

واقعیت افزوده در گجت های همراه

تشخیص و دنبال کردن بی نقص اشیاء در فضای زندگی به منظور

نمایش زنده اطلاعات و عبارات متنی بر مبنای آن ها

توبایس لنگلاتز، کلاوس دیچندوفر، آلسندرو مولونی، گرهارد شال، گرهارد ریتمیر، دیتر اشمالستینگ.

کاری از دانشکده فنی گرتز اتریش - انستیتوی بینایی و گرافیک کامپیوتری

اطلاعات مقاله:

تاریخ دریافت: ۳۰ ژانویه ۲۰۱۱	تاریخ دریافت نسخه بازبینی شده: ۲۶ آپریل ۲۰۱۱
تاریخ پذیرش: ۲۷ آپریل ۲۰۱۱	قرار گیری بر روی اینترنت: ۷ مه ۲۰۱۱
کلید واژه ها: حقیقت افزوده، عبارات متنی، دنبال کردن، تلفن همراه،	

www.farzad.im

ترجمه و تدوین از: فرزاد عباس پور

چکیده

یکی از اهداف مشترک تولید گجت هایی که با استفاده از حقیقت افزوده کار میکنند، پردازش عبارات متنی است که به اشیای کلیدی موجود در محیط زندگی ما مربوط میگردند. ما در اینجا یک روش جدید برای تشخیص اشیا و دنبال کردن آنها در هنگام تغییر زاویه دید مطرح می کنیم. این روش بر روی تلفن های همراه هوشمند نسل جدید به راحتی قابل پیاده سازی و استفاده است. همچنین کاربران می توانند از متدهای استفاده شده در این روش برای ساخت هر گجت دیگری که در زندگی عادی استفاده می شود بهره ببرند. این روش تولید و تهیه عبارات متنی بر اساس پردازش تعداد زیادی عکس Panorama به دست آمده است. این عکس ها با استفاده از دوربین های عکاسی عادی و به صورت ۳۶۰ درجه تهیه شده اند. این عکس ها در ابتدا به صورت ثابت و در حالتی که شخص بیننده در جای خود ایستاده باشد مورد بررسی و تحلیل قرار گرفتند. در مرحله بعدی عکس ها در حالت حرکت چرخشی بررسی شدند و این پروسه تا جایی ادامه داده شده است که حرکت ها به صورت آزادانه در یک دامنه پویا انجام شوند. در مرحله بعدی میزان نورپردازی محیط تغییر داده شد و تهیه عکس در شرایط نوری مختلف مورد بررسی قرار گرفت. همچنین تشخیص و دنبال کردن اشیا در هر مرحله مجدداً انجام شد تا الگوریتم تشخیص و دنبال کردن اشیا پس از طی مراحل به تدریج بهبود پیدا کرد و توانست در تست های نهایی

در شرایط پیچیده‌تر و سخت‌تر کلیه اشیا و آیتم‌های موجود در محیط را با صحت بالایی تشخیص دهد. شایان ذکر است در مرحله آخر یکپارچگی نتایج تشخیص داده‌شده با استفاده از اضافه کردن اطلاعات آماری به الگوریتم به طور کامل بهبود داده شد به طوری که قابلیت تشخیص الگوریتم در شرایط مختلف به درستی ثابت گردید. ما سیستم خود را بر روی اپلیکیشن حقیقت افزوده "کمپس" به عنوان یک نرم‌افزار نمونه تست کردیم و نتایج آن را بر روی یک تلفن همراه هوشمند میان رده در بازار نیز مورد استفاده قرار دادیم تا امکان استفاده از این الگوریتم را توسط اپلیکیشن های هادی که توسط کاربران در محیط زندگی استفاده میشود مورد ارزیابی قرار دهیم. نتایج حاکی از آنند که بیش از ۹۰٪ از کلیه اشیا و شرایط نوری مختلف توسط این الگوریتم به درستی مورد پردازش قرار می‌گیرد و بیننده به راحتی می‌تواند اشیا مختلف را ببیند، اطلاعاتی را پیرامون آنها به صورت عبارات متنی در مقابل چشمانش دریافت کند و در عین حال به راحتی و آزادانه راه برود و یا بدود.

۱. مقدمه:

تکنولوژی تشخیص اشیا با استفاده از حقیقت افزوده یک دسته جدید از کاربردهای طراحی شده برای تلفن‌های هوشمند هستند. کاربرد اصلی گجت های تشخیص اشیا که از تکنولوژی حقیقت افزوده بهره می‌برند نمایش عبارات (عمدتاً) متنی می‌باشد که به اشیا و مکان‌های موجود در جهان واقعی نسبت داده می‌شوند. در اولین مرحله از یک سری نقاط که در جهان واقعی قرار دارند برای تشخیص جهانی استفاده می‌شود. امروزه به منظور تشخیص مکان و زاویه دید از تکنولوژی‌های دیگری همچون GPS و سنسور الکترومغناطیسی (ژایروسکوپ) نیز بهره گرفته می‌شود تا مکان دقیق نمایش عبارات متنی در کنار اشیا و مکان‌ها مشخص گردد.

شایان ذکر است که کارایی چنین سنسورهایی بسیار پایین است و دقت مورد نیاز را ارائه نمی‌کند. سنسورهای الکترومغناطیسی نویز بسیار زیادی دریافت می‌کنند که سبب می‌شود زاویه دید به صورت دقیق ارائه نشود و اطلاعات عددی ارائه شده توسط سنسور در هنگام پردازش به سرعت تغییر کند و به مقادیر نادرستی تعبیر شود. حتی اگر صحت مکان‌یابی GPS را در سطح بالایی تصور کنیم، این دقت عملاً برای استفاده عادی توسط کاربر عامی در فضای شهری و در مقیاس بزرگ مناسب نخواهد بود و سبب خواهد شد تا عبارات متنی در مکان‌های نادرستی نمایش داده شوند.

در طرف دیگر دوربین تعبیه شده در یک تلفن همراه هوشمند می‌تواند چالش‌های موجود را برطرف سازد و مسئله بومی‌سازی مکان‌ها و اشیا را با بهره‌گیری از سایر تکنولوژی‌های موجود در خود تصحیح کند. هنوز تشخیص اشیا در فضای بیرونی و مناظر ممکن است چالش برانگیز به نظر برسد زیرا در فضای آزاد بیرونی تغییرات

نوری به شدت و با سرعت بالایی اتفاق می‌افتد. همچنین شایان ذکر است که این مشکل و چالش‌های مربوط به آن هنگام رویارویی با شرایط مختلف آب و هوایی و هنگام بررسی آماری مکان‌هایی که فرد می‌تواند در آنها قرار گیرد چالش بر انگیز تر تر به نظر می‌رسد. اما کلیه مشکلات مذکور با استفاده از تکنولوژی‌های موجود و قدرت پردازشی یک تلفن همراه هوشمند کاملاً برطرف می‌شود. تکنولوژی تشخیصی مبتنی بر حقیقت افزوده همچنین نیاز دارد تا عبارات متنی که به اشیا نسبت داده می‌شوند علاوه بر تشخیص اولیه در سایر مواقع و در هنگام مشاهده شدن توسط افراد دیگر هم در همان جا باقی بمانند. این قابلیت عمومی‌سازی و محدودیت‌های مرتبط با آن با استفاده از قدرت پردازشی تلفن همراه هوشمند برطرف شده است.

در این روش به منظور بهبود مشکلات مذکور ما از توانایی‌ها و ویژگی‌های عادی تلفن همراه هوشمند و اپلیکیشن‌های نرم افزاری تشخیصی حرکت افزوده که بر روی آن نصب میشوند بهره می‌بریم. و لذا ترکیب تکنولوژی‌های تلفن همراه هوشمند و دوربینی که بر روی آن نصب شده است به همراه سنسورهای فیزیکی موجود در آن همانند GPS امکان تشخیص دقیق مکان و اشیا را به هنگام حرکت و در شرایط مختلف ممکن می‌سازد. در کارهای تحقیقاتی قبلی که پیرامون این موضوع توسط محققان انجام شده است تنها از تصاویر ۳۶۰ درجه در حالت حرکت وضعی برای پردازش اطلاعات استفاده شده است. اما در کارهای تحقیقاتی اخیر این تصاویر ۳۶۰ درجه (Panorama) در شرایط مختلف و در همه زوایای دید - نه تنها در شرایط حرکت وضعی به چپ و راست جلو و عقب - مورد پردازش قرار گرفتند تا تشخیص و قرار دادن اعتبارات متنی در حول اشیا، و مکان‌ها را ممکن سازند و یک رابط تجربه کاربری بی‌نقص، بدون مزاحمت‌های تکنولوژیکی را برای کاربر عامی فراهم سازند. این مقاله تحقیقاتی یک متد تصریح شده و پیشرفت داده شده پیرامون این موضوع را مورد بررسی و تحلیل دقیق قرار می‌دهد. این متد کارایی و قدرت پردازش و تشخیص اشیا و همچنین امکان دنبال کردن آنها را به سیستم قدیمی اضافه کرد است. در مرحله آخر قدرت تشخیص الگوریتم با استفاده از افزودن داده‌های آماری به آن بهبود داده شده و امکان ایجاد خطا تا حد امکان کاهش یافت است. در این روش ابتدا نقاط ثابت موجود در فضای زندگی با استفاده از سنسورها تشخیص داده شده و سپس علاوه بر حرکت وضعی در چهار جهت، امکان پردازش اطلاعات از فضای سه‌بعدی با استفاده از یک سنسور حرکتی با پشتیبانی از سه محور (Z, Y, Z) فراهم شده است تا کاربر بتواند به صورت آزادانه به همه زوایای دید احاطه داشته باشد و پردازش اطلاعات با استفاده از ترکیبی از تکنیک‌های مذکور همچنان با کارایی بالایی ادامه داشته باشد.

ما سیستم خود را بر روی اپلیکیشن حقیقت افزوده واحد دانشگاهی خود به عنوان یک نرم‌افزار نمونه تست کردیم و نتایج آن را بر روی یک تلفن همراه هوشمند میان رده در بازار نیز مورد استفاده قرار دادیم تا امکان استفاده از این الگوریتم را توسط اپلیکیشن‌های عادی که توسط کاربران در محیط زندگی استفاده میشود مورد ارزیابی قرار

دهیم. نتایج حاکی از آنند که بیش از ۹۰٪ از کلیه اشیا و شرایط نوری مختلف توسط این الگوریتم به درستی مورد پردازش قرار می‌گیرد و بیننده به راحتی می‌تواند اشیا و مختلف را ببیند و اطلاعاتی را پیرامون آنها به صورت عبارات متنی در مقابل چشمانش دریافت کند.

۲. کارهای مرتبط:

کارهای تحقیقاتی که پیش تر در این زمینه انجام شده است را می‌توان به دو دسته تقسیم کرد. در دسته اول سیستم‌هایی وجود دارند که کاربر را قادر می‌سازند تا عبارات متنی را به صورت حباب‌هایی تولید و آنها را با استفاده از رابط کاربری اپلیکیشن‌های حقیقت افزوده به اشیا موجود در فضای بیرون نسبت دهد.

فعالیت‌هایی که پیش تر در زمینه تولیدی عبارات متنی انجام شده است در برخی پروژه‌ها همچون پروژه "MARS" فایزر به کار گرفته شده است و بنیان اصلی تولید اپلیکیشن‌های تجاری مبتنی بر حقیقت افزوده برای کاربری‌های روزانه را تشکیل می‌دهد. به منظور کاربری‌های تجاری روزانه اپلیکیشن‌ها و رابط‌های کاربری خاصی کار دنبال کردن و تشخیص را انجام می‌دهند. این رابط‌های کاربری با استفاده از حقیقت افزوده تجهیز شده‌اند. از جمله آنها میتوان به اپلیکیشن‌های Wikitude و Layer اشاره کرد که بر روی تلفن‌های همراه هوشمند نیز قابلیت نصب دارند. این اپلیکیشن‌ها از پایگاه‌های اطلاعاتی آنلاین استفاده می‌کنند. در مقابل فعالیت‌های تحقیقاتی اخیر انجام این کار را به صورت آنلاین پیشنهاد می‌کند.

در روش تولید عبارات متنی با استفاده از پایگاه‌های اطلاعاتی آن لاین روش‌های مختلفی پیشنهاد شده است. برای مثال در روش ریتمایر از یک مدل سه بعدی از فضایی که کاربر در آن واقع شده است برای محاسبه موقعیت عبارات متنی استفاده می‌شود. این روش همچنین علاوه بر نمایش عبارات، حدود فضای زندگی کاربر را با استفاده از خطوطی مشخص و کاربر را نسبت به وجود محرک‌های فیزیکی مطلع می‌سازد. روش ریتمایر همچنین تکنیک‌های را برای تسهیل درج اطلاعات پیرامون اشیا و موقعیت‌های مختلف مطرح کرده است که عمل بومی سازی و تهیه نقشه‌های مکانی در این اپلیکیشن‌ها را بهبود می‌بخشد.

در مقابل روش پیشنهاد شده توسط توماس و پیکارسکی از ترسیم مثلث برای تشخیص مکان نمایش عبارات متنی استفاده می‌کند. در این روش خطوطی از موقعیت‌های تشخیص داده شده توسط الگوریتم در فضای بیرون به سمت یکدیگر رسمی می‌شوند و سپس از تحلیل نقاط برخورد این خطوط و اشیا موجود در فضای درونی این مثلث‌ها برای تشخیص موقعیت قرارگیری عبارات متنی بهره برده می‌شود.

در طرف دیگر در روش "ویدر" از عکس‌های هوایی برای پشتیبانی اطلاعات موجود و پردازش عبارات متنی استفاده می‌شود. پس از مشخص شدن حدود فیزیکی فضایی که کاربر در آن قرار گرفته است و تشخیص اشیا توسط رابط کاربری اپلیکیشن‌های حقیقت افزوده، یک نمای جانبی، تصویری هوایی (برای مثال گرفته شده توسط ماهواره) از مکانی که کاربر در آن واقع شده است را نمایش می‌دهد که به کاربر این امکان را می‌دهد مکان عبارات متنی را در فضا بهتر درک کند و یا جای آنها را در فضا تغییر دهد. در این پروسه تحقیقاتی، ویدر بعدها، استفاده از لیزر را برای تشخیص دامنه حرکت کاربر پیشنهاد کرد.

این فعالیت تحقیقاتی همچنین یک روش دیگر را برای کاربر پیشنهاد می‌کند تا بتواند با استفاده از آن عبارات متنی را در یک نمای ۳۶۰ درجه از فضایی که در آن قرار گرفته است قرار دهد. این روش که در واقع بنیان این مقاله تحقیقاتی نیز می‌باشد، در بخش سوم مفصلاً توضیح داده شده است. عمده‌ترین ایده این تکنیک کارایی ضعیف آن در تشخیص مجدد اشیا در گذر زمان می‌باشد.

بسیاری از اپلیکیشن‌های تجهیز شده با حقیقت افزوده برای استفاده بر روی کامپیوترهای شخصی در گذشت تولید شده‌اند. این اپلیکیشن‌ها از ترکیبی از روش‌ها همچون تصاویر دو بعدی، GPS و همچنین سنسورهای درونی برای تشخیص اشیا و نسبت دادن عبارات متنی به آنها در موقعیت‌های مختلف استفاده می‌کنند. این چنین سنسورهای امروز در کلیه تلفن‌های هوشمند هم به کار می‌روند و در عین حال ارزان قیمت بوده و توان مصرفی کمتری هم دارند. اما شایان ذکر است که کارایی این سنسورها در مقایسه با سنسورهای صنعتی پایین تر می‌باشد.

در همه این گجت‌ها سنسور GPS موقعیت سه بعدی اشیا را بر روی کره زمین تعیین می‌کند و سپس از قطب نما ها و سنسورهای الکترومغناطیسی همچون ژایروسکوپ برای تعیین ارتفاع اجسام در فضا یا موقعیت قرارگیری دقیق آن‌ها استفاده می‌شود. در بهترین حالت هم می‌بایست به نویزهایی که این سنسورهای الکترومغناطیسی تولید می‌کنند و تأثیر آنها در عمل تشخیص اشیا توجه ویژه‌ای معطوف شود. زیرا این نویزها می‌توانند عمل تشخیص و مکان یابی را با خطاهایی همراه سازند. چنانچه اشیا تشخیص داده شده دور باشند این خطاها بسیار عمده تر و از نظر عددی بزرگتر هم خواهند بود.

روش‌های متعددی برای غلبه کردن بر محدودیت‌های مربوط به سنسورها وجود دارد. کالیبره کردن دقیق سنسور ها در هنگام تولید و در نظر گرفتن تأثیرات مغناطیسی آهن آلات مورد استفاده در ساختار فیزیکی سنسورها و همچنین محاسبه جهت، راستا و مهندسی دقیق تولیدات می‌تواند خطاهای محاسباتی را تا حد زیادی کاهش دهد. کالیبره کردن سنسور ها می‌تواند براساس عملکرد این سنسورها در فضای واقعی زندگی کاربران انجام شود تا دقت محاسبات بصورت عملی تا حد زیادی افزایش یابد. کالیبره کردن دستی سنسورها و استفاده از مقایسه

عملکرد محاسباتی آنها با فاکتورهای موجود در فضای زندگی سبب می‌شود تا دقت سنسورها در وضعیت‌های مختلف افزایش یابد و لذا تأکید می‌کنیم که تنها یک بار کالیبره کردن برای این چنین سنسورهایی به تنهایی کافی نیست. به همین دلیل استفاده از یک روش آنلاین برای کالیبره کردن این سنسورها پیشنهاد می‌شود تا کارکرد و عملکرد این سنسورها در موقعیت‌های مشابه به صورت خودکار صورت پذیرد. مقایسه متعدد عملکرد سنسورها در موقعیت‌های مشابه و توسط کاربران مختلف دقت محاسبات را افزایش می‌دهد. روش دنبال کردن اشیا با استفاده از روش هیبریدی که در بخش پنجم مورد بررسی قرار گرفته است از روش‌های کالیبره کردن آن لاین می‌باشد.

در طرف دیگر استفاده از دوربین‌های عکاسی برای تشخیص و دنبال کردن اشیا صحت بالاتری را نسبت به روش‌های مستقل از بینایی (Vision-based) که تنها از سنسور برای تشخیص استفاده می‌کنند ارائه می‌کند زیرا این چنین سنسورهایی هیچ تصویر قابل پردازشی از فضای بیرونی به صورت سه‌بعدی ندارند. اما در این روش چرخش و شیوه قرارگیری دوربین و مکان دقیق تصویربرداری در پردازش اطلاعات استفاده می‌شود. در این راستا، کلاین و موری یک دنبال کننده طراحی کردند که یک مدل سه بعدی از حدود محیطی که کاربر در آن واقع شده است تهیه می‌کند. همچنین مدل "آرث" روشی را برای بومی‌سازی موقعیت قرارگیری دوربین تلفن همراه کاربر در هنگام تصویربرداری پیشنهاد می‌کند که در یک محیط سه بعدی با پردازش موقعیت محدودیت‌ها (دیوارها) وضعیت و زاویه دیده کاربر را به صورت کامپیوتری مشخص می‌کند. واضح است که ترکیب روش‌های مبتنی بر بینایی به علاوه روش‌های مبتنی بر سنسورها این امکان را فراهم کند تا اطلاعات به صورت دقیق‌تر پردازش شوند و تحت شرایط خاص همچون در هنگام راه رفتن و یا دویدن نتایج بهتری را به دست دهند. اطلاعات تولید شده با استفاده از ترکیب این روش‌ها می‌تواند اطلاعات خام اولیه برای استفاده شدن در کامپوننت‌های سخت افزاری و نرم افزاری تجهیز شده با حقیقت افزوده باشند. البته باید توجه داشت که به علت تحقیقات اندک انجام شده بر روی سنسورهای به کار رفته در تلفن‌های همراه هوشمند ممکن است کارایی این روش به علت کیفیت پایین سنسورهای به کار رفته در تلفن‌های همراه هوشمند یا کمبود تلفن‌های همراه هوشمند با کیفیت ساخت بالا کاهش می‌یابد.

ممکن است در هنگام قرارگیری کاربر در محیط‌های ناشناخته تشخیص و دنبال کردن اشیا به صورت دقیق صورت نپذیرد. اما چنین موقعیت‌هایی می‌توان از اطلاعات آنلاین برای مقایسه موقعیت کاربر با سایر کاربرانی که در گذشته در آن موقعیت قرار گرفتن استفاده کرد و تجربه کاربری را بهبود بخشید. در این راستا "آزوما" از حرکات چرخشی دو بعدی برای تشخیص نویزهای الکترومغناطیس دریافتی از فضا و تشخیص اشیا استفاده کرده. همچنین در گذشته پروژه‌های تحقیقاتی بسیاری برای غلبه بر نویزهای اندک انجام شده است که می‌تواند کارآیی

سنسورهای مبتنی بر بینایی را در کنار سنسورهای مبتنی بر قطب نما بهبود بخشد. هرگونه عاملی که سبب ایجاد مغایرت بین اطلاعات دریافت شده از سنسورهای فیزیکی و سنسورهای بینایی شود می‌بایست توسط یک واحد منطبق مورد بررسی قرار بگیرد تا یک نتیجه قابل اعتماد را به کاربر ارائه کند. اگرچه شایان ذکر است که در این روش اجازه صرف نظر کردن از نویز های موجود را نمی‌دهد، لذا روش بررسی شده در این مقاله تحقیقاتی خطاهای محاسباتی را با یکدیگر مقایسه کرده و تا حد امکان سعی در کاهش مغایرت‌های به وجود آمده دارد.

۳. عبارات متنی ۳۶۰ درجه (Panoramic) به صورت حقیقت افزوده

در ادامه ما به بررسی مختصر پیشینه کاری خود در زمینه نقشه‌کشی و تشخیص اشیا با استفاده از تصاویر ۳۶۰ درجه می‌پردازیم. شایان ذکر است بنیان این مقاله تحقیقاتی بر اساس تشخیص اشیا و تولید عبارات متنی با استفاده از تصاویر ۳۶۰ درجه Panorama می‌باشد. در این روش از تصاویر ۳۶۰ درجه استفاده شده تا عملکرد اپلیکیشن نهایی کاملاً بی نقص و بصورت بلادرنگ در حین حرکت و تغییر زاویه دید کاربر ممکن شود. همچنین ما در اینجا به بررسی دورنمایی از استفاده از تصاویر ۳۶۰ درجه Panorama به منظور بررسی و تشخیص اشیا با استفاده از قالب‌های از پیش تولید شده که نقاط خاصی در آنها مشخص و توسط تجهیزات حقیقت افزوده قابل تشخیص می‌باشند می‌پردازیم. این روش در مقایسه با روش‌های قبلی کمک می‌کند تا یک رابط کاربری بی نقص و به دور از خطا داشته باشیم که میتواند به صورت بلادرنگ فعالیت کند.

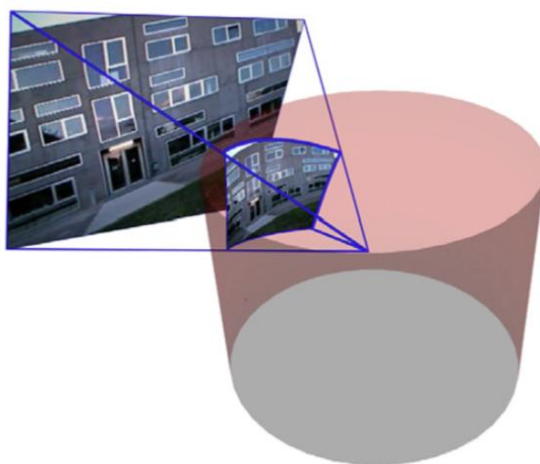
۳-۱. نقشه‌کشی و دنبال کردن آیتم‌ها در حالت ۳۶۰ درجه (Panoramic)

نقشه‌کشی و تشخیص اشیا با استفاده از تکنولوژی تصاویر Panorama با این تصور انجام می‌شود که حرکات کاربر تنها چرخشی باشند و برای مثال دوربین تلفن همراه هوشمند در یک نقطه ثابت گردد و با چرخش این دوربین به اطراف یک تصویر دوبعدی از فضای اطراف به صورت استوانه‌ای ایجاد شود.

تشخیص بخش‌های از محیط که توسط دوربین به درستی تصویربرداری نشده‌اند به الگوریتم ما کمک خواهد کرد تا فضای اطراف را بهتر تشخیص دهد چرا که نسخه کوچک‌تری از هر فریم (با ضلع کمتر از ۱۰۰۰ پیکسل) توسط الگوریتم مورد پردازش قرار می‌گیرد.

پس از پایان مراحل به‌روزرسانی تصاویر ۳۶۰ درجه، الگوریتم کلیه اشیا ی موجودیت های درون هر فریم را پردازش می‌کند و به تشخیص اشیا می‌پردازد. در این مرحله از یک پروسه جستجوی برای تشخیص استفاده می‌شود که بر پایه علم مدل سازی سه بعدی و حرکتی استوار است. نقاط کلیدی مربوط به هر فریم توسط الگوریتم تشخیص

داده می شود و سپس این نقاط با نقشه کامپیوتری تولید شده از محیط مقایسه می گردد. در مرحله بعدی نقشه کامپیوتری به موزاییک هایی تقسیم بندی می گردد که هر موزاییک در واقع یک تصویر مستطیل شکل است. عمل پردازش برای هر یک از موزاییک ها به صورت جداگانه انجام می شود و برای مثال در تصاویری که شامل پیش زمینه و پس زمینه می باشند اطلاعات در زمینه به صورت جداگانه پردازش خواهند شد. این الگوریتم به صورت بلادرنگ با سرعتی معادل ۳۰ فریم بر ثانیه فعالیت می کند و به راحتی بر روی بسیاری از تلفن های همراه



شکل ۱: نمایش یک موزاییک بعنوان بخشی از ساختار استوانه ای نقشه های تهیه شده از تصاویر پانوراما.

هوشمند میان رده همچون HTC HD2 با همین کارایی قابل پیاده سازی می باشد. شایان ذکر است بخش عمده ای از قدرت پردازشی این الگوریتم به تعداد تصاویری که دوربین تلفن همراه می تواند در هر لحظه به صورت Panorama یا همان ۳۶۰ درجه ثبت کند وابسته است. همچنین سرعت واکنش لنز دوربین تلفن های همراه هوشمند در هنگام تغییر زاویه دید در اینجا اهمیت بسزایی دارد.

۲-۳. تشخیص عبارات متنی با استفاده از قالب های آماده

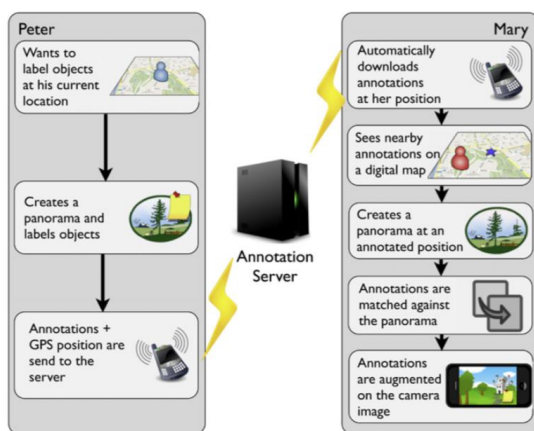
ما در این روش متدی را معرفی می کنیم که از تصاویر Panorama به منظور تشخیص محیط با فاکتورهای حقیقت افزوده به صورت زنده بهره می برد. به محض اینکه یک کاربر جدید به فضایی که حاوی اشیایی می باشد نگاه می کند، اپلیکیشن کلیه تصاویر آن اشیا را به همراه عبارات متنی مربوطه دانلود می کند. شرح عملکرد این الگوریتم در بخشی ۱-۳ مورد بررسی قرار گرفته است. سپس با مقایسه اشیا موجود در تصاویر و اطلاعات آنها با آنچه در زاویه دیده کاربر قرار گرفته است، آنچه در دید کاربر می بایست با عبارات متنی تجهیز شود تشخیص

داده می شود. این عمل با استفاده از پردازش موزاییک ها انجام می شود. از موزاییک ها برای تشخیص نقاط ثابت موجود در محیط که توسط الگوریتم به عنوان نقاط ثابت مورد استفاده قرار میگیرند هم بهره برده می شود.

به منظور تسهیل و افزایش سرعت پردازشی این روش می توان از یک تست درختی استفاده کرد. بدین منظور می بایست ابتدا با استفاده از انجام یک تبدیل "والش" دور نمای مشاهده شده را نسبت به وجود اشیا بررسی کرد تا در مرحله بعدی نیازی به پردازش های سنگین برای هر فریم بصورت جداگانه نباشد.

تشخیص مکان نمایش عبارات متنی با استفاده از مقایسه های مکرر با قالب های از پیش آماده شده به ما کمک می کند تا سرعت پردازش و پاسخ گویی را به میزان مورد نیاز و رضایت بخش بالا ببریم و در عین حال دقت بالایی را در سیستم داشته باشیم: توجه کنید که هر موزاییک تشخیص داده شده در تصاویر ۳۶۰ درجه بلافاصله پس از تصویربرداری مورد پردازش قرار نمی گیرند زیرا بودجه زمانی محدود است و پردازش تنها تا پایان بودجه زمانی مربوط به هر یک از تصاویر موجود در صف ادامه پیدا می کند.

بنابراین با تمهیدات در نظر گرفته شده، سیستم ما می تواند با سرعت نسبتاً مناسبی بر روی هر یک از تلفن های همراه هوشمند عملکردی بلادرنگ و در عین حال سریع داشته باشد تا تصاویر ۳۶۰ درجه به درستی نقشه برداری شده و اشیا موجود در آنها به درستی تشخیص داده شود. چنانچه تلفن همراه هوشمند ما دارای پردازشگر سریع تری باشد، عبارات متنی سریع تر نمایش داده خواهند شد برای مثال: در تلفن همراه هوشمند HTC HD2 تشخیص مکان ۱۲ عبارت متنی در یک زاویه دید خاص تقریباً ۲۸ میلی ثانیه، تصویر برداری برای هر فریم ۵۰ میلی ثانیه و برای تشخیص هر شی هم بودجه زمانی ۱۰ میلی ثانیه در نظر گرفته می شود. در شکل ۲ چرخش اطلاعات در سیستم مذکور بصورت کامل نمایش داده شده است: "پیتر" ابتدا یک تصویر سه بعدی به صورت Panorama تولید می کند و سپس عمل تشخیص اشیا انجام می شود. عبارات متنی درج شده پیتر، موقعیت



شکل ۲: چرخش اطلاعاتی عبارات متنی مبتنی بر تصاویر پانورامای سیستم حقیقت افزوده از دید دو کاربر. - پیتر عبارات متنی را می سازد و در یک زمان دیگر این عبارات ساخته شده توسط ماری فراخوانی و مشاهده میگردند.

GPS وی و همچنین جزئیات اطلاعاتی پیرامون اشیایی که در اطراف خود مشاهده می کند ابتدا به یک سرور منتقل می شوند. سپس برای مثال "ماری" می خواهد اطلاعات درج شده توسط پیتز را مشاهده کنند. در این صورت چنانچه ماری در موقعیت مکانی پیتز قرار بگیرد یک پیام اطلاع رسانی برای وی ارسال می شود و به او اطلاع می دهد که دوستش پیتز قبلاً در آن مکان قرار گرفته بوده و اطلاعاتی را پیرامون حواشی آن جا درج کرده است. ماری حالا پس از نگاه کردن به یک منظره تکراری که پیتز قبلاً آنرا دیده است الگوریتم را فراخوانی میکند. الگوریتم از تصاویر ۳۶۰ درجه جدید برای قرار دادن عبارات متنی استفاده می کند تا مکان قرارگیری عبارات در فضا دقیق باشند. هم اکنون ماری هم خواهد توانست تا خود به درج عبارات متنی جدید پیرامون حواشی خود پردازد.

عمده ترین ایراد این روش آن است که این روش عمدتاً بر پایه روش های تشخیصی مبتنی بر بینایی استوار است. و لذا ممکن است در صورت تغییر شرایط نوری محیط، سیستم نتایج نادرستی را به کاربر تحویل دهد و یا صحت و عملکرد درستی نداشته باشد. همچنین روش تشخیص اشیا با استفاده از قالب های از پیش طراحی شده کل تصاویر ۳۶۰ درجه را بدون در نظر گرفتن اطلاعات پیشین مورد پردازش قرار می دهد. حال آنکه در صورت استفاده از اطلاعات قبلی می توان سرعت پردازش را تا حد زیادی بهینه تر کرد تا هر بار کل فاکتور های محیطی مجدداً نیاز به پردازش نداشته باشند.

در نهایت این روش امتیازات بسیار خوبی را در آزمایشات ما کسب کرد و نتایج مفیدی را که تا ۹۰٪ موفقیت آمیز بودند تولید نمود. تست ها و جستجوهای تشخیصی در شرایط آب و هوایی و نوری مختلف همچون آفتابی و بارانی انجام شدند. البته شایان ذکر است اگر شرایط محیطی متفاوت باشند عمل تشخیص و نتایج تست های انجام شده ممکن است تا ۵۶ درصد کاهش داشته باشد که این سبب می شود این روش، روش مناسبی برای استفاده شدن در اپلیکیشن هایی که کاربرد انبوه دارند نباشد و در بلند مدت کنار گذاشته شود.

۴. روش تشخیص توسعه یافته عبارات متنی در نقشه های ۳۶۰ درجه (Panoramic)

تمرکز اصلی این بخش از فعالیت تحقیقاتی ما بهبود پروسه تشخیص مجدد تصاویر و مکان عبارات متنی در صورت تغییر زاویه دید با استفاده از روش مطرح شده در بخش ۲-۳ می باشد. بنا بر این در این روش هنگامی که کاربران به نقطه ای که از پیش با عبارات متنی علامت گذاری شده است هدایت می گردند، ابتدا GPS فعال شده و سپس هنگامی که آن ها تلفن همراه هوشمند خود را به طرف محیط های مختلف می گیرند، ما از تصاویر ۳۶۰ درجه و پردازش آنها برای تشخیص دقیق مکان عبارات متنی استفاده می کنیم. تفاوت این روش با کار های قبلی، سه متد است که برای بهبود الگوریتم قبلی و افزایش پایداری آن به کار رفته است:

اولاً کیفیت نقشه‌های تولید شده از روی تصاویر ۳۶۰ درجه می‌بایست افزایش یابد تا در هنگام مشاهده شدن همان زوایای دید توسط افراد دیگر، الگوریتم توانایی کنار آمدن با تغییرات ظاهری عکس را داشته‌باشد. بدین منظور اپلیکیشن مربوطه در هنگام تصویربرداری می‌بایست اطلاعات بیشتری را به همراه تصاویر ۳۶۰ درجه ذخیره نمایند تا نهایتاً در تهیه نقشه به صورت نرم‌افزاری به کار برود و نقشه‌ای با جزئیات و فاکتورهای اطلاعاتی بیشتر و دقیق‌تری تولید شود.

دوماً تکنولوژی‌های سخت‌افزاری به کار رفته در تلفن‌های همراه هوشمند می‌بایست به صورت بهینه‌تری مورد استفاده قرار بگیرند. بدین منظور سنسورهای فیزیکی موجود در آنها که از تکنولوژی‌های مبتنی بر بینایی استفاده نمی‌کنند می‌بایست منطقه جستجو و دنبال کردن اشیا را به صورت دقیق مشخص نمایند تا سیستم ما در هنگام پردازش دچار خطا نگردد.

سوماً ما از یک تبدیل جهانی به نام T برای منطبق کردن تصاویر جدید و تصاویر پانورامای از پیش ذخیره شده استفاده می‌کنیم. این تکنیک از اطلاعات آماری از پیش ذخیره شده به منظور هماهنگ کردن نقاط ثابت موجود در محیط با قالب‌های ذخیره شده به کار می‌رود. در واقع ما با اعمال کردن تبدیل T، مکان نمایش کلیه عبارات متنی را که در گذشته در قالب‌هایی بر روی پایگاه‌های اطلاعاتی ما، روی سرور ذخیره شده‌اند را بر روی محیط زندگی خود بصورت دقیق مشخص می‌کنیم. در ادامه جزئیات این سه متد مفصلاً بررسی شده است.

۱-۴. نقشه‌های ۳۶۰ درجه پهن شده با دامنه دید متغیر

تکنولوژی تشخیص قالب‌ها و منطبق کردن آنها با تصاویر و زوایای جدیدی که ممکن است کاربر در آن‌ها قرار بگیرد به صورت مستقیم با کیفیت تصاویر ۳۶۰ درجه و دقت تصویربرداری آنها بستگی دارد.

یکی از مشکلات اصلی در هنگام تصویربرداری، پروسه تشخیص خودکار میزان دریافت نور و حفظ تعادل رنگی توسط لنز دوربین در نسل حاضر تلفن‌های همراه هوشمند می‌باشد. البته شایان ذکر است که دوربین‌های تلفن‌های همراه هوشمند بدون نیاز به هیچ‌گونه تنظیمات خاصی نهایت تلاش خود را برای ثبت یک عکس خوب می‌کنند. اگرچه این پروسه تصویربرداری تأثیر زیادی بر روی پروسه تشخیص و دنبال کردن اشیا ندارد اما ممکن است سبب شود که محدودیتی در تعداد تشخیص‌های ما ایجاد گردد و یا برخی اشیا در یک زاویه دید خاص تشخیص داده نشوند. برای مثال در شکل ۳ مقایسه دو تصویر با میزان دریافت نور انجام شده است. به هنگام تشخیص این دو تصویر و پردازش آنها تیرگی برخی نقاط یکی از تصویرها سبب می‌شود تا برخی بخش‌های آن تصویر به صورت دقیق توسط الگوریتم نقشه‌برداری نشود. شایان ذکر است که چنانچه اطلاعات جدید دریافتی در پروسه

به روزرسانی جایگزینی اطلاعات ذخیره شده در قالب های قدیمی شود، این مشکل اهمیت دوچندانی می یابد. در واقع در نتیجه این تشخیص اشتباه می تواند بر روی قالب های از پیش ذخیره شده تأثیرات منفی بگذارد و نقاط ثابت موجود در قالب که برای پردازش ضروری هستند را مخدوش کند.

بهترین روش برای مقابله با این مشکل استفاده از دوربینی می باشد امکان تغییر دستی میزان دریافت نور را به برنامه نویس بدهد. استفاده از چنین دوربینی این امکان را فراهم می آورد که یک تصویر دقیق با دامنه گسترده تولید کنیم که کلیه نقاط تصویر را با روشنایی مورد نظر آشکار می سازد در حال حاضر تنها تلفن همراه هوشمندی



شکل ۳: (بالا) یک تصویر پانورامای ۳۶۰ درجه که با عناصر مختلف که نمای منظره را از نظر نوری متفاوت میکنند تجهیز شده. در این تصویر میزان دریافت نور بصورت خودکار توسط دوربین تلفن همراه هوشمند تنظیم شده. (پایین) یک تصویر پانورامای ۳۶۰ درجه که با بهره گیری از روش تصویر برداری با دامنه گسترده ثبت شده، سپس نقشه برداری شده و درخشندگی هر یک از بخش های تصویر پس از تصویر برداری مجدداً بهینه گشته است.

که قادر به تصویربرداری با امکان تغییر دستی میزان دریافت نور می باشد گوشی تلفن همراه نوکیا مدل N900 می باشد که با Frankenkam API تجهیز شده است. البته باید توجه داشت که به احتمال زیاد استفاده از دوربین هایی که امکان تغییر دستی کلیه تنظیمات تصویربرداری را برای کاربر فراهم سازند در آینده به صورت گسترده رایج نخواهد شد.

در همین راستا ما روشی را برای تهیه تصاویر ۳۶۰ درجه با دامنه گسترده طراحی و خلق کردیم تا بتوان با پیاده سازی آن بر روی هر تلفن همراه هوشمندی بدون نیاز به تغییر تنظیمات تصویربرداری چنین تصاویری (EDR) را با دقت بالا خلع کرد. اگرچه در این روش از تخمین های ساده استفاده شده است اما این روش آسان می تواند بسیاری از مشکلات مطرح شده در هنگام استفاده از تعیین خودکار دریافت نور را هنگام عکاسی برطرف نماید. به طور دقیق تر، ما در این روش از عکس اولیه برای تهیه تصاویر ۳۶۰ درجه خام استفاده می کنیم و سپس میزان

درخشندگی هر یک از نقاط تصویر را پس از اعمال تغییرات بر روی میزان دریافت و بر روی تصویر قبلی اعمال می کنیم که سبب تولید تصویر نهایی می شود.

ما برای دست یافتن به این تکنولوژی و الگوریتم، از مفهوم "نقاط کلیدی" در پروژه FAST استفاده کردیم. در این روش از مقایسه تصاویر ساده عکس برداری شده با استفاده از دوربین ها با نقشه های تولید شده از روی تصاویر ۳۶۰ درجه استفاده می شود. این نقاط کلیدی از پیش برای استفاده شدن در پروسه دنبال کردن اشیا تولید شده اند، لذا هیچ عملیات پردازشی به سیستم تحمیل نمی کنند. ما میزان درخشندگی و نورانی بودن هر یک از نقاط تصویر را تنها با مقایسه آنها با نزدیک ترین نقاط کلیدی تصویر که تعدادشان محدود هم هستند محاسبه می کنیم. و سپس اطلاعات آن را به نقشه تهیه شده قبلی اضافه می کنیم. این روش سبب هماهنگی دقیقی بر روی میزان درخشانی تصویر می شود و کمک میکند تا کیفیت تصویربرداری در نهایت در سطح بالایی باقی بماند. سپس نقشه ۳۶۰ درجه از روی تصویر Panorama با کانال رنگی ۱۶ بیتی تولید می شود. در تست های انجام شده کانال رنگی ۱۶ بیتی یک کیفیت مناسب برای جلوگیری از هرگونه خطا در عملیات پردازشی تشخیص داده شد. همچنین پردازش تصاویری با این کیفیت میزان حافظه زیادی در هنگام پردازش از تلفن همراه هوشمند طلب نمی کند. در ادامه بررسی تصاویر ۳۶۰ درجه با دامنه گسترده با استفاده از نقشه برداری با روش خطی صورت می پذیرد. بنابراین تصویر نهایی که قرار است پردازش شود دارای تعادل رنگی بالا و تغییرات شدید اندکی بوده و در نتیجه توسط الگوریتم به راحتی قابل آزمایش و تعبیر می باشد.

۲-۴. ادغام چند سنسور برای تشخیص بهتر

تلفن های همراه هوشمند جدید معمولاً با استفاده از چند سنسور فیزیکی همچون GPS، قطب نما، شتاب سنج و حتی سنسور تشخیص راستا نسبت به راستای شتاب جاذبه (ژایروسکوپ) تجهیز شده اند. دقت و صحت اطلاعات دریافتی از این سنسورها معمولاً آنچنان خوب نیست اما باید توجه داشت که ما از فاکتورهای دیگری هم برای دریافت اطلاعات نقشه و افزایش دقت تشخیص بهره می بریم تا یک پردازش کامل و بی نقص را انجام دهیم. بنابراین ما در این روش از کاربرد قطب نما و شتاب سنج موجود در تلفن های همراه هوشمند برای بهبود پروسه تشخیص مجدد مکان عبارات متنی استفاده کردیم.

به منظور بهبود عملیات تشخیص مجدد مکان عبارات متنی در زوایای دید مشاهده شده از پیش، منطقه اسکن شده توسط سنسور های مبتنی بر بینایی همچون دوربین در ابتدا محدود شده سپس به منظور تشخیص اشیا جستجو



شکل ۴: (پایین) تصویر پانورامای مرجع که برای ساختن عبارات متنی مورد استفاده قرار گرفته. (بالا) یک تصویر پانورامای جدید با بهترین کاندید های مشخص شده برای علامت گذاری با استفاده از عبارات متنی که موقعیت آنها هم از قالب های از پیش آماده الهام گرفته شده است. برای هر یک از نقاط ثابت که میتواند به یک عبارت متنی نسبت داده شود ما نهایتاً سه مچ در نظر میگیریم. نقطه های سبز در تصویر بالا بهترین امتیاز ها را جهت مچ شدن بدست آورده اند و لذا برای علامت گذاری مورد استفاده قرار میگیرند. نقاط قرمز هم در امتیاز گیری از سیستم در رده های دوم و سوم هستند که سبب میشوند کاندید هایی برای مچ شدن در نظر گرفته شوند.

می شود. منطقه جستجو شده سپس با قالب های موجود مقایسه می گردد. در آخر مکانی که قرار است یک عبارت متنی احتمالاً در آنجا قرار بگیرد با استفاده از مسیریابی ترکیبی با بهره گیری از همه سنسورهای موجود تعیین می گردد.

نقشه ۳۶۰ درجه به همراه اطلاعات مربوط آن از هر تصویر با رزولوشن ۳۲۰ در ۲۴۰ نقشه ای با رزولوشن ۲۰۸۰ در ۵۱۲ پیکسل تولید می کند. یک دوربین به صورت عادی ناحیه ای معادل ۶۰ درجه را می تواند تصویربرداری جمعی بنابراین نسبت طول به عرض نقشه تولید شده با استفاده از الگوریتم ما با این نسبت همخوانی نزدیکی دارد: ۳۲۰ پیکسل / ۶۰ درجه (برای $360 = 1920$ پیکسل). بنابراین رزولوشن مورد نیاز برای هر درجه معادل $2080/360 = 5.77$ درجه محاسبه می شود. اگر حداکثر میزان خطای قطب نما را ± 10 در نظر بگیریم می توانیم مکان قرار گیری عبارات متنی را در یک بازه ± 5.7 درجه محاسبه کنیم. البته دقتی معادل ۵۷ عملاً مورد نیاز نیست زیرا ما مکانی با سه برابر طول و عرض ۵۷ پیکسل را برای قرار گیری یک عبارت متنی در نظر گرفته ایم. همچنین شایان ذکر است که چنانچه عملیات تشخیص با دقت بالایی صورت می گیرد، عبارات متنی به نمایش در نخواهد آمد، و تنها در صورت تشخیص صحیح پردازش انجام خواهد شد و یا ما تنها عبارات متنی موجود در فاصله دور (در صورت دقت پایین محاسبه) را مشاهده خواهیم کرد.

۳-۴. تشخیص عبارات متنی با استفاده از یک تبدیل کلی و جهان شمول

در روش‌های پیشین در هر بار رصد کردن یک منظره تکراری مکان نمایشی عبارات متنی مجدداً پردازش و تشخیص داده می‌شدند. لذا تشخیص داده شدن مکان عبارات پس از اولین رویت در تشخیص مکان عبارات متنی در رویت‌های بعدی تأثیری نمی‌گذاشت. بنابراین آنالیز اطلاعاتی یک منظره به صورت آماری لزوماً بر قدرت تشخیص الگوریتم در آینده تأثیر مثبتی نداشت و موجب می‌شد تا هر بار امکان ایجاد نتایج اشتباه حتی با وجود قالب‌های از پیش تولید شده وجود داشته باشد. به همین دلیل استفاده از قالب‌ها به عنوان یک فاکتور قابل اعتماد تلقی نمی‌شود. چنین موقعیتی سبب شد تا نیاز به فاکتورهای هندسی جدیدی که بتوانند اطلاعات عددی ارائه کنند در محاسبات ما احساس شود.

بنابراین ما این مشکل را با محاسبه کردن فاکتورهای هندسی (طول و عرض و ارتفاع) مکان قرارگیری عبارات متنی بر روی تصاویر ۳۶۰ درجه تحلیل کردیم. بنابراین این محاسبات با استفاده از یک تبدیل جهانی به نام T گسترده و پر بار داده شدند. تبدیل مذکور عملاً کمک می‌کند تا مکان هندسی دقیق قرارگیری اشیا با مقایسه قالب از پیش تولید شده و تصویر دو بعدی که کاربر در اطراف خود مشاهده می‌کند قابل محاسبه باشد. البته ما در پروسه تحقیقات، این طور در نظر گرفتیم که تصویر Panorama در شرایط برابری نسبت به قالب‌های تولید شده تصویربرداری شود.

برای محاسبه تصاویری که در نتیجه تغییر زاویه دیده کاربر به دست می‌آید، نقاط ثابت موجود در قالب که پیش‌تر مورد بررسی قرار گرفتند، استفاده کرده‌ایم (Anchor points). چه نقاط در تصاویر Panorama ی تصویر برداری شده و همچنین در قالب اصلی موجود است. مقایسه موقعیت این نقاط در قالب و تصویر دیده شده در زاویه دید کاربر کمک می‌کند که عمل مطابق سازی با استفاده از یک وکتور سه بعدی (3D) به درستی انجام شود. ما چرخه اطلاعاتی این روش را در بخش ۲-۳ مفصلاً بررسی کرده و همچنین گفتیم که هر یک از بخش‌های تصویر ۳۶۰ درجه به صورت مجزا و در قالب موزاییک‌هایی مورد بررسی قرار می‌گیرد. در هر صورت اطلاعات مربوط به عبارات متنی و عددی آن‌ها در ابتدا به یک سرور منتقل می‌شود و سپس با موقعیت‌های مکانی مربوطه علامت‌گذاری می‌گردد (شکل ۲). شایان ذکر است که ما هیچ‌گاه تصاویر سه بعدی را مستقیماً به سرور آپلود نمی‌کنیم و تنها از اطلاعات متنی و عددی دیتاست‌ها استفاده می‌کنیم. از آنجایی که ساینز دیتاست برای هر موقعیت مکانی که ممکن است کاربر در آن قرار بگیرد بسیار اندک و در حد چند کیلو بایت است، اطلاعات ارسالی و دریافتی به راحتی می‌تواند با استفاده از یک ارتباط 3G (اینترنت همراه) انتقال یابد.

هنگامی که یک کاربر به مکانی برسد که توسط عبارات متنی علامت گذاری شده باشد، تلفن همراه هوشمند به صورت خودکار برای دریافت اطلاعات مربوط به آن موقعیت مکانی اقدام می کند. ما در این روش چنان تصور می کنیم که اطلاعات GPS به تنهایی از سلامت و دقت کافی برخوردار نباشد و بنابراین کلیه دیتاست های مربوط به آن مکان تا شعاع ۵۰ متر دانلود خواهند شد. پس از دانلود دیتاست ها، نقاط ثابت موجود در محیط به سرعت تشخیص داده خواهند شد و سپس با قالب های از پیش تولید شده مقایسه می شوند. سپس تشخیص کلیه اشیای موجود در تصویر انجام می گردد. به جای استفاده از تنها یک نقطه ثابت برای تشخیص سایر مکان ها از سه نقطه تشخیص داده شده با دقت بالا استفاده خواهیم کرد تا مکان سایر اشیای را در سه بعد طول، عرض و ارتفاع با بالاترین دقت تشخیص به و سپس عبارات متنی را در جای صحیح نمایش دهیم.

اگرچه این پروسه تشخیص اشیای و دنبال کردن آنها به صورت آنلاین ادامه پیدا می کند اما به روزرسانی مکان دقیق قرارگیری عبارات متنی هم در زمینه انجام خواهد شد. یک پردازش براساس RANSAC هم انجام می شود تا مکان قرارگیری عبارات متن پس از انجام هر تبدیل T به روزرسانی گردند. این تعبیر عمدتاً کمک می کند تا عمل مطابقت تصویر جدید با قالب های از پیش تولید شده با دقت بالایی انجام شود.

ما در این روش از دو نقطه، از سه کاندیدای در نظر گرفته شده استفاده و آنها را پردازش می کنیم. ما از آنها به عنوان اطلاعات اولیه پردازش RANSAC بهره میبریم. به منظور تشخیص میزان تفاوت قالب با منظره ای که کاربر در حال مشاهده آنست، ابتدا تفاوت زاویه ای محاسبه میشود. سپس متناسب با این تفاوت زاویه ای (تفاوت زاویه دید تصویر از پیش گرفته شده با دید کاربر)، دو بردار برای هر یک از عبارات متنی ترسیم و به نام های A_1 و A_2 و B_1 و B_2 نامگذاری می گردد. سپس بردار L که بر هر جفت از بردارها عمود است محاسبه میشود. در مکان یابی دقیق ترسیم این سه بردار یعنی طول عرض و ارتفاع دقیق ترین تخمین مکانی را به ما ارائه میکند. همه نتایج به دست آمده با اطلاعات قالب های از پیش آماده ی ۳۶۰ درجه مقایسه می گردند و سپس نتایج حاصل بر روی تصویر دو بعدی پیاده سازی می شوند تا مکان قرارگیری عبارات به درستی بر روی نقشه نهایی مشخص شود. در نهایت عبارات متنی تنها در صورتی نمایش داده می شوند که دقت محاسبه از میزان اولیه (THRSHHOLD) بیشتر باشد. شایان ذکر است کلیه محاسبات بر روی تصاویر دو بعدی انجام شده و اندازه گیری ها به صورت خطی در محاسبات لحاظ میگردند.

به عنوان یک نظریه جانبی می توان گفت که میزان خطای نرمال این الگوریتم با تقسیم میزان خطای کل بر تعداد اشیای موجود در هر فریم تصویر قابل محاسبه است. در این صورت امتیاز محاسبه شده می تواند جایگزین فاکتور محاسباتی خطا در روش قبلی گردد. پروسه تبدیل T و تشخیص مجدد آنقدر انجام می شود تا میزان خطا به کمتر

از ۵۰ درصد میزان ورودی (Threshold) برسد. در تحقیقات ما تعداد ۱۰ پیکسل به عنوان ورودی (Threshold) در نظر گرفته شده است که یک میزان مناسب و قابل چشم پوشی در هنگام محاسبات محسوب میشود و نتایج قابل اعتمادی را بدست می دهد.

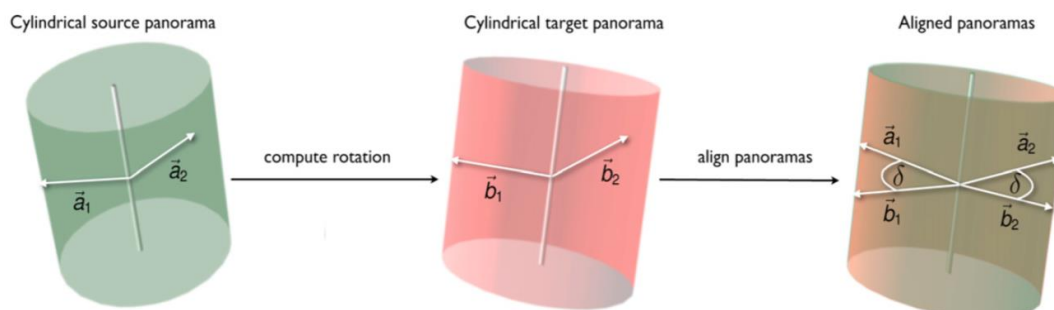
پیدا کردن بهترین موقعیت دو بردار برای تشخیص یک سه محور طول عرض و ارتفاع به هنگام مقایسه تصاویر Panorama نیازمند پردازشی می باشد که تقریباً ۳۰ میلی ثانیه به طول می انجامد. این پردازش می تواند هشت عبارت متنی را به درستی تشخیص دهد حتی اگر کلیه این هشت عبارت در یک موزاییک از فریم تصویر برداری شده قرار نداشته باشند. شایان ذکر است که عمل بررسی برای یافتن سه بردار طول و عرض و ارتفاع برای هر منظره تنها یک بار انجام میشود و سپس نقاط ثابت به سرعت با استفاده از سنسورهای مبتنی بر بینایی (تصویر دریافتی از لنز دوربین) در محیط تشخیص داده میشوند و موقعیت موزاییک های دیگر فریم بصورت خودکار نسبت به موزاییک حاوی بردارها محاسبه میشود.

۵. تشخیص زاویه نسبت به محورهای طول، عرض و ارتفاع به روش هیبریدی

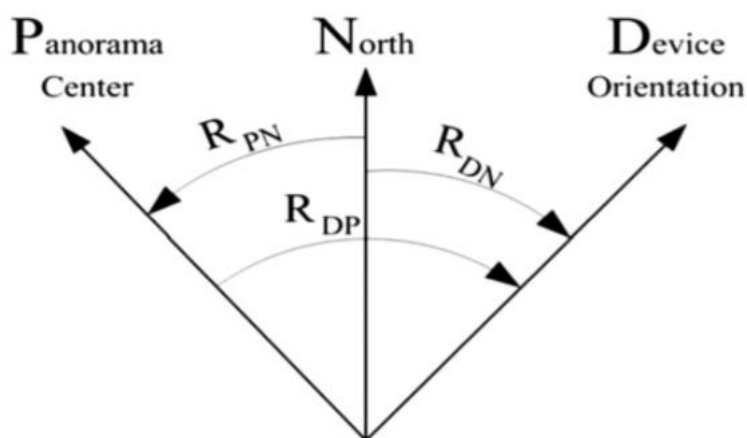
به منظور تجهیز محیط با استفاده از فاکتورهای حقیقت افزوده، ما پس از اینکه نقاط ثابت مربوط به عبارات متنی را در محیط با موفقیت تشخیص دادیم، نیاز داریم تا زاویه دید را نسبت به محورهای طول عرض و ارتفاع با دقت بالایی محاسبه نماییم. شرح مختصری از این روش در بخش ۲-۴ بررسی شد. همچنین ذکر کردیم که برای محاسبات مربوط به این بخش از قطب نما و شتاب سنج های خطی بهره می بریم زیرا استفاده از سنسورهای مبتنی بر بینایی به دقت بالای مورد نیاز ما را ارائه نمی کند. همچنین شایان ذکر است که پردازش اطلاعات دریافتی از سنسورهای مبتنی بر بینایی در هنگام حرکت سریع کاربر به صورت بلادرنگ و بدون تحمیل بار پردازشی سنگین به پردازنده قابل پردازش نیستند و سبب کاهش کارایی الگوریتم و میزان دقت آن می گردند. بنابراین ما از ترکیب این دو روش یعنی استفاده از سنسورهای مبتنی بر بینایی و سنسورهای فیزیکی برای محاسبه زاویه نسبت به محورها استفاده می کنیم.

البته باید متذکر شویم که استفاده از سنسورهای مبتنی بر بینایی برای تشخیص زاویه نسبت به محورها کافی خواهد بود اما اطلاعات دریافتی از این سنسورها نسبی می باشد. بنا بر این این اطلاعات برای تشخیص مکان در موقعیت های جهانی کافی نمی باشد. لذا از سنسورهای دیگری همچون GPS همانطور که گفته شد برای تشخیص موقعیت برای کره زمین استفاده می کنیم.

به عبارت دیگر ما علاوه بر استفاده از فاکتورهای سنسورهای فیزیکی و همچنین سنسورهای مبتنی بر بینایی، از اطلاعات دریافتی به صورت آنلاین هم بهره خواهیم برد تا بصورت مداوم تشخیص های نادرست را تصحیح و دقیق ترین نتیجه را در سیستم داشته باشیم.



شکل ۵: تصویرها موقعیت سه استوانه که در واقع همان تصاویر پانوراما هستند را در موقعیت های مختلف نشان می دهند. این سه موقعیت نسبت به مکان فرضی سه نقطه ثابت (Anchor point) که مکان نمایش عبارات متنی را مشخص میکنند در نظر گرفته می شوند a_1 و a_2 به مکان دو عبارت متنی روی فضای درونی استوانه (روی عکس ۳۶۰ درجه) اشاره میکنند. استوانه وسطی موقعیت عکس قالب گرفته شده با دوربین تلفن هوشمند را نشان می دهد. بردارهای b_1 و b_2 به دو نقطه احتمالی در تصویر جدید اشاره میکند. چرخش دو استوانه در یکدیگر برای انطباق بر شرایط تصویر اصلی در محیط واقعی صورت میگیرد. این کار با استفاده از محاسبه خطای سیکما که بعداً در محاسبه RANSAC مورد استفاده قرار میگیرد انجام میشود تا نهایتاً میزان خطا به کوچکترین مقدار برسد.



شکل ۶: دورنمای چرخش در موقعیت جهانی نسبت به راستای N یا همان راستای شمال جنوب. مرجع سیستم ما محور D و مرجع تصاویر پانوراما با محور P نمایش داده شده است.

ما در ابتدا با استفاده از سنسور الکترو مغناطیسی موجود، راستای شمال، جنوب را بصورت دقیق و بدون نیاز به سنسورها و ابزارهای جانبی تشخیص می دهیم. سپس همین سنسورها به بررسی میزان شتاب جاذبه و در نتیجه زاویه انحراف از آن می پردازند که در نتیجه زاویه R_{DN} محاسبه میشود که همان زاویه انحراف از راستای جاذبه است (شکل ۶).

بدین منظور ابتدا دستگاہ ما میزان انحراف تصویر مشاهده شده D را نسبت به تصویر مرجع P (همان قالب دریافت شده از سرور) محاسبه میکند. هدف ما در این مرحله محاسبه میزان انحراف نسبت به شمال یا همان R_{PN} می باشد. زیرا تصویر مرجع بصورت صاف و در راستای صحیح شمال جنوب (بدون انحراف) گرفته شده است. با در نظر گرفتن R_{DP} و آنچه گفته شد میزان انحراف نسبت به راستای شمال جنوب بدست می آید:

$$R_{DP}R_{PN} = R_{DN}$$

$$R_{PN} = R_{DP}^{-1}R_{DN}.$$

با استفاده از معادله ۲ زاویه R_{PN} محاسبه می شود. ما در هر بازه زمانی t ، فاکتور شتاب جاذبه یعنی g ، و m_t را به منزله میزان میدان مغناطیسی موجود محاسبه میکنیم. m و g به ما کمک میکنند تا جهان اطراف را به درستی از نظر اطلاعاتی تعریف کنیم و بتوانیم سه زاویه انحراف نسبت به سه محور را با دقت بالایی بصورت بردار $R_{DN} = [r_x \ r_y \ r_z]$ داشته باشیم، بصورتی که:

$$g_t = R_{DN}g, \quad \text{and}$$

$$m_t r_z = 0.$$

بنا بر این ما با محاسبه فاکتور های دو بردار R_{DN} و R_{PN} کلیه اطلاعات لازم برای تشخیص موقعیت هندسی خود در فضا را بدست آوردیم. این هدف، که با بهره گیری از میزان میدان مغناطیسی و اطلاعات شتاب سنج بدست آمده ممکن است دارای خطای زیادی باشد زیرا محاسبه میدان مغناطیسی همیشه با نویز زیادی همراه است. شایان ذکر است که ما به منظور جلوگیری از محاسبات مجدد در هر بار محاسبه R_{PN} از فیلتر Kalman یا همان (EKF) استفاده میکنیم. به منظور در نظر گرفتن حالت فیلتر ما حالات چرخش زاویه دید نسبت به سه پارامتر طول و عرض و ارتفاع را در گذر زمان t مدل میکنیم. در این عمل از اطلاعات نقشه 3 Lie در $SO(3)$ استفاده میکنیم. ما وضعیت فیلتر را در هر بازه زمانی t بررسی میکنیم و تغییر وضعیت نسبت به سه محور (μ) را مدام به دست می آوریم. این پارامتر میزان خطای R_{PN} را در هر چرخش زاویه دید بدست می دهد. μ در واقع پخش نرمال $\mu \sim N(0, P_t)$ است و عکس واریانس P می باشد. این فاکتور در واقع R_t تخمین زده شده را به R_{PN} واقعی با مقادیر دقیق طبق محاسبه زیر نسبت می دهد:

$$R_{PN} = \exp(\mu)\hat{R}_t$$

در اینجا از \exp برای محاسبه میزان چرخش زاویه ای R از روی جبر $so(3)$ Lie استفاده می شود. در مقابل از $\log(R)$ چرخشی از $SO(3)$ را از طریق جبر Lie محاسبه میکند. از آنجایی که ما همواره در حال محاسبه یک

سری ثوابت هستیم تغییرات این ثوابت را در هر لحظه از طریق مدل حرکتی که تولید کردیم پیش بینی میکنیم، بطوری که در اینجا μ ثابت بوده و اندازه عکس واریانس رشد میکند (دقت بالا تر میرود). این رشد از محاسبه دقیق نویز ناشی میشود.

اندازه گیری های ما در معادله در حالت فیلتر μ در نهایت به وضوح بیان میکند که فاکتور محاسباتی ما برابر مقدار چرخش R در لحظه حاضر میباشد و تفاوت اصلی در میزان چرخش حول راستای شمال جنوب است زیرا انحراف از این راستا در مرحله قبل محاسبه شده است و تنها میزان چرخش مورد نیاز است:

$$\log(R_{DN}\hat{R}_t^{-1}) = \mu$$

ژاکوبین معادله ۶ در واقع همان ماتریس میزان چرخش است که مقادیر چرخش را ثبت میکند. این فاکتور مدام برای بروز رسانی حالت μ در فریم ورک فیلتر Kalman بکار میرود. در نهایت ما با لحاظ کردن خطای محاسبه شده در بردار قدیمی، پارامترهای دقیق بردار R را محاسبه میکنیم. این کار با ضرب چپ $\exp(\mu)$ در آن انجام میشود. در نهایت میزان خطا مجدداً توسط الگوریتم ۰ در نظر گرفته شده و پردازش از سر گرفته می شود.

محاسبه موقعیت دقیق زاویه دید کاربر در جهان و نحوه رویت منظره با استفاده از موقعیت دریافتی از سنسور های فیزیکی همچون GPS و فاکتور های مبتنی بر بینایی محاسبه شده یعنی RPN و RDN بصورت دقیق همانطور که گفته شده محاسبه میگردد. در واقع RPN با بررسی نتایج سنسور های فیزیکی و RDN با بررسی نتایج دریافتی از مقایسه تصاویر و قالب ها (سنسور های مبتنی بر بینایی) بدست آمد. RDP هم راستای شمال جنوب محاسبه شده اولیه طبق معادله ۱ بود. سپس در نهایت با ترکیب نتایج نسبی با مدل سازی علمی بر پایه فیلتر مذکور نتایج دقیق نهایی بدست آمدند. الگوریتم نهایی یک پروسه بازگشتی، سریع و کارا است که به قدرت پردازشی و حافظه کمی نیاز دارد.

۶. آزمایش ها و نتایج

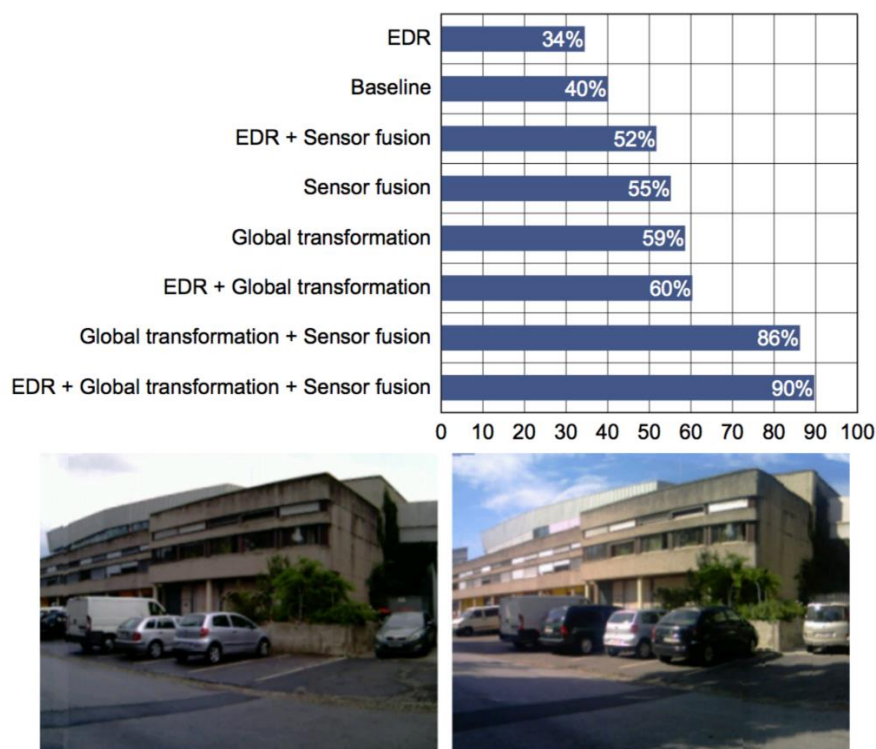
ما پس از پیاده سازی پروژه، آن را بر روی یک تلفن همراه هوشمند میان رده یعنی HTC HD2 اجرا و به ارزیابی همه جانبه آن پرداختیم. در تست های انجام شده دو معیار بیش از هر چیز مد نظر بودند: اولاً قابلیت تشخیص مجدد و سرعت پردازش الگوریتم برای تولید عبارات متنی (پس از اولین رویت) با استفاده از نقاط ثابت موجود در تصاویر، و دیگری صحت روش هیبریدی در تشخیص اشیا و دنبال کردن آنها. در ادامه به بررسی هر یک از این موارد به صورت مفصل می پردازیم.

۱-۶. کارایی مند تشخیص مجدد اشیا (پس از اولین رویت)

ما به منظور بررسی کارایی عملیات تشخیص مجدد اشیا پس از اولین رویت، ۱۲ تصویر Panorama به صورت ۳۶۰ درجه تهیه کردیم. ما با این ۱۲ عکس شرایط جغرافیایی، آب و هوایی و همچنین موقعیت‌های نوری متفاوت را در تست‌های خود گنجانده‌ایم و گسترده‌ای از شرایط محیط زیستی مختلف را مورد ارزیابی قرار دادیم. این تصاویر از فاصله پنجاه متری تهیه شدند و هر کدام دارای ۴ الی ۶ عبارت متنی بودند که در نقاط مختلف به اشیای گوناگون نسبت داده شدند. لذا مجموعاً ۵۸ عبارت متنی در سیستم قرار داده شد. این تصاویر با استفاده از روش دامنه پویا با کانال رنگی ۸ بیتی که در بخش ۱-۴ بررسی شدند، تصویر برداری گردیدند. ما همچنین یک ویدیوی گرفته شده بصورت ۳۶۰ درجه را در تست‌های خود به عکس تبدیل و نتایج آن را در تست‌های خود مورد ارزیابی و تحلیل قرار دادیم.

کارایی سیستم تشخیص مجدد در شرایط نوری مختلف مطابق شکل ۷ مورد بررسی قرار گرفت. ما تصاویر ۳۶۰ درجه Panorama را یک بار، یک ساعت پیش از غروب آفتاب در یک روز آفتابی تهیه کردیم و در دفعه بعدی سعی کردیم عبارات متنی را در هنگام رویت مجدد منظره به هنگام ظهر مجدداً توسط سیستم تشخیص دهیم. در هنگام تشخیص مجدد، شرایطی وجود داشت که سبب شده بود تا برخی نقاط همچون قسمت‌های داخلی ساختمان، سایه‌های تیره‌ای را دربرداشته باشد. ما همچنین شرایط نوری مختلف را در قسمت‌های مختلف تصاویر با استفاده از ابزارهایی مختلف همچون لنزها و نقاط نورانی ایجاد کردیم. این اقدامات را انجام دادیم تا شرایط تشخیص مکان نمایش عبارات متنی را برای سیستم طراحی شده خود سخت تر کنیم.

از آنجایی که روش طراحی شده ما نیاز دارد تا در هنگام رویت مجدد منظره، کاربر در مکان قبلی حاضر شده باشد، مادر هنگام رویت مجدد، مکان قرارگیری کاربر را دو متر تغییر دادیم. با توجه به اینکه پارامترهای ارائه



شکل ۷: عملیات ارزیابی پروسه تشخیص مجدد اشیا. (بالا) دورنمای نتایج مجدد تشخیص داده شده. (پایین) بخش هایی از دو تصویر پانوراما که شرایط مختلف محیطی را در هنگام ارزیابی نمایش می دهند.

شده توسط سنسور GPS گاهی با خطاهایی همراه بودند، گاهی مشاهده می کردیم که علاوه بر اطلاعات مربوط به یک منطقه مشخص در محیط دید کاربر، دیتاست های منطقه همسایه، یعنی ۵۰ متر دورتر هم توسط سیستم دانلود می شد. اگرچه شایان ذکر است که اطلاعات تنها یکی از این نواحی توسط سیستم پردازش می شد و به نمایش در می آمد زیرا از نتایج و امتیازات دقیق تری در سیستم برخوردار بود.

عملیات بررسی و تحلیل پردازش مجدد اشیا در رویت های بعدی به گونه ای طراحی شده بود که کلیه بخش های الگوریتمهای تشخیص مجدد به صورت خودکار و سیستماتیک مورد ارزیابی دقیق قرار بگیرند تا صحت کارایی آنها تشخیص داده شود. در تست های انجام شده قدرت تشخیص معادل ۴۰٪ ارزیابی شد که البته این میزان از میزان تئوری پیش تر گزارش شده کمتر است. علت این امر هم شرایط سخت تری است که در هنگام تست الگوریتم و پردازش های آن بکار رفته است. نتایج تست ها در شرایط مختلف در شکل ۷ نمایش داده شده اند. استفاده از ترکیبی از این روش ها، قدرت تشخیص را تا ۱۵٪ افزایش می دهد تا جایی که روش RANSAC تعداد نقاط کافی را به منظور پردازش نتایج صحیح در محیط به درستی تشخیص می دهد. نهایتاً تست های انجام شده به منظور تشخیص زاویه انحراف نسبت به محور ها و مکان یابی بین المللی صحت عملکردی معادل ۸۶٪ را ارائه

کردند. همچنین استفاده از روش EDR در تست ها تغییر زیادی در نتایج نداد و به نظر میرسد که حتی در موقعیت های سخت استفاده از این روش ممکن است نتایج صحیح بدست آمده را دچار خطا هایی بکند. در هر صورت استفاده از ترکیبی از همه این روش های بهبود بخش، سبب بالارفتن میزان صحت تشخیص مجدد تصاویر در رویت های بعدی تا میزان ۹۰٪ شد که این میزان از دو برابر کارایی پیش بینی شده از نظر آماری بالاتر است و برای کاربری روزانه توسط کاربر عادی مناسب در نظر گرفته می شود.

۲-۶. صحت عملکرد روش هیبریدی در دنبال کردن اشیا

با استفاده از ترکیبی از سنسورهای فیزیکی و مبتنی بر بینایی ما میزان انحراف زاویه دید دوربین تلفن همراه هوشمند را نسبت به سه محور طول، عرض و ارتفاع با دقت بالایی در یک ماتریس ۳ در ۳ ثبت می کنیم. همان طور که گفته شد موقعیت نسبت به محور ها با مقایسه زاویه دید حاضر و تصاویر مرجع که در قالب ها گنجانده شده اند تشخیص داده می شود. همچنین شایان ذکر است که فاصله دوربین تلفن همراه هوشمند تا نقاط مرجع موجود در قالب ها بین ۲۶ تا ۹۲ متر در حالت عادی می باشد.

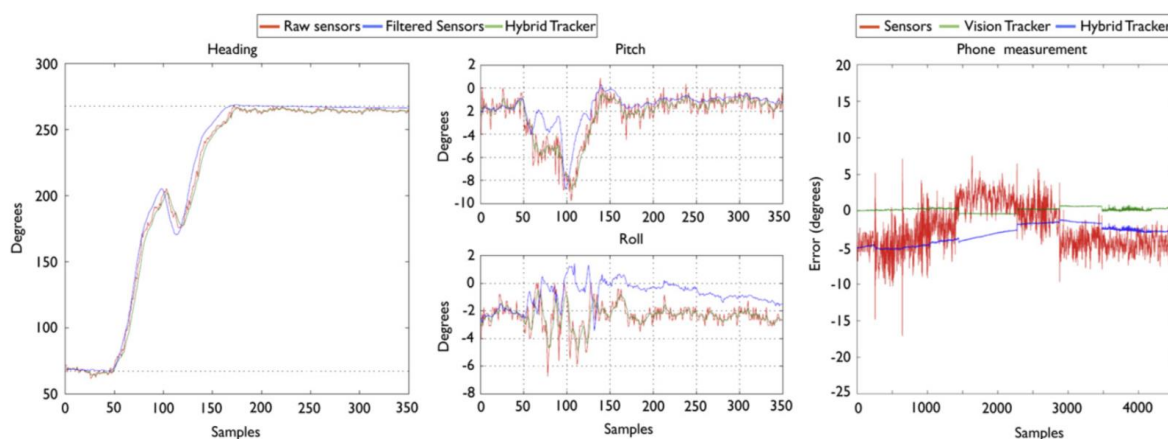
به منظور افزایش دقت تصویربرداری در هنگام تهیه قالب های مرجع، دوربین بر روی یک سه پایه قرار داده می شد. ما به منظور ایجاد یکپارچگی در صحت اندازه گیری اطلاعات، پس از اطمینان از صاف بودن موقعیت دوربین نسبت به منظره، آن را به همان حالت چرخانده و در مقابل نقاط مرجع دیگر قرار می دادیم و همچنین میزان انحراف نسبت به محورهای طول، عرض و ارتفاع را در تمام مدت ثبت می کردیم.

اندازه گیری میزان نویز در هنگام ارزیابی ها و تست های انجام شده با استفاده از مشاهده تمام وقت سنسورها و عملکرد آنها انجام می شد. ما در واقع از دو تابع مختلف برای محاسبه مقادیر استفاده می کردیم. یکی از توابع برای محاسبه عددی فاکتور مورد نظر، و دیگری برای لحاظ کردن تأثیر نویز در محاسبات. در عمل، سنسور های مبتنی بر بینایی خطا و نویز کمتری دارند که در بلند مدت کاهش هم میباید. زیرا این میزان خطا در هنگام تبدیل شدن داده های سه بعدی به دو بعدی برای هر عکس، در واقع به توان میرسد و کوچک تر میشود. در شکل ۸ (چپ و وسط)، نمایی از یک جلسه اندازه گیری نمایش داده شده است. در عین اینکه دوربین تلفن همراه هوشمند در جهت عقربه های ساعت می چرخد، دید دوربین از موقعیت دید یک نقطه ثابت به مکان نقطه ثابت دیگر تغییر موقعیت می دهد. در شکل ۸ (راست) خطای اندازه گیری مربوط به نزدیک ترین نقطه ثابت نمایش داده شده است که به تدریج در حال کاهش می باشد.

نتایج نشان دهنده دو روش بهبود بخش به هنگام استفاده صرف از سنسورها برای تشخیص انحراف زاویه ای میباشند: اولاً نویز با فراکانس بالا در روش ما به سرعت با یک لگ زمانی کوچک تصحیح میشود و صحت عملکرد آن حتی در هنگام حرکات سریع نیز کاهش نمی یابد. علت این امر آن است که تشخیص موقعیت بر اساس فاکتورهای بینایی بر تشخیصهای فیزیکی احاطه دارد و لغزشهای کوچک در حرکت، بر محاسبات و تخمینها تأثیری نمی گذارد. همچنین چرخشهای کوچک توسط الگوریتم فیلتر شده و تأثیرات منفی نویزهای مربوط به قطب نما را خنثی می کند. استفاده ترکیبی از این دو روش نهایتاً به یک عملکرد بدون وقفه و بدون نویزهای قابل مشاهده، منجر می شود. سلامت این عملکرد در هنگام حرکت سریع هم صحت و کیفیت خود را حفظ میکند. علت آن هم استفاده از ترکیب دو نوع سنسور است که هر کدام به صورت مستقل فعالیت میکنند و همکاری تنگاتنگ دارند.

۷. نتیجه گیری و کارهای آتی

ما روشی را برای تشخیص و دنبال کردن مکان نمایش عبارات متنی ارائه کردیم که بر روی هر اپلیکیشن تجهیز شده با استفاده از حقیقت افزوده که بر روی تلفنهای همراه هوشمند هم قابل دسترسی و پیاده سازی می باشد، ارائه کردیم.



شکل ۸: (سمت چپ و وسط) نمای نوک جسم، چرخش و مسیر حرکت برای یک حرکت آزادانه دست به همراه تلفن همراه هوشمند بین دو نقطه ثابت مرجع. ما عمل ثبت اطلاعات را بر اساس اطلاعات خام دریافتی از سنسورها، تخمین فیلتر شده و با روش ترکیبی (هیبریدی) رسم میکنیم. (راست) شرایط تست که خطاها را در مسیر رو به شمال بر روی تلفن همراه ثبت کرده است. تشخیص با استفاده از سنسورهای مبتنی بر بینایی تنها بعنوان مرجع برای بهبود اطلاعات استفاده می شود زیرا موقعیت دقیق نسبت به محورهای طول عرض و ارتفاع معمولاً در آنها مشخص نیست.

روش مطرح شده این امکان را فراهم می آورد تا پس از یک بار بررسی کردن یک منطقه محیطی و قرار دادن عبارات متنی با استفاده از اپلیکیشنهای حقیقت افزوده، در رویت های بعدی کلیه آن عبارات متنی در هنگام

نگاه کردن به زوایا و مناظر مختلف به صورت زنده نمایش داده شوند. عبارات متنی نوشته شده توسط کاربر اول، توسط کاربران دیگر در هنگام رویت مجدد منظره فراخوانی می شوند. همچنین شایان ذکر است مناظر و محیط زندگی کاربران با استفاده از تهیه تصاویر ۳۶۰ درجه Panorama تصویربرداری و به نقشه های قابل پردازش کامپیوتری تبدیل می گردند. به منظور محدود کردن منطقه جستجوی اشیا ما از پارامترهای هندسی بهره بردیم. پس از تشخیص عبارات متنی و محدود کردن آنها به اشیا مختلف در محیط دید کاربر، از روش هیبریدی (ترکیب سنسورهای فیزیکی و مبتنی بر بینایی) برای محاسبه دقیق مکان و موقعیت کاربر بر روی زمین و انحراف زاویه دید کاربر نسبت به محورهای اصلی استفاده می شود. پس از بررسی های انجام شده، متوجه می شویم که عملکرد این روش از روش های پیشین سالم تر و از صحت بالاتری برخوردار است. همچنین استفاده از ترکیبی از روش ها که در این مقاله تحقیقاتی مورد بررسی قرار گرفت، کلی ضمن حل مشکلات روش های پیشین دقت بالایی را در هنگام تشخیص مجدد محیط ارائه می کند که میزان آن بالغ بر ۹۰٪ می باشد. در روش ترکیب سنسورها، اطلاعات دریافتی در هنگام تشخیص میزان انحراف نسبت به محورهای طور عرض و ارتفاع، با دقت بالاتری ارائه میگردد و کیفیت محاسبات تخمینی هم افزایش پیدا می کند. شایان ذکر است که روش های مطرح شده به صورت گسترده و عمومی می توانند بر روی هر گجت یا اپلیکیشن دیگر مبتنی بر حقیقت افزوده پیاده سازی شوند. اما به صورت تخصصی تر کارایی این روش و عملکرد آن بر روی تلفن های همراه هوشمند میان رده با کیفیت پایین سنسورهای بکار رفته در آنها و قدرت پردازشی محدودشان مفصلاً مورد ارزیابی قرار گرفته است. کارهای تحقیقاتی آتی که در این زمینه انجام خواهد شد می بایست بصورت تخصصی تر به بررسی و تحلیل مشکل پیدا کردن دقیق تر و بهینه تر نقاط ثابت (Anchor points) در قالب ها و محیط دید کاربر بپردازد.

اگر چه ما در این روش از ذخیره کردن موزاییک های تصویری برای تشخیص نقاط ثابت موجود در محیط استفاده کردیم، اما تشخیص اشیا با استفاده از صفات واقعی آنها در آینده کمک بسزایی به عمل تشخیص مجدد اشیا به صورت بلادرنگ خواهند کرد. متأسفانه قابلیت تشخیص اشیا در این روش اگرچه از دقت بالایی برخوردار است، اما هنوز کاملاً قابل اعتماد نیست و به صورت عمده به کیفیت تصاویر گرفته شده بستگی دارد. این مسئله می تواند موضوع تحقیقاتی سایرین را در آینده تشکیل دهد. بررسی های بیشتری هم می تواند به منظور بهبود ساختار دیتاست های اطلاعات متنی علامت گذاری شده با اطلاعات GPS انجام شود که این ساختارها را با صفات برگرفته از تصاویر بهبود بخشد. همچنین بررسی های دیگری نیز در این زمینه مورد نیاز است تا رابطه میان تصاویر دامنه گسترده با عمل تشخیص مجدد اشیا بهتر درک شود.

سایر فعالیت‌های تحقیقاتی میبایست پیرامون عمل دنبال کردن اشیا انجام شود. از آنجایی که اطلاعات دریافتی از سنسورهای مبتنی بر بینایی برای تشخیص میزان انحراف از محورها نسبی است، خطاهای محاسبه شده در هنگام تصویر برداری با استفاده از لنز دوربین‌های مختلف (غیر از دوربین استفاده شده در این پروژه) که پارامترهای عددی مختلفی دارند می‌تواند پردازش‌ها را با خطاهایی مواجه سازد و نتایج دریافتی را تغییر دهد.

در نهایت استفاده صرف از یک سیستم پردازشی برای محاسبه دقیق میزان‌های خطا کاملاً ایده آل نیست زیرا باید توجه شود که فیلترهای ریاضی تولیدی همگی بر پایه اطلاعات و اندازه‌گیری‌های عددی انجام شده فعالیت می‌کنند و خطاهای دریافتی را تنها با استفاده از میانگین‌گیری نرمال‌سازی یا از بین می‌برند. بنابراین اندازه‌گیری و بررسی یک محیط خاص در طی زمان‌های طولانی ممکن است سبب افزایش خطاها و همچنین کاهش دقت اندازه‌گیری سنسورها شود (از حالت کالیبره خارج شود). بنابراین تهیه یک مدل مفهومی جهان شمول به منظور محاسبه میزان خطاهای اندازه‌گیری ضمن بهبود شیوه اندازه‌گیری قبلی می‌تواند قابلیت اعتماد به اعداد دریافتی را هم در روش دینامیک و هم در هنگام کالیبره کردن سنسورها افزایش دهد.

تقدیر:

کلیه هزینه‌های این کار تحقیقاتی توسط آزمایشگاه حقیقت افزوده کریستین داپلر و با کمک آژانس تحقیقاتی اتریش، به شماره قرارداد FIT-IT 820922 تامین گردیده است. همچنین فاند تحقیقاتی بنیاد علوم اتریش هم (به شماره W1209) به این پروژه اختصاص داده شده است. همچنین لازم است از Stadtvermessungsamt Graz بخاطر ارائه دیتاست‌های مطابق با جهان واقعی شان تشکر کنیم.