی به منظور بهبود توانایی ربات در تحلیل صحنه‌های واقعی

دانشجو :

محمد رضا غریب سامی

استاد راهنما :

دکتر حسین خسروی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

زمستان ۱۳۹۳

دانشگاه صنعتی شهرود

دانشکده: برق و رباتیک

گروه: رباتیک

پایان نامه کارشناسی ارشد آقای محمدرضا غریب سامی

تحت عنوان:

طراحی و پیاده‌سازی الگوریتم بلاذرنگ تطابق استریو در سطح ویژگی به منظور بهبود توانایی ربات در
تحلیل صحنه‌های واقعی

در تاریخ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و
با درجه مورد پذیرش قرار گرفت.

امضاء	استاد راهنما
	دکتر حسین خسروی

امضاء	نماينده تحصيلات تمكيلي	امضاء	اساتيد داور
	دکتر مسعود مهدی زاده رخی		دکتر علیرضا احمدی فرد
			دکتر هادی گرایلو

تقدیم به

پدر و مادر مهربانم

و تمام کسانی که دوستشان دارم

ستایش از آن اوست

شکر و سپاس ایزد منان را که به من توفیق کسب علم و شناخت گوشه‌ای از شگفتی‌های
عالم هستی را عطا فرمود.

برخود لازم می‌دانم که از استاد محترم، جناب دکتر خسروی به پاس زحمات و راهنمایی‌هایی که در
تدوین و پیشبرد این تحقیق داشتند، تقدیر و تشکر نمایم. از دیگر اساتید گروه مهندسی برق و رباتیک
دانشگاه صنعتی شاهروд که طی این دوره راهنمای من بوده‌اند، به خاطر تمامی زحماتشان سپاس‌گزارم.

محمد رضا غریب سامی

تعهد نامه

اینجانب دانشجوی دوره کارشناسی ارشد رشته دانشکده

..... دانشگاه صنعتی شاهرود نویسنده پایان نامه

تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.

در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.

مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی

در هیچ جا ارائه نشده است.

کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام «

دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ

خواهد رسید.

حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات

مستخرج از پایان نامه رعایت می گردد.

در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده

است ضوابط و اصول اخلاقی رعایت شده است.

در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا

استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.

استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

روند بیرون کشیدن اطلاعات عمقی از تصاویر دوبعدی دوتایی که از صحنه‌ای واحد برداشته شده‌اند تحت عنوان بینایی استریو شناخته می‌شود. جابجایی مکانی، اختلاف موقعیت دو پیکسل متناظر از تصویرهای چپ و راست صحنه، با در نظر گرفتن یکی از تصاویر به عنوان مرجع می‌باشد. در چند دهه گذشته شاهد توسعه صدها الگوریتم بینایی استریو مختلف بوده‌ایم، این امر طبقه‌بندی و مقایسه رویکردهای مختلف نسبت به این مسئله را مشکل ساخته است. در این تحقیق ما یک مرور اجمالی از انواع رویکردهایی ارائه می‌کنیم که برای حل مسئله بینایی استریو وجود دارند. به دلیل ارائه یک روش بلادرنگ در استخراج نقشه جابجایی مکانی، از میان الگوریتم‌های استریوی محلی و سراسری، روش محلی را استفاده نمودیم. در آزمایشات انجام شده طی سال‌های اخیر، مهم‌ترین بخش از الگوریتم‌های استریو محلی، شیوه جمع‌بندی هزینه‌های تطبیق آن بیان شده است.

در روش موردنظر ما، ابتدا یک الگوریتم استریو محلی بر مبنای ویژگی با استفاده از الگوریتم گوشیابی هریس معرفی می‌شود و سپس با استفاده از شیوه‌ی جمع‌بندی هزینه‌های تطبیق پیشنهادی در پنجره‌ی اطراف پیکسل مورد نظر، روشی برای بازیابی نقشه جابجایی مکانی گستته پیشنهاد می‌گردد که نسبت به روش‌های پیشین دارای دقت عملکرد بهتری می‌باشد. سپس یک الگوریتم استریوی محلی بر مبنای پیکسل، برای بازیابی نقشه جابجایی مکانی متراکم با استفاده از تصحیح گاما در فضای رنگ RGB بر روی جفت تصاویر ورودی، در مرحله پیش‌پردازش و فیلتر دوطرفه در مرحله پس‌پردازش و استفاده از شیوه جمع‌بندی هزینه‌های تطبیق به دست آمده از حالت گستته، ارائه شده است که دارای میانگین خطای ۱.۵٪ در پایگاه داده معرفی شده در مقالات می‌باشد و دارای سرعت مطلوبی است.

کلمات کلیدی- بینایی استریو، نقشه جابجایی مکانی، گوشیابی هریس، تصحیح گاما، فیلتر دوطرفه

فهرست مطالب

صفحه

عنوان

فصل اول

شناخت مسئله

۲	۱-۱) مقدمه و پیشینه
۴	۲-۱) تعریف اصطلاحات مورد استفاده
۱۰	۳-۱) اصول بنیادین بینایی استریو
۱۰	۱-۳-۱) بینایی استریو چیست؟
۱۰	۲-۳-۱) بینایی استریو انسان‌ها در مقابل دستگاهها
۱۱	۱-۳-۲) نقشه جابجایی مکانی
۱۲	۴-۳-۱) مشکلاتی که الگوریتم‌های استریو با آن مواجه هستند
۱۵	۴-۳-۲) مقدمه‌ای بر تحقیق‌های مشابه
۱۶	۱-۵) نتیجه‌گیری

فصل دوم

مروارید الگوریتم‌ها

۲۰	۱-۲) مقدمه
۲۱	۲-۲) الگوریتم‌های تطبیق استریو محلی
۲۲	۳-۲) روش‌های محاسبه هزینه
۲۳	۱-۳-۲) مجموع اختلاف‌های مربعی (SSD)
۲۴	۲-۳-۲) مجموع اختلاف‌های مطلق (SAD)
۲۴	۳-۳-۲) همبستگی نرمالیزه (NCC)
۲۵	۴-۳-۲) تبدیل رتبه و آمارگیری

۲۶ ۵-۳-۲) کاربرد اطلاعات رنگ
۲۷ ۴-۲) روش‌های تجمع هزینه
۲۷ ۱-۴-۲) روش تجمع هزینه پنجره ثابت
۲۸ ۲-۴-۲) روش تجمع هزینه وزن انطباقی
۳۰ ۳-۴-۲) تجمع هزینه بر مبنای بخش‌بندی
۳۲ ۴-۴-۲) تجمع هزینه پشتیبان بخش‌بندی
۳۲ ۵-۴-۲) اختلاف میان این رویکردهای تجمعی
۳۴ ۵) محاسبه اختلاف و بهینه‌سازی
۳۴ ۶-۲) روش‌های اصلاح
۳۵ ۱-۶-۲) فیلتر میانه برای نقشه‌های جابجایی مکانی
۳۶ ۲-۶-۲) تطبیق دو-جهته
۳۸ ۷-۲) مرور اجمالی الگوریتم‌های تطبیق استریو سراسری
۴۰ ۸-۲) قیود مورد استفاده در تطابق استریو
۴۲ ۹-۲) ارزیابی استریو میدلبوری
۴۴ ۱۰-۲) نتیجه‌گیری

فصل سوم

الگوریتم پیشنهادی

۴۶ ۱-۳) مقدمه
۴۷ ۲-۳) الگوریتم تطابق استریو با استفاده از گوشه‌یابی
۵۰ ۳-۳) الگوریتم استخراج نقشه جابجایی مکانی متراکم
۵۰ ۱-۳-۳) مرحله اول
۵۲ ۱-۱-۳-۳) تصحیح گاما

۵۲ مرحله دوم ۳-۳-۲
۵۳ ۳-۲-۱) محاسبه هزینه
۵۴ ۳-۳-۳) مرحله سوم
۵۴ ۳-۳-۱) فیلتر دوطرفه بازگشتی

فصل چهارم

طرح آزمایش

۶۲ مقدمه ۴-۱)
۶۲ ۴-۲) مجموعه داده تصویر استریو میدل بوری
۶۳ ۴-۳) تجهیزات مورد استفاده ما در پیاده‌سازی الگوریتم
۶۴ ۴-۴) ارزیابی الگوریتم پیشنهادی تطابق در سطح ویژگی
۶۶ ۴-۵) معیارهای کیفیتی تطبیق استریو
۶۷ ۴-۶) ارزیابی نقشه‌های جابجایی مکانی متراکم
۶۸ ۴-۷) سایر نتایج
۶۹ ۴-۸) نتیجه‌گیری

فصل پنجم

نتیجه‌گیری و پیشنهادات

۷۴ ۵-۱) نتیجه‌گیری
۷۶ ۵-۲) پیشنهاد برای تحقیقات آینده
۷۷ مراجع

فهرست شکل‌ها

<u>صفحه</u>	<u>عنوان</u>
۵	شکل ۱-۱: هندسه تشکیل تصویر با صفحه تصویر در جلو
۶	شکل ۱-۲: هندسه تصویربرداری با دو دوربین در حالت کلی
۷	شکل ۱-۳: هندسه تصویربرداری استریوی دو دوربینه با صفحات تصویر موازی
۸	شکل ۱-۴: انسداد در دو حالت مختلف
۱۱	شکل ۱-۵: نمونه یکسوزده تصویر استریو جفت
۱۲	شکل ۱-۶: نقشه جابجایی مکانی متراکم
۱۳	شکل ۱-۷: انسداد نمونه در یک جفت تصویر استریو
۱۵	شکل ۱-۸: یک تصویر پارازیتی گرفته شده با دوربین پاناسونیک LX1
۲۲	شکل ۲-۱: (الف) مثالی از سایز پنجره و (ب) مثالی از پنجره‌های جستجو
۲۳	شکل ۲-۲: فضای جستجو تطبیق استریو
۲۶	شکل ۲-۳: تبدیل رتبه (چپ) و تبدیل آمارگیری (راست)
۲۸	شکل ۲-۴: نمونه پنجره مربعی ثابت با سایز ۵
۲۹	شکل ۲-۵: نمونه پنجره تطبیقی وزن انطباقی
۳۱	شکل ۲-۶: مثال از یک پنجره تجمع پویا با بخش‌بندی رنگ
۳۳	شکل ۲-۷: مناطق پشتیبانی وزن برای روش‌های تجمع هزینه مختلف
۳۵	شکل ۲-۸: محاسبه مقدار متوسط برای یک پیکسل
۳۵	شکل ۲-۹: دسته‌بندی شدت پیکسل برای تعیین میانه
۳۶	شکل ۲-۱۰: نقشه جابجایی مکانی نویزی، قبل و بعد از استفاده از یک فیلتر میانه ۵×۵

شکل ۱۱-۲: (الف) برقراری محدودیت ترتیب و (ب) عدم برقراری محدودیت ترتیب ۴۱	۴۱
شکل ۱۲-۲: عدم برقراری قید یکتایی در دو حالت مختلف ۴۲	۴۲
شکل ۱۳-۲: هشت الگوریتم برتر توسط بستر ارزیابی میدلبوری ۴۳	۴۳
شکل ۱-۳: تصویر Tsukuba : الف- تصویر چپ، ب- تصویر راست ۴۷	۴۷
شکل ۳-۲: ویژگی‌های استخراج شده در تصویر Tsukuba ۴۹	۴۹
شکل ۳-۳: تطبیق ویژگی‌های استخراج شده در تصویر Tsukuba ۵۰	۵۰
شکل ۳-۴: بلوك دياگرام يافتن تطابق ویژگی‌ها در جفت تصاویر استرييو ۵۱	۵۱
شکل ۳-۵: تصویر Tsukuba : الف- قبل از تصحیح گاما، ب- بعد از تصحیح گاما ۵۲	۵۲
شکل ۳-۶: تصویر جابجایی مکانی استخراج شده از الگوریتم، قبل از اعمال فیلتر دوطرفه ۵۴	۵۴
شکل ۳-۷: تصویر خروجی نهایی مربوط به Tsukuba ۵۸	۵۸
شکل ۳-۸: بلوك دياگرام استخراج نقشه جابجایی مکانی متراکم ۵۹	۵۹
شکل ۴-۱: مجموعه داده میدلبوری ۶۳	۶۳
شکل ۴-۲: تطبیق ویژگی‌های جفت تصاویر استرييو ۶۵	۶۵
شکل ۴-۳: تولید نقشه‌های جابجایی مکانی با استفاده از روش پنجره ثابت ۶۸	۶۸
شکل ۴-۴: تولید نقشه‌های جابجایی مکانی با استفاده از روش وزن- انطباقی ۶۸	۶۸
شکل ۴-۵: تولید نقشه‌های جابجایی مکانی با استفاده از روش بخش‌بندی تصویر ۶۹	۶۹
شکل ۴-۶: نقشه جابجایی مکانی استخراج شده به وسیله روش پیشنهادی برای تصاویر ۶۹	۶۹
شکل ۴-۷: مقایسه نقشه خطای اختلاف مطلق پیکسل‌های غیرانسدادی در نقشه‌های صحیح(ردیف بالا) و روش پیشنهادی(ردیف پایین) ۷۰	۷۰

فهرست جداول

<u>عنوان</u>	<u>صفحه</u>
جدول ۴-۱: مقایسه الگوریتم‌ها ۶۴	
جدول ۴-۲: محدوده تغییرات جابجایی مکانی و ضریب تبدیل مناسب برای تصاویر استریو ۶۷	
جدول ۴-۳: مقایسه بین روش پیشنهادی و الگوریتم‌های مبتنی بر روش‌های محلی ۷۰	
جدول ۴-۴: مقایسه بین روش پیشنهادی و الگوریتم‌های مبتنی بر روش‌های سراسری ۷۱	

فصل اول

شناخت مسئله

۱-۱ مقدمه و پیشینه

از قرن‌ها پیش مکانیزم سیستم بینایی انسان در درک سه بعدی محیط مورد توجه بوده است. تا اوایل قرن ۱۷ این تئوری پذیرفته شده بود که از چشم‌های انسان پرتوهایی در محیط ساطع می‌شود. در سال ۱۶۲۵ شاینر به صورت تجربی، تشکیل تصاویر در پشت کره چشم را مشاهده نمود. کشفیات جدید این سوال را که "ادراك عمق چگونه صورت می‌گیرد"، برجسته‌تر نمود. چگونه با تصاویر تشکیل شده بر روی شبکیه چشم، انسان به درک سه بعدی از محیط دست می‌یابد؟ و بالاخره با پیشرفت کامپیوترها در قرن بیستم سؤال جامع‌تری مطرح گردید: چگونه با ابزار بینایی می‌توان به استخراج عمق محیط پرداخت؟ پاسخ به این سؤالات علاوه بر حل نسبی معماه بینایی انسان، پیامدهای عملی مفید بسیاری به دنبال داشت که از آن جمله می‌توان اتوماسیون صنعتی، ناوبری خودکار رباتها و تعامل انسان با ماشین را ذکر نمود [۱].

دهه گذشته شاهد پیشرفت‌های فناوری در حوزه تصویربرداری دیجیتال بوده و این باعث شده تا تجهیزات تصویربرداری دیجیتال با کیفیت خوب عرضه شود. یکی از ابعاد دید کامپیوتری هم که به طور گسترده مورد تحقیق قرار گرفته، توانایی بیرون‌کشیدن اطلاعات عمقی از تصاویر دو بعدی جهت بازسازی صحنه‌های سه بعدی است. در تحقیقات این موضوع را با اصطلاحات متراffد بسیاری نامیده‌اند، که برخی از آنها عبارتند از: بینایی استریو، نقشه‌برداری عمق استریو، بازسازی سطح و برآورد سطح. بینایی استریو کاربردهای متنوعی دارد که برخی از آن‌ها عبارتند از:

- هوش مصنوعی (ربات‌های) سیار: نقشه‌برداری و شناسایی موانع [۲]؛

- ردیابی شیء؛ و

- مدل‌سازی از محیط‌های سه بعدی؛

بازسازی شکل‌های سه بعدی از تصاویر دیجیتال دو بعدی معمولاً با استفاده از تصاویر استریو صورت می‌گیرد، این تصاویر از یک صحنه واحد هستند، ولی از زوایای متفاوتی گرفته شده‌اند. الگوریتم‌های تطابق

استریو بر روی جفت تصویر به کار می‌روند تا اطلاعات عمقی را از این تصاویر بیرون بکشند، که عموماً خروجی این الگوریتم‌ها، نقشه جابجایی مکانی یا اختلاف^۱ نام دارد [۳]. یک نقشه جابجایی مکانی معمولاً یک تصویر خاکستری است که در آن شدت پیکسل‌ها نشانگر یک ارزش فاصله‌ای بوده، هرچه پیکسل‌ها روشن‌تر باشند، نقطه مربوطه عمق کمتری خواهد داشت.

اصل به کار رفته برای تعیین عمق یک نقطه در یک تصویر تعیین جابجایی مکانی آن نقطه از تصویر در مقایسه با تصویر دیگر است. هرچه یک تصویر به دوربین‌های استریو نزدیک‌تر باشد، این جابجایی بیشتر خواهد بود. این اصل به حد کافی ساده است؛ مشکل اصلی آن است که برای تعیین این جابجایی، یک سیستم باید بتواند یک نقطه خاص از یک تصویر را با یک نقطه در تصویر دوم تطبیق دهد. این مسئله در بسیاری از الگوریتم‌هایی که تاکنون در تحقیقات توصیف شده‌اند، به طرق مختلفی حل شده است. تمامی الگوریتم‌های شناخته شده برای تطابق استریو را می‌توان در دو دسته اساسی دسته‌بندی نمود، الگوریتم‌های ویژگی- مبنا^۲ و الگوریتم‌های مساحت- مبنا^۳.

الگوریتم‌های تطابق ویژگی- مبنا، موقعیت ویژگی‌های خاص در تصاویر را در نظر می‌گیرند، همچون لبه‌ها، مقطع‌ها، خطوط کناره‌نما و سپس نتایج را با ویژگی‌های مربوطه در تصویر دوم از جفت استریو، تطبیق می‌دهند. بیرون‌کشیدن چنین ویژگی‌هایی از لحاظ محاسباتی پرهزینه است؛ البته، مرحله تطبیق‌دهی چنین الگوریتم‌هایی عموماً سریع است چرا که نه تمامی پیکسل‌ها، بلکه تنها شکل‌های مشخصی در تصاویر را شامل می‌شوند. الگوریتم‌های تطابق استریو شاید نتوانند از پس صحنه‌هایی با سطوح درخشش متغیر یا صحنه‌هایی که قادر یک تعداد کافی از مشخصه‌های قابل استخراج هستند، برآیند.

الگوریتم‌های تطابق استریو مساحت- مبنا براساس این فرضیه عمل می‌کنند که دو تصویر در یک جفت استریو از لحاظ موقعیتی شبیه یکدیگر هستند و بنابراین می‌توان همبستگی قوی میان نواحی

۱- Disparity Map

۲- Feature-based Algorithms

۳- Area-based Algorithms

کوچک در تصاویر را یافت. چنین الگوریتم‌هایی به ویژه برای صحنه‌های دارای بافت مؤثر هستند، اما عموماً نمی‌توان از آنها برای حدود انسدادی و نواحی وسیع بدون بافت استفاده کرد. هر دو دسته از الگوریتم‌ها اغلب اوقات با فنون ویژه‌ای مانند تطبیق دو-جهته^۱، یافتن نقاط انسداد به کار رفته‌اند تا عملکرد تطبیق‌دهی بهبود یابد [۲].

با این فنون می‌توان نقشه‌های جابجایی مکانی از صحنه‌های سه‌بعدی تهیه کرد. رویکردهای مختلف نسبت به تطبیق استریو تحت شرایط محیطی مختلف همچون نورپردازی حساس هستند. الگوریتم‌ها معمولاً ترکیبی از روش‌های مختلف را جهت تولید نقشه‌های جابجایی مکانی دقیق تحت شرایط مختلف و حل مسائلی چون نواحی بدون بافت و نقاط انسداد به کار می‌گیرند.

۱-۲ تعریف اصطلاحات مورد استفاده

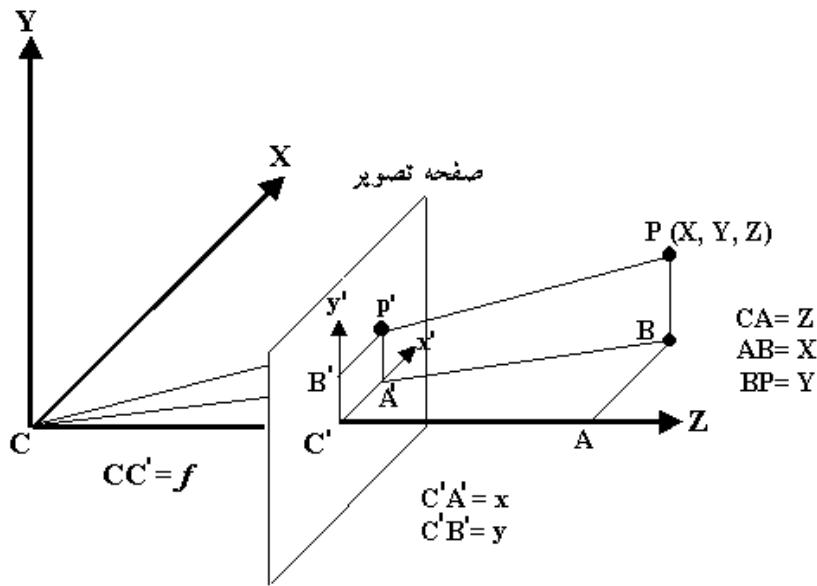
در این بخش، برخی از اصطلاحاتی که در این نوشتار استفاده می‌شوند، با جزئیات لازم جهت وضوح مطالب، تعریف می‌شوند [۱,۴]:

فاصله کانونی^۲: در یک سیستم تصویربرداری مبتنی بر لنز، مجموعه لنز که متشکل از چندین عدسی محدب و مقعر است با یک عدسی محدب به فاصله کانونی f ، مدل می‌شود.
صفحه تصویر^۳: محل تشکیل تصویر که آن را صفحه تصویر می‌گویند، صفحه‌ای عمود بر محور عدسی و به فاصله f در پشت آن می‌باشد.

روابط پرسپکتیو: اگر P نقطه‌ای در فضای سه بعدی باشد، $'P$ یعنی تصویر P در صفحه تصویر، از برخورد CP با صفحه تصویر بدست می‌آید. اگر (X,Y,Z) مختصات P در فضای سه بعدی و (x,y) مختصات $'P$ در صفحه تصویر باشد، با توجه به شکل ۱-۱ می‌توان رابطه (۱-۱) را نوشت:

۱- Bi-Directional
۲- Focus Center
۳- Image Plane

$$x = f \frac{X}{Z}, y = f \frac{Y}{Z} \quad (1-1)$$



شکل ۱-۱: هندسه تشکیل تصویر با صفحه تصویر در جلو

استریوی دو دوربینه^۱: شکل ۱-۲ هندسه تصویربرداری با دو دوربین را در حالت کلی نشان می‌دهد.

در این شکل I و I' صفحات تصویر چپ و راست و همچنین C و C' مراکز نوری دو دوربین هستند. u

و u' تصاویر نقطه سه بعدی X در صفحه‌های تصویر چپ و راست فرض می‌شوند.

تطابق استریو^۲: u و u' که تصاویر یک نقطه از فضا در صفحات تصویر چپ و راست می‌باشند، نقاط متناظر یا منطبق نامیده می‌شوند. تطابق استریو، یافتن منطبق هر نقطه از یک تصویر در تصویر دیگر می‌باشد. نشان داده می‌شود که در حالت کلی و به علت وجود سطوح دارای تغییر عمق تدریجی، تطابق تعداد پیکسل‌های نامساوی از دو تصویر با یکدیگر نیز امکان‌پذیر است.

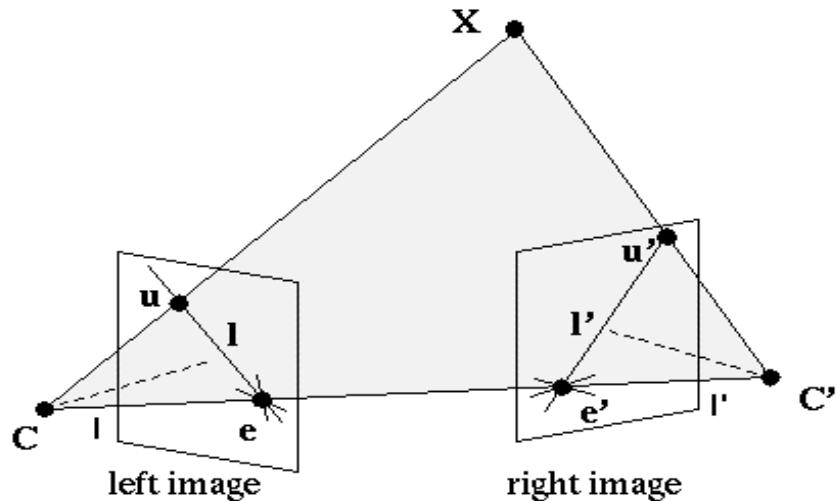
صفحه همقطبی^۳: مطابق با شکل ۱-۲، برای هر نقطه سه بعدی X در محیط، صفحه‌ای که از این نقطه و مراکز دو دوربین یعنی C و C' می‌گذرد، صفحه همقطبی نامیده می‌شود.

۱- Binocular Stereo

۲- Stereo Matching or Stereo Correspondence

۳- Epipolar

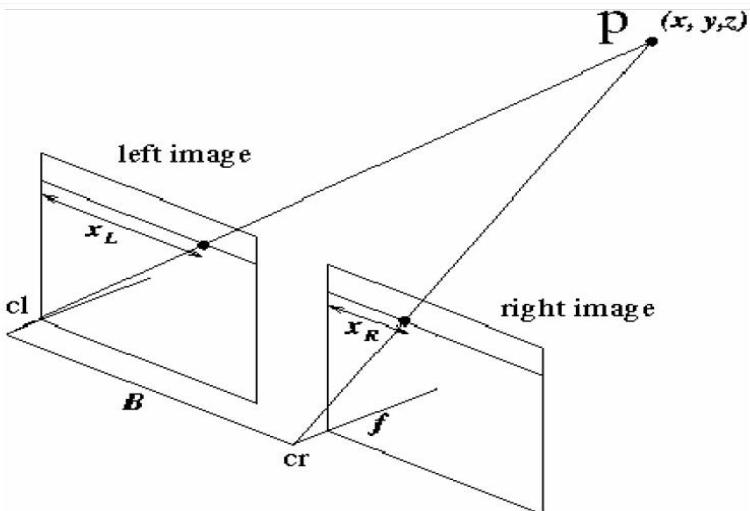
خطوط همقطبی: در صفحه تصویر هر دوربین، خطی که از تقاطع صفحه همقطبی با صفحه تصویر تشکیل می‌گردد، خط همقطبی نامیده می‌شود. در شکل ۱-۲، خط همقطبی متناظر با نقطه X در صفحه تصویر چپ با l و در صفحه تصویر راست با l' نشان داده شده است.



شکل ۱-۲: هندسه تصویربرداری با دو دوربین در حالت کلی

استریوی موازی یا کانونی^۱: اگر محور نوری تصاویر چپ و راست مطابق با شکل ۱-۳ با یکدیگر موازی باشند و تنها با یک انتقال افقی بر هم منطبق شوند، استریوی موازی یا کانونی ایجاد می‌گردد.
کالیبراسیون سیستم استریو: ایجاد استریو موازی یا کانونی در عمل به سادگی میسر نمی‌باشد. بنابراین یک مدل تصحیح کننده برای رسیدن به شرایط کانونی، اعمال می‌گردد. پارامترهای این مدل با کالیبراسیون سیستم تعیین می‌گردند. کالیبراسیون سیستم عبارت است از فرآیند تعیین پارامترهای داخلی و خارجی سیستم تصویربرداری. پارامترهای داخلی، مشخصات نوری و هندسی داخلی دوربین‌ها را بیان می‌کنند و شامل فاصله کانونی مؤثر، ضرایب اعوجاج لنزها و پارامترهای اسکن تصویر، نظیر مرکز تصویر می‌باشند. پارامترهای خارجی، موقعیت و جهت سه بعدی دوربین‌ها را نسبت به همدیگر و نیز نسبت به دستگاه مختصات مرجع مشخص می‌کنند.

تصاویر یکسوشده^۱: تصاویر یکسوشده، تصاویری هستند که منطبق با شرایط استریوی موازی یا کانونی بدست آمده باشند. معمولاً این تصاویر با اعمال مدل تصحیح‌کننده بدست آمده از فرآیند کالیبراسیون بر تصاویر واقعی، بدست می‌آیند.



شکل ۱-۳: هندسه تصویربرداری استریوی دو دوربینه با صفحات تصویر موازی

جابجایی مکانی: در سیستم استریوی موازی چنانچه P یک نقطه در فضا یا مختصات واقعی باشد، تصویر این نقطه در صفحه تصویر چپ و (x_l, y_l) متناظر این نقطه در تصویر راست باشد، ضمن آنکه رابطه $y_l = y_r$ برقرار است، جابجایی مکانی d بین این دو نقطه بر حسب پیکسل، از رابطه (۲-۱) به دست می‌آید.

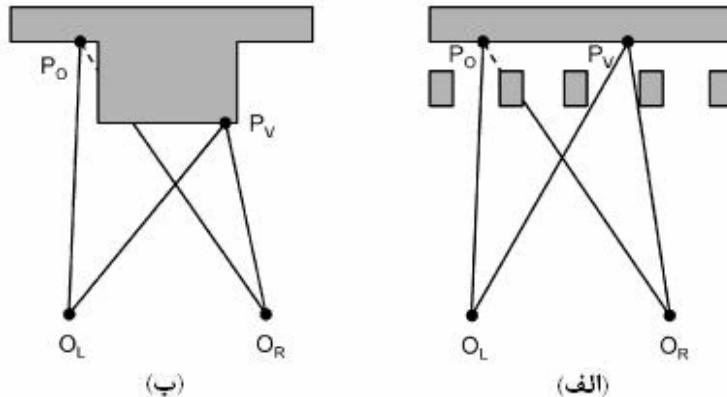
$$d = x_l - x_r \quad (2-1)$$

انسداد^۲: انسداد در استریو از این واقعیت نتیجه می‌شود که به علت آنکه دوربین‌ها در دو موقعیت مکانی متفاوت قرار گرفته‌اند، بخشی از نقاط صحنه در یک تصویر مشاهده شده اما در تصویر دیگر دیده نمی‌شوند. شکل ۱-۴ دو حالت مختلف برای انسداد را نشان می‌دهد. نقطه P_v توسط هر دو

۱- Rectified Image

۲- Occlusion

دوربین مشاهده می‌شود ولی نقطه P_o تنها در محدوده دید دوربین چپ قرار گرفته است. در حالت (الف)، انسداد ناشی از سد شدن جسم عقب‌تر به وسیله جسم جلوتر می‌باشد و در حالت (ب)، انسداد در مجاورت ناپیوستگی عمودی عمق، یعنی ناپیوستگی در راستای لبه‌های موازی با صفحه تصویر، روی داده است.



شکل ۱-۴: انسداد در دو حالت مختلف

نقشه جابجایی مکانی گسسته^۱: در این حالت، مقدار جابجایی مکانی تنها برای نقاط خاصی از تصویر که متناظر با ویژگی‌های مشخصی از صحنه مانند گوشها و لبه اجسام می‌باشند، بدست می‌آید. نقشه جابجایی مکانی گسسته محصول روش‌های استریو مبتنی بر ویژگی است. کاربرد نقشه جابجایی مکانی گسسته در کاربردهای ناوبری ربات سیار، به عنوان مثال در موقعیت‌یابی ربات سیار در محیط، می‌باشد.

نقشه جابجایی مکانی متراکم^۲: مقدار جابجایی مکانی، با مرجع قرار دادن یکی از تصاویر، برای کلیه نقاط آن تصویر تعیین می‌شود. از آنجا که به علت انسداد، متناظر بخشی از پیکسل‌های تصویر مرجع در تصویر دیگر وجود ندارد، عدم امکان تعیین جابجایی مکانی در این نقاط نیز می‌باشد در نقشه جابجایی مکانی متراکم به درستی تشخیص داده شود.

۱- Sparse Disparity Map
۲- Dense Disparity Map

نقشه جابجایی مکانی متراکم، که محصول روش‌های استریو مبتنی بر پیکسل^۱ است، در بازسازی سه‌بعدی صحنه و نیز انتقال اطلاعات تصاویر استریو به کار می‌رود. به طور معمول نمایش نقشه جابجایی مکانی متراکم، به صورت تصویری، با سطوح روشنایی خاکستری صورت می‌گیرد. ضریب ثابتی برای تبدیل مقادیر جابجایی مکانی به مقادیر متناظر شدت روشنایی منظور می‌گردد. این ضریب ثابت با توجه به محدوده تغییرات جابجایی مکانی در تصویر مورد نظر تعیین می‌شود.

استخراج عمق^۲ : کاربرد استریو در بدست آوردن اطلاعات عمق محیط با بازسازی سه‌بعدی صحنه می‌باشد. سیستم استریو موازی، از دو دوربین با صفحات موازی، که به فاصله B از هم قرار گرفته‌اند^۳ و فاصله کانونی هر کدام f است، تشکیل شده است. چنانچه d مقدار جابجایی مکانی پیکسل(x,y) از تصویر مرجع چپ، متناظر با نقطه P در فضای مختصات واقعی (x,y,z) باشد و دستگاه مختصات در مرکز دوربین چپ فرض شود، عمق نقطه P ، یعنی Z ، از رابطه (۳-۱) محاسبه می‌گردد .

$$Z = \frac{Bf}{d} \quad (3-1)$$

استخراج Z با استفاده از رابطه (۳-۱) کار سراسرتی است، اما دو مشکل تطابق و وجود خطای کالیبراسیون دوربین‌ها، معمولاً یافتن عمق را با مشکل همراه می‌سازد. با استفاده از نقشه جابجایی مکانی گسسته، امکان تعیین عمق محیط در نقاط محدودی از صحنه فرآهم می‌گردد، بنابراین برای بازسازی کامل صحنه معمولاً از روش برازش صفحات بر این نقاط استفاده می‌گردد. با استفاده از نقشه جابجایی مکانی متراکم امکان بازسازی کل صحنه به طور مستقیم امکان‌پذیر می‌باشد، هر چند در این حالت نیز چون هر پیکسل تصویر معادل با بخشی از صحنه می‌باشد، با استفاده از روش درون‌یابی می‌توان مقدار عمق را در نقاط صحنه با مقدار دقیق‌تری تخمین زد.

۱- Pixel-Based Stereo

۲- Depth Extraction

۳- Base Line

۱-۳ اصول بنیادین بینایی استریو

این بخش با یک نگاه نزدیک‌تر به بینایی استریو و مقایسه آن با بینایی انسان‌ها آغاز می‌شود، سپس یک مرور اجمالی از مشکلاتی که الگوریتم‌های بینایی استریو با آن مواجه هستند را ارائه می‌کند.

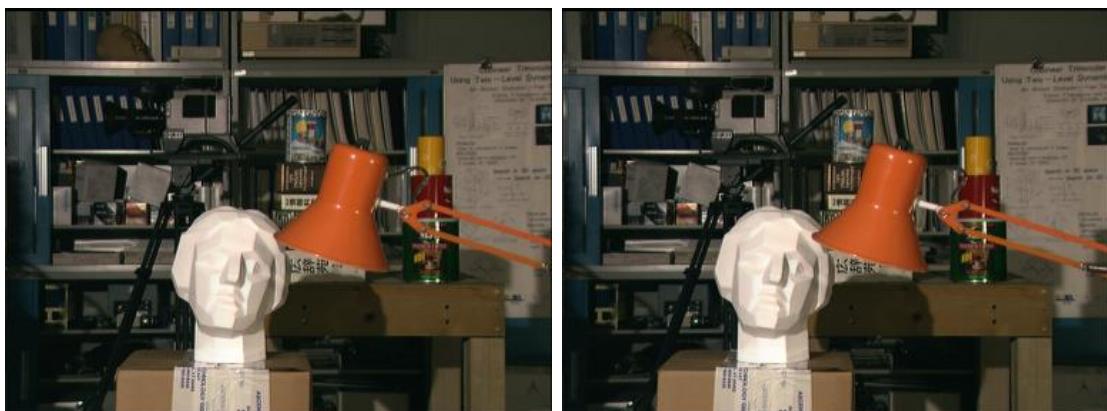
۱-۱ بینایی استریو چیست؟

روندهای موجود در دریافت بصری که به دریافت عمق می‌انجامد تحت عنوان بینایی استریو شناخته شده و اولین بار توسط چارلز ویتسون در ۱۸۳۳ کشف شد [۵]. وی اشاره داشت که جدایی افقی چشم‌مان انسان به هر چشمی اجازه می‌دهد تا جهان را از موقعیت اندکی متفاوت بنگرد و آنکه موقعیت اشیاء در هر منظری بسته به فاصله آن از شخص نظاره‌گر اندکی متفاوت است. این را می‌توان به سادگی و تنها با نگاه کردن به جابجایی افقی انگشت خود که در نزدیکی بینی نگهداشته و به ترتیب در حین بسته نگهداشتن یکی از چشم‌مان با دست دیگر، مشاهده نمود. هرچه انگشت شما به چشم‌مان شما نزدیک‌تر باشد، جابجایی افقی بزرگ‌تر خواهد بود. با اندازه‌گیری جابجایی مکانی هر نقطه در یک منظره می‌توان یک صحنه سه‌بعدی را از دو تصویر دو بعدی بازسازی نمود.

۲-۳ بینایی استریو انسان‌ها در مقابل دستگاه‌ها

بینایی استریو ایجاد شده توسط مغز انسان کاملاً با آنچه که برای بینایی ماشینی به دست آمده، تفاوت دارد. علت اصلی این امر پیچیدگی مکانیکی چشم‌مان انسان است. مغز انسان از حرکات ظریف چشم‌مان که به سمت یک نقطه بر روی یک شیء تقارب می‌یابد به عنوان مقیاسی برای تعیین فاصله آن شیء تا ناظر استفاده می‌کند. دریافت عمق توسط انسان همچنین علل دیگری را شامل می‌شود که از دید دوچشمی فراتر می‌روند، همچون اختلاف دید حرکت، ادراک و رنگ تا اطلاعاتی را درباره شکل و نورپردازی اشیاء در یک صحنه ایجاد کند.

شکل ۱-۵ مثالی از یک جفت تصویر اصلاحی است (که از مجموعه داده میدلبوری^۱ گرفته شده است [۶]، دو تصویر بروی قاب دوربینی یکسان قرار گرفته‌اند، اما از موقعیت‌های افقی اندکی متفاوت برداشته شده‌اند. چنین تصاویری نقطه آغازی را برای طیف گستردگی‌های الگوریتم‌های تطابق استریو تشکیل داده‌اند.

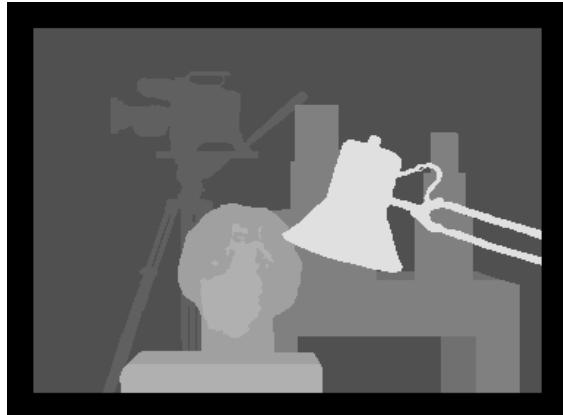


شکل ۱-۵: نمونه یکسوزده تصویر استریو جفت [۶]

۱-۳-۳ نقشه جابجایی مکانی

اکثر الگوریتم‌های بینایی استریو به عنوان خروجی، نقشه جابجایی مکانی متراکم تولید می‌کنند، که در قالب (x, y, d) نوشته شده است. یک نقشه جابجایی مکانی تصویری است که در آن شدت هر پیکسلی دارای تناسبی وارون با فاصله از شیء قرار گرفته در مختصات‌های آن تصویر است، یعنی هر چه نقطه در یک تصویر به دوربین نزدیک‌تر باشد، شدت مربوطه در نقشه جابجایی مکانی بیشتر خواهد بود. یک مثال از نقشه جابجایی مکانی در شکل ۱-۶ نشان داده شده است.

۱- Middlebury



شکل ۱-۶: نقشه جابجایی مکانی متراکم [۶]

چنین نقشه‌های جابجایی مکانی را می‌توان با یکی از دو تصویر شکل ۱-۵ در جفت تصویر استریو مربوط دانست. وقتی نقشه جابجایی مکانی به یکی از تصاویر استریو مربوط باشد، مختصات‌های (x, y) از فضای اختلاف با مختصات‌های پیکسل تصویر مرجع مربوط هستند. رابطه میان یک پیکسل (x, y) در تصویر مرجع r و یک پیکسل (x', y') در تصویر تطبیقی m را می‌توان طبق رابطه (۱-۵) توصیف نمود:

$$x' = x + s d(x, y), \quad y' = y \quad (۱-۵)$$

که در آن $s = \pm 1$ تنها یک علامت است که برای اطمینان از همیشه مثبت بودن اختلاف‌ها انتخاب شده است [۲].

۴-۳-۱ مشکلاتی که الگوریتم‌های استریو با آن مواجه هستند

همانطور که قبل تر مطرح شد، درک مفهوم کلی برای تعیین اطلاعات عمقی از یک جفت تصویر استریو نسبتاً سرراست است، البته در زمان به کارگیری چنان سیستمی، برخی از مشکلات به سرعت ظاهر می‌شوند. برای آنکه بتوانیم عمق یک موقعیت را در یک تصویر تعیین کنیم، ضرورت دارد که یک موقعیت تطبیقی را در تصویر دوم از جفت تصویری یافته و سپس جابجایی مکانی یا اختلاف از

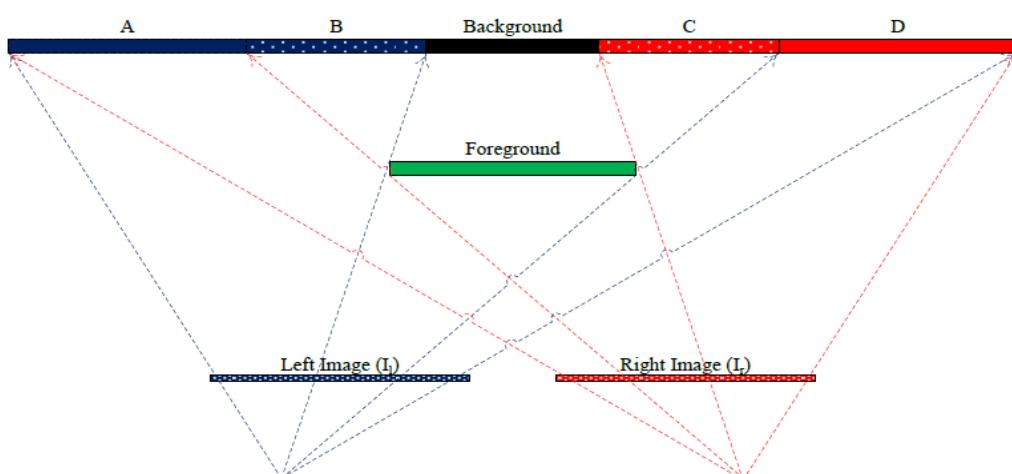
یک تصویر به دیگری را تعیین کنیم. یافتن این موقعیت تطبیقی را مسئله تطابق می‌نامند و به‌ویژه در زمان مواجهه با پیچیدگی‌های اضافی مثل نقاط انسداد، نواحی بدون بافت وسیع و نویز تصویر، خودکارسازی این روند آسان نیست.

این مفهوم را می‌توان به‌سادگی به مسائلی با بیش از دو نما تعمیم بخشید. می‌توان با N تعداد دوربین مختلف از N نقطه در فضا یا از تنها یک دوربین که نسبت به صحنه حرکت می‌کند، تصاویری تهیه کرد. مسئله تطابق حتی وقتی اشیاء در صحنه درحال حرکت هستند، پیچیده‌تر هم می‌شود.

■ نقاط انسداد

همانطور که در بخش پیش‌هم اشاره شد، یکی از مشکلات اصلی که الگوریتم‌های تطابق استریو با آن مواجه هستند، حضور نقاط انسداد در تصاویری است که از یک صحنه اما از زوایای مختلف ضبط شده‌اند. این یعنی یک نقطه تطبیقی برای یک تصویر استریو همیشه در تمامی تصاویر وجود ندارد.

شکل ۱-۷ در زیر نشان می‌دهد که یک شیء در پس‌زمینه از یک صحنه می‌تواند وضوح پیش‌زمینه را در یکی از تصاویر جفتی مسدود سازد. برای مثال، قسمت پیش‌زمینه در تصویر که دارای برجسب B است، در تصویر چپ مشهود بوده اما در تصویر راست چنین نیست.



شکل ۱-۷: انسداد نمونه در یک جفت تصویر استریو [۷]

در تحقیقات برای حل این مسئله روش‌های متعددی معرفی شده است [۷، ۸، ۹]. یک روش نمونه عبارت است از محاسبه نقشه جابجایی مکانی، یکبار نسبت به تصویر چپ و بار دیگر نسبت به تصویر راست و سپس استفاده از اطلاعات، جهت تشخیص وجود نواحی انسداد. چنین فنونی را عموماً تطبیق دو-جهته^۱ یا مضاعف نامیده‌اند [۸].

▪ نواحی بدون بافت

با نواحی بدون بافت و همگن، مسئله دیگری مطرح است. به جای عدم توانایی یافتن یک تطبیق آنطور که در مورد انسدادها وجود داشت، در نواحی بدون بافت اغلب تطبیق‌های متعددی یافت می‌شود، چرا که نواحی تصویر برای یافتن یک تطبیق به حد کافی منحصر به فرد نیستند. برای برآمدن از پس چنین شرایطی، اغلب فرضیاتی درباره پیوستگی چنین نواحی بدون بافتی مطرح شده، که روش‌هایی چون درون‌یابی خطی را ممکن می‌سازد، که باید برای تولید نقشه‌های جابجایی مکانی هموار در چنین نواحی به کار رود.

▪ نویز تصویر

نویز در تصاویر یک مشکل عمده بوده که تقریباً تمامی سیستم‌های پردازشی با آن مواجهه‌اند. نویز را می‌توان در قالب تنوع ناخواسته تصادفی در درخشندگی و اطلاعات رنگ از یک تصویر در نظر گرفت. نویز تصویر در تصویربرداری دیجیتال از طیفی از منابع مختلف ناشی می‌شود: نویز گاوی که بطور طبیعی رخ می‌دهد؛ نویز القاء شده با حسگر که می‌تواند نویز شمارش فوتونی باشد؛ نویز از پردازش‌های مختلفی چون خطای تبدیل سیگنال و انتقال سیگنال.

در کل، نویز تصویر در قسمت‌های با نورکمتر بیشتر از همه مشهود است، نمونه‌ای از تصویر نویزی در شکل ۱-۸ نشان داده شده است.

۱- Bi-Directional
۲- Colour-Weighting



شکل ۱-۸: یک تصویر نویزی گرفته شده با دوربین پاناسونیک LX1

برای آنکه بتوانیم تأثیر نویز را بکاهیم، ضرورت دارد که مدل‌های نویز را فرمول‌بندی کنیم. می‌توان یک تصویر را به دو مؤلفه مجزا تقسیم نمود، یعنی تصویر مطلوب و نویز. در اغلب موارد این دو مؤلفه جمع‌پذیر بوده و می‌توان آنها را اینگونه نوشت:

$$f(\cdot) = g(\cdot) + q(\cdot) \quad (6-1)$$

که در آن $f(\cdot)$ همان تصویر اصلی است، $g(\cdot)$ مؤلفه مطلوب از تصویر و $q(\cdot)$ نویز موجود در تصویر است.

۴-۱ مروری بر تحقیقات مشابه

در این بخش ما تحقیقات قبلی در مورد تطابق استریو را بطور مختصر مرور می‌کنیم. در سال ۲۰۰۲، شارستاین و زلیسکی مقاله‌ای را تحت عنوان طبقه‌بندی و ارزیابی الگوریتم‌های تطبیق استریو دو-جفتی متراکم منتشر ساختند [۳]. این مقاله طیف گسترده‌ای از رویکردهای مختلف نسبت به مسئله تطابق استریو را پوشش داده و توانست روش‌های مختلف را طبقه‌بندی کند. همانطور که از عنوان مقاله حدس می‌زنیم، این مقاله تنها الگوریتم‌های استریویی را درنظر می‌گیرد که نقشه‌های

جابجایی مکانی متر acumalی را از تصاویر ورودی استریو دو- جفتی تولید می کنند. الگوریتم هایی که در این مقاله مورد تحقیق قرار گرفتند عبارتند از برخی روش های محلی و همچنین روش هایی برمبنای بهینه سازی سراسری. از پایان سال ۲۰۰۸ بهترین الگوریتم های عملیاتی عمدتاً براساس کمینه سازی انرژی هستند، مانند آنهایی که در [۱۰] و [۱۱] توصیف شده اند. این الگوریتم ها نقشه های جابجایی مکانی را تولید می کنند که در آن درصد اختلاف های غیر دقیق با تصاویر سنجشی ارائه شده توسط [۶] کمتر از ۶٪ بوده است. یقیناً بزرگترین سهمی که می توان در [۳] یافت، صفحه آزمایش استانداردی است که کارشناسان برای مقایسه دقت الگوریتم های موجود تهیه کردند. آنها این صفحه آزمایش را بطور عمومی در دسترس قرار دادند و هم اکنون به عنوان یک مرکز سراسری برای سنجش الگوریتم های تطبیق استریو به کار می رود. صفحه آزمایش متشکل از چهار جفت تصویر استریو مختلف و نقشه های جابجایی مکانی صحیح آنهاست. با استفاده از این صفحه، دقت یک الگوریتم استریو را می توان تنها با محاسبه تفاوت میان نقشه جابجایی مکانی تولید شده با برخی الگوریتم ها و نقشه صحیح، مقایسه کرد.

در [۱۲] روش های جمع بندی هزینه مختلف برای روش های استریوی محلی، توسط مینگلان گانگ و همکارانش در سال ۲۰۰۶ منتشر شد. کارشناسان در این مقاله شش فن جمع بندی هزینه مختلف را به کار گرفتند و عملکرد آنها را هم از لحظه سرعت و هم دقت مقایسه کردند. آنها از اختلاف های مطلق در شدت ها به عنوان هزینه های تطابقی برای هر یک از فنون جمع بندی هزینه استفاده کردند.

۱-۵ نتیجه گیری

این فصل با یک مرور اجمالی از مفهوم بینایی استریو و هندسه آن آغاز شد، همچنین یک مرور اجمالی از تحقیقات مشابه و یک بحث از چگونگی مقایسه این مقاله ها ارائه کرد. مرور اجمالی

تحقیقات که در اینجا ارائه شد، مبنایی از دانش را ایجاد می‌کند که در چند فصل بعدی به کار خواهد رفت.

هدف ما در این پایان‌نامه، بررسی ایده‌های جدید و تلفیق آنها برای بهبود در استخراج نقشه جابجایی مکانی گستته و متراکم به صورت بلادرنگ می‌باشد. این پایان‌نامه در پنج فصل تدوین شده است. در فصل جاری، پیشینه کار و تعریف اصطلاحات مورد استفاده ارائه شد. در فصل دوم الگوریتم‌های رایج برتر و مباحث تئوری مربوط به آنها ارائه می‌شود. در فصل سوم، روش پیشنهادی در حل مسئله مورد بررسی قرار می‌گیرد و بر اساس آن، اهداف و روش‌های پیشنهادی و همچنین فعالیت‌هایی که در طی انجام پایان‌نامه صورت گرفته، ارائه می‌شود. فصل چهارم شامل معیارها و روش‌های ارزیابی الگوریتم‌ها و بانک تصاویر می‌باشد. در فصل پنجم، نتیجه‌گیری از تحقیق و پیشنهادات برای کارهای آینده ارائه می‌شود. با این مقدمه، حال می‌توان بررسی را با توصیف برخی از الگوریتم‌های تطابق استریوی رایج در فصل بعد ادامه داد.

فصل دوم

مروار اجمالي الگوریتم‌ها

۱-۲ مقدمه

در این فصل پس از دسته‌بندی کلی روش‌ها و معرفی برخی الگوریتم‌های برجسته و پایه، به بررسی قیود و محدودیت‌های مورد استفاده در تطابق استریو می‌پردازیم.

الگوریتم‌های تطابق استریو را می‌توان در دو دسته اصلی دسته‌بندی نمود: روش‌های محلی و سراسری. روش‌های محلی بر رنگ پیکسل یا ارزش‌های شدت در یک پنجره محدود مرکز دارند تا ارزش‌های اختلاف را برای هر پیکسلی تعیین کنند. روش‌های سراسری تمامی اختلاف‌ها را در آن واحد با استفاده از فنون کمینه‌سازی انرژی عمومی محاسبه می‌کنند. برخی از این فنون کمینه‌سازی انرژی عبارتند از [۱۰]:

- برش‌های نموداری [گرافی]؛

- برنامه‌نویسی پویا^۱؛

بسیاری از این فنون کمینه‌سازی انرژی سراسری زمانی که با نواحی انسدادی برخورد کنند، به خوبی عمل نمی‌کنند. اخیراً، یک نسل تازه از الگوریتم‌ها ظاهر شده‌اند تا به این مسئله بپردازنند. این الگوریتم‌ها روش‌های مقطع-مبنا^۲ را شامل شده و به دلیل عملکرد خوب خود جلب توجه کرده‌اند. این الگوریتم‌ها، بهترین نتایج را روی پایگاه داده میدلبوری به دست آورده‌اند [۱۰، ۱۱].

الگوریتم‌های تطبیق استریو محلی را می‌توان به چهار مرحله متوالی با برخی زیرمجموعه‌ها تقسیم نمود، این چهار مرحله عبارتند از:

- ۱- محاسبه هزینه تطبیق؛

- ۲- جمع‌بندی هزینه؛

- ۳- محاسبه اختلاف و بهینه‌سازی؛ و

۱- Dynamic Programming
۲- Segment-Based

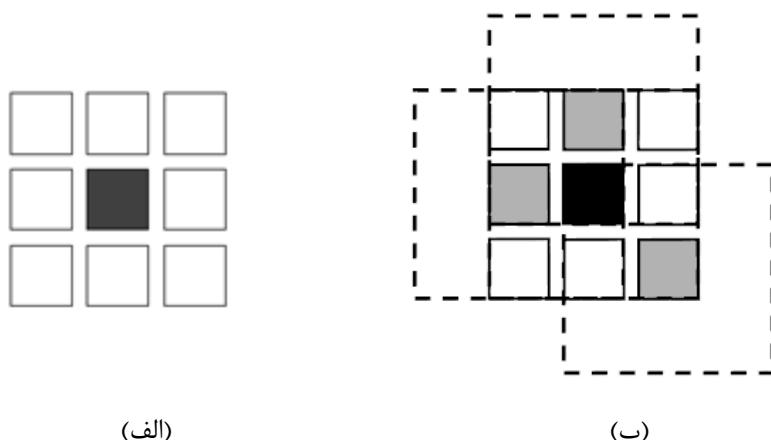
۴- اصلاح اختلاف^۱.

۲-۲ الگوریتم‌های تطبیق استریو محلی

روش‌های تطبیق استریوی محلی، با وجود اینکه اغلب به اندازه رویکردهای سراسری دقیق نیستند، اما دارای سرعت مطلوبی هستند. در روش‌های انطباق محلی، روش حل مسئله بر اساس شباهت مقدار شدت روشناهی اجزای مشابه در دو عکس، پایه‌ریزی شده است. در این روش، انطباق با انتخاب پنجره‌ای کوچک از عکس حاصل از دوربین سمت چپ و جستجوی پنجره‌ای مشابه در عکس سمت راست، انجام می‌شود. در این روش ابتدا تصویر سمت چپ به عنوان مرجع انتخاب می‌شود و تصویر سمت راست به عنوان محیط جستجو برای یافتن پنجره‌ی متناظر با پنجره‌ی انتخابی در تصویر سمت چپ انتخاب می‌شود. شکل ۱-۵ نشان‌دهنده تصویر سمت راست و چپ می‌باشد که با استفاده از شکل ۲-۱(الف) به پنجره‌های مساوی تقسیم می‌شوند. اندازه‌ی این پنجره‌ها می‌تواند کوچک یا بزرگ در نظر گرفته شود. کوچک بودن این پنجره‌ها سبب کاهش دقت و افزایش سرعت می‌شود و بزرگ بودن این پنجره‌ها تا حدی سبب افزایش دقت و کاهش سرعت می‌شود.

شکل ۲-۱(ب) پنجره‌های جستجو در تصویر سمت راست را نشان می‌دهد. مبدأ پنجره‌های جستجو در تصویر سمت راست برابر با محل پنجره‌ی انتخابی در تصویر سمت چپ می‌باشد. از آنجایی که دو دوربین مربوط به دوربین‌های استریو در یک محور افقی قرار دارند، اجسام در دو تصویر فقط در راستای افق جایه جا می‌شوند. این عامل سبب می‌شود جستجو فقط در راستای محور افقی محدود شود و نیازی به جستجو در محور عمودی نداشته باشیم. در بخش‌های بعد مروری بر روش‌های مورد استفاده در مراحل مختلف الگوریتم‌های تطبیق استریو می‌اندازیم.

۱- Disparity Refinement

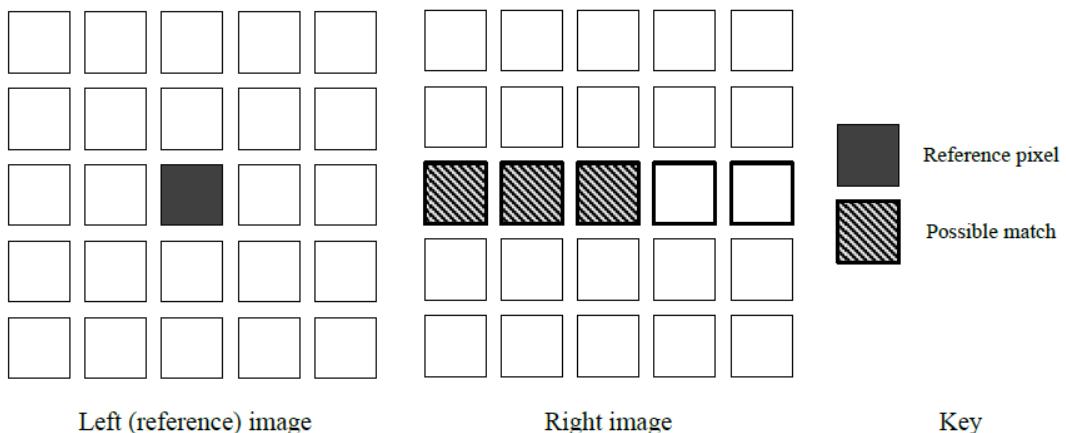


شکل ۲-۱: (الف) مثالی از اندازه پنجره و (ب) مثالی از پنجره‌های جستجو

۲-۳ روش‌های محاسبه هزینه

اولین گام در هر الگوریتم تطبیق استریو، یافتن راهی برای محاسبه هزینه تطبیق میان دو پیکسل در یک جفت تصویر استریو است. هدف کلی از یک الگوریتم تطبیق استریو آن است که هر پیکسلی در تصویر مرجع را با یک پیکسل در تصویر دیگر تطبیق دهد. در صورتی که تصاویر یکسو شده باشند، برای یافتن یک تطبیق برای یک پیکسل در تصویر مرجع، باید تنها درامتداد ردیف افقی از پیکسل‌ها در تصویر تطبیقی جستجو نمود. اگر دوربین مسافتی را به سمت راست جابجا شده باشد، در آن صورت تمامی پیکسل‌ها در تصویر جدید تا حدی به سمت چپ جابجا می‌شوند، این در شکل ۲-۲ به تصویر کشیده شده است.

همانطور که در شکل ۲-۲ نشان داده شده، برای یافتن پیکسل در تصویر دوم با کوچکترین هزینه تطبیق با پیکسل مرجع، تنها باید هزینه تطبیق را برای یک تعداد اندک از پیکسل‌های انتخابی در تصویر دوم محاسبه نمود.



شکل ۲-۲: فضای جستجو تطبیق استریو

البته، هزینه تطبیق اطلاعاتی را درباره نه تنها پیکسل‌های مورد تطبیق، بلکه همچنین یک ناحیه پشتیبان از پیکسل‌ها در حول آنها تلفیق می‌سازد. اختلاف (d) عبارت است از میزان جابجایی افقی و واحد آن تعداد پیکسل‌هاست. در تصویر، اگر پیکسل مرجع با پیکسل میانی در تصویر راست تطبیق یابد، اختلاف صفر خواهد بود. اختلاف همچنین دارای تنشی میانی است که با وارون عمق است (هرچه ارزش d بالاتر باشد، پیکسل به دوربین نزدیک‌تر است). تنها خروجی الگوریتم‌های تطبیق استریو همین ارزش اختلاف، یعنی d، برای هر پیکسلی در تصویر مرجع است. سپس مجموعه حاصل از این ارزش‌های اختلاف را می‌توان برای ساخت نقشه جابجایی مکانی به کار برد. اندازه، شکل و وزن منطقه پشتیبان با روش تجمع انتخابی تعیین می‌شود. هدف از مرحله محاسبه هزینه تطبیق تنها تعریف آن است که هزینه تطبیق میان دو منطقه پشتیبان چگونه محاسبه شود. روش‌های مختلف برای محاسبه هزینه تطبیق، عموماً از ارزش شدت پیکسل‌هایی استفاده می‌کنند که در مناطق پشتیبان قرار دارند.

۱-۳-۲ مجموع اختلاف‌های مربعی^۱ (SSD)

رویکرد SSD برای محاسبه هزینه‌های تطبیق، یکی از پرکاربردترین روش‌های است. می‌توان آن را

اینگونه نوشت:

^۱- Sum Of Square Differences

$$C_{SSD}(x, y, d) = \sum_{(i, j) \in N(x, y)} (I_1(i, j) - I_2(i - d, j))^2 \quad (1-2)$$

که در آن $C_{SSD}(x, y, d)$ مجموع هزینه SSD از دو تصویر و I_1 و I_2 عبارتند از $N(x, y)$ همان منطقه پشتیبان از پیکسل‌ها پیرامون موقعیت (x, y) است. بنابراین $I_1(i, j)$ شدت پیکسل در تصویر اول است که در موقعیت (i, j) واقع شده است. ارزش‌های شدت معمولاً ارزش‌هایی صحیح میان صفر و ۲۵۵ بوده، که می‌توانند بازه صفر و یک نرمال شوند.

۲-۳-۲ مجموع اختلاف‌های مطلق^۱ (SAD)

یک معیار بسیار نزدیک به SSD از هزینه SAD است. می‌توان آن را اینگونه نوشت:

$$C_{SAD}(x, y, d) = \sum_{(i, j) \in N(x, y)} |I_1(i, j) - I_2(i - d, j)| \quad (2-2)$$

که در آن I_1 و I_2 دو تصویر ورودی و $N(x, y)$ منطقه پشتیبان پیرامون موقعیت (x, y) است.

۳-۳-۲ همبستگی نرمالیزه^۲ (NCC)

یک روش پیچیده‌تر اما در عین حال قوی‌تر برای سنجش هزینه تطبیق همبستگی نرمالیزه است. این هزینه نسبت به تغییرات در سطح و قوت سیگنال در میان تصاویر استریو نامتفاوت است و این بدان خاطر است که هزینه با توجه به میانگین و انحراف معیار پیکسل‌ها در نواحی پشتیبان نرمالیزه شده است. متأسفانه NCC از لحاظ محاسباتی از رویکردهای ساده‌تری چون SSD پیچیده‌تر است. دو منطقه پشتیبان را می‌توان اینگونه نوشت [۱۳]:

^۱- Sum Of Absolute Differences
^۲- Normalised Cross-Correlation

$$C_{NCC}(x, y, d) = \frac{S_{xy,d}(I_1, I_2)}{A_{xy}(I_1) \times B_{xy,d}(I_2)} \quad (3-2)$$

که در آن،

$$S_{xy,d}(I_1, I_2) = \sum_{(i,j) \in N(x,y)} I_1(i, j) I_2(i-d, j) \quad (4-2)$$

$$A_{xy}(I_1) = \sum_{(i,j) \in N(x,y)} I_1(i, j)^2 \quad (5-2)$$

$$B_{xy,d}(I_2) = \sum_{(i,j) \in N(x,y)} I_2(i-d, j)^2 \quad (6-2)$$

در این معادلات، d نشانگر جابجایی درامتداد خطوط همقطبی است. $N(x, y)$ منطقه پشتیبان مورد ارزیابی است که همان اندازه پنجره همبستگی است.

به دلیل محدودیت همقطبی معلوم است که تطبیق‌ها برای یک پیکسل در تصویر مرجع در خط افقی معادل در تصویر راست قرار دارد. برای یک نقطه در تصویر مرجع می‌توان چنین فرض کرد که پنجره جستجو به یک طیف $[\omega, +\omega]$ محدود شده است. بنابراین، ارزش d در معادلات بالا تنها در طیف $[-\omega, +\omega]$ تغییر می‌کند. روش‌های مختلفی برای بهبود بیشتر راندمان این محاسبات وجود دارد که محاسبه همبستگی در حوزه فرکانس را شامل می‌شود [۱۴، ۱۳].

۴-۳-۲ تبدیل رتبه و تبدیل آماری

روش دیگر برای محاسبه هزینه‌های تطبیق، که نسبت به تغییرات در جهت‌گیری دوربین قوی است، تبدیل رتبه^۱ می‌باشد [۱۶، ۱۵]. تبدیل رتبه یک منطقه پشتیبان در قالب تعداد پیکسل‌های در آن ناحیه تعریف می‌شود که دارای یک ارزش شدت کمتر از ارزش پیکسل مرکزی است.

^۱- Rank Transform

یک نسخه دیگر از تبدیل رتبه ، تبدیل آماری^۱ است. بیشتر بهجای ثبت تنها تعداد پیکسل‌های با یک ارزش شدت که کمتر از پیکسل مرکزی است، تبدیل آماری اطلاعات را در یک رشته بیتی کدگذاری می‌کند. سپس هزینه تطبیق برای دو ناحیه پشتیبان را می‌توان با استفاده از مسافت همینگ^۲ (تعداد بیت-هایی که اختلاف دارند) برای رشته‌های بیتی تبدیل آماری ارزیابی نمود. برای مثالی از تبدیل رتبه و تبدیل آماری، شکل ۳-۲ را ملاحظه کنید.

89	63	72		89	63	72	
67	55	64	$\Rightarrow 2$	67	55	64	$\Rightarrow 00000011$
58	51	49		58	51	49	

شکل ۳-۲: تبدیل رتبه (چپ) و تبدیل آماری (راست) [۱۶]

۳-۳-۵ کاربرد اطلاعات رنگ

تمامی محاسبات هزینه تطبیق که در بالا مطرح شد، تنها از ارزش‌های شدت پیکسل‌ها در تصاویر استفاده نموده و از اطلاعات رنگی استفاده نمی‌کنند، این ضرورتاً نباید مسئله‌ساز باشد. به سادگی می‌توان تصور کرد که چگونه بهجای شدت‌ها، این روش‌ها می‌توانستند بطور انفرادی هزینه تطبیق نواحی پشتیبان را برای کانال‌های قرمز، سبز و آبی از یک تصویر تعیین کنند. همچنین می‌توان از اطلاعات رنگ در مرحله تجمع هزینه برای الگوریتم‌های تطبیق استریو محلی استفاده نمود. در مرحله تجمع هزینه، اطلاعات رنگ را می‌توان به طرق مختلفی به کار برد تا با تغییر تعریف از نواحی پشتیبان بطور پویا بر هزینه‌های تطبیق اثر گذاشته شود. اطلاعات بخش‌بندی رنگ را می‌توان برای تنظیم پویای اندازه و شکل نواحی پشتیبان به کار برد، به گونه‌ای که از نواحی که از بخش‌های رنگی متعدد

۱- Census Transform
۲- Hamming

عبور می‌کند، اجتناب شود. همچنین می‌توان از آن برای تنظیم وزن پیکسل‌هایی خاص در یک ناحیه پشتیبان، استفاده کرد.

۴-۲ روش‌های تجمع هزینه

مرحله دوم برای الگوریتم‌های تطبیق استریو محلی تعریف نواحی پشتیبانی است که هزینه تطبیق براساس آن محاسبه می‌شود. ویژگی‌های مختلف بسیاری از یک ناحیه پشتیبان وجود دارد که باید در زمان تعریف آن مدنظر قرار گیرد. فهرست زیر نشانگر برخی از این ویژگی‌هاست:

- اندازه منطقه،

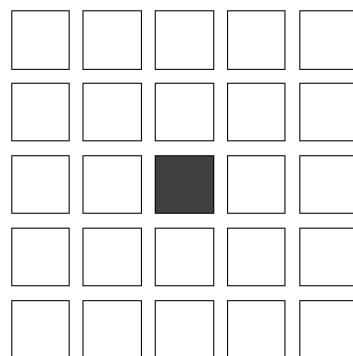
- شکل منطقه، و

- وزن پیکسل‌ها در منطقه.

این ویژگی‌ها همچنین می‌توانند برای هر پیکسل موردنظری بطور پویا تغییر کنند. در این صورت، ناحیه پشتیبان بطور پویا و برمبنای اطلاعات بیرون کشیده شده از پیکسل‌های پیرامون پیکسل مرجع تعریف می‌شود. تقسیم‌بندی تصویر مرجع به نواحی با رنگ مشابه اغلب می‌تواند اطلاعات مفیدی برای تعیین مرزهای منطقه پشتیبان به دست دهد. تعریف منطقه پشتیبان که هزینه‌های تطبیق براساس آن تجمع می‌یابند، نقشی حیاتی در توانایی الگوریتم‌های تطبیق استریو برای رفع مشکلاتی چون نواحی بدون بافت، نقاط انسداد و ناپیوستگی‌های عميق ایفا می‌کند.

۱-۴-۲ روش تجمع هزینه پنجره ثابت

ساده‌ترین روش ممکن برای تعریف نواحی پشتیبان کاربرد یک پنجره مربعی ساده است که برای تمامی پیکسل‌ها در تصویر مرجع ثابت می‌ماند. وقتی چنان ناحیه‌ای تعریف می‌شود، شایان توجه است که اندازه پنجره باید یک عدد فرد باشد تا اطمینان یابیم که در پنجره یک پیکسل مرکزی واحد وجود دارد. شکل ۴-۲ نشانگر یک پنجره ثابت مربعی با اندازه ۵ پیکسل است.



شکل ۲-۴: نمونه پنجره مربعی ثابت با اندازه ۵

اندازه پنجره بهینه به محتوای جفت تصویر ورودی بستگی دارد. معمولاً یک پنجره بزرگتر، تطبیق‌های غلط کمتری ارائه نموده و همیشه اطمینان می‌بخشد که تطبیق‌های صحیح در نواحی بدون بافت، یافت شده‌اند. در نواحی بافتدار، بهویژه در نزدیکی ناپیوستگی‌های عمق، اندازه‌ی پنجره کوچکتر اغلب نتایج بهتری ارائه می‌کنند. همچنین می‌دانیم که اندازه‌ی پنجره بزرگتر، یعنی محاسبه هزینه باید پیکسل‌های بیشتری را شامل شده و بنابراین مستلزم زمان پردازشی بیشتری است.

۲-۴-۲ روش تجمع هزینه وزن انطباقی^۱

یک بهبود در روش تجمع پنجره ایستا آن است که در ناحیه پشتیبان مربعی، وزن‌هایی بر پیکسل‌ها اعمال شود. این به سیستم اجازه می‌دهد تا تأثیری را که هر پیکسل در ناحیه پشتیبان بر هزینه تطبیق تجمعی دارد، کنترل کند. این وزن‌ها بطور پویا تخصیص یافته، برای هر پیکسل مرجعی تغییر می‌کنند. همانطور که توسط نویسنده‌گان [۱۷] حدس زده شد، وزن‌های تخصیصی می‌باشند که دو عامل بستگی داشته باشند؛ شباهت رنگ و فاصله استریو از پیکسل در مرکز منطقه. شکل ۵-۲ نشانگر یک ناحیه پشتیبان نمونه با یک وزن است که به هریک از پیکسل‌ها تخصیص یافته است.

۱- Adaptive Weight

0.5	0.6	0.7	0.6	0.5
0.6	0.8	0.9	0.8	0.6
0.7	0.9		0.9	0.7
0.6	0.8	0.9	0.8	0.6
0.5	0.6	0.7	0.6	0.5

شکل ۲-۵: نمونه پنجره تطبیقی وزن انطباقی

همانطور که مطرح شد، وزن برای هر پیکسلی در منطقه پشتیبان برمبنای دو عامل محاسبه شده است، اختلاف رنگ (Δc_{ij}) و فاصله استریو (Δg_{ij}) با توجه به پیکسلی که در مرکز منطقه پشتیبان قرار دارد. هرچه اختلاف رنگ کوچکتر باشد، وزن تخصیص یافته به پیکسل باید بزرگتر باشد. بطور مشابه، هرچه پیکسل به پیکسل مرکزی نزدیکتر باشد، وزن تخصیص یافته به پیکسل پشتیبان باید بزرگتر باشد. پس می‌توان وزن‌ها را اینگونه نوشت:

$$w(i, j) = f_s(\Delta c_{ij}) \cdot f_p(\Delta g_{ij}) \quad (7-2)$$

در (7-2)، f_s و f_p نشانگر روابط از لحاظ شباهت رنگ و مجاورت است. می‌توان رابطه شباهت

رنگ را اینگونه نوشت:

$$f_s(\Delta c_{ij}) = \exp\left(-\frac{\Delta c_{ij}}{\gamma_s}\right) \quad (8-2)$$

بطور مشابه، می‌توانیم رابطه را برای مجاورت اینگونه تعریف کنیم:

$$f_p(\Delta g_{ij}) = \exp\left(-\frac{\Delta g_{ij}}{\gamma_p}\right) \quad (9-2)$$

که در آن Δg_{ij} عبارت است از فاصله اقلیدسی میان پیکسل‌های i و j . ثابت γ دارای تناسبی مستقیم با اندازه ناحیه پشتیبان است [۱۹].

حال می‌توانیم با استفاده از معادلات (8-۲) و (۹-۲)، معادله (7-۲) را بازنویسی کنیم:

$$w(i, j) = \exp\left(-\left(\frac{\Delta c_{ij}}{\gamma_s} + \frac{\Delta g_{ij}}{\gamma_p}\right)\right) \quad (10-2)$$

وزن‌های $w(i, j)$ همیشه ارزشی بین صفر و یک خواهد بود. با استفاده از اطلاعات فضای رنگ، معادله (10-2) را می‌توان برای محاسبه یک وزن برای هر پیکسل در نواحی پشتیبان به کار برد. کارشناسان [۱۷] این روش را طراحی کردند تا از نواحی پشتیبان متقارن بهره ببرند، یعنی وزن‌ها باید برای ناحیه پشتیبان هم در تصویر مرجع و هم تصاویر هدف از جفت استریو ورودی محاسبه شود. با استفاده از این وزن‌ها، می‌توان هزینه‌ها را مجتمع ساخت تا ارزشی برای هزینه مجتمع به دست آید:

$$A(x, y, d) = \frac{\sum_{i, j \in N(x, y)} w_r(i, j) \times w_t(i, j) \times C(i, j, d)}{\sum_{i, j \in N(x, y)} w_r(i, j) \times w_t(i, j)} \quad (11-2)$$

در معادله ۱۱-۲، w_r و w_t وزن‌های محاسبه شده به ترتیب برای تصاویر مرجع و هدف هستند و C از رابطه زیر به دست می‌آید:

$$C(i, j, d) = \min\left(\sum_{c \in \{r, g, b\}} |I_c(i, j) - I'_c(i, j, d)|, T\right) \quad (12-2)$$

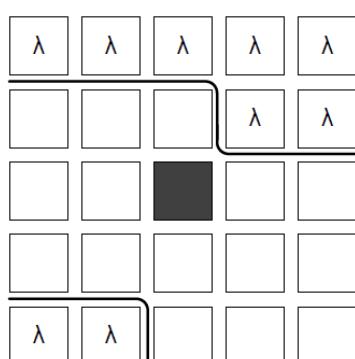
I_c شدت روشنایی رنگ و T مقدار آستانه‌ای است که حد هزینه انطباق را تعیین می‌کند. این به تجمع هزینه‌ای می‌انجامد که به خوبی از تأثیر پیکسل‌های کمارزش در نواحی پشتیبان می‌کاهد.

۳-۴-۲ تجمع هزینه بر مبنای بخش‌بندی

روش دیگر برای ترکیب اطلاعات رنگ در روند تجمع هزینه، تعریف پویای شکل و اندازه مناطق پشتیبان با استفاده از بخش‌بندی رنگ است. تمامی پیکسل‌ها در نواحی پشتیبان که دارای اختلاف‌هایی هستند که به میزان زیادی با پیکسل مرجع تفاوت دارد، نقاط کمارزش بوده و نباید در روند تجمع محسوب گردند. متأسفانه، اختلاف‌ها نامعلوم هستند بنابراین شکل نواحی پشتیبان را می‌توان در عوض به گونه‌ای تعریف کرد که تنها مقاطعی از تصاویر ورودی را پوشش دهند که به نظر

دارای رنگی مشابه است. این ناشی از آن فرض است که عمق به آرامی در میان بخش‌ها با رنگ یکسان در تصویر مرجع تغییر می‌یابد. پس در طی روند تجمع، ناپیوستگی‌های عمق در مرزهای این بخش‌بندی‌های رنگی قرار می‌گیرند [۱۵]. اندازه نواحی پشتیبان نقش بزرگی در دقت الگوریتم تطبیق استریو محلی ایفا می‌کند. وقتی اندازه پنجره خیلی کوچک باشد، این نواحی حاوی اطلاعات کافی نبوده و مستعد ابهاماتی می‌شوند، که این به نتایج نویزی می‌انجامد. از طرف دیگر، وقتی اندازه پنجره خیلی بزرگ باشد، آنها در نزدیکی ناپیوستگی‌های عمق حاوی پیکسل‌هایی با عمق‌های متفاوت می‌شوند.

ولین‌گام در این روش تجمع، بخش‌بندی تصویر مرجع به بخش‌هایی با رنگ مشابه از طریق روندی است که تفکیک رنگ جابجایی-میانگین^۱ نام دارد [۱۸]. با استفاده از این اطلاعات می‌توان از تأثیر پیکسل‌ها در ناحیه پشتیبانی که در بخش رنگی مشابه با پیکسل مرکزی از آن منطقه قرار نمی‌گیرند کاست. همانطور که در شکل ۶-۲ نشان داده شده، یک وزن کوچک (λ) بر هر پیکسلی در ناحیه پشتیبان اعمال شده که در بخش رنگی یکسانی با پیکسل مرکزی قرار نمی‌گیرد. در تصویر، خط پرنگ میان برخی از پیکسل‌ها نشانگر مرزهای بخش‌بندی رنگ است، همانطور که در طی روند تفکیک جابجایی میانگین محاسبه شد.



شکل ۶-۲: مثال از یک پنجره تجمع پویا با بخش‌بندی رنگ

۱- Mean-Shift

علت آنکه پیکسل‌های خارجی بطور کامل از نظر نیافتاده‌اند، حفاظت از تجمع در مقابل بخش‌بندی بیش‌ازحد است.

۴-۴-۲ تجمع هزینه پشتیبان بخش‌بندی^۱

فن دیگر برای تجمع هزینه تطبیق توسط نویسنده‌گان [۱۹] توسعه یافت تا روش‌هایی که پیش‌تر مطرح شدند، بهبود یابند. روش پشتیبانی بخش‌بندی با ترکیب عقاید روش‌های وزن انطباقی و برمبانای بخش‌بندی تهیه شد. آنها همچنین از وزن‌بندی نرمالیزه به همان شکلی که در [۱۷] استفاده و در معادله (۱۱-۲) نشان داده شد، استفاده کردند. آنها بحث کردند که عامل مجاورت در وزن‌های به کار رفته توسط [۱۷] دارای تأثیری منفی بر دقت بوده و به جای آن وزن‌ها را به‌گونه‌ای تعریف کردند که تنها حاوی عبارت اختلاف رنگ باشند (Δc_{ij}). وزن‌ها چنانچه پیکسل (j, i) در بخش رنگی یکسانی با پیکسل مرجع باشد، برابر با یک بوده، در غیراینصورت وزن اینگونه تعریف می‌شود:

$$w(i, j) = \exp\left(-\frac{\Delta c_{ij}}{\gamma_s}\right) \quad (13-2)$$

۴-۵-۲ اختلاف میان این رویکردهای تجمعی

تمامی روش‌های تجمع مطرح شده در اینجا با یک منطقه پشتیبان مربعی حول پیکسل موردنظر آغاز می‌شوند. سپس آنها وزن‌هایی را به تمامی پیکسل‌ها در ناحیه تخصیص می‌دهند تا از تأثیر آنها بر هزینه تطبیق بکاهند. آنها از لحاظ فنون به کار رفته برای تعیین این وزن‌ها تفاوت دارند. یک روش برای بصری‌سازی اختلافات در نواحی پشتیبان وزنی برای طرح‌های تجمعی مختلف، جستجوی نواحی است که برای نقاطی ویژه در یک تصویر استریو تولید شده‌اند. شکل ۷-۲ نشانگر نواحی پشتیبان با برنامه‌های تجمعی مختلف برای تعداد کمی نقاط انتخابی در تصویر عروسکی از مجموعه‌داده استریو میدل‌بوری است، مانند شکل ۷-۲ (الف). هرچه شدت پیکسل‌ها در ناحیه پشتیبان روشن‌تر باشد، تأثیر آن نقطه بر روند تجمع بیشتر است.

۱- Segmented Support Cost Aggregation



شکل ۲-۲: مناطق پشتیبانی وزن برای روش‌های تجمع هزینه مختلف

(الف) استخراج نقاط نمونه، (ب) روش وزن انطباقی، (ج) روش بخشندی، (د) روش پشتیبان بخشندی

شکل ۲-۷ (ب) نشانگر نواحی تولید شده با استفاده از روش وزن انطباقی است که نسبتاً پراکنده می‌باشد. از شکل ۲-۷ (ج)، مشخص است که چگونه روش برمبنای بخشندی، تنها هزینه پیکسل‌ها در یک بخش رنگی را به کار می‌برد. روش پشتیبان بخشندی در شکل ۲-۷ (د)، فنون وزن انطباقی و بخشندی هزینه را با استفاده از وزن‌های ۱، برای پیکسل‌های موجود در بخش رنگی ترکیب می‌کند، اما همچنین از وزن‌های پویا برای پیکسل‌ها، در خارج از مرز بخشندی استفاده می‌کند. هر دو روش برمبنای بخشندی، به نظر هزینه را بیشتر نسبت به منطقه پشتیبان مربعی تجمع می‌سازند تا روش وزن انطباقی. به همین دلیل است که نواحی پشتیبان بخشندی می‌توانند به طرز بهتری با ابهامات در نواحی بافت پراکنده کنار آیند.

۵-۵ محاسبه اختلاف و بهینه‌سازی

مرحله سوم در جایی است که ارزش‌های اختلاف محاسبه شده‌اند. برای الگوریتم‌های تطبیق استریو محلی، این مرحله حیاتی بوده و در بسیاری موارد صرفاً انتخاب کمترین ارزش تجمع هزینه را برای هر پیکسل شامل می‌شود. این روش معمولاً یک محاسبه محلی (WTA)^۱ در هر پیکسل است.

برای روش‌های سراسری این مرحله غیرحساس بوده و اغلب روش‌های سراسری از این مرحله برای اجرای کمینه‌سازی انرژی نسبت به کل تصویر استفاده می‌کنند، تا یک تابع اختلاف تعیین شود که انرژی سراسری را کمینه سازد. اختلاف اصلی میان روش‌های سراسری در رویه کمینه‌سازی به کار رفته برای این مرحله است.

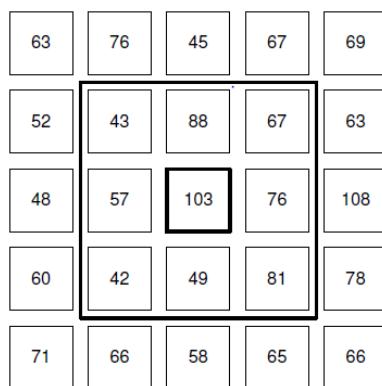
۶-۶ روش‌های اصلاح^۲

مرحله نهایی در یک الگوریتم تطبیق استریو محلی مرحله پس‌پردازشی است. هدف از این مرحله، همانطور که از نامش پیداست، تنها اصلاح نتایج به دست آمده از الگوریتم‌های است. انتخاب روش اصلاح، از انتخاب‌های صورت‌گرفته برای مراحل قبلی مستقل است. ورودی برای این مرحله یک نقشه جابجایی مکانی است و خروجی یک نسخه اصلاح شده از همان نقشه جابجایی مکانی است. همچنین می‌توان از یک روش اصلاح بطور ترتیبی بر یک نقشه جابجایی مکانی استفاده کرد. هدف از این مرحله در یک الگوریتم تطبیق استریو محلی اصلاح نتایج از طریق فنون مختلف است، روش‌های توصیف شده در زیر بخشی از آنها را تشکیل می‌دهند.

۱- Winner Takes All
۲- Refinement Methods

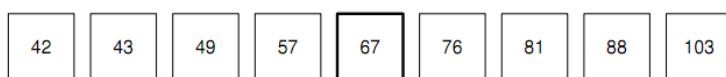
۱-۶-۲ فیلتر میانه برای نقشه‌های جابجایی مکانی

یک روش ساده برای بهبود یک نقشه جابجایی مکانی کاربرد یک فیلتر میانه بر تصویر جابجایی مکانی است. این به حذف برخی تطبیق‌های اشتباه کمک می‌کند. یک فیلتر میانه یک فیلتر پردازش تصویر با کاربرد متداول است که می‌تواند یک تصویر را هموار ساخته، در عین حال همچنان حاشیه‌ها را حفظ کند. این با جایگزین‌سازی هر پیکسلی در یک تصویر با ارزش میانه از پیکسل‌های مجاور آن عمل می‌کند. پیکسل‌های مجاور همگی پیکسل‌هایی در یک پنجره هستند که پیکسل مرتع را احاطه می‌کند. چنین پنجره‌ای معمولاً دارای یک اندازه ثابت از دست‌کم 3×3 پیکسل بوده، پیکسل مورد بررسی در مرکز آن قرار دارد. هرچه پنجره فیلتر بزرگ‌تر باشد، هموارسازی بیشتری صورت می‌گیرد.



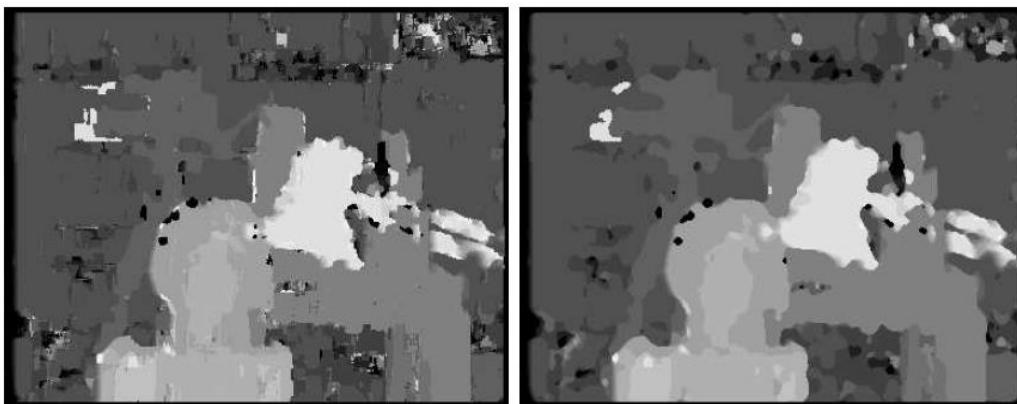
شکل ۸-۲: محاسبه مقدار متوسط برای یک پیکسل

در شکل ۸-۲، ارزش‌ها نشانگر شدت پیکسل اولیه و مربع نشانگر یک پنجره فیلتر مشابه از 3×3 پیکسل است. ارزش جدید برای پیکسل مرکزی با یافتن میانه تمامی پیکسل‌های موجود در پنجره فیلتر محاسبه شده است. شکل ۹-۲ در زیر نشانگر شدت پیکسل ذخیره شده است.



شکل ۹-۲: دسته‌بندی شدت پیکسل برای تعیین میانه

این شدت پیکسل جدید در موقعیت معادل در نقشه جابجایی مکانی اصلاحی خواهد بود. همین روند برای هر پیکسلی در تصویر جابجایی مکانی اولیه تکرار می‌شود. شکل ۲-۱۰ نشانگر یک مثال از نقشه جابجایی مکانی نویزی در قبل و پس از فیلترشدن با یک فیلتر میانه با یک پنجره 5×5 است.



شکل ۲-۱۰: نقشه جابجایی مکانی نویزی، قبل و بعد از استفاده از یک فیلتر میانه 5×5

یقیناً مهارت‌انگیزترین بخش از به کارگیری یک فیلتر میانه پرداختن به حاشیه‌های تصویری است که تحت فیلتر قرار می‌گیرد. این مستلزم یافتن میانه برای یک پیکسل درجایی است که تمامی پیکسل‌ها در پنجره فیلتر دارای یک ارزش نیستند. روش‌های مختلف برای حل این مسئله وجود دارد. یک رویکرد، عدم اصلاح حاشیه‌های تصویر در جایی است که این مسئله رخ می‌دهد. متداول‌ترین رویکرد برای این مسئله پرکردن ارزش‌های ناپیدا در پنجره فیلتر با رقم صفر است.

۲-۶-۲ تطبیق دو- جهته

تطبیق دو- جهته یک روش اصلاح ساده و مؤثر برای الگوریتم‌های تطبیق استریوو است. این روش به‌ویژه در نواحی انسدادی مؤثر است. این روش با اغلب روش‌های اصلاح تفاوت دارد، از آن جهت که تنها نقشه جابجایی مکانی اصلاح‌نشده را به عنوان ورودی دریافت نمی‌کند. در عوض، شامل تولید یک نقشه جابجایی مکانی دوم با استفاده از تصویر مقابل جفت تصویر ورودی به عنوان تصویر مرجع است. درنتیجه، دو نقشه جابجایی مکانی تولید شده، یک نقشه جابجایی مکانی چپ به راست و یک نقشه

جابجایی مکانی راست به چپ. سپس یک نقشه جابجایی مکانی اصلاح شده نهایی با ترکیب نتایج این دو نقشه جابجایی مکانی تولید می‌شود.

نواحی نیمانسدادی در یک جفت تصویر استریو، که در آن یک نقطه در یک تصویر به دلیل انسداد در تصویر دوم وجود ندارد، یک مسئله عمدۀ برای الگوریتم‌های تطبیق استریو است. وقتی چنان نقاطی صریحاً رفع نشده باشند، یک الگوریتم تطبیق در امتداد یک خط پیمایش بیشتر از حد لازم جستجو می‌کند تا نقطه‌ای را در پس زمینه بباید که شبیه به نقطه نیمانسدادی در تصویر مرجع باشد. این باعث می‌شود که پیکسل‌ها از یک ارزش اختلاف برخوردار می‌شوند که غیرضرورتاً بزرگ است. با بررسی عرضی دو نقشه جابجایی مکانی تولید شده، چنان نقاطی را می‌توان یافته و اصلاح نمود.

برای انجام این بررسی ضرورت دارد که اختلاف‌ها از یکی از نقشه‌های جابجایی مکانی تبدیل شوند، به‌گونه‌ای که با نقشه جابجایی مکانی دیگر، همتراز گردند. برای مثال، تصویر استریو چپ را با یک نقطه ویژه x_L درنظر بگیرید. پس نقطه تطبیقی در تصویر راست باید $x'_R = x_L + d_L(x_L)$ باشد که در آن $d_L(x_L)$ ارزش اختلاف در نقشه جابجایی مکانی چپ به راست در موقعیت x_L است. بطور مشابه، نقطه x_R از تصویر استریو راست به تصویر چپ می‌دهد: $x'_L = x_R + d_R(x'_R)$. حال می‌توان ارزش حقیقی x_L را با نقطه پیش‌بینی شده x'_L مقایسه کرد تا خطای میان نقشه‌های جابجایی مکانی چپ و راست تعیین شود. این خطای می‌توان اینگونه نوشت:

$$error = |x_L - x'_L| \quad (14-2)$$

وقتی این خطای بیش از یک آستانه معین باشد (معمولًاً ۲ پیکسل)، می‌توان فرض کرد که نقطه x_L یک نقطه نیمانسدادی است. کارشناسان [۲۰] نشان دادند که تطبیق دو-جهته مؤثرترین روش برای تشخیص نواحی نیمانسدادی است. با علامت‌گذاری تمامی پیکسل‌های نیمانسدادی، مرحله بعدی درون‌یابی ارزش‌های جدید برای آنهاست. این را می‌توان با استفاده از یک روش درون‌یابی خطی انجام داد. ساده‌ترین روش برای پرکردن این شکاف‌ها جستجوی شکاف‌ها در خطوط پیمایش و پرکردن آنها

با کمترین مقدار از دو اختلاف مجاور است. این بدان خاطر است که نقاط نیمانسدادی در جایی ایجاد شده‌اند که ارزش‌های اختلاف بالاتر (اشیاء پیش‌زمینه)، نقاط با اختلاف کمتر (پس‌زمینه) را مسدود ساخته‌اند.

۷-۲ مرور اجمالی الگوریتم‌های تطبیق استریو سراسری

الگوریتم‌های سراسری مرحله تجمع در تقسیم چهار مرحله‌ای الگوریتم‌های محلی را انجام نداده، اما در عوض سعی می‌کنند تا یک مشتق اختلاف را بیابند که یک تابع هزینه سراسری را به حداقل می‌رساند. چنین تابع هزینه سراسری معمولاً یک عبارت داده‌ای و عبارت همواری را ترکیب می‌کنند. بُعدی که الگوریتم‌های تطبیق سراسری را متمایز می‌سازد عمدتاً رویه کمینه‌سازی ویژه‌ای است که بر آن اتکا دارند. برخی از این روش‌ها عبارتند از برنامه‌نویسی پویا [۱۳] و برش‌های نموداری [۲۱، ۲۲، ۲۳]. در این پایان‌نامه به دلیل تمرکز بر روش‌های محلی و ارائه یک الگوریتم بلاذرنگ، روش‌های انطباق استریو سراسری به طور مختصر معرفی شده است.

اولین گام در تمامی الگوریتم‌های سراسری که خیلی شبیه به الگوریتم‌های محلی بوده، محاسبه هزینه تطبیق است. این با استفاده از روش‌هایی که در بخش قبل توصیف شد، همچون اختلاف شدت مربعی و اختلاف شدت مطلق انجام می‌شود. هدف از این گام تنها ایجاد روشی برای ارزیابی هزینه‌های تطبیق است. همانطور که پیش‌تر مطرح شد، الگوریتم‌های سراسری معمولاً مرحله دوم را اجرا نمی‌کنند (مرحله تجمع)، که الگوریتم‌های محلی اکثریت محاسبات را در آن انجام می‌دهند. گام سوم، یعنی محاسبه اختلاف، همان‌جایی است که الگوریتم‌های سراسری عمدۀ محاسبات خود را انجام می‌دهند. اغلب روش‌های سراسری در یک چهارچوب کمینه‌سازی انرژی فرمول‌بندی شده‌اند، هدف از این مرحله یافتن یک تابع اختلاف d است که انرژی سراسری $E(d)$ را کمینه می‌سازد:

$$E(d) = E_{data}(d) + \lambda E_{smooth}(d) \quad (15-2)$$

که در آن عبارت $E_{data}(d)$ همان عبارت همواری و λ یک وزن ثابت است که بر عبارت همواری اعمال شده است. عبارت داده به ارزیابی آن می‌پردازد که تابع اختلاف d تا چه میزان با جفت تصویر ورودی انطباق می‌یابد. عبارت داده را می‌توان اینگونه نوشت:

$$E_{data}(d) = \sum_{(x,y)} C(x,y,d(x,y)) \quad (16-2)$$

در این نمونه نمایشی، $C()$ عبارت است از هزینه تطبیق. عبارت همواری $E_{smooth}(d)$ یک شیوه برای ارائه فرضیه همواری است که براساس الگوریتم ویژه تهیه می‌شود. برای ساده‌سازی روند بهینه‌سازی، عبارت همواری اغلب تنها به اختلافات پیکسل مجاور محدود شده است:

$$E_{smooth}(d) = \sum_{(x,y)} \rho(d(x,y) - d(x+1,y)) + \rho(d(x,y) - d(x,y+1)) \quad (17-2)$$

دراینجا ρ نشانگر تابعی از تفاوت اختلاف است که می‌بایست بطور همنواخت افزایش یابد. اگر ρ در قالب یک تابع مرتبی تعریف می‌شد، d درامتداد کل تصویر هموار بود که این به دقت ضعیف در مزهای شیء می‌انجامید. توابع انرژی که از این مسئله پرهیز می‌کنند را عموماً توابع حفظ‌کننده ناپیوستگی^۱ نامیده‌اند [۳].

یک رویکرد متداول دیگر فرمول‌بندی عبارت همواری به‌شیوه‌ای است که به اختلافات شدت وابسته باشد [۲۲]. این رویکرد مؤثر است چرا که باعث می‌شود تا اختلاف ادامه یافته، با حاشیه‌ها در شدت تصویر یا لبه‌های رنگی انطباق یابد. اثبات شده که چنین روش‌هایی بسیار مؤثر هستند و بسیاری از الگوریتم‌های برتر از آن استفاده می‌کنند. با تعریف تابع انرژی، مرحله بعدی یافتن یک حداقل (محلی) است. بسته به تابع انرژی فهرست‌شده، روش‌های مختلفی برای انجام آن وجود دارد، مانند برنامه‌نویسی پویا و بطور جدیدتر روش‌های برش نموداری، که نشان داده‌اند بسیار مؤثر هستند.

^۱- Discontinuity-Preserving

۸-۲ قیود مورد استفاده در تطابق استریو

جهت کاهش فضای جستجو می‌توانیم از قیودی در تصاویر استریو استفاده نماییم. در [۲۶، ۲۵، ۲۴]

این قیود ذکر شده‌اند، که در زیر پاره‌ای از آنها به اختصار بیان شده است:

▪ **قید تشابه رنگ^۱ یا شدت روشنایی:** از آنجا که هر دو بخش تطابق یافته دو تصویر،

متناظر با یک ناحیه از صحنه می‌باشند، این دو بخش در حالت ایده‌آل از نظر رنگ

یا شدت روشنایی در دو تصویر یکسان ظاهر می‌شوند.

▪ **قید همقطبی:** با توجه به شکل ۲-۱ اگر u' و u تصاویر نقطه سه بعدی X در صفحات

تصویر چپ و راست باشند، این دو نقطه روی خطوط همقطبی متناظر یعنی l ، l' واقعند. به

عبارت دیگر برای هر نقطه مثل u در صفحه تصویر چپ، نقطه متناظرش در صفحه تصویر

راست می‌بایست روی خطی باشد که از محل تقاطع صفحه همقطبی و صفحه تصویر ایجاد

می‌شود. به طور خلاصه در هندسه تصویر برداری استریو، محدودیت همقطبی، ناحیه جستجو

برای یافتن نقطه متناظر در تصویر دیگر را، از دو بعد به یک بعد کاهش می‌دهد. در استریوی

موازی با تصاویر یکسوزده، خطوط همقطبی، خطوط افقی متناظر دو تصویر می‌باشند.

▪ **قید حداقل-حداکثر مقدار جابجایی مکانی:** اجسام موجود در صحنه به طور معمول

دارای حداقل فاصله مشخصی از دوربین می‌باشند. بنابراین طبق رابطه (۳-۱) مقدار حداکثر

جابجایی مکانی، d_{max} محدود می‌گردد. از آنجا که مقدار عمق همواره مثبت است مقدار

جابجایی مکانی، بزرگتر از صفر بوده ضمن آنکه در صحنه‌های داخلی بر خلاف صحنه‌های

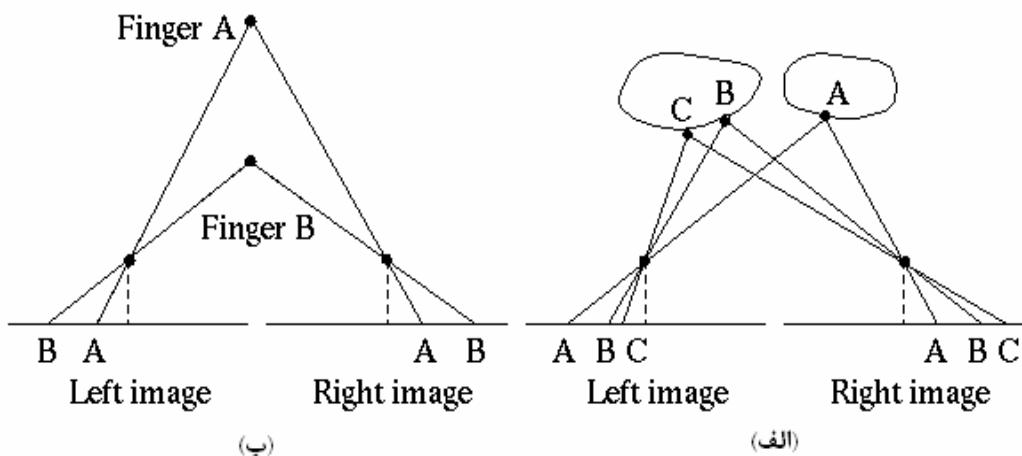
خارجی حداکثر عمق صحنه که متناظر با دورترین نقطه زمین می‌باشد، محدود بوده و

بنابراین مقدار جابجایی مکانی از یک مقدار مشخص مثبت d_{min} ، بزرگتر می‌گردد. حداقل و

حداکثر جابجایی مکانی، فضای جستجو را در فرایند تطابق، محدود می‌کند.

۱- Color Constancy

▪ **قید ترتیب^۱:** این محدودیت بیان می‌کند که ترتیب مکانی نقاط در روی تصویر B حفظ می‌شود. به عبارت دیگر اگر در تصویر چپ نقطه A در سمت راست نقطه B قرار داشته باشد آنگاه متناظر A در تصویر راست هم در سمت راست متناظر B قرار دارد. این محدودیت چنانچه در صحنه، ناپیوستگی‌های بزرگ وجود نداشته باشد قابل اعمال است. شکل ۱۱-۲، در قسمت (الف)، برقراری و در قسمت (ب)، عدم برقراری این محدودیت را در دو صحنه مختلف نشان می‌دهد.



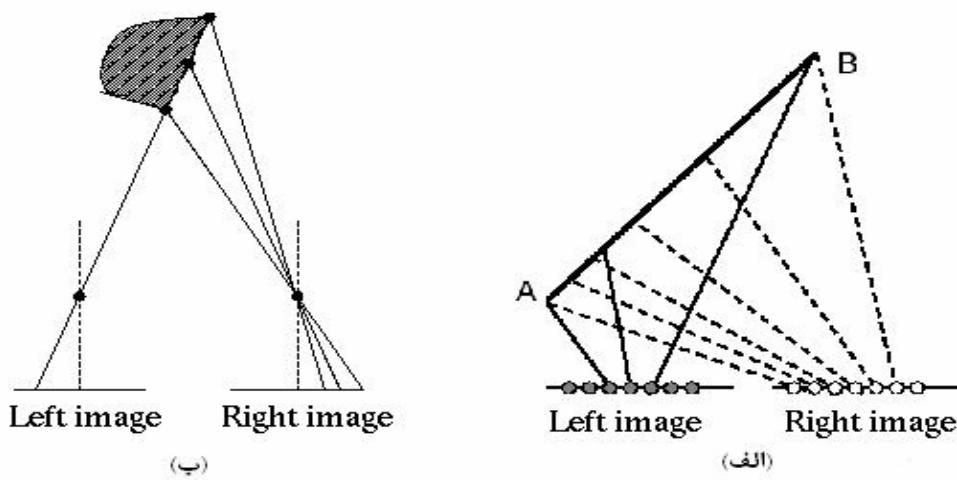
شکل ۱۱-۲: (الف) برقراری محدودیت ترتیب و (ب) عدم برقراری محدودیت ترتیب

▪ **قید یکتایی^۲:** این محدودیت بیان می‌کند که متناظر هر نقطه در صفحه تصویر چپ، یک و تنها یک نقطه از تصویر راست می‌باشد. گرچه این محدودیت برای بسیاری از صحنه‌های طبیعی موجود برقرار است اما می‌توان حالات خاصی را یافت که این محدودیت برقرار نباشد. شکل ۱۲-۲، قسمت (الف)، نمونه‌ای از این حالت را نشان می‌دهد. علاوه بر این در [۲۷] نشان داده شده است که بر روی سطوح شیبدار

۱- Ordering Constraint
۲- Uniqueness

امکان انطباق تعداد نامساوی از پیکسل‌ها وجود دارد. چنین حالتی در قسمت (ب) از شکل ۱۲-۲ نشان داده شده است.

▪ **قید پیوستگی جابجایی مکانی**: صرفنظر از ناپیوستگی‌های عمق در مرز اشیاء، اجسام موجود در محیط غالباً دارای سطوحی با تغییرات عمق ناچیز یا پیوسته می‌باشند. از آنجا که لبه‌ها یا مرزها، بخش ناچیزی از نقاط صحنه را شامل می‌شوند جابجایی در اغلب نقاط تصویر پیوسته تغییر می‌کند. تشخیص تغییرات ناگهانی جابجایی مکانی نقاط واقع بر روی مرزهای ناپیوستگی عمق و در عین حال، حفظ پیوستگی جایجایی مکانی در سایر نقاط تصویر منجر به نتایج بهتر می‌شود.



شکل ۱۲-۲: عدم برقراری قید یکتاوی در دو حالت مختلف

۹-۲ ارزیابی استریو میدلبوری

در حال حاضر بهترین مقایسه الگوریتم‌های تطبیق استریو، ارزیابی استریو میدلبوری است که در [۶] معرفی شد. بستر آزمایش برخط نتایج بهترین الگوریتم‌های تطبیق استریو جهان را با استفاده از چهار مجموعه از تصاویر مقایسه می‌کند که حاوی تمامی مشکلاتی هستند که چنین

۱- Continuity of Disparity Constraint

سیستم‌هایی در جهان حقیقی با آن مواجه می‌شوند. بستر آزمایش همچنین حاوی نقشه‌های جابجایی مکانی صحیح می‌باشد، به‌گونه‌ای که نقشه‌های اختلاف تولید شده را می‌توان مقایسه نمود. الگوریتم‌ها با نقشه‌های جابجایی مکانی، با کوچکترین درصد از پیکسل‌های بد، بهترین رتبه را دریافت نموده و سپس رتبه‌بندی کلی الگوریتم‌ها با محاسبه رتبه میانگین در بین چهار تصویر آزمایشی تعیین شده است. شکل ۱۳-۲ نشانگر ۸ الگوریتم تطبیق استریوو است که براساس بستر آزمایش میدلبوری بهترین رتبه را دریافت کرده‌اند. با بررسی می‌توان فهمید که چگونه یک سهم گستردگی از الگوریتم‌ها در بین ۸ الگوریتم برتر در طی سه سال اخیر منتشر شدند و آنکه بسیاری از آنها نوعی از روند تفکیک را شامل می‌شوند به‌گونه‌ای که اطلاعات رنگ را برای تعیین ناپیوستگی‌های جابجایی مکانی ترکیب می‌سازند.

ارزیابی روش‌ها در شکل ۱۳-۲، در نواحی غیرانسدادیافته (*nonocc*)، کل تصویر (*all*) و مجاورت ناپیوستگی‌های عمق (*disc*) صورت گرفته است.

Algorithm	Avg.	Tsukuba ground truth			Venus ground truth			Teddy ground truth			Cones ground truth			Average percent of bad pixels (explanation)		
		Rank	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc		
TSGO [143]	10.5		<u>0.87</u> 4	<u>1.13</u> 1	4.66 6	<u>0.11</u> 7	0.24 9	1.47 10	<u>5.61</u> 38	8.09 17	13.8 31	<u>1.67</u> 1	<u>6.16</u> 1	4.95 1		4.06
JSOSP+GCP [151]	11.7		<u>0.74</u> 1	1.34 8	3.98 1	<u>0.08</u> 2	<u>0.16</u> 1	1.15 2	<u>3.96</u> 14	10.1 33	11.8 17	<u>2.28</u> 15	7.91 28	6.74 18		4.18
ADCensus [82]	14.1		<u>1.07</u> 19	1.48 17	5.73 22	<u>0.09</u> 3	0.25 12	1.15 2	<u>4.10</u> 16	6.22 7	10.9 13	<u>2.42</u> 21	7.25 15	6.95 22		3.97
AdaptingBP [16]	17.8		<u>1.11</u> 22	1.37 10	5.79 24	<u>0.10</u> 5	0.21 8	1.44 9	<u>4.22</u> 18	7.06 15	11.8 18	<u>2.48</u> 25	7.92 30	7.32 30		4.23
CoopRegion [39]	18.2		<u>0.87</u> 8	1.16 2	4.61 5	<u>0.11</u> 6	0.21 6	1.54 14	<u>5.16</u> 30	8.31 20	13.0 25	<u>2.79</u> 42	7.18 14	8.01 49		4.41
CCRADAR [152]	22.5		<u>1.15</u> 25	1.42 15	6.23 38	<u>0.15</u> 19	0.27 15	1.89 24	<u>5.39</u> 33	10.6 38	14.7 42	<u>2.01</u> 2	7.37 17	5.88 2		4.75
RDP [87]	23.7		<u>0.97</u> 11	1.39 12	5.00 11	<u>0.21</u> 40	0.38 30	1.89 24	<u>4.84</u> 22	9.94 32	12.6 22	<u>2.53</u> 28	7.69 21	7.38 31		4.57
MultiRBF [129]	24.1		<u>1.33</u> 49	1.56 22	6.02 33	<u>0.13</u> 11	0.17 3	1.84 21	<u>5.09</u> 28	6.36 8	13.4 29	<u>2.90</u> 50	6.76 8	7.10 27		4.39

شکل ۱۳-۲: هشت الگوریتم برتر توسط بستر ارزیابی میدلبوری

۱۰-۲ نتیجه‌گیری

هدف از این فصل تعریف روش‌های مختلف و مورد استفاده در مقالات گوناگون، در محاسبه مراحل مختلف الگوریتم‌های تطبیق استریو محلی بود. همچنین بطور مختصر در مورد الگوریتم‌های استریو سراسری بحث شد. بزرگترین مزیت روش‌های محلی آن است که از لحاظ توان محاسباتی پیچیدگی کمتری داشته و بنابراین برای مصارف بلادرنگ مناسب هستند.

فُصْل سوم

الگوريتم پيشنهادي

۱-۳ مقدمه

با توجه به هندسه تصویربرداری، مسائل اصلی هر الگوریتم استریو عبارتند از [۲۷]:

۱- انتخاب و استخراج اجزای تصویر: در این مرحله اجزای مهم و مورد نظر مثل لبه‌ها، پاره‌خط‌ها و مرزها استخراج می‌شوند.

۲- تطبیق و یافتن اجزای متناظر : در این مرحله اجزای انتخاب شده و استخراج شده از دو تصویر، با استفاده از توابع معیار مناسب تطبیق داده شده و اجزای متناظر در دو تصویر مشخص می‌شوند.

۳- بازسازی: در این مرحله با داشتن اجزای متناظر در دو تصویر چپ و راست و اطلاعات مربوط به دوربین‌ها و هندسه سیستم تصویربرداری، یعنی وضعیت دوربین‌ها نسبت به هم، موقعیت ۳D اجزاء بدست می‌آیند. یکبار تصویر چپ را مرجع قرار می‌دهند و نقاط متناظر را در تصویر راست محاسبه می‌کنند و در مرحله بعد تصویر راست را مرجع قرار می‌دهند و نقاط متناظر از روی تصویر سمت چپ محاسبه می‌شود.

۴- پس پردازش^۱: در این مرحله نقاط انسدادیافته و نامعتبر شناسایی و در نقشه جابجایی مکانی نهايی تصحیح می‌شود. معمولاً در این قسمت ابتدا روی نقشه‌های جابجایی مکانی بدست آمده از مرحله قبل یک فیلتر میانه وزن دار [۲۸] می‌گذارند و سپس نقاط انسدادیافته که درست تشخیص داده نشده است، مشخص و تصحیح می‌شود.

همان‌طور که در قبل اشاره شد، الگوریتم‌های استریو به دو روش کلی تقسیم‌بندی می‌شوند: روش‌های محلی و روش‌های سراسری که تمرکز ما در اینجا بیشتر بر روی روش‌های محلی می‌باشد. در این فصل ابتدا در بخش ۲-۳ الگوریتم تطابق استریو پیشنهادی در حالت گسسته که تنها برای نقاط خاصی از تصویر مانند گوشه‌ها و لبه‌ها محاسبه می‌شود و در ناوبری ربات سیار در محیط کاربرد

۱- Post Processing

دارد و سپس در بخش ۳-۳ الگوریتم بازیابی نقشه جابجایی مکانی متراکم که محصول روش‌های مبتنی بر پیکسل بوده و برای تمامی نقاط تصویر محاسبه می‌گردد، ارائه می‌شود.

۲-۳ الگوریتم تطابق استریو با استفاده از گوشه‌یابی

در این قسمت از بانک تصاویر میدل‌بوری استفاده می‌نماییم. تصویر سمت چپ را به عنوان تصویر مرجع و تصویر سمت راست را به عنوان تصویر هدف در نظر می‌گیریم. این تصاویر در فضای رنگ RGB و یکسو شده هستند. جفت تصویر چپ و راست مربوط به Tsukuba در شکل ۱-۳ نشان داده شده است. در قسمت (الف) و (ب) از شکل مذکور، به ترتیب تصویر چپ و راست ارائه گردیده است.



(ب)



(الف)

شکل ۱-۳: تصویر Tsukuba : الف- تصویر چپ، ب- تصویر راست

به دلیل متداول بودن روش گوشه‌یابی هریس در مقالات کنونی، ما نیز روش گوشه‌یابی هریس را بر روی دو تصویر ورودی برای یافتن ویژگی‌های موردنظر، که گوشه‌ها می‌باشند اعمال می‌کنیم. برای تشخیص نقاط گوشه در تصویر، به انتخاب نقاطی که در آن سطح خاکستری حداقل در دو جهت مختلف تغییر قابل توجهی دارند، نیاز داریم. مراحل الگوریتم به شرح زیر است [۲۹، ۳۰]:

۱) مشتق را در راستای x و y با استفاده از عملگر سوبل بدست می‌آوریم. که از کانولوشن تصاویر ورودی با دو ماتریس رابطه (۱-۳) بدست می‌آید.

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & -2 & -1 \end{bmatrix} \quad (1-3)$$

۲) پنجره‌ای با اندازه $5*5$ اطراف هر پیکسل تصویر انتخاب کرده و سپس ماتریس C را به صورت معادله (۲-۳) محاسبه می‌کنیم که d_x و d_y مشتق در راستای محورهای x و y می‌باشد:

$$C = \begin{bmatrix} \sum d_x^2 & \sum d_x d_y \\ \sum d_x d_y & \sum d_y^2 \end{bmatrix} \quad (2-3)$$

توجه داریم که تمام پیکسل‌های داخل پنجره در مجموع رابطه بالا محسوب می‌شوند.
۳) اگر نقطه مورد نظر یک گوشه واقعی باشد، ماتریس C یک ماتریس رتبه کامل خواهد بود. در نتیجه از مقادیر ویژه ماتریس C برای تعیین اینکه آن نقطه گوشه است یا نه، استفاده می‌کنیم.

۴) در هر پیکسل، قدرت گوشه را از معادله (۳-۳) بدست می‌آوریم:

$$\text{Cornerstrength} = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (3-3)$$

که k بطور معمول بین 0.06 تا 0.04 انتخاب می‌شود. λ_1 و λ_2 مقادیر ویژه ماتریس C می‌باشند. در عمل، تریس C برابر با $(\lambda_1 + \lambda_2) \lambda_1 \lambda_2$ و دترمینان آن $\lambda_1 \lambda_2$ می‌شود. از این رو محاسبه ^۱ ماتریس C نیازی نیست.

۵) سپس مقدار آستانه‌ای را برای قدرت گوشه در نظر می‌گیریم تا تعیین کنیم که نقطه مورد نظر گوشه است یا نه. مقدار قدرت گوشه برای نقاط گوشه واقعی بزرگ، برای لبه‌ها منفی و در سطوح هموار خیلی کوچک خواهد بود. نمونه‌ای از ویژگی‌های استخراج شده در جفت تصاویر شکل ۲-۳ با نقاط سفید رنگ نشان داده شده است. برای وضوح بهتر، تعداد ۱۰ ویژگی را در

^۱- Singular Value Decomposition

جفت تصاویر استریو ورودی، در نظر گرفته‌ایم.



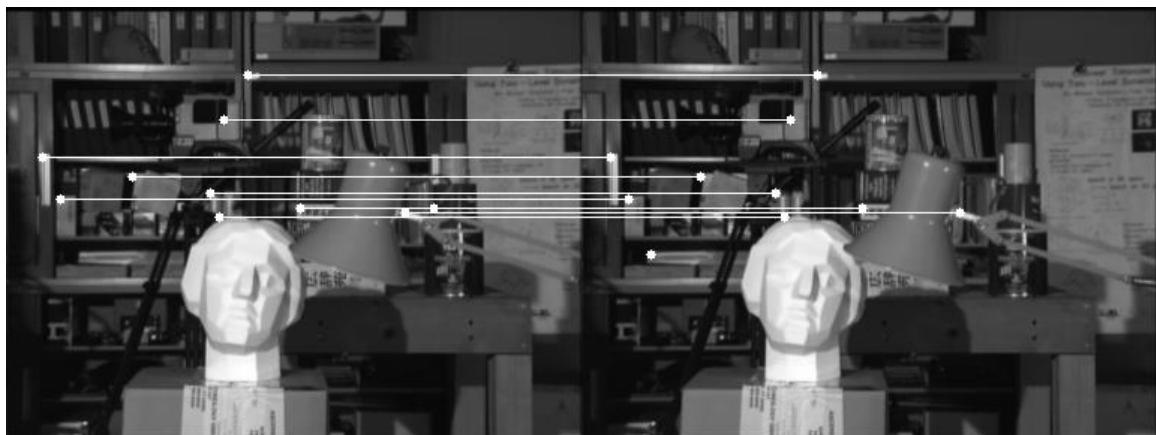
شکل ۲-۳: ویژگی‌های استخراج شده در تصویر Tsukuba

۶) سپس برای تعیین تطبیق بین پنجره‌های دو تصویر ورودی ما به مقایسه سطوح خاکستری در اطراف نقاط گوشه انتخابی در پنجره مورد نظر در تصویر چپ با سطوح خاکستری در اطراف پیکسل متناظر در تصویر راست با استفاده ازتابع هزینه پیشنهادی در معادله (۴-۳) می‌پردازیم، برای این کار از یک پنجره مربعی با اندازه ۳۵ استفاده نمودیم.

$$T = \sum_{i=n} \sum_{j=m} \min[|(I_1(u+i, v+j) - m_1) - (I_2(u+d, v+j) - m_2)|, \eta_1] + \sum_{i=n} \sum_{j=m} \min[|d_x I_1(u+i, v+j) - d_x I_2(u+d, v+j)|, \eta_2] \quad (4-3)$$

که در این معادله، I_1 و I_2 مقدار شدت پیکسل‌ها در تصاویر چپ و راست، m_1 و m_2 مقادیر میانگین شدت پیکسل‌ها در پنجره انتخاب شده در تصاویر چپ و راست می‌باشد و $d_x I_1$ و $d_x I_2$ مشتق در راستای محور x پیکسل‌ها در پنجره مفروض است، T تابع تطابق پیشنهادی در دو پنجره تصویر چپ و راست می‌باشد، η_1 و η_2 مقادیر برش می‌باشند که در پیاده‌سازی، مقادیر ۷ و ۲ در نظر گرفته شده است. شکل ۳-۳ خروجی حاصل از اعمال الگوریتم پیشنهادی را بر روی جفت تصویر گرفته شده از مجموعه داده میدل‌بوری نشان می‌دهد، همانطوری که ملاحظه می‌کنید تعداد ۹ تطابق از ۱۰ ویژگی

یافته شده در هر جفت تصویر ورودی را، به درستی تشخیص داده است. شکل ۳-۴ بلوک دیاگرام روش پیشنهادی را نشان می‌دهد، n در شکل ۳-۴ بیان کننده تعداد ویژگی‌ها می‌باشد.



شکل ۳-۳: تطبیق ویژگی‌های استخراج شده در تصویر Tsukuba

۳-۳ الگوریتم استخراج نقشه جابجایی مکانی متراکم

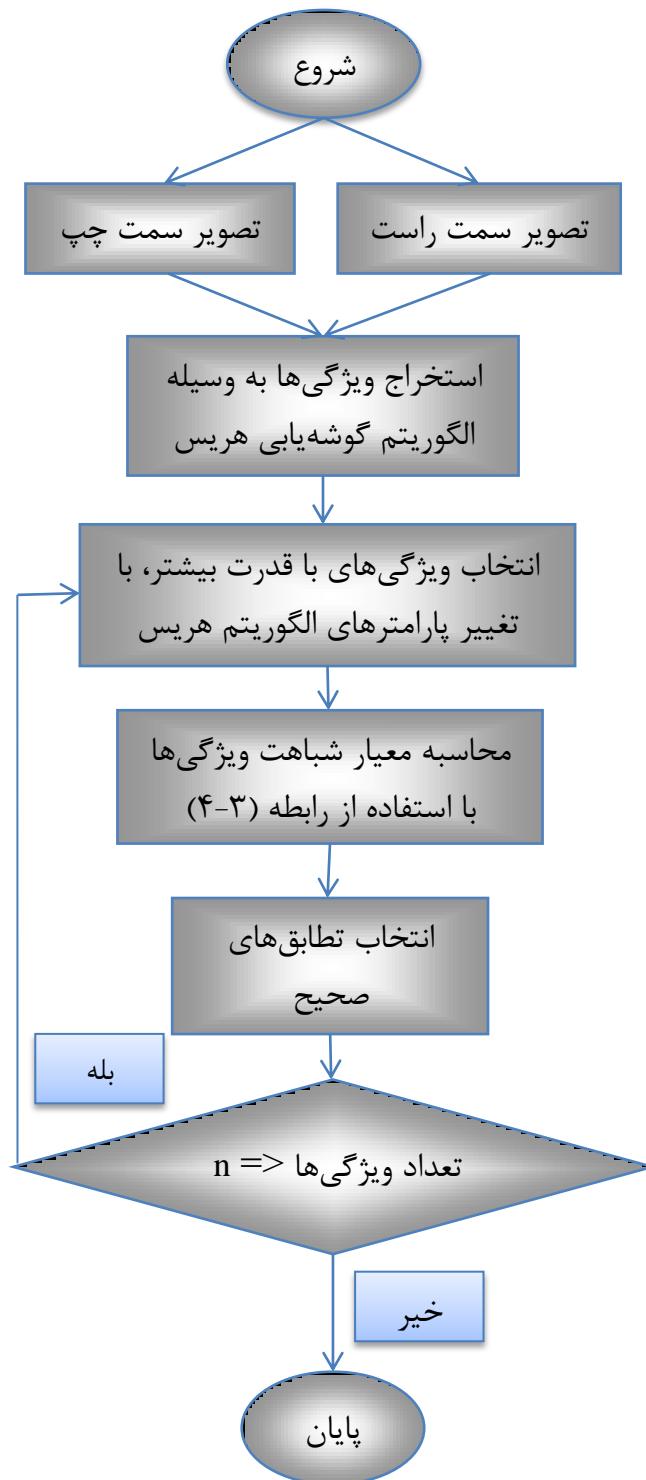
اکنون که دورنمای کلی کار معرفی شد، در این بخش روش پیشنهادی برای استخراج نقشه جابجایی مکانی را با جزئیات بیشتری بیان می‌نماییم. استخراج نقشه جابجایی مکانی مهمترین قسمت هر الگوریتم استریوو است و اکثر محققان به آن توجه دارند، چرا که افزایش سرعت اجرا و دقت الگوریتم به این مرحله وابسته است. الگوریتم پیشنهادی، جهت یافتن نقشه جابجایی مکانی به شرح زیر است:

۱-۳-۳ مرحله اول

این بخش شامل افزایش وضوح تصویر است. در اولین گام، تصحیح گاما^۱ را بر روی هر دو تصویر ورودی اعمال می‌نماییم. سپس فیلتر گوسی با اندازه 3×3 را روی دو تصویر اعمال می‌کنیم، این کار

۱- Gamma Correction

سبب می شود تصویر در مرزها و لبه ها صاف تر شود و تغییرات ناگهانی شدت روشنایی تاثیر کمتری بر روی محاسبه تابع هزینه داشته باشد، روش مورد نظر در بخش ۳-۱-۱ شرح داده شده است.



شکل ۳-۴: بلوک دیاگرام یافتن تطابق ویژگی ها در جفت تصاویر استریو

۱-۳-۳ تصحیح گاما

در این تبدیل رابطه سطوح روشنایی به صورت معادله (۵-۳) می‌باشد:

$$s = T(r) = c \times r^\gamma \quad (5-3)$$

پارامترهای رابطه مثبت می‌باشند. c مقدار شدت روشنایی پیکسل ورودی و γ یک مقدار ثابت می‌باشد. به صورت کلی توابع لگاریتمی و نمایی را می‌توان در این روش نشان داد، برای مثال اگر $\gamma < 1$ باشد، به تبدیل لگاریتمی تبدیل می‌شود. در اینجا برای گاما، از مقدار 0.7 برهه برده‌ایم. همان‌طور که در شکل ۳-۵ ملاحظه می‌شود، بعد از اعمال عملیات بالا بر روی تصاویر ورودی شکل ۳-۱، تصاویر روشن‌تر می‌شود که شاید نتوان در نگاه اول تغییر زیادی را مشاهده نمود، ولی در نتیجه خروجی، عملکرد را بطور مطلوبی افزایش می‌دهد.



(ب)



(الف)

شکل ۳-۵: تصویر Tsukuba : الف - قبل از تصحیح گاما، ب - بعد از تصحیح گاما

۲-۳-۳ مرحله دوم

همان‌طور که در قبلاً گفته شده، اولین گام در هر الگوریتم تطبیق استریوو یافتن راهی برای محاسبه هزینه تطبیق میان دو پیکسل در یک جفت تصویر استریوو است. هدف کلی از یک الگوریتم تطبیق استریوو آن است که هر پیکسلی در تصویر مرجع را با یک پیکسل در تصویر دیگر تطبیق دهد. به دلیل

یکسو بودن تصاویر ورودی، برای یافتن یک تطبیق برای یک پیکسل در تصویر مرجع، باید تنها درامتداد ردیف افقی از پیکسل‌ها در تصویر تطبیقی عملیات جستجو را انجام داد.

۱-۲-۳ محاسبه هزینه

مقدار هزینه بیان کننده این است که پیکسل i در تصویر I چقدر با پیکسل مشابه در تصویر ' I' که با بردار η تغییر مکان داده می‌شود، مشابه می‌باشد. ما در اینجا از هزینه تطبیق پیکسل- مبنا با اختلاف مطلق رنگ و گرادیان مانند معادله (۳-۶) در نقاط تطبیق، با تغییر معادله (۳-۴) استفاده می‌کنیم:

$$T = (1 - \alpha) \times \sum_{i=n} \sum_{j=m} \min[|(I_1(u+i, v+j) - m_1) - (I_2(u+d, v+j) - m_2)|, \eta_1] + (\alpha) \times \sum_{i=n} \sum_{j=m} \min[|d_x I_1(u+i, v+j) - d_x I_2(u+d, v+j)|, \eta_2] \quad (3-6)$$

که α در رابطه بالا، برای ایجاد توازن بین رنگ و گرادیان به معادله (۳-۴) اضافه شده است. در این قسمت از بالا به پایین و از چپ به راست شروع به جدا کردن پنجره‌های $n \times n$ حول پیکسل‌های تصاویر سمت چپ و راست می‌نماییم. سپس بلوک اول از تصویر سمت چپ را انتخاب می‌کنیم و در تصویر سمت راست با استفاده از دو قید هم قطبی و حداقل-حداکثر مقدار جابجایی مکانی، بلوک‌هایی را که می‌توانند جهت تطابق کاندید شوند را انتخاب می‌نماییم. سپس در این قسمت تابع هزینه را از رابطه (۳-۶)، برای تمام بلوک‌های کاندید تطابق محاسبه می‌نماییم و در نهایت با استفاده از قاعده WTA، شبیه‌ترین بلوک را به بلوک تصویر سمت چپ انتخاب می‌نماییم و مقدار جابجایی مکانی را بدست می‌آوریم.

تصویر خروجی حاصل از این بخش الگوریتم در شکل ۳-۶ نشان داده شده است، همانطور که مشاهده می‌شود نتیجه در نقاط ناپیوستگی عمق دچار مقداری ابهام می‌باشد.



شکل ۳-۶: تصویر جابجایی مکانی استخراج شده از الگوریتم، قبل از اعمال فیلتر دوطرفه

۳-۳-۳ مرحله سوم

از اجرای مراحل قبل یک نقشه جابجایی مکانی اولیه بدست می‌آید. در این کار جهت بهبود این نقشه در لبه‌ها از فیلتر دوطرفه بازگشتی^۱ استفاده می‌نماییم. در ادامه این بخش به معرفی این فیلتر می‌پردازیم و سپس نتیجه حاصل از اعمال آن را نشان می‌دهیم:

۳-۳-۱ فیلتر دوطرفه بازگشتی

فیلتر دوطرفه یک فیلتر مناسب برای حفظ لبه‌ها در تصاویر دارای توسط توماس و ماندوچی در [۳۱] معرفی شد. فیلتر دوطرفه دارای دو هسته فیلتر مکانی^۲ و دامنه‌ای^۳ برای اندازه‌گیری فاصله مکانی و دامنه‌ای بین پیکسل مرکزی و پیکسل‌های همسایه می‌باشد، که معمولاً بر اساس توزیع گوسی پیاده‌سازی می‌شوند. ما در این تحقیق از فیلتر دوطرفه بازگشتی به دلیل پیچیدگی محاسباتی کمتر آن نسبت به فیلتر دوطرفه معمولی بهره برداریم، برای مثال اگر N تعداد پیکسل‌های موجود در تصویر و D تعداد کانال‌های تصویر باشد، پیچیدگی محاسباتی روش بازگشتی $O(ND)$ و روش معمولی

۱- Recursive Bilateral Filter

۲- Spatial Filtering

۳- Range Filtering

$O(ND^2)$ خواهد بود که نشان از خطی بودن محاسبات در روش بازگشتی دارد. خلاصه‌ای از فیلترهای دوطرفه [۳۱، ۳۲، ۳۳] و بازگشتی در بخش‌های (الف و ب و پ) و پیاده‌سازی فیلتر دوطرفه بازگشتی در بخش (ت) ارائه شده است.

❖ الف- فیلتر دوطرفه

فیلتر دوطرفه ترکیبی از فیلتر کردن مکانی و دامنه‌ای است. فرض کنید x نشانه پیمایش خطی یک تصویر سیاه و سفید باشد، مقدار فیلتر دوطرفه یک بعدی x در پیکسل i به صورت زیر محاسبه می‌شود:

$$y_i = \sum_{k=0}^i R_{k,i} S_{k,i} x_k \quad (7-3)$$

که ($R_{k,i} = R(x_k, x_i)$ در معادله بالا، هسته فیلتر برای اندازه‌گیری شباهت دامنه‌ای پیکسل‌های k و i می‌باشد و $S_{k,i} = S(x_k, x_i)$ هسته فیلتر برای محاسبه شباهت مکانی دو پیکسل است).

❖ ب- فیلتر بازگشتی یک بعدی

فرض کنید x نشان‌دهنده سیگنال ورودی یک بعدی از یک سیستم بازگشتی از مرتبه n و y نشان‌دهنده خروجی باشد، پس:

$$y_i = \sum_{l=0}^{n-1} (a_l \cdot x_{i-l}) - \sum_{k=1}^n (b_k \cdot y_{i-k}) \quad (8-3)$$

این سیستم بازگشتی سپس توسطتابع انتقال زیر مشخص می‌گردد:

$$H^a(Z) = \frac{\sum_{l=0}^{n-1} a_l Z^{-l}}{1 + \sum_{k=1}^n b_k Z^{-k}} \quad (9-3)$$

$$= \sum_{k=0}^{+\infty} h_k^a Z^{-k} \quad (10-3)$$

به معنی پاسخ ضربه سیستم بازگشتی که تبدیل \mathcal{Z} آن $H^a(Z)$ است، می‌باشد. مشکل پیاده‌سازی h_k^a بازگشتی این فیلتر، سر و کار داشتن برای تعیین ضرایب a_k و b_k در معادله (۹-۳) است، به طوری که در معادله (۱۰-۳) بطور کامل یا بهترین تخمین برای h_k با به حداقل رساندن معادله زیر باشد:

$$E = \sum_{k=0}^{+\infty} (h_k^a - h_k)^2 \quad (11-3)$$

❖ پ- فیلتر بازگشتی مرتبه اول

ساده‌ترین فیلتر بازگشتی، فیلتر بازگشتی مرتبه اول است. بر طبق معادله (۸-۳) خروجی فیلتر بازگشتی مرتبه اول به صورت زیر نوشته می‌شود:

$$y_i = a_0 \cdot x_i - b_1 \cdot y_{i-1} \quad (12-3)$$

بر اساس [۳۴]، $a_o = 1 - a$ و $b_1 = -a$ ، پس رابطه (۱۲-۳) به صورت زیر نوشته می‌شود:

$$y_i = (1 - a) \cdot x_i + a \cdot y_{i-1} \quad (13-3)$$

$a \in [0, 1]$ در معادله بالا ضریب بازخورد می‌باشد.

❖ ت- فیلتر دوطرفه بازگشتی

در این بخش، پیاده‌سازی بازگشتی فیلتر دوطرفه با محدود کردن هسته فیلتر دامنه‌ای با مشخصه زیر بیان می‌گردد:

$$R_{k,i} = R_{i,k} = R_{k,k+1} R_{k+1,k+2} \dots R_{i-2,i-1} R_{i-1,i} = \prod_{j=k}^{i-1} R_{j,j+1} \quad (14-3)$$

در تحقیقات [۳۱، ۳۳، ۳۵، ۳۶]، هسته فیلتر دامنه اغلب گوسی می‌باشد:

$$R_{j,j+1} = R_{j+1,j} = \exp\left(-\frac{|x_j - x_{j+1}|^2}{2\sigma_R^2}\right) \quad (15-3)$$

که در آن σ_R انحراف معیار توزیع گوسی است و $|x_j - x_{j+1}|^2$ هزینه رفتن از پیکسل j به $j+1$

می‌باشد و $R_{j,j} = 1$ ، پس:

$$R_{k,i} = \prod_{j=k}^{i-1} R_{j,j+1} = \prod_{j=k}^{i-1} \exp\left(-\frac{|x_j - x_{j+1}|^2}{2\sigma_R^2}\right) = \exp\left(-\frac{\sum_{j=k}^{i-1} |x_j - x_{j+1}|^2}{2\sigma_R^2}\right) \quad (16-3)$$

همان‌طور که در معادله (14-3) مشاهده می‌شود، هسته فیلتر جدید، $R_{k,i}$ ، فاصله بین پیکسل k و i را با جمع‌آوری فاصله بین هر دو پیکسل مجاور در مسیر بین k و i محاسبه می‌کند. با استفاده از هسته فیلتر دامنه‌ای جدید، پیاده‌سازی بازگشتی فیلتر دوطرفه با تغییر اندازه ضرایب (a_l و b_k) در سیستم بازگشتی تعریف شده به وسیله‌ی هسته فیلتر مکانی در مکان هر پیکسل به دست می‌آید:

(17-3)

$$y_i = \sum_{l=0}^{n-1} (a_l^{new} \cdot x_{i-1}) - \sum_{k=1}^n (b_k^{new} \cdot y_{i-k}) = \sum_{l=0}^{n-1} (R_{i,i-1} \cdot a_l \cdot x_{i-1}) - \sum_{k=1}^n (R_{i,i-k} \cdot b_k \cdot y_{i-k})$$

که $n \geq 1$ می‌باشد. خروجی سیستم بازگشتی تغییر داده شده به صورت زیر است:

$$y_i = \sum_{k=0}^i R_{i,k} \left(\sum_{m=0}^{n-1} \lambda_{i-m-k} a_m \right) x_k \quad (18-3)$$

که در آن،

$$\lambda_i = \begin{cases} 1 & , \quad i = 0 \\ \sum_{k=1}^{\min(i,n)} -b_k \lambda_{i-k} & , \quad i > 0 \\ 0 & , \quad i < 0 \end{cases} \quad (19-3)$$

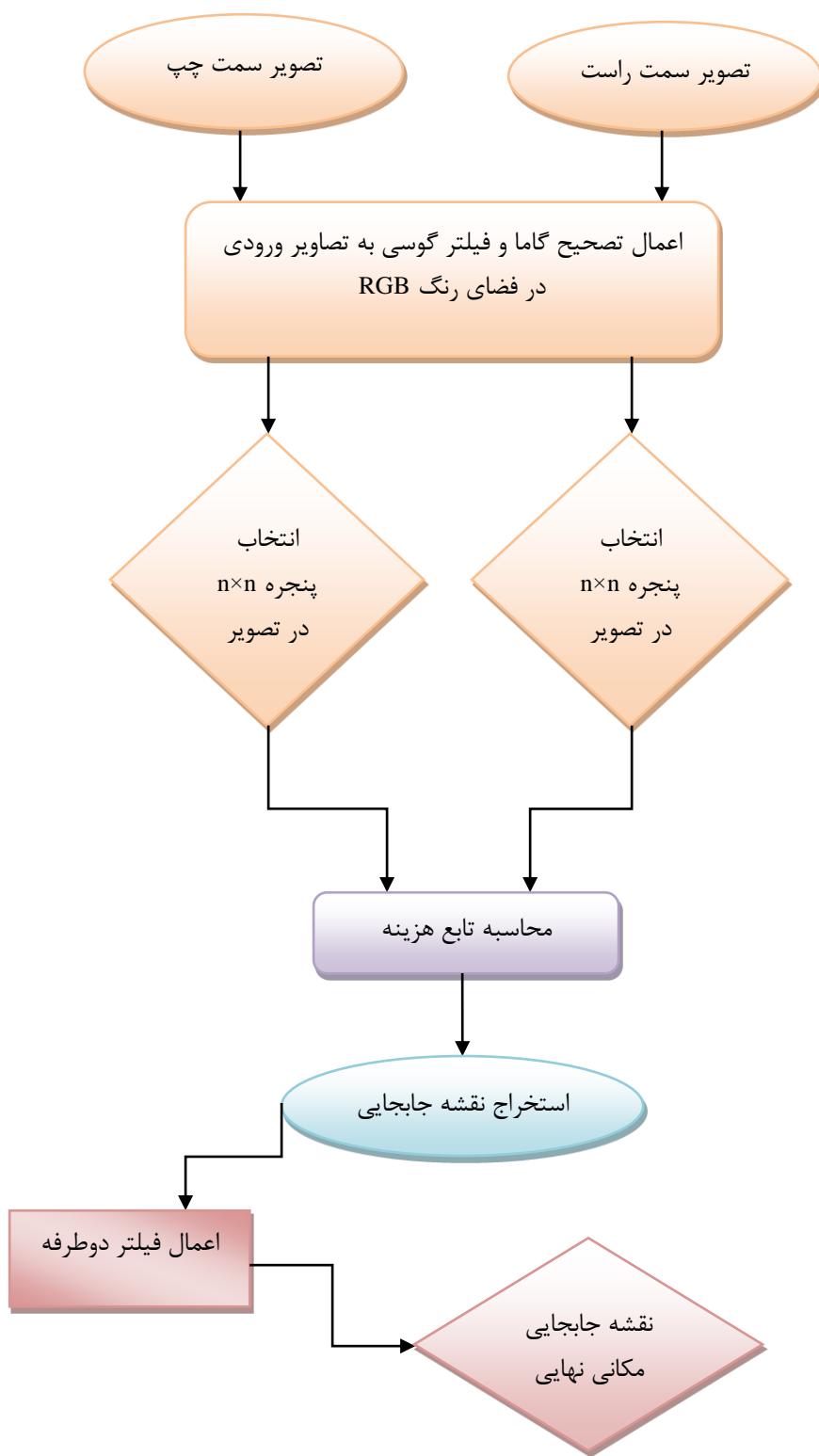
با شرط اولیه $R_{k,i}$ هسته فیلتر دامنه است می‌باشد. که در آن $i < 0$ وقتی که $x_i = o$ و $y_0 = a_0x_0$ هسته فیلتر مکانی به صورت زیر خواهد بود:

$$S_{i,k} = \sum_{m=0}^{n-1} \lambda_{i-m-k} a_m \quad (20-3)$$

پس از اعمال عملیات بالا در راستای افقی و سپس در راستای عمودی تصویر ورودی، خروجی این بخش به ما تصویر جابجایی مکانی متراکم نهایی را می‌دهد، نتیجه خروجی بر روی تصویر Tsukuba در شکل ۷-۳ و بلوک دیاگرام روش پیشنهادی در شکل ۸-۳ نشان داده شده است:



شکل ۷-۳: تصویر خروجی نهایی مربوط به Tsukuba



شکل ۳-۸: بلوک دیاگرام استخراج نقشه جابجایی مکانی متراکم

فصل چهارم

طرح آزمایش

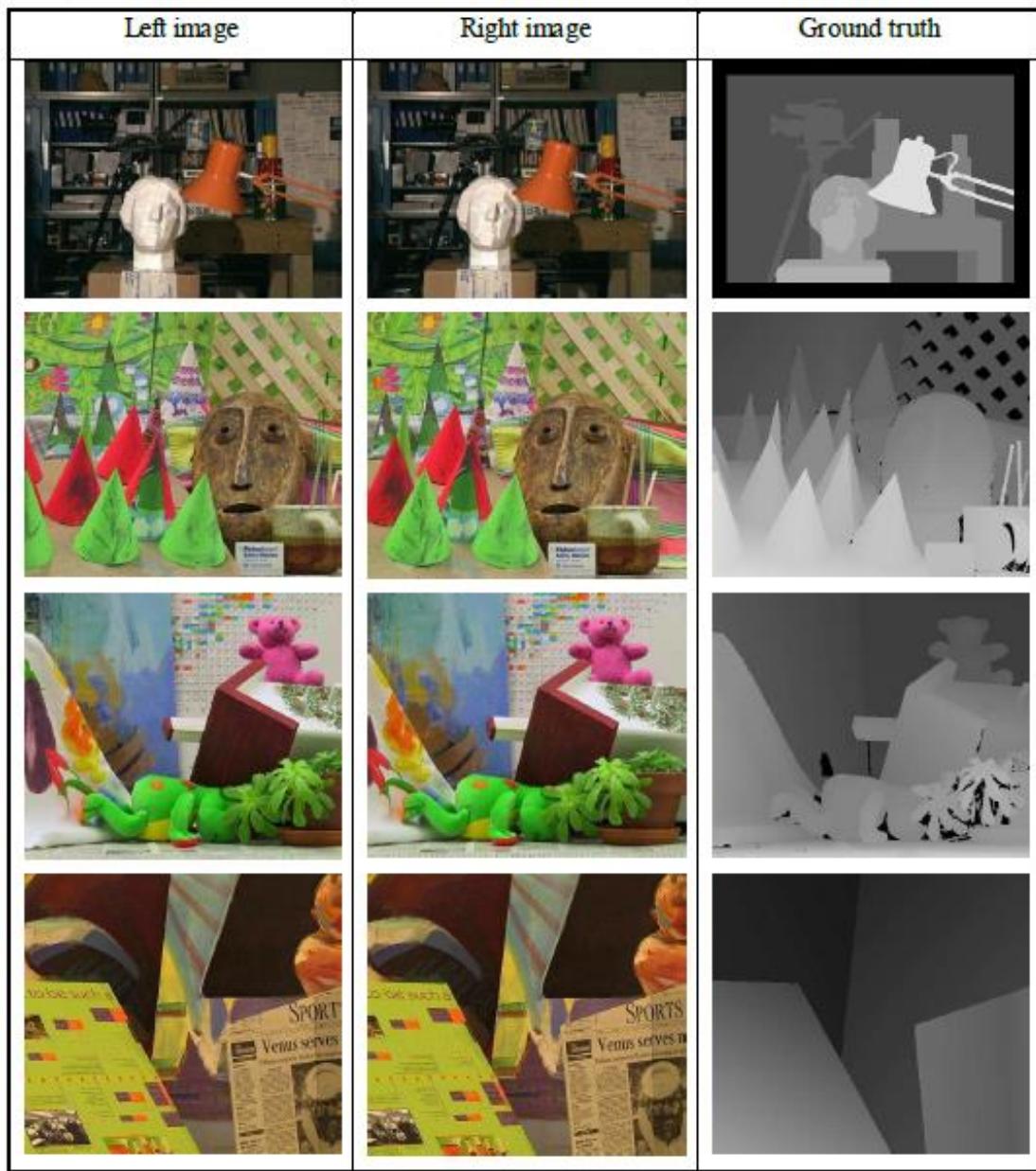
۱-۴ مقدمه

در این فصل رویه‌های را توصیف می‌کنیم تا الگوریتم‌های فصل قبلی را بتوان تحلیل کرد. طرح تجربی با دنبال کردن رهنمودها و پیشنهادهای ارائه شده توسط منابع مختلفی چون [۳۷] و [۳۸] توسعه یافت. ما از این فصل برای توصیف دقیق آزمایشاتی استفاده می‌کنیم که هدف آنها سنجش و مقایسه عینی عملکرد الگوریتم‌های مختلف است. مقایسه مشکل است از دو بعد اصلی: کیفیت نقشه‌های جابجایی مکانی حاصله و سرعت تولید این نقشه‌های جابجایی مکانی.

ابتدا MATLAB برای پیاده‌سازی استفاده شد. متأسفانه به کارگیری MATLAB بسیار کند بود، در نتیجه با استفاده از ویژوال استودیوی ۲۰۰۸ زبان C++ را برگزیدیم تا با تلفیق کتابخانه OpenCv در آن برنامه نهایی را ایجاد کنیم. استفاده از زبان C++ یک انتخاب خوب بود که عملکرد خیلی بهتری از نظر سرعت اجرا دارد.

۲-۴ مجموعه‌داده تصویر استریو میدلبوری

به منظور محاسبه معیارهای کیفیت برای یک الگوریتم تطبیق استریو، ضرورت دارد که برای جفت تصویر استریو، دارای یک نقشه جابجایی مکانی صحیح باشیم. این به شخص اجازه می‌دهد تا نقشه جابجایی مکانی محاسبه شده را با نقشه جابجایی مکانی صحیح مقایسه نموده و دقیقاً تعیین کند که در کجا و تا چه میزان این دو اختلاف دارند و بنابراین نقاط ضعف یک الگوریتم تطبیق استریو را تعیین کند. مجموعه داده تصویر استریو میدلبوری یک مجموعه کوچک از تنها چهار جفت تصویر استریو است، که برای هر جفتی نقشه‌های جابجایی مکانی صحیح، بسیار دقیقی همراه آن می‌باشد. این مجموعه داده مشهور بوده و با بستر آزمایش [۶] ترکیب شده است، این مجموعه داده برای تعیین کیفیت الگوریتم‌های تطبیق استریو مورد استناد قرار می‌گیرد. شکل ۱-۴ نشانگر مجموعه پایگاه داده میدلبوری است.



شکل ۴: مجموعه داده میدلبوری [۶]

۳-۴ تجهیزات مورد استفاده ما در پیاده‌سازی الگوریتم

در این تحقیق، الگوریتم‌ها را با استفاده از یک سیستم با مشخصات زیر پیاده‌سازی نمودیم:

- CPU: Intel(R) Core(TM) 2 Duo T9300 @ 2.5 GHz
- RAM: 1GB Hynix Value Ram DDR2
- Motherboard: Dell Vostro 1510
- Graphics: NVIDIA GeForce 8400M with 1GB RAM

۴-۴ ارزیابی الگوریتم پیشنهادی تطابق در سطح ویژگی

برای آزمایش الگوریتم‌ها، پس از اعمال الگوریتم گوشیابی هریس بر روی دو شکل ورودی و استخراج ویژگی‌های مطلوب، به تعیین تطابق بین پنجره‌ی ایجاد شده در اطراف هر ویژگی می‌پردازیم. پس از اعمال معیارهای تطابق مجموع اختلافهای مربعی، مجموع قدر مطلق تفاضل‌ها و همبستگی نرمالیزه که روابط آنها در معادلات (۱-۲)، (۲-۲) و (۳-۳) بیان شده و مقایسه آن‌ها با معیار تطابق پیشنهادی در معادله (۴-۳)، مشاهده می‌کنیم که دقت الگوریتم در یافتن تعداد تطبیق‌های درست در جفت تصویر ورودی، بطور قابل ملاحظه‌ای افزایش یافته است. جدول ۱-۴ نتایج مقایسه‌ها و شکل ۴-۲ خروجی حاصل از اعمال الگوریتم پیشنهادی که در شکل ۴-۳ نشان داده شده است، بر روی جفت تصاویر گرفته شده از مجموعه داده میدل‌بوری را نشان می‌دهد.

جدول ۱-۴: مقایسه الگوریتم‌ها

الگوریتم	جفت تصویر	تعداد تطبیق‌های درست	زمان(ثانیه)
SSD	Tsukuba	۷۷	۰.۰۶۴۹
	Teddy	۷۵	۰.۰۶۷۰
	Venus	۷۹	۰.۰۶۳۵
	Cones	۷۰	۰.۰۶۰۵
SAD	Tsukuba	۷۷	۰.۰۸۹۰
	Teddy	۷۰	۰.۰۹۰۴
	Venus	۷۷	۰.۰۸۶۱
	Cones	۶۸	۰.۰۸۱۱
NCC	Tsukuba	۶۶	۰.۱۵۴۹
	Teddy	۶۰	۰.۱۵۳۳
	Venus	۷۴	۰.۱۵۷۶
	Cones	۶۵	۰.۱۵۲۸
روش پیشنهادی	Tsukuba	۸۶	۰.۲۰۵۰
	Teddy	۸۰	۰.۲۰۹۲
	Venus	۸۹	۰.۲۰۲۶
	Cones	۷۴	۰.۲۰۱۱



(الف)



(ب)



(ج)



(د)

شکل ۲-۴: تطبیق ویژگی‌های جفت تصاویر استریو

Cones (د) Venus (ب) Tsukuba (ج) (الف)

توجه داریم در جدول ۱-۴ تعداد تطبیق‌های درست از میان ۱۰۰ ویژگی استخراج شده به وسیله‌ی الگوریتم گوشیابی هریس، در هر جفت شکل ورودی به دست آمده است و برای بقیه ویژگی‌ها تطبیق مناسبی یافت نشده است. همانطور که در جدول ۱-۴ مشاهده نمودید، با وجود اینکه میزان زمان مصرفی در محاسبه تطبیق بین پنجره‌ها افزایش یافته است، اما دقت الگوریتم با توجه به تغییرات اعمال شده در محاسبه تطبیق، به ترتیب در جفت تصاویر استریو Venus, Teddy, Tsukuba و Cones به میزان ۹٪، ۷.۵٪، ۱۱٪ و ۶.۳۳٪ افزایش یافته است.

۴-۵ معیارهای کیفیتی تطبیق استریو

روش‌های مختلفی برای تعیین کیفیت یک نقشه جابجایی مکانی تولیدشده وجود دارد، که اغلب آنها بر داشتن یک نقشه جابجایی مکانی صحیح تکیه دارند. نویسنده‌گان از روشی استفاده کردند که شامل تعیین تعداد پیکسل‌هایی است که نقشه‌های جابجایی مکانی تولیدشده برای آنها با نقشه جابجایی مکانی صحیح تا بیش از یک آستانه معین تفاوت دارد. آنها سنجش کیفیت را در قالب درصدی از پیکسل‌های بد اشاره می‌کنند که در آن یک درصد پایین‌تر، به معنای یک نقشه جابجایی مکانی با کیفیت بهتر است.

همچنین می‌توان نواحی نقشه‌های جابجایی مکانی موردنظر را محدود ساخت به‌گونه‌ای که درصد پیکسل‌های بد را بتوان تنها برای انواع خاصی از نواحی تعیین نمود. چنان نواحی معمولاً شامل نواحی هستند که حاوی ناپیوستگی‌های عمق می‌باشند و دارای نواحی که حاوی نقاط انسدادی و نواحی بدون بافت وسیع هستند، نمی‌باشند. جدول ۲-۴ محدوده تغییرات جابجایی مکانی و ضریب تبدیل مناسب برای نمایش نقشه‌ها را در مجموعه داده میدل‌بوری نشان می‌دهد.

جدول ۲-۴: محدوده تغییرات جابجایی مکانی و ضریب تبدیل مناسب برای تصاویر استریو [۶]

نام تصویر	محدوده تغییرات جابجایی مکانی	ضریب تبدیل به شدت روشنایی
Tsukuba	۰...۱۵	۱۶
Venus	۰...۱۹	۸
Teddy	۰...۵۹	۴
Cones	۰...۵۹	۴

۶-۴ ارزیابی نقشه‌های جابجایی مکانی متراکم

برای بررسی عملکرد الگوریتم‌های استریو و مشاهده اثر پارامترهای مختلف، معیارهای کمی مورد نیاز است. با فرض در اختیار داشتن نقشه جابجایی مکانی صحیح، دو معیار کمی از رابطه ۱-۴ و ۲-۴ پیشنهاد شده است. $d_C(x, y)$ مقدار جابجایی مکانی محاسبه شده در پیکسل (x, y) تصویر و $d_T(x, y)$ مقدار جابجایی مکانی صحیح می‌باشد. N تعداد پیکسل‌های تصویر است.

۱- معیار خطای RMS

$$R = \left(\frac{1}{N} \sum_{(x, y)} (d_C(x, y) - d_T(x, y))^2 \right)^{1/2} \quad (1-4)$$

۲- درصد پیکسل‌های انطباق نیافته

$$B = \frac{1}{N} \sum_{(x, y)} (|d_C(x, y) - d_T(x, y)| > \delta_d) \quad (2-4)$$

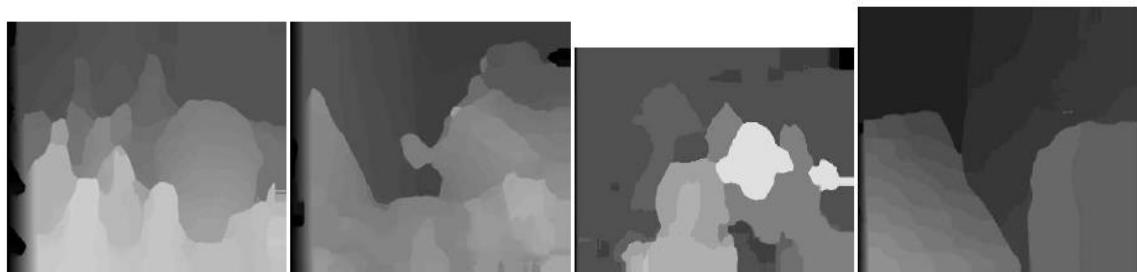
δ_d حد خطای جابجایی مکانی بر حسب پیکسل می‌باشد و به طور معمول یک فرض می‌شود. معیار RMS وابستگی شدیدی به میزان خطای بدست آمده در هر پیکسل دارد، در حالیکه آنچه اهمیت دارد و در معیار خطای دوم نمود یافته است درصد پیکسل‌هایی است که دارای خطای بوده‌اند.

۱- Root Mean Square

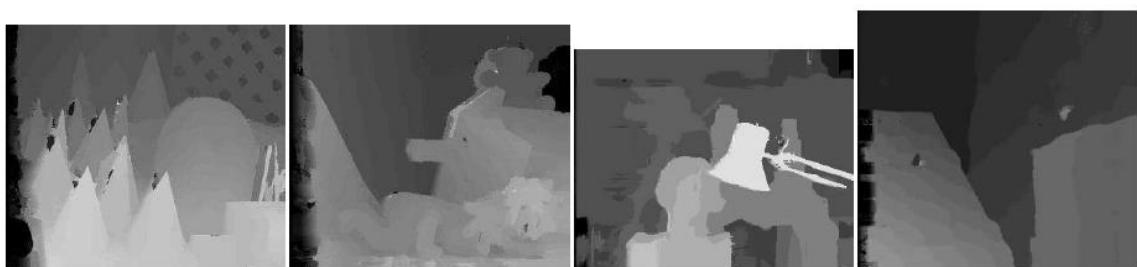
هر یک از این معیارها علاوه بر آنکه بر روی کل تصویر پیاده‌سازی می‌شوند، برای ارزیابی عملکرد الگوریتم‌ها در نواحی مختلف تصویر، به عنوان مثال نواحی بدون بافت یا مجاور ناپیوستگی‌های عمق، نیز قابل محاسبه می‌باشند. در جدول ارزیابی ارائه شده که بخشی از آن در شکل ۱۳-۲ نشان داده شده است، ارزیابی روش‌ها با معیار دوم (B) و در نواحی غیرانسدادیافته، کل تصویر و مجاورت ناپیوستگی‌های عمق صورت می‌گیرد.

۷-۴ سایر نتایج

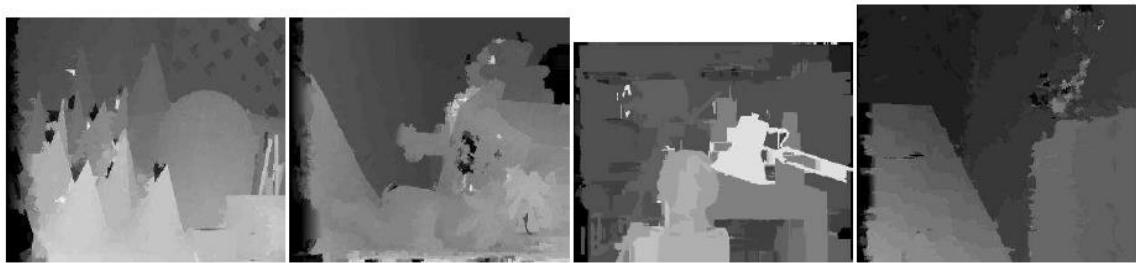
این بخش شامل به تصویر کشیدن نتایج الگوریتم‌های دیگر است، در واقع نقشه‌های جابجایی مکانی تولید شده توسط روش‌های دیگر در شکل های (۳-۴)، (۴-۴) و (۴-۵) به تصویر کشیده شده است.



شکل ۴-۳: تولید نقشه‌های جابجایی مکانی با استفاده از روش پنجره ثابت



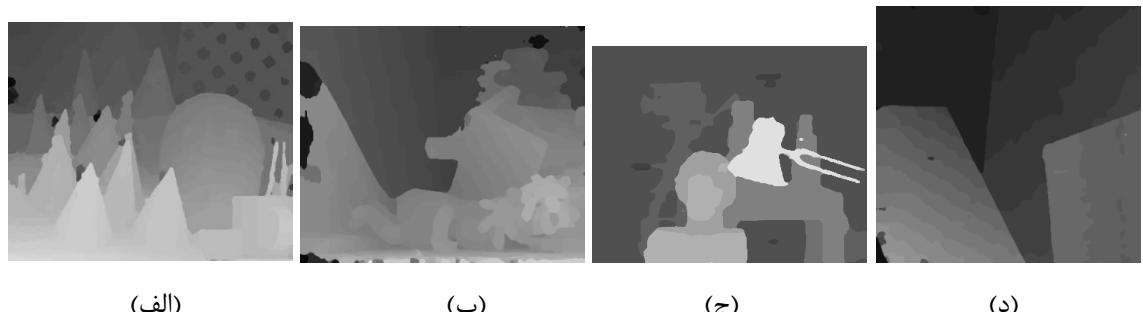
شکل ۴-۴: تولید نقشه‌های جابجایی مکانی با استفاده از روش وزن- انطباقی



شکل ۴-۵: تولید نقشه‌های جابجایی مکانی با استفاده از روش بخش‌بندی تصویر

۸-۴ نتیجه‌گیری

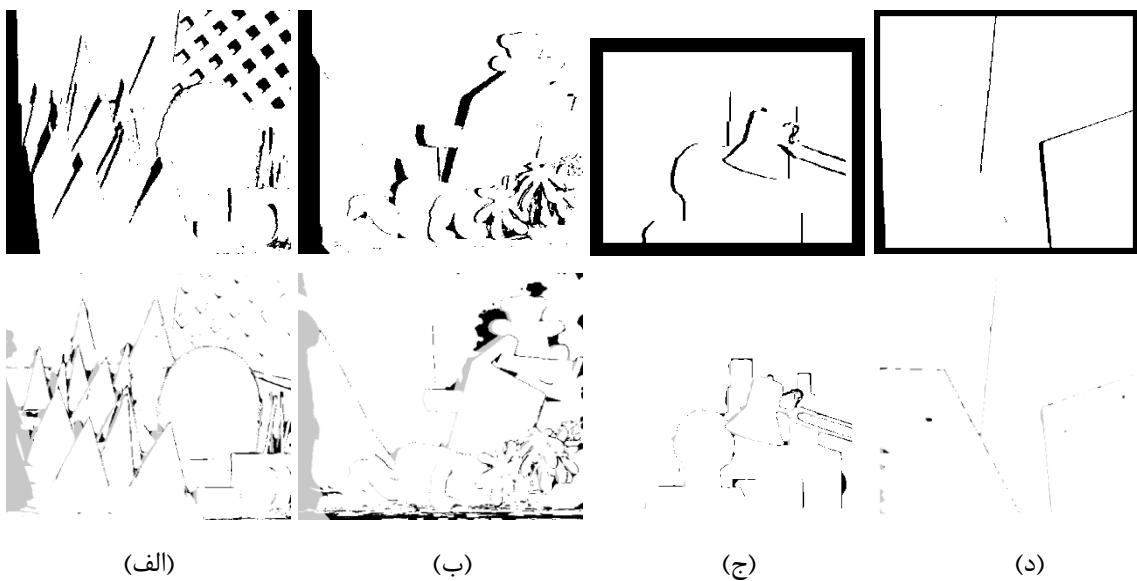
جهت انجام مقایسه بهتر، الگوریتم مطرح شده را روی چهار تصویر استریو از پایگاه داده میدلبوری تست نمودیم. در شکل ۴-۶(الف، ب، ج و د)، نقشه‌های جابجایی مکانی روش پیشنهادی نشان داده شده است و در شکل ۴-۷، نقشه مقایسه خطای اختلاف مطلق در نقاط غیرانسدادیافته روش پیشنهادی و نقشه‌های صحیح نشان داده شده است.



شکل ۴-۶: نقشه‌های جابجایی مکانی استخراج شده به وسیله روش پیشنهادی برای تصاویر

الف (۱)، ب (۲)، ج (۳)، د (۴)

در جدول ۳-۴ مقایسه‌ای بین روش پیشنهادی ما و روش‌های ارائه شده در [۱۷، ۳۹، ۴۰، ۴۱] انجام شده است. همانطور که مشاهده می‌شود، نتایج در تصاویر مربوطه بطور میانگین بهبود خوبی پیدا کرده است.



شکل ۴-۷: مقایسه نقشه اختلاف مطلق پیکسل‌های غیرانسدادی در نقشه‌های صحیح(ردیف بالا) و روش پیشنهادی(ردیف پایین) (الف) ، (ب) ، (ج) ، (د) Teddy ، Cones ، Venus

جدول ۴-۳: مقایسه بین روش پیشنهادی و الگوریتم‌های مبتنی بر روش‌های محلی

Venus	Cones	Teddy	Tesukuba	الگوریتم
۰.۴۸	۸.۳۸	۱۲.۲	۱.۹۶	روش ما
۰.۸۱	۹.۳۴	۱۳.۳	۱.۹۹	روش [۳۹]
۱.۴۵	۸.۵۷	۱۳.۸	۲.۷۵	روش [۴۰]
۱.۱۹	۹.۷۹	۱۳.۳	۱.۸۵	روش [۱۷]
۰.۴۵	۱۰.۱	۱۲.۶	۳.۶۸	روش [۴۱]

با توجه به درصد خطأ ملاحظه می‌شود که روش پیشنهادی نسبت به روش‌های دیگر از عملکرد مطلوب‌تری برخوردار است. این روش‌ها، از روش‌های محلی الگوریتم استریو می‌باشند، برای مقایسه کامل‌تر، در جدول ۴-۴، الگوریتم پیشنهادی را با چند الگوریتم دیگر که از روش‌های سراسری

استفاده نموده‌اند، مقایسه کردیم. روش [۴۲] هم‌اکنون در بستر آزمایش میدل‌بوری دارای رتبه یک از نظر میانگین خطای کمتر می‌باشد.

جدول ۴-۴: مقایسه بین روش پیشنهادی و الگوریتم‌های مبتنی بر روش‌های سراسری

Venus	Cones	Teddy	Tesukuba	الگوریتم
۰.۲۴	۶.۱۶	۸.۰۹	۱.۱۳	روش [۴۲]
۰.۴۸	۸.۳۸	۱۲.۲	۱.۹۶	روش ما
۱.۶۷	۹.۶۸	۱۱.۹	۱.۶۲	روش [۴۳]
۱.۶۵	۱۲.۷	۱۴.۲	۲.۴۵	روش [۴۴]
۲.۷۲	۱۴.۲	۱۴.۵	۴.۴۰	روش [۴۵]

فصل پنجم

نتیجه‌گیری و پیشنهادات

۱-۵ نتیجه‌گیری

یکی از موضوعاتی که در دید کامپیوتری مورد تحقیق قرار گرفته، مسئله استخراج اطلاعات عمق از تصاویر دو بعدی است تا یک صحنه سه بعدی بازسازی شود. این را در تحقیقات با اصطلاحات متراծی نامیده‌اند که برخی از آنها عبارتند از بینایی استریو، نقشه‌برداری عمق استریو، بازسازی سطح، برآورده سطح. دید استریو با مقایسه دو تصویر مختلف از صحنه‌ای واحد عمل می‌کند که از موقعیت‌های اندکی متفاوت در فضا برداشته شده‌اند.

صدها الگوریتم مختلف در تحقیقات وجود دارند که یک نقشه جابجایی مکانی را از تصاویر استریو تهیه می‌کنند و همگی نقاط قوت و ضعف خود را دارند. رویکردها را می‌توان بطور کلی در دو دسته طبقه‌بندی نمود، فنون محلی و سراسری. روش‌های محلی پیکسل‌های مرتبط را بر روی یک خط پیمایش جستجو نموده، هزینه تطبیق را (یعنی یک سنجش عدم شباهت) برای یک ناحیه یا پنجره پشتیبان محدود حول پیکسل‌های مورد مقایسه جمع‌بندی می‌کند. فنون سراسری عموماً یک مسئله بهینه‌سازی سراسری را حل می‌کنند تا بهترین تطبیق‌ها را بیابند. در کل روش‌های محلی نتایجی با دقت کمتر به دست داده، اما در عوض آنها معمولاً از لحاظ محاسباتی کم‌هزینه‌تر بوده و بنابراین برای مصارف بلادرنگ مناسب‌تر هستند.

این تحقیق بر روش‌های محلی تمرکز داشت؛ کارشناسان همچنین نشان دادند که چگونه اغلب روش‌های بینایی استریو محلی را می‌توان به چهار مرحله مجزا تقسیم نمود:

- ۱- محاسبه هزینه، ۲- تجمع هزینه، ۳- محاسبه اختلاف، ۴- اصلاح اختلاف.

الگوریتم‌های تطبیق استریو محلی به میزان زیادی بسته به شیوه‌ای که هزینه‌های تطبیق را جمع‌بندی می‌کنند و چگونگی تعریف نواحی پشتیبان به کار رفته در آنها تفاوت دارند. در این تحقیق ما تعداد کمی از روش‌های تجمع هزینه برتر را به کار گرفته و ارزیابی کردیم که در تحقیقات اخیر ارائه شدند. ما همچنین یک مجموعه از روش‌های مختلف را برای انجام محاسبه هزینه و مراحل اصلاح به

کار بردیم. عملکرد الگوریتم‌ها از لحاظ دقت نقشه‌های جابجایی مکانی در قالب درصدی از پیکسل‌های بد در مقایسه با نقشه‌های صحیح ارزیابی شدند. ما از مجموعه داده استریو میدلبوری برای تمامی سنجش‌های عملکرد استفاده کردیم.

برای مرحله محاسباتی هزینه، دریافتیم که استفاده از مجموع اختلاف‌های مطلق مؤثرترین نتایج را، در روش ما به دست می‌دهد. برای فنون تجمع هزینه ما اجرای رویکردهای مختلف را با استفاده از پارامترهای ثابت برگزیدیم. این به ما اجازه داد تا تحلیل کنیم که فنون تجمع هزینه مختلف با استفاده از یک منطقه پشتیبان مربعی ثابت بدون هیچ رویه اصلاحی تا چه حد مؤثر است.

ما دو فن اصلاح نقشه جابجایی مکانی متفاوت را ارزیابی کردیم که می‌توان از آن برای پاکسازی نتایج بسیاری از روش‌های تجمع هزینه استفاده نمود. اولی کاربرد یک فیلتر میانه است. این روش نشان داد که در فیلترسازی سوء‌طبیق‌ها بسیار مؤثر است، بهویژه در نواحی بافت‌دار پراکنده از تصاویر ورودی. روش دومی که ما ارزیابی کردیم، بررسی تطبیق دو-جهته بود [۴۶]. این فن بسیار مؤثر است و قادر است تا نواحی انسدادی را تعیین موقعیت نموده و از یک فن درون‌یابی استفاده می‌کند تا بطور هوشمندی، ارزش‌های اختلاف مناسب را برای چنین نواحی حدس بزنند. متأسفانه این فن مستلزم محاسبه یک نقشه جابجایی مکانی، از لحاظ هردو تصویر ورودی است، که این زمان پردازش را افزایش می‌دهد.

ما همچنین ترکیب این دو فن اصلاح را آزموده و دریافتیم که کاربرد هردوی آنها بطور متوالی حتی به اصلاح بهتری می‌انجامد. با اطلاعات به‌دست آمده از آزمایش‌ها ما یک الگوریتم استریو محلی جدید را بر مبنای الگوریتم محلی پیشنهاد کرده و به کار گرفتیم. هدف از تحقیق، ایجاد یک الگوریتم مؤثرتر بود که بتواند نتایجی مشابه با نتایج نوین به‌دست دهد. ما از روشی مشابه SAD برای محاسبه هزینه تطبیق استفاده کردیم و از فیلتر دوطرفه بازگشتی برای اصلاح نتایج استفاده نمودیم، که در نهایت نقشه‌های جابجایی مکانی را با میانگین پیکسل بد ۵.۱٪ در مجموعه داده میدلبوری تولید نمود.

۲-۵ پیشنهاد برای تحقیقات آینده

بسیاری از مصارف برای تطبیق استریو مستلزم آن است که یک سیستم بتواند بطور بلادرنگ اطلاعات عمقی را در کمترین زمان ممکن تعیین کند، یک راه اثبات شده برای بهبود عملکرد چنین الگوریتم‌هایی، طراحی و اجرای آنها بر روی واحد پردازش گرافیکی^۱ (GPU) است [۱۲,۴۷]. همچنین بررسی تصمیمات طراحی بیشتری را شامل می‌شود که در میان مراحل مختلف الگوریتم صورت می‌گیرد؛ همچون گزینه‌هایی مانند نگاه‌کردن به یک طیف وسیع‌تر از فضاهای رنگ برای تعیین اختلاف رنگ میان دو پیکسل.

در آخر، به اعتقاد ما یک روش ارزشمند برای بهبود الگوریتم‌های استریو می‌توانست با طراحی و خلق دقیق یک مجموعه داده استریو ترکیبی حاصل شود. یک چنین مجموعه داده‌ای را می‌شد برای سنجش انواع ویژه‌ای از نواحی و مسائلی به کار برد که الگوریتم‌های استریو با آن مواجه هستند و بنابراین شیوه‌ای بهتر برای تحلیل روش‌های موجود به دست آید. چنین نواحی مهمی که باید بررسی شوند عبارتند از: نواحی بافت تکراری، نواحی بافت‌دار پراکنده، نویز تصویر ورودی، انواع مختلف از ناپیوستگی‌های عمق، سطوح مسطح موازی و صفحه‌های اریب. یک جفت استریو ترکیبی جداگانه باید برای هر یک از این ناحیه‌ها خلق شود، به گونه‌ای که توانایی یک الگوریتم استریو برای پرداختن به آنها را بتوان به آسانی سنجیده و بطور جداگانه‌ای مقایسه نمود.

۱- Graphics Processing Unit

مراجع

- [1] سعیدی م، (۱۳۹۲)، پایان نامه ارشد، ”**تطابق استریو با استفاده از ویژگی های بافت و رنگ**“، دانشکده برق و رباتیک، دانشگاه صنعتی شاهرود.
- [2] Moravec H., (1980). **“Obstacle avoidance and Navigation in the Real World by Seeing Robot Rover”**, Stanford University, Computer Science.
- [3] Scharstein D., Szeliski R., (2002). **“A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”**, International Journal of Computer Vision, 47(1-3), 7-42.
- [4] Mattoccia S., (2011). **“Stereo Vision: Algorithms and Applications”**, DEIS, University of Bologna.
- [5] Kessel M., (2006). **“Surface Estimation from Multiple Images”**, Doctoral dissertation, Technical University of Denmark.
- [6] Scharstein D., Szeliski R., (2002). Retrieved 01-26-2011, from Middlebury Stereo Vision Page: <http://vision.middlebury.edu/stereo/>
- [7] Sun J., Li Y., Kang S. B., Shum H. Y., (2005). **“Symmetric Stereo Matching for Occlusion Handling”**, In Computer Vision and Pattern Recognition. CVPR 2005. Computer Society Conference on IEEE.
- [8] Yang Q., Wang L., Yang R., Stewénius H., Nistér D., (2008). **“Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling”**, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31(3), 492-504.
- [9] Zitnick L. C., Kanade T., (1999). **“A Cooperative Algorithm for Stereo and Occlusion Detection”**, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(7), 675-684.
- [10] Klaus A., Sormann M., Karner K., (2006). **“Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure”**, In Pattern Recognition, ICPR 2006. 18th International Conference on IEEE.
- [11] Wang Z., Zheng Z., (2008). **“A region based stereo matching algorithm using cooperative optimization”**, In Computer Vision and Pattern Recognition, CVPR 2008. IEEE.
- [12] Gong Minglun., Yang R., Wang L., Gong Mingwei., (2006). **“A Performance Study on Different Cost Aggregation Approaches Used in Real-Time Stereo Matching”**, International Journal of Computer Vision, 75(2), 283-296.

- [13] Sun C., (1997). ***A Fast Stereo Matching Method***, In Digital Image Computing: Techniques and Applications, 95-100.
- [14] Lewis J. P., (1995). ***Fast Normalized Cross-Correlation Vision Interface*** , In Vision interface, 10(1), 120-123.
- [15] Gerrits M., Bekaert P., (2006). ***Local Stereo Matching with Segmentation-based Outlier Rejection***, In Computer and Robot Vision, The 3rd Canadian Conference on IEEE.
- [16] Brown M. Z., Burschka D., Hager G. D., (2003). ***Advances In Computational Stereo***, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 25(8), 993-1008.
- [17] Kuk-Jin Y., So K.. (2006). ***Adaptive Support-Weight Approach for Correspondence Search***, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(4), 650-656.
- [18] Comaniciu D., Meer P., (2002). ***Mean Shift: A Robust Approach toward Feature Space Analysis***, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(5), 603-619.
- [19] Tombari F., Mattoccia S., Di Stefano L., (2007). ***Segmentation-Based Adaptive Support for Accurate Stereo Correspondence***, In Advances in Image and Video Technology, Springer Berlin Heidelberg.
- [20] Egnal G., Wildes R. P., (2002). ***Detecting Binocular Half-Occlusions: Empirical Comparisons of Five Approaches***, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(8), 1127-1133.
- [21] Worby J. A., (2007). ***Multi-Resolution Graph Cuts for Stereo-Motion Estimation***, Doctoral dissertation, University of Toronto.
- [22] Boykov Y., Veksler O., Zabih R., (2001). ***Fast Approximate Energy Minimization via Graph Cuts***, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(11), 1222-1239.
- [23] Kolmogorov V., Zabih R., (2001). ***Computing Visual Correspondence with Occlusions via Graph Cuts***, In Computer Vision, ICCV 2001. Proceedings. Eighth IEEE International Conference on.
- [24] moallem P., faez K., haddadnia J., (2002). ***Fast Edge-Based Stereo Matching Algorithms through Search Space reduction***, IEICE Transactions on Information and Systems, 85(11), 1859-1871.
- [25] moallem P., faez K., (2005). ***Effective Parameters in Search Space Reduction Used in a Fast Edge-Based Stereo Matching***, Journal of Circuits, Systems, and Computers, 14(02), 249-266.

- [26] Moallem P., Ashorian M., Mirzaeian B., Ataei M., (2006). “**A Novel Fast Feature Based Stereo Matching Algorithm with Low Invalid Matching**”, WSEAS Transactions on Computers, 5(3), 469-476.
- [27] Trucco E., Verri A., (1998). “**Introductory Techniques For 3D Computer Vision**”, (Vol. 201). Englewood Cliffs: Prentice Hall.
- [28] He K., Sun J., Tang X., (2013). “**Guided image filtering**”, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35(6), 1397-1409.
- [29] Bleyer M., Gelautz C.M., Rhemann C., (2012). “**A New Stereo Matching Method Based on Edges and Corners**”, Journal of Computational Information Systems, 8(14), 6041-6048.
- [30] Zhao D., Liu D., Yang Y., (2008). “**A New Stereo Matching Method Based on Sub-pixel Corner Detection**”, In Proceedings of the 2008 International Conference on Computer Science and Software Engineering. IEEE Computer Society.
- [31] Tomasi C., Manduchi R., (1998). “**Bilateral filtering for gray and color images**”, In Computer Vision, Sixth International Conference on IEEE.
- [32] Paris S., Kornprobst P., Tumblin J., Durand F., (2009). “**Bilateral filtering: Theory and applications**”, Foundations and Trends in Computer Graphics and Vision.
- [33] Yang Q., (2012). “**Recursive bilateral filtering**”, In Computer Vision–ECCV 2012 (pp. 399-413). Springer Berlin Heidelberg.
- [34] Gastal E., Oliveira M., (2011). “**Domain transform for edge-aware image and video processing**”, In ACM Transactions on Graphics, 30(4), 69-76.
- [35] Paris S., Durand F., (2009). “**A fast approximation of the bilateral filter using a signal processing approach**”, International Journal of Computer Vision, 81(1), 24-52.
- [36] Durand F., Dorsey J., (2002). “**Fast bilateral filtering for the display of high-dynamic-range images**”, In ACM Transactions on Graphics, 266 21(3), 257-266.
- [37] Wu C. F., Hamada M., (2000). “**Experiments: Planning, Analysis, and Parameter Design Optimization**”, Technometrics, 43(3), 368-369.
- [38] Fang K., Sudjianto A., (2006). “**Design and Modelling for Computer Experiments**”, Chapman and Hall, CRC 2006.
- [39] Kowalcuk J., Psota E., Pérez L., (2013). “**Real-time stereo matching on CUDA using an iterative refinement method for adaptive support-weight correspondences**”, Circuits and Systems for Video Technology, IEEE Transactions on, 23(1), 94-104.
- [40] De-Maeztu L., Villanueva A., Cabeza R., (2011). “**Stereo matching using gradient similarity and locally adaptive support-weight**”, Pattern Recognition Letters, 32(13), 1643-1651.

- [41] Wang L., (2014). ***“Feature based stereo matching using two-step expansion”***, Mathematical Problems in Engineering.
- [42] Mozerov M., Van Weijer J., (2014). ***“Accurate stereo matching by two step global optimization”***, Submitted to IEEE Transactions.
- [43] Min D., Sohn K., (2008). ***“Cost aggregation and occlusion handling with WLS in stereo matching”***, Image Processing, IEEE Transactions on, 17(8), 1431-1442.
- [44] Yu W., Chen T., Franchetti F., Hoe J., (2010). ***“High performance stereo vision designed for massively data parallel platforms”***, Circuits and Systems for Video Technology, IEEE Transactions on, 20(11), 1509-1519.
- [45] Michael M., Salmen J., Stallkamp J., Schlippling M., (2013). ***“Real-time stereo vision: optimizing semi-global matching”***, In Intelligent Vehicles Symposium (IV), IEEE.
- [46] Zigh E., Belbachir M. F., (2012). ***“Soft computing strategy for stereo matching of multi spectral urban very high resolution IKONOS images”***, Applied Soft Computing, 12(8), 2156-2167.
- [47] Lang D., Little J. J., (2004). ***“GLStereo: Stereo Vision Implemented in Graphics Hardware”***, European Conference on Computer Vision.

Abstract

The process of extracting depth information from multiple two-dimensional images taken of the same scene is known as stereo vision. Spatial displacement is the difference between corresponding pixels from the left and right images, by taking one of the images as reference. In the past few decades have witnessed the development of hundreds of different stereo vision algorithms, this classification made it difficult to compare the various approaches to this problem. In this thesis we provide an overview of everything that approaches to solving stereo vision problem. Because we provide a method to extract real-time disparity map, between local and global stereo algorithms, we use local methods. In tests carried out over recent years, the most important element of the local stereo algorithms, matching cost aggregation methods expressed.

Our preferred method, first, a feature-based local stereo algorithm using harris corner detection are presented and then with using proposed matching cost aggregation methods around the window of interest pixel, method is proposed to recover the sparse disparity map that accuracy is better than previous methods. Then provided a local stereo algorithm based on pixel to recover a dense disparity map using gamma correction in RGB color space on the pair of input images in the preprocessing, bilateral filter in the postprocessing and using matching cost aggregation methods expressed in sparse methods that a mean error at the presented database in the literature is %5.1 and speed is desired.

Keywords- Stereo Vision, Disparity Map, Harris Corner Detection, Gamma Correction, Bilateral Filtering.



**Shahrood University of Technology
Faculty electricity and Robotics**

**Design and Implementation of a Real Time
Stereo Matching Algorithm to Improve the
Ability of a Robot in Scene Analysis**

Mohammad Reza Gharib Sami

Supervisor:

Dr. Hossein Khosravi

Date: 2014