دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

پیدا کردن هدف در تصاویر هوایی با استفاده از توصیف‌گرهای جدید DOOS

پایان‌نامه کارشناسی ارشد مهندسی برق- مخابرات

غلامرضا رمضانیان

اساتید راهنما

دکتر بهزاد نظری

دکتر رسول امیرفتاحی

۱۳۹۲

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

پیدا کردن شی در تصاویر هوایی با استفاده از توصیف‌گرهای جدید DOOS

پایان‌نامه کارشناسی ارشد مهندسی برق- مخابرات

غلامرضا رمضانیان

اساتید راهنما

دکتر بهزاد نظری

دکتر رسول امیرفتاحی

۱۳۹۲

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

پایان‌نامه‌ی کارشناسی ارشد رشته‌ی مهندسی برق – مخابرات آقای غلامرضا رمضانیان

تحت عنوان

پیدا کردن هدف در تصاویر هوایی با استفاده از توصیف‌گرهای جدید DOOS

در تاریخ توسط کمیته‌ی تخصصی زیر مورد بررسی و تصویب نهایی قرار گرفت.

۱- اساتید راهنمای پایان‌نامه دکتر بهزاد نظری

دکتر رسول امیرفتاحی

سرپرست تحصیلات تکمیلی دانشکده دکتر مسعود عمومی

بر خود لازم می‌دانم که از توجه، راهنمایی و تشویق دوستان مخصوصاً آقایان کاظمی،کریمی‌مهر و مزروعی سپاسگزاری کنم.

کلیه‌ی حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه (رساله) متعلق به دانشگاه صنعتی اصفهان است.

تقدیم به:

پدر و مادر مهربانم

...

فهرست مطالب

عنوان **صفحه**

فهرست مطالب ..... هشت

چکیده ۱

فصل اول: مقدمه

۱-۱ بیان مسئله ۲

۱-۲ کارهای انجام‌شده ۴

۱-۳ نوآوری پایان‌نامه ۵

۱-۴ روند نگارش پایان‌نامه ۵

فصل دوم: شناسایی شی

۲-۱ چالش‌های شناسایی شی ۷

۲-۱-۱ تغییرات خطی و غیرخطی روشنایی ۸

۲-۱-۲ تغییر مقیاس ۸

۲-۱-۳ چرخش ۹

۲-۱-۴ شلوغی صحنه ۹

۲-۱-۵ پوشیده شدن قسمتی از شی ۹

۲-۱-۶ تغییر نقطه‌ی دید ۹

۲-۲ روش‌های جستجوی تصاویر الگو ۱۱

۲-۲-۱ روش همبستگی ۱۱

۲-۲-۲ بردارهای ویژگی سراسری ۱۷

۲-۲-۳ آنالیز مؤلفه‌های اصلی ۲۰

۲-۲-۴ جستجو در فضای انتقال ۲۱

۲-۲-۵ روش‌های دیگر ۲۶

فصل سوم: شناسایی اشیاء با استفاده از ویژگی‌های محلی

۳-۱ کلیات مسئله‌ی شناسایی شی با استفاده از ویژگی‌های محلی ۲۸

۳-۲ روش SIFT ۲۹

۳-۲-۱ آشکارساز نقاط کلیدی SIFT ۳۰

۳-۲-۲ پیدا کردن زاویه‌ی نقاط کلیدی ۳۴

۳-۲-۳ توصیف‌گرهای SIFT ۳۵

۳-۲-۴ شناسایی اشیاء با استفاده از SIFT ۳۶

۳-۳ روش‌های دیگر آشکارساز ویژگی‌های محلی ۴۱

۳-۳-۱ آشکارسازهای نقاط کلیدی بر مبنای هریس و هسین ۴۲

۳-۳-۲ آشکارساز FAST ۴۴

۳-۳-۳ آشکارساز MSER ۴۵

۳-۳-۴ مقایسه‌ی آشکارسازها ۴۶
۳-۴ انواع توصیف‌گرها ۴۷
۳-۴-۱ توصیف‌گرهای PCA-SIFT ۴۷
۳-۴-۲ توصیف‌گرهای GLOH ۴۷
۳-۴-۳ توصیف‌گرهای BRISK ۴۸
۳-۴-۴ توصیف‌گرهای FREAK ۵۰
فصل چهارم: توصیف‌گرهای جدید
۴-۱ توصیف‌گرهای DOOS ۵۲
۴-۲ انطباق توصیف‌گرهای DOOS ۵۴
۴-۳ نحوه‌ی ارزیابی توصیف‌گرهای DOOS ۵۷
۴-۴ ارزیابی توصیف‌گرهای DOOS ۶۰
۴-۴-۱ تعیین پارامتر اول توصیف‌گرهای DOOS ۶۲
۴-۴-۲ تعیین پارامترهای دوم و سوم توصیف‌گرهای DOOS ۶۶
۴-۵ یافتن آستانه گذاری مناسب بر روی نزدیک‌ترین همسایه‌ها ۷۱
فصل پنجم: ارائه‌ی الگوریتم پیدا کردن هدف و معرفی پایگاه داده
۵-۱ الگوریتم پیشنهادی برای پیدا کردن هدف در تصاویر هوایی ۷۵
۵-۱-۱ مرحله‌ی آموزش ۷۷
۵-۱-۲ مرحله‌ی جستجو ۷۷
۵-۲ معرفی پایگاه داده ۷۸
فصل ششم: بررسی نتایج و ارائه‌ی پیشنهادات
۶-۱ نتایج الگوریتم ارائه‌شده در فصل قبل ۸۱
۶-۲ مقایسه‌ی توصیف‌گرهای DOOS با دیگر روش‌ها ۸۵
۶-۳ نتیجه‌گیری ۸۷
۶-۴ پیشنهادات برای پژوهش‌های بعدی ۸۸
مراجع ۸۹

چکیده

پیدا کردن هدف در تصاویر هوایی، یکی از موضوعاتی است که با گسترش هواپیماهای بدون سرنشین کاربردهای زیادی پیداکرده است. در این پایان‌نامه الگوریتمی برای پیدا کردن هدف بر اساس توصیف نقاط کلیدی تصویر هدف و تصویر صحنه (تصویری که باید هدف در آن شناسایی شود) با توصیف‌گرهای جدید DOOS ارائه می‌شود. در این الگوریتم پس از آشکار سازی نقاط کلیدی دو تصویر و توصیف آن با توصیف‌گرهای DOOS، توصیف‌گرها باهم تطبیق داده می‌شوند و با آستانه گذاری بر روی آن‌ها بهترین تطبیق نقاط کلیدی بدست می‌آید. پس از آن با اعمال الگوریتم RANSAC بر روی این تطبیق‌ها و حذف موارد اشتباه، پارامترهای تبدیل مستوی بین تصویر هدف و تصویر صحنه تخمین زده می‌شود. در آخر با استفاده از این تبدیل مستوی، محل هدف در تصویر صحنه پیدا می‌شود. نتایج آزمایش‌ها نشان داد که توصیف‌گرهای DOOS بسیار قدرتمند هستند و قادرند با توصیف‌گرهای قدرتمند دیگر مثل SIFT رقابت کنند. به علاوه به علت ابعاد کم و ساختار این توصیف‌گرها، انطباق آن‌ها سریع‌تر از توصیف‌گرهایی مانند SIFT و SURF است و پیاده‌سازی مرحله‌ی انطباق توصیف‌گرهای DOOS بر روی سخت‌افزارهایی مثل DSPها و FPGAها ساده‌تر است.

کلمات کلیدی: ۱- پیدا کردن هدف ۲- توصیف‌گرهای DOOS ۳- روش SIFT ۴- تصاویر هوایی

فصل اول
مقدمه

امروزه استفاده از سیستم‌های تصویربرداری برای کاربردهای نظامی و نظارتی در پرنده‌های بدون سرنشین[1] گسترش پیداکرده است و به همین جهت مقالات زیادی به پردازش تصاویر هوایی پرداخته‌اند. پایدارسازی دنباله‌ی تصاویر[2]، موزاییک کردن تصاویر[3]، ثبت تصاویر[4]، به هم پیوستن تصاویر[5] و تولید نقشه‌های هوایی، پیدا کردن جاده‌ها یا رودخانه‌ها و غیره از موضوعاتی هستند که در این مقالات به آن‌ها پرداخته شده است [۵-۱]. یکی از موضوعات دیگر پیدا کردن یک مکان خاص در تصاویر هوایی است که ما در این پایان‌نامه به آن می‌پردازیم [۶].

۱-۱ بیان مسئله

فرض می‌کنیم که یک تصویر هوایی (تصویر زمین که با دوربینی از بالاگرفته شده) از یک مکان خاص که در این مقاله به عنوان هدف از آن یاد می‌شود، موجود باشد؛ این تصویر را تصویر الگو می‌نامیم. حال قرار است تا هدفِ موردِ نظر در تصاویر هوایی (تصویر صحنه) که در شرایطی متفاوت از شرایط تصویر الگو بدست آمده است،

[1] Unmanned Aerial Vehicle (UAV)
[2] Image\video Stabilization
[3] Image Mosaicking
[4] Image Registration
[5] Image Stitching

جستجو شود. این شرایط متفاوت باعث تغییر در روشنایی، اندازه، جهت و ... هدف در تصویر صحنه می‌شود. معمولاً تکرار تصویربرداری این تغییرات را به همراه دارد به این دلیل که در بیشتر مواقع ارتفاع دوربین و جهت آن با ارتفاع و جهت دوربین به هنگام تصویربرداری قبلی یکسان نیست و نور صحنه نیز معمولاً نسبت به زمان تصویربرداری قبلی متفاوت است. به علاوه ممکن است در اثر گذشت زمانِ زیاد بین دو تصویربرداری، دو تصویر به مقدار زیادی باهم متفاوت باشند؛ برای مثال تغییر فصل باعث عوض شدن پوشش گیاهی منطقه شود و یا ساخت‌وسازهایی در منطقه صورت گیرد.

مسئله در حالت کلی می‌تواند بسیار گسترده باشد؛ برای همین ما شرایط و فرض‌های منطقی در نظر می‌گیریم تا مسئله ساده‌تر شود. این شرایط و فرض‌ها در ادامه آورده شده است:

- هدف در حالت کلی ممکن است اشکال متفاوتی و یا حتی بدون شکل خاصی باشد (مثلاً نقطه‌ای در وسط یک خیابان باشد که باید از اطراف آن، محل هدف شناسایی شود).
- فرض می‌شود ابعاد هدف در حدود ۰.۰۵ ابعاد تصویر است و محلی که برای مرکز هدف بدست می‌آید نباید از محدوده‌ی هدف بیرون باشد.
- تصویر الگو شامل منطقه‌ای گسترده‌تر از هدف می‌شود به گونه‌ای که از مرکز هدف تا شعاع حداقل ۳ برابری ابعاد آن در تصویر مشاهده می‌شود.
- حداکثر تغییر مقیاس تصویر الگو و تصویر صحنه، ۲.۵ و محدوده‌ی چرخش تصویر صحنه نسبت به تصویر الگو ۰ تا ۳۶۰ درجه می‌باشد. تغییرات زاویه‌ی دید دو تصویر کمتر از ۱۰ درجه فرض می‌شود.
- به دلیل گذشت فاصله‌ی زمانی بین تصویربرداری تصویر الگو و تصویر صحنه، ممکن است تصویر منطقه‌ی هدف علاوه بر داشتن تغییراتی مثل روشنایی، از نظر وضعیت سایه‌ها نیز باهم متفاوت باشند و یا حتی تغییراتی در ظاهر هدف به وجود آمده باشد.

هدف از این پایان‌نامه مطالعه‌ی روش‌های مختلف برای این منظور و ارائه‌ی الگوریتم خودکار و مناسب برای یافتن هدف است. الگوریتم پیشنهادی باید متناسب با کاربرد باشد و نسبت به تغییرات رایج حساسیت نداشته باشد یا لااقل حساسیت کمی داشته باشد.

۱-۲ کارهای انجام‌شده

برای پیدا کردن هدف در تصاویر هوایی، می‌توان از روش‌های بازشناسی شی[۶] کمک گرفت، در واقع می‌توان هدف را شی‌ای تصور کرد که در تصویر صحنه قرار دارد و باید آن را پیدا کرد. روش‌های مختلفی برای بازشناسی اشیاء معرفی شده، اما در استفاده از این روش‌های برای پیدا کردن هدف در تصاویر هوایی باید به نکاتی توجه داشت؛ شکل اهداف ممکن است بسیار متنوع باشند و از روش‌هایی که برای استخراج یک شکل خاص طراحی می‌شوند، به تنهایی نمی‌توان استفاده کرد. معمولاً پس‌زمینه‌ی تصویر صحنه (کل تصویر به‌جز قسمت هدف)، شلوغ است، یعنی به جز هدف اشیاء دیگری در تصویر صحنه وجود دارد و همین سبب می‌شود روش‌هایی که براساس جداسازی هدف از پس‌زمینه کار می‌کنند، ناکارآمد شوند. اگرچه وجود اشیاء دیگر در تصویر صحنه، پیدا کردن هدف را دشوارتر می‌کند، ولی از آن‌جاکه موقعیت آن‌ها نسبت به هدف ثابت است، می‌توان از وجود آن‌ها برای یافتن هدف بهره برد؛ برای مثال ممکن است هدف را به دلایلی مثل تغییرات زیاد آن یا پوشیده شدن آن نتوان پیدا کرد، در اینجا یافتن اشیاء اطراف آن می‌تواند ما را در پیدا کردن محل هدف راهنمایی کند.

یکی از رویکردهای اخیر شناسایی اشیاء، استفاده از ویژگی‌های محلی شی مانند روش SIFT[۷] است [۷،۸]. این رویکرد را می‌توان به سه قسمت تقسیم کرد: آشکارسازی ویژگی‌های محلی، توصیف ویژگی‌های محلی و تطبیق این ویژگی‌ها که در فصل‌های بعدی به شکل جامع‌تر هر قسمت را توضیح می‌دهیم. قابلیت تطبیق‌پذیری این رویکرد با شرایط گوناگون و توانایی‌های آن سبب شده که افراد زیادی روی آن تحقیق کنند. عده‌ای سعی در ارائه‌ی روشی برای بهبود جنبه‌های مختلف این رویکرد داشتند، مانند روش SURF[۸] که سریع‌تر از SIFT کار می‌کرد یا روش‌های ASIFT و ASURF[۹] که برای شناسایی شی حتی در حالت‌هایی که شی تغییرات مستوی زیاد داشته باشد، ارائه شد [۹-۱۱]. برخی دیگر سعی در به‌کارگیری این روش‌ها برای کاربردهای خود ازقبیل پایدارسازی، به هم پیوستن تصاویر، موزاییک کردن و ایجاد تصاویر با نمای گسترده‌تر و ردیابی اهداف کردند [۱۲-۱۴].

مطالعات اولیه‌ی ما نشان داد که می‌توان از این رویکرد برای کاربرد خود بهره برد و در نهایت نیز ما همین رویه را در پیش گرفتیم.

---

[6] Object Recognition
[7] Scale Invariant Feature Transform
[8] Speeded Up Robust Feature
[9] Affine SIFT and Affine SURF

۱-۳ نوآوری پایان‌نامه

در این پایان‌نامه الگوریتمی برای یافتن هدف در تصاویر هوایی بر اساس توصیف‌گرهای جدید ارائه شده است. ما این توصیف‌گرها که حاصل کنار هم گذاشتن عدد مربوط به "جهت‌گیری غالب زیر ناحیه‌ها"[10] است، به اختصار DOOS می‌نامیم. این توصیف‌گرهای جدید پارامتری برای تنظیم دارند که با تغییر آن، می‌توان توصیف‌گرها را از حالت ساده به پیچیده که قابلیت رقابت با توصیف‌گرهای قدرتمندی مثل SIFT را دارد تبدیل کرد. برای مثال توصیف‌گرهای DOOS با ابعاد ۱۶ مناسب برای کاربردهایی که دو تصویر فصل مشترک قابل‌توجهی دارند، است مانند کاربردهایی ازقبیل پایدارسازی تصاویر و به هم پیوستن دنباله‌ای از تصاویر. از ابعاد بالاتر این توصیف‌گرها مثلاً ۶۴ برای کاربردهایی نظیر جستجو در شرایط سخت (مثل کاربرد فعلی) که نیاز به تمایز بالاتر توصیف‌گرها است، می‌توان استفاده کرد.

خصوصیت این توصیف‌گرها همان طور که بعداً نیز به آن اشاره می‌کنیم، این است که مقادیر این توصیف‌گرها شامل مجموعه‌ی محدودی از اعداد طبیعی (صحیح مثبت) می‌شود و مناسب برای پیاده‌سازی بر روی پردازشگرهای DSP و FPGA (به خصوص قسمت انطباق‌دهنده‌ی توصیف‌گرها) می‌باشد. انطباق این توصیف‌گرها حتی در پیاده‌سازی معمولی نیز نسبت به روش SIFT سریع‌تر صورت می‌گیرد به دلیل این‌که ابعاد توصیف‌گرهای آن کوچک‌تر از ابعاد توصیف‌گرهای SIFT است (انطباق توصیف‌گرهای DOOS با ابعاد ۱۶ و ۶۴ به ترتیب ۸ و ۲ برابر سریع‌تر از انطباق توصیف‌گرهای SIFT با ابعاد ۱۲۸ انجام می‌شود).

علاوه بر این ما در مرحله‌ی جدا کردن تطبیق‌های درست و نادرست، الگوریتم RANSAC[11] را به گونه‌ای تغییر دادیم که این کار سریع‌تر انجام شود.

۱-۴ روند نگارش پایان‌نامه

روند نگارش این پایان‌نامه به این صورت است که در فصل دوم پس از بررسی چالش‌های پیش‌روی شناسایی شی، برخی از روش‌های پیشین شناسایی شی مانند همبستگی و تعمیم تبدیل هاف، شرح داده شده است. ما در فصل سوم به شناسایی شی با استفاده از ویژگی‌های محلی پرداخته‌ایم و پس از بیان روش SIFT، به انواع آشکارسازهای نقاط کلیدی و توصیف‌گرهایی که می‌تواند در این رویکرد مورد استفاده قرار گیرد، اشاره‌ای

[10] Dominate Orientation of Subregions (DOOS)
[11] Random Sample Consensus

می‌کنیم. نوآوری این پایان‌نامه در فصل چهارم آورده شده است؛ در این فصل توصیف‌گرهای DOOS معرفی می‌شوند و مقدار مناسب برای پارامترهای آن بدست می‌آید؛ همچنین این توصیف‌گرها با توصیف‌گرهای SIFT مقایسه می‌شوند. در فصل پنجم الگوریتمی برای پیدا کردن هدف در تصاویر هوایی با استفاده از توصیف‌گرهای DOOS ارائه می‌شود. در همین فصل پایگاه داده‌ای که برای آزمایش الگوریتم استفاده‌شده معرفی می‌گردد. نتایج آزمایش الگوریتم در فصل ششم قرار دارد. در فصل ششم نتایج الگوریتم بررسی‌شده و با حالتی که در این الگوریتم از توصیف‌گرهای SIFT و BRISK استفاده شود، مقایسه می‌شود؛ در انتها پس از نتیجه‌گیری به کارهایی که در آینده می‌توان انجام داد، اشاره می‌شود.

فصل دوم
شناسایی شی

شناسایی اشیاء یکی از بحث‌های متنوع و گسترده‌ی پردازش تصویر است. تنوع بحث را می‌توان از طیف گسترده‌ی کاربردهای آن دریافت. از جمله کاربردهای آن می‌توان به موارد اندازه‌گیری مختصات، بازرسی، مرتب کردن، شمارش، آشکارسازی شی، دسته‌بندی صحنه‌ها و بازیابی تصویر اشاره کرد. علاوه بر آن وجود محدودیت-ها و الزامات خاص هر کاربرد به تنوع الگوریتم‌های آن افزوده است. محدودیت زمانی، دقت، قابلیت اطمینان شناسایی و تغییراتی که شی در تصویر ممکن است داشته باشد، از مواردی هستند که در طراحی الگوریتم شناسایی اشیاء باید به آن توجه نمود. در ادامه‌ی این فصل ابتدا در مورد مسائل و چالش‌های شناسایی شی بحث می‌کنیم و پس از آن، چند مورد از روش‌های پایه‌ای برای شناسایی شی بیان و قابلیت‌های آن‌ها بررسی می‌شود.

۲-۱ چالش‌های شناسایی شی

در این قسمت ما چالش‌های روبروی شناسایی شی را بررسی می‌کنیم. تغییرات خطی و غیرخطی روشنایی[1]، تغییرات مقیاس[2]، چرخش[3]، شلوغی صحنه[1]، پوشیده شدن[2] قسمتی از شی و تغییر در نقطه‌ی دید[3] از جمله‌ی این

[1] Illumination
[2] Scale
[3] Rotation

موارد هستند. شناخت این مسائل و اثرات آن بر تصویر می‌تواند ما را در یافتن الگوریتم مناسب برای کاربرد خود راهنمایی کند.

۲-۱-۱ تغییرات خطی و غیرخطی روشنایی

هرچقدر هم که شرایط تصویربرداری از صحنه کنترل شود، تفاوت‌هایی در روشنایی شی در تصویر الگو و تصویر صحنه وجود خواهد داشت. این تفاوت‌ها در حالتی که شرایط تصویربرداری کنترل نشود یا اساساً غیرقابل‌کنترل باشد، بیشتر می‌شود. اگرچه این تغییرات در روشنایی- چه به صورت خطی و چه غیرخطی- کار پیدا کردن هدف یا شی را سخت‌تر می‌کنند، ولی تغییرات غیرخطی اثر بیشتری دارند.

فرض کنید شدت روشنایی در تصویری با تابع $I(x, y)$ بیان شده باشد. تغییرات روشنایی باعث می‌شود تابع روشنایی آن به شکل $I'(x, y) = a(x, y).I(x, y) + b(x, y)$ درآید. در تغییر روشنایی غیرخطی a و b در حالت کلی تابعی از مکان هستند ولی در تغییرات خطی آن‌ها ثابت‌اند. بنابراین در تغییرات خطی $I'(x, y) = a.I(x, y) + b$ می‌شود. اثرات خطی را به سادگی با کم کردن میانگین شدت روشنایی‌ها و سپس نرمالیزه‌کردن آن‌ها می‌توان از بین برد یا می‌توان از روش‌هایی که نسبت به این پارامترها حساس نیستند مثل نرمال شده‌ی گرادیان‌ها (گرادیان‌ها چون با تفاضل کار می‌کنند ضریب b را حذف کرده و با نرمالیزه‌کردن آن‌ها ضریب a نیز حذف می‌شود) استفاده کرد. در تغییرات غیرخطی اگر تغییرات نرم باشد، یا اینکه مقادیر a و b در یک ناحیه مقدارهای ثابتی باشند، استفاده از ویژگی‌های محلی تا حدودی می‌تواند این مسئله را حل کند.

۲-۱-۲ تغییر مقیاس

استفاده از دوربین‌های متفاوت و یا اختلاف در فاصله شی از دوربین، بین تصویر شی و تصویر صحنه موجب می‌شود که شی در دو تصویر با اندازه (تعداد پیکسل) متفاوتی دیده شود. بسیاری از الگوریتم‌ها به تغییر مقیاس شی حساس‌اند و این تغییر موجب افت عملکرد و در مواردی عدم کارایی آن‌ها می‌شود. در برخی موارد این اختلاف اندازه مشخص است و به یکی از تصاویر الگو یا صحنه (اغلب تصویر الگو) اعمال می‌شود و پس از آن جست‌وجو انجام می‌شود. برای جست‌وجو در حالتی که اختلاف اندازه‌ی شی مشخص نباشد، یکی از تصاویر را در

---

[1] Background clutter
[2] Occlusion
[3] Viewpoint change

محدوده‌ی قابل انتظار تغییر مقیاس، بسته به نوع الگوریتم، تغییر اندازه می‌دهند و چندین تصویر تولید می‌شود و برای تمامی آن‌ها عمل شناسایی انجام می‌شود، بنابراین حجم پردازش‌ها به شدت افزایش می‌یابد.

۲-۱-۳ چرخش

در اکثر مواقع شی در تصویر صحنه نسبت به تصویر الگوی آن، چرخیده است. چرخش شی نیز مانند تغییرات اندازه‌ی آن یا باید به یکی از تصاویر اعمال شود (در صورت معلوم بودن) و یا این‌که حجم پردازش را به شدت افزایش می‌دهد.

۲-۱-۴ شلوغی صحنه

در مواردی تصویر صحنه تنها شامل شی مورد نظر نیست و اشیاء دیگری نیز حضور دارند و به اصطلاح پس‌زمینه که کل تصویر غیر از شی را شامل می‌شود، شلوغ است. این شلوغی باعث می‌شود روش‌هایی که با جداسازی شی از پس‌زمینه (در مرحله نخست) به شناسایی آن می‌پردازند، دچار مشکل شوند و در کل احتمال خطا را افزایش می‌دهد.

۲-۱-۵ پوشیده شدن قسمتی از شی

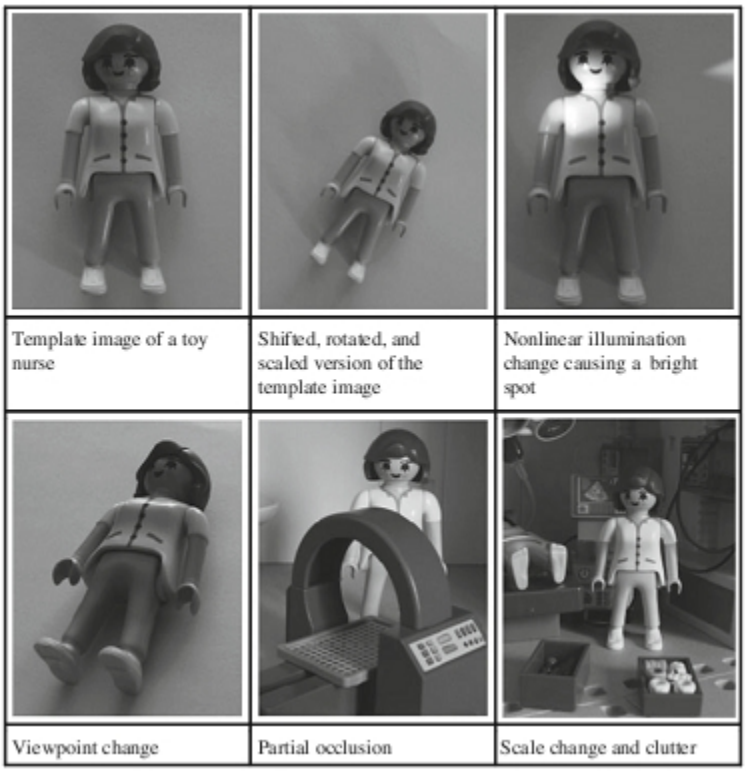
گاهی همه‌ی شی در تصویر صحنه مشخص نبوده و قسمت‌هایی از آن پوشیده شده است. این مسئله الگوریتم‌هایی که به شی به صورت سرتاسری نگاه می‌کنند (مانند همبستگی و PCA[1])، به شدت با مشکل دچار می‌کند تا جایی که بعضی از آن‌ها در شرایطی که این مسدود شدگی از یک اندازه زیادتر شود، عملاً ناکارآمد می‌شوند.

۲-۱-۶ تغییر نقطه‌ی دید

تغییر نقطه‌ی دید باعث تغییر در اندازه، زاویه و در حالت کلی تغییر انعکاسی[2] (در حالت ساده‌تر تغییر مستوی[3]) تصویر جسم می‌شود. با این شرح مشخص است که این تغییر نیز به دشواری کار می‌افزاید.

---

[1] Principal Component Analysis
[2] Perspective
[3] Affine



شکل ۲-۱- نمایش مشکلات و چالش‌های پیش روی شناسایی اشیاء [۱۵].

این دشواری‌ها زمانی شدت می‌یابد که تغییر نقطه‌ی دید زیاد باشد و شی مورد نظر، ارتفاع قابل‌ملاحظه‌ای نسبت به ارتفاع دوربین داشته باشد. مثلاً برای پیدا کردن یک ساختمان در تصویر شهری که از بالا تهیه‌شده، اگر ارتفاع دوربین زیاد نباشد، اختلاف نقطه‌ی دید ممکن است باعث دیده شدن دو سمت مختلف ساختمان شود، در این حالت انطباق دو تصویر حتی برای انسان هم دشوار است. در اکثر اوقات نقطه‌ی دید متفاوت است، هرچند اندک (اختلاف در yaw و pitch)،

بنابراین یکی از روش‌ها این است که چند تصویر الگو از نقطه دیدهای متفاوت تهیه و تک‌تک آن‌ها را جست‌وجو کرد؛ در این موارد اگر نقطه دید صحنه حدوداً مشخص باشد، کار ساده‌تر می‌شود. در شرایطی که نقطه‌ی دید عوض می‌شود، استفاده از روش‌هایی که بر اساس توصیف ویژگی‌های محلی کار می‌کنند، مناسب‌تر است؛ چراکه بعضی از ویژگی‌ها به دلیل خاصیت محلی بودن، از نقطه دیدهای مختلف یکسان هستند.

شکل۲-۱ مشکلات و چالش‌های مختلف پیش روی شناسایی اشیاء را نشان می‌دهد.

۲-۲ روش‌های جستجوی تصاویر الگو

در این قسمت ما به چند روش پیدا کردن تصویر الگو اشاره می‌کنیم. لازم به ذکر است که روش‌های جدید معمولاً بر اساس ترکیب یک و یا چند مورد از این روش‌هاست. فرض بر این است که در ابتدا یک یا چند تصویر از شی مورد نظر وجود دارد (تصویر الگو) و از این تصویر برای شناسایی شی مورد نظر در تصویر صحنه (یا به صورت خام یا این‌که به عنوان داده‌ی آموزش برای یک سیستم یادگیری) استفاده می‌شود.

۲-۲-۱ روش همبستگی

معمولاً در این روش ابتدا تصاویر رنگی به خاکستری تبدیل می‌شود و پس از آن همبستگی تصویر الگو با تصویر صحنه محاسبه می‌شود. برای از بین بردن اثرات تغییرات خطی روشنایی از روش همبستگی متقابل نرمالیزه[1] با رابطه‌ی زیر استفاده می‌شود.

$$\rho(a,b)=\frac{\sum_{x=0}^{W}\sum_{y=0}^{H}(I_S(x+a,y+b)-\bar{I}_S).(I_T(x,y)-\bar{I}_T)}{\sqrt{\sum_{x=0}^{W}\sum_{y=0}^{H}(I_S(x+a,y+b)-\bar{I}_S)^2.\sum_{x=0}^{W}\sum_{y=0}^{H}(I_T(x,y)-\bar{I}_T)^2}} \qquad ۱-۲$$

در این رابطه $\rho(a,b)$ ضریب همبستگی نرمالیزه در جابجایی (a,b) بین تصویر صحنه و الگو است. $I_S(x,y)$ و $I_T(x,y)$ به ترتیب روشنایی تصاویر صحنه[2] و الگو[3] در مختصات $[x,y]$ است. $\bar{I}_S$ و $\bar{I}_T$ میانگین شدت روشنایی و W و H عرض و ارتفاع تصویر الگو هستند. عبارت مخرج باعث نرمالیزه شدن مقدار ° می‌شود. بنابراین مقدار ° عددی بین ۱- تا ۱ است. مقادیر نزدیک به ۱ بیانگر شباهت بالای تصویر الگو با تصویر صحنه و مقادیر نزدیک صفر از نا همبسته بودن محتویات تصاویر حکایت می‌کند. سرانجام هر چه مقدار ρ منفی‌تر باشد یعنی محتویات دو تصویر معکوس هم هستند.

[1] Normalized Cross Correlation(NCC)
[2] Scene
[3] Template