دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

تشخیص اشیاء متحرک در تصاویر هوائی دوربین متحرک

پایان‌نامه کارشناسی ارشد مهندسی برق- مخابرات (سیستم)

رضا لالچینی

استاد راهنما

دکتر محمّدرضا احمدزاده

۱۳۹۲

بسم الله الرحمن الرحيم

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

تشخیص اشیاء متحرک در تصاویر هوائی دوربین متحرک

پایان‌نامه کارشناسی ارشد مهندسی برق- مخابرات (سیستم)

رضا لالچینی

استاد راهنما

دکتر محمدرضا احمدزاده

۱۳۹۲

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

پایان‌نامه‌ی کارشناسی ارشد رشته‌ی مهندسی برق – مخابرات (سیستم) آقای رضا لالچینی

تحت عنوان

تشخیص اشیاء متحرک در تصاویر هوائی دوربین متحرک

در تاریخ ۹۲/۱۱/۰۹ توسط کمیته‌ی تخصصی زیر مورد بررسی و تصویب نهایی قرار گرفت.

۱- استاد راهنمای پایان‌نامه	دکتر محمدرضا احمدزاده
۲- استاد داور	دکتر رسول امیرفتاحی
۳- سرپرست تحصیلات تکمیلی دانشکده	دکتر محمدعلی خسروی فرد

بر خود لازم می‌دانم که از راهنمایی‌های موثر دکتر محمدرضا احمدزاده و دوست عزیزم آقای مهندس سعید کاظمی سپاسگزاری نمایم.

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

که همواره حامی و همراهم هستند.

فهرست مطالب

عنوان **صفحه**

فهرست مطالب ..... هشت

چکیده ۱

فصل اول:مقدمه

۱-۱ بیان مسئله..... ۳

۱-۱-۱ انواع حرکت‌های محتمل در دنباله تصاویر..... ۴

۱-۱-۲ فرآیند تشخیص شی متحرک در تصاویر دوربین متحرک..... ۵

۱-۲ نوآوری پایان نامه..... ۸

۱-۳ ساختار پایان نامه..... ۸

فصل دوم: تشخیص اشیاء متحرک در تصاویر دوربین متحرک

۲-۱ جبران سازی خود حرکتی و تشخیص شی متحرک..... ۱۰

۲-۱-۱ تفاضل زمانی..... ۱۰

۲-۱-۲ تفریق پس زمینه..... ۱۱

۲-۱-۳ شار نوری..... ۱۲

۲-۲ تقسیم بندی روش‌ها بر اساس چگونگی تخمین خود حرکتی و تشخیص شی متحرک..... ۱۲

۲-۲-۱ رویکرد موزائیک پس زمینه..... ۱۲

۲-۲-۲ رویکرد شار نوری..... ۱۳

۲-۲-۳ رویکرد تطبیق ویژگی..... ۱۴

۲-۲-۴ رویکرد مدل هندسی دوربین..... ۱۴

۲-۳ مدل دوربین..... ۱۵

۲-۴ مدل ریاضی خود حرکتی..... ۱۸

۲-۴-۲ میدان حرکتی ناشی از حرکت انتقالی دوربین..... ۱۸

۲-۴-۲ میدان حرکتی ناشی از حرکت دورانی دوربین..... ۲۰

۱.دوران حول محور Y..... ۲۰

۲.دوران حول محور X..... ۲۱

۳.دوران حول محور Z ۲۲
۲-۴-۳ میدان حرکتی ناشی از حرکت دورانی و انتقالی دوربین ۲۳
۲-۴-۴ چند نکته ۲۴
۲-۵ تقسیم بندی روش‌های تشخیص شی متحرک در تصاویر دوربین متحرک ۲۴
۲-۶ مواردی که می‌توان در آن‌ها خود حرکتی را با یک تبدیل دو بعدی تقریب زد ۲۵
۲-۶-۱ صحنه تخت ۲۵
۲-۶-۲ صحنه از راه دور ۲۵
۲-۶-۳ حرکت دورانی خالص ۲۶
۲-۶-۴ بزرگنمایی دوربین ۲۶
۲-۷ حرکت انتقالی دوربین در صحنه‌های سه بعدی ۲۶
۲-۷-۱ تشخیص شی متحرک بر اساس تشخیص اپی پل ۲۷
۲-۷-۲ تشخیص شی متحرک بر اساس تخمین خود حرکتی در صحنه‌های سه بعدی ۲۷
۲-۷-۳ مدل استریو LONGUET-HIGGINS برای حرکت ایجاد شده بر اثر حرکت دوربین ۲۸
۲-۷-۴ شار نوری ۲۹
۲-۷-۵ شار نوری همراه با هرم ۳۱
۲-۷-۶ تخمین خود حرکتی بر اساس شار نوری ۳۳
۲-۸ تقریب خود حرکتی به وسیله تبدیل دو بعدی پارامتری ۳۴
۲-۹ تبدیل مستوی ۳۵
۲-۱۰ تبدیل انعکاسی ۳۶
۲-۱۱ خلاصه و نتیجه گیری ۳۷

فصل سوم: رویکرد تطبیق ویژگی

۳-۱ تخمین خود حرکتی با رویکرد تطبیق ویژگی ۳۹
۳-۲ نوع ویژگی مورد استفاده و تطبیق ویژگی‌ها ۳۹
۳-۳ تخمین پارامترهای تبدیلات دو بعدی ۴۱
۳-۴ رگرسیون کامل و رگرسیون استوار ۴۲

۳-۴-۱ انواع نقاط پرت و نقطه شکست ۴۲
۳-۴-۲ روش حداقل مربعات ۴۴
۳-۴-۳ تخمین گر حداقل میانه مربع (LMEDS) ۴۶
۳-۴-۴ تخمین گر تبدیل هاف ۴۸
۳-۴-۵ تخمین گر RANSAC ۴۹
۳-۵ مروری بر کارهای انجام شده در تشخیص شی متحرک ۵۰
۳-۶ نتیجه گیری و خلاصه ۵۲

فصل چهارم: پایه‌های روش پیشنهادی

۴-۱ تخمین تبدیل بین دو فریم متوالی ۵۵
۴-۲ استخراج ویژگی‌های متناظر یک بعدی ۵۸
۴-۳ تخمین پارامترهای تبدیل ۶۰
۴-۴ خلاصه و نتیجه گیری ۶۲

فصل پنجم: روش پیشنهادی

۵-۱ دوران حول محور Y ۶۴
۵-۲ دوران حول محور X ۶۵
۵-۳ چگونگی خود حرکتی موجود در تصاویر هوایی و خصوصیات آن ۶۷
۵-۳-۱ تقریب خود حرکتی به دست آمده با افکنش ارتوگرافیک در صحنه تخت ۶۷
۵-۴ تبدیل فرض شده برای تقریب خود حرکتی ۶۸
۵-۵ تخمین گر مورد استفاده ۶۸
۵-۶ پیاده سازی و نتایج ۶۹
۵-۷ خلاصه و نتیجه گیری ۷۵

فصل ششم: نتیجه گیری و پیشنهادها

۶-۱ نتیجه گیری ۸۲
۶-۲ پیشنهادها ۸۴
مراجع ۸۵

فهرست شکل‌ها

عنوان **صفحه**

شکل۱-۱ فرآیند تشخیص شی متحرک در تصاویر دوربین متحرک..... ۵

شکل۲-۱ نمونه‌ای از یک موزائیک تصویر ۱۳

شکل۲-۲ تصویر شبح و تصویری که اثر شبح در آن حذف شده است..... ۱۵

شکل۲-۳ هندسه افکنشی..... ۱۶

شکل۲-۴ ایجاد دستگاه پیکسل با گسسته سازی و انتقال مبدأ صفحه تصویر..... ۱۷

شکل۲-۵ شناسایی جسم متحرک بر اساس تشخیص اپی پل..... ۲۷

شکل۲-۶ مدل استریو تصویر برداری..... ۲۹

شکل۲-۷ ساخت هرم گوسی از تصاویر به منظور کوچک کردن میزان حرکت یک پیکسل در دو تصویر متوالی.. ۳۱

شکل۲-۸ ساخت هرم گوسی از دو تصویر متوالی به منظور محاسبه دقیق شار نوری..... ۳۲

شکل۲-۹ دسته متناظر برای محاسبه شار نوری در دو تصویر متوالی..... ۳۲

شکل۲-۱۰ انتخاب دسته متناظر برای محاسبه شار نوری با استفاده از شار نوری به دست آمده از سطح بالاتر هرم.... ۳۳

شکل۲-۱۱ نمونه‌ای از تبدیل مستوی..... ۳۶

شکل۲-۱۲ نمونه‌ای از تبدیل انعکاسی..... ۳۷

شکل۳-۱ فرآیند تخمین حرکت کلی..... ۳۹

شکل۳-۲ انواع نقطه پرت..... ۴۳

شکل۳-۳ تخمین پارامترهای تبدیل انعکاسی به روش حداقل مربع خطا بدون وجود نقطه پرت..... ۴۶

شکل۳-۴ تخمین پارامترهای تبدیل انعکاسی به روش حداقل مربع خطا با وجود نقطه پرت..... ۴۶

شکل۴-۲ تقسیم بندی تصویر به بخش‌های عمودی و افقی که قرار است ویژگی‌ها از این بخش‌ها استخراج شوند. (A) و (B) تقسیم بندی به بخش‌های افقی. (C) و (D) تقسیم بندی به بخش‌های عمودی..... ۵۸

شکل۴-۳ فرآیند استخراج ویژگی‌های متناظر: (A) بخش عمودی سوم از جفت تصویر شکل ۲-۵ (B) پروفایل‌هایی افکنش سطح روشنایی جفت تصویر قسمت (A) که بیشینه‌ها و کمینه‌های آن استخراج شده‌اند . (C) ویژگی‌های متناظر همه بخش‌های عمودی جفت تصویر شکل ۲-۵..... ۶۰

شکل۴-۴(A) نتیجه حذف نقاط پرت (B) خط تخمین زده شده با استفاده از ویژگی‌های متناظر قسمت (A)..... ۶۱

شکل۴-۵ نتایج حذف خود حرکتی: (A) و (B) جفت تصویر اصلی. (C) تبدیل یافته تصویر (A). (D) قدر مطلق تفاضل تصویر (A) و (B). (E) قدر مطلق تفاضل تصاویر (B) و (C). (F) تصویر باینری، تصویر (E)..... ۶۱
شکل۵-۱ خود حرکتی ایجاد شده در اثر حرکت PAN..... ۶۵
شکل۵-۲ (الف) تصویر اول. (ب) تصویر دوم. (ج) و (د) پروفایل افکنش افقی تصویر اول و تصویر دوم. (ه) و (ز) پروفایل افکنش عمودی تصویر اول و تصویر دوم..... ۷۱
شکل۵-۳ (الف) و (ب) بیشینه‌ها و کمینه‌های محلی افکنش عمودی تصویر اول و دوم. (د) و (ز) بیشینه‌ها و کمینه‌های افکنش افقی تصویر اول و دوم. (ج) ویژگی‌های منطبق در افکنش های عمودی. (ه) ویژگی‌های منطبق در افکنش های افقی..... ۷۲
شکل۵-۴ پارامترهای تخمین زده شده از روی ویژگی‌های عمودی و افقی به لحاظ هندسی یک خط را مشخص می‌کنند. (الف) پارامترهای $\boldsymbol{a11}$ و $\boldsymbol{b1}$ از روی ویژگی‌های متناظر عمودی تخمین زده می‌شوند و به لحاظ هندسی یک خط را تشکیل می‌دهند. (ب) پارامترهای $\boldsymbol{a22}$ و $\boldsymbol{b2}$ از روی ویژگی‌های متناظر افقی تخمین زده می‌شوند و به لحاظ هندسی یک خط را نشان می‌دهند..... ۷۳
شکل۵-۵ (الف) تفاضل زمانی دو فریم متوالی بعد از جبران سازی خود حرکتی. (ج) تصویر باینری ساخته شده از تصویر قسمت (الف) با آستانه ۷۰. (ب) نتیجه اعمال عملگرهای مورفولوژی کی بر روی تصویر قسمت (ج). (د) تفاضل دو فریم متوالی بدون جبران سازی خود حرکتی. (ه) معکوس تصویر قسمت (ب)..... ۷۴
شکل۵-۶ (الف) فریم ۱ پایگاه داده. (ب) فریم ۲ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۷۵
شکل۵-۷ (الف) فریم ۳۰ پایگاه داده. (ب) فریم ۳۱ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۷۶
شکل۵-۸ (الف) فریم۲۷۴ پایگاه داده. (ب) فریم ۲۷۵ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۷۷
شکل۵-۹ (الف) فریم ۳۷۰ پایگاه داده. (ب) فریم ۳۷۱ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۷۸
شکل۵-۱۰ (الف) فریم ۴۱۵ پایگاه داده. (ب) فریم ۴۱۶ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۷۹
شکل۵-۱۱ (الف) فریم ۴۹۰ پایگاه داده. (ب) فریم ۴۹۱ پایگاه داده. (ج) اشیاء متحرک شناسایی شده..... ۸۰

فهرست جدول

عنوان **صفحه**

جدول۳-۱ مقایسه ویژگی‌های نقطه‌ای ٤١

چکیده

آشکارسازی هدف متحرک یکی از مباحث مطرح در چند ساله اخیر در مسائل بینایی کامپیوتر بوده است. دقت و درجه اطمینان الگوریتم‌های آشکارسازی اهداف متحرک تأثیر مستقیم در مسائلی از جمله ردیابی و موقعیت یابی اهداف متحرک دارد. یکی از مسائلی که مشکلات مختلفی را در آشکارسازی و ردیابی جسم متحرک ایجاد می‌کند، حرکت دوربین می‌باشد. در اصطلاح علمی به حرکت دوربین خود حرکتی[1] می‌گویند و به معنای ایجاد تغییرات دو بعدی در دنباله تصاویر بر اثر حرکت سه بعدی دوربین می‌باشد. برای تشخیص اشیاء متحرک در این تصاویر باید بتوانیم خود حرکتی و حرکتی که در اثر حرکت شی به وجود آمده است را از یکدیگر متمایز کنیم. در اکثر روش‌های موجود ابتدا خود حرکتی موجود در دنباله تصاویر را تخمین می‌زنند، سپس خود حرکتی را از دنباله تصاویر حذف می‌کنند. آنگاه با استفاده از روش‌هایی مانند تفاضل زمانی و یا تفریق پس زمینه که برای تشخیص شی متحرک در تصاویر دوربین ثابت توسعه داده شده‌اند، اشیاء متحرک شناسایی می‌شوند. در این پژوهش ابتدا مدل ریاضی خود حرکتی را معرفی می‌کنیم سپس با استفاده از این اطلاعات دسته بندی روش‌ها را عنوان می‌نماییم. از بین روش‌های معرفی شده رویکرد تطبیق ویژگی[2] به دلیل پایین بودن هزینه زمانی و محاسباتی و همچنین عدم وابستگی‌اش به سخت افزار نسبت به روش‌های دیگر مورد توجه قرار گرفته است. در رویکرد تطبیق ویژگی، از ویژگی‌های متفاوتی جهت تخمین خود حرکتی استفاده می‌شود. در بین این ویژگی‌ها، ویژگی‌های یک بعدی به جهت پایین بودن هزینه زمانی و محاسباتی نسبت به سایر ویژگی‌ها، مورد توجه قرار گرفته‌اند. در این پژوهش معیاری را معرفی می‌کنیم که نیاز و یا عدم نیاز به بخش بندی تصویر به نوارهای عمودی و افقی به جهت استخراج ویژگی‌های یک بعدی را مشخص می‌کند و با استفاده از این معیار نشان می‌دهیم نیازی نیست که تصاویر هوایی را به نوارهای افقی و عمودی برای استخراج ویژگی یک بعدی بخش بندی نماییم. از این‌رو روشی موثر برای تشخیص اشیاء متحرک در تصاویر هوایی توسعه داده‌ایم که نسبت به سایر روش‌های موجود، هزینه زمانی و محاسباتی بسیار پایین تری دارد و می‌تواند به تشخیص شی متحرک در صحنه‌هایی بپردازد که پس زمینه یکنواختی دارند و همچنین اشیاء متحرک دارای حرکت اندک هستند. با پیاده سازی روش پیشنهادی توسط نرم افزار *Matlab* در سیستمی با پردازشگر *CORE i3* با فرکانس ۲.۵۳ گیگا هرتز و حافظه اصلی چهار گیگابایت با سیستم عامل ویندوز ۷ و اعمال آن بر تصاویری با دقت (رزولوشن) ۶۴۰×۴۸۰ در مدت زمان ۰/۶ ثانیه اشیاء متحرک شناسایی می‌شوند.

کلمات کلیدی: ۱- تشخیص شی متحرک، ۲- خود حرکتی، ۳- جبران سازی خود حرکتی، ۴- دوربین متحرک.

---

[1] Ego-motion
[2] Feature matching approach

فصل اول

مقدمه

در چند دهه اخیر، به دلیل نیاز جوامع به حفاظت بیشتر، سیستم‌های پایش محیط به صورت پیوسته رشد کرده‌اند. این سیستم‌ها در کاربردهای نظامی و شهری مورد استفاده قرار گرفته‌اند. در بررسی‌هائی که چند گروه حافظ حقوق بشر در سال ۲۰۰۵ انجام داده‌اند، عنوان کرده‌اند که ۴ میلیون دوربین پایش محیط در انگلستان وجود دارد که به نسبت جمعیت گفته شده است برای هر ۱۷ نفر یک دوربین وجود دارد. آن‌ها عنوان کرده‌اند که تنها بخش کوچکی از این ویدئوها مورد ارزیابی قرار می‌گیرند زیرا در اکثر موارد، ارزیابی توسط انسان صورت می‌گیرد. علاوه بر این بسیاری از این ویدئوها در هیچ صفحه نمایشی نشان داده نمی‌شوند و تنها ضبط می‌شوند و در صورت ضرورت مورد بررسی قرار می‌گیرند. ارزیابی توسط انسان مشکلات خاص خود را دارد، علاوه بر آن ممکن است بعضی از وقایع مهم در تصاویر که می‌بایست به صورت بلادرنگ تشخیص داده شوند، از دست بروند. این مشکل و سایر مشکلات ارزیابی توسط انسان اهمیت خود کار سازی پایش محیط توسط دوربین را مورد توجه قرار می‌دهد [۱].

هدف اصلی از خودکار سازی سیستم‌های پایش محیط استخراج اطلاعات سطح بالا از تصاویر ویدئویی بدون دخالت انسان می‌باشد. سیستم‌های خودکار پایش محیط از چهار هسته تشکیل شده‌اند: تشخیص اشیاء[1]، دسته بندی اشیاء[2]، ردگیری اشیاء[3] و تشخیص هویت (شناسایی) و بررسی رفتار[4] اشیاء. در ابتدا اشیاء مورد نظر در تصاویر ویدئویی به وسیله خصوصیات ظاهری و یا حرکتیشان شناسایی می‌شوند. در مرحله دوم اشیاء شناسایی شده در چندین دسته متفاوت مانند

[1] Object detection
[2] Object classification
[3] Object tracking
[4] Behavior and identity recognition

انسان و یا وسیله، دسته بندی می‌شوند. در مرحله سوم اشیاء دسته بندی شده به منظور تخمین مکانشان و همچنین تخمین تغییرات ظاهریشان، ردگیری می‌شوند. در آخرین مرحله هویت و رفتار اشیاء با بررسی خصوصیات حرکتی و بیو متریک (زیست سنجی)[1] آن‌ها تشخیص داده می‌شود [۲].

۱-۱ بیان مسئله

در این پژوهش فرض ما بر این بوده است که تنها منبع اطلاعاتی که در اختیار داریم، دنباله تصاویر تهیه شده به وسیله یک دوربین متحرک می‌باشد. سؤالی که می‌توان مطرح کرد، این است که در یک دنباله از تصاویر چه اطلاعاتی وجود دارد؟ اطلاعات موجود در دنباله تصاویر را می‌توان در دو دسته اطلاعات مکانی و زمانی دسته بندی نمود. اطلاعات مکانی از چگونگی سطح روشنایی الگوهای نوری موجود در یک تصویر و یا به طور کلی از خصوصیات ظاهری الگوهای نوری استخراج می‌شوند. اطلاعات زمانی، از جابجایی مکانی الگوهای نوری در طول زمان، استخراج می‌شوند. اطلاعات مکانی ویژگی‌هایی مانند گوشه و لبه و یا انواع ویژگی‌های دیگر است که یک جزء خاص از صحنه را توصیف می‌کند و از اجزای صحنه استخراج می‌شود. جابجایی الگوهای نوری در طول زمان را شار نوری می‌نامند[۳]. از شار نوری و یا چگونگی جابجایی الگوهای نوری نیز می‌توان اطلاعاتی را استخراج نمود. با استفاده از این اطلاعات قصد داریم به حل مسئله تشخیص شی متحرک در تصاویر دوربین متحرک بپردازیم.

در میان چهار هسته تکنولوژی خودکار پایش محیط، پایین‌ترین سطح، تشخیص شی می‌باشد و عملکرد این بخش به طور مستقیم بر مراحل بعدی آن تأثیر گذار می‌باشد. تشخیص شی را می‌توان در دو رویکرد دسته بندی نمود: ۱. تشخیص شی متحرک[2] ۲. تشخیص شی بر اساس ظاهر[3]. رویکرد تشخیص شی متحرک، اشیاء مورد نظر را بر اساس تغییراتی که در اثر حرکت آن‌ها در دنباله تصاویر ایجاد می‌شود، شناسایی می‌کند. این رویکرد اشیاء متحرک را فارغ از شکل و اندازه آن‌ها شناسایی می‌کند و نیازی به نمونه‌های زیاد و زمان طولانی جهت فرآیند آموزش ندارد و به منابع محاسباتی پایینی نیازمند است. اما این رویکرد نسبت به تکان‌های کوچک فرکانس بالای[4] دوربین، حرکت پس زمینه و تغییرات روشنایی حساس می‌باشد. رویکرد تشخیص شی بر اساس ظاهر، اشیاء مورد نظر را بر اساس شکل فیزیکی آن‌ها شناسایی می‌کند. از آنجایی که این رویکرد از اطلاعات مکانی استفاده می‌کند نسبت به رویکرد تشخیص شی متحرک که از اطلاعات زمانی استفاده می‌کند، نسبت به حرکت دوربین و پس زمینه پایدار می‌باشد و نسبت به تغییرات روشنایی محیط حساسیت کمتری دارد. اما این رویکرد نسبت به تغییر ظاهر اشیاء که به دلیل تغییر زاویه دید دوربین ایجاد می‌شود و چگونگی قرار گرفتن شی حساس می‌باشد. همچنین این رویکرد به نمونه‌های زیاد و زمان آموزش طولانی نیازمند است و

---

[1] Biometric
[2] Moving object detection
[3] Appearance-based object detection
[4] Jitters

هزینه محاسباتی بالایی برای پیدا کردن اشیاء در مکان‌ها و مقیاس‌های متفاوت دارد. از این‌رو، روش تشخیص شی متحرک به طور وسیع تری در سیستم‌های بلادرنگ پایش محیط نسبت به رویکرد تشخیص شی بر اساس ظاهر مورد استفاده قرار گرفته است. به همین دلیل، در این پژوهش تمرکز ما بر تشخیص شی متحرک خواهد بود. تشخیص شی متحرک در تصاویر دوربین ثابت در دهه‌های گذشته مطرح بوده است و تحقیقات فراوانی در این موضوع انجام پذیرفته است. استفاده از دوربین‌های متحرک به دلیل پوشش یک ناحیه وسیع تر نسبت به دوربین ثابت مورد توجه قرار گرفته است. همچنین در مواردی طبیعت مسئله ایجاب می‌کند که دوربین متحرک باشد مانند استفاده از فناوری تشخیص شی متحرک در تصاویر ویدئویی تهیه شده توسط یک هواپیمای بدون سرنشین و یا تشخیص شی متحرک توسط دوربینی که بر روی یک ربات یا یک وسیله نقلیه قرار دارد. اما در تصاویر دوربین متحرک به دلیل تغییر پیوسته خط دید دوربین، پس زمینه دارای یک حرکت ظاهری خواهد بود که تشخیص شی متحرک را با مشکل مواجه می‌نماید.

۱-۱-۱ انواع حرکت‌های محتمل در دنباله تصاویر

کلیدی‌ترین مشکل در تشخیص شی متحرک در تصاویر دوربین متحرک، حرکت دوربین می‌باشد که در اصطلاح علمی به آن خود حرکتی[1] می‌گویند [۴]. خود حرکتی به معنای ایجاد تغییرات دوبعدی در دنباله تصاویر بر اثر حرکت سه بعدی دوربین می‌باشد. از آنجا که حرکت دوربین باعث حرکت در تمامی اجزای تصویر می‌شود، برای آشکار سازی شی متحرک نمی‌توانیم از روش‌های توسعه داده شده جهت تشخیص شی متحرک در تصاویر دوربین ثابت استفاده نماییم [۵].

حرکت دو بعدی مشاهده شده در دنباله تصاویر می‌تواند به وسیله چند عامل ایجاد شده باشد، این عوامل عبارتند از[۶]:

۱- حرکت دوربین (خود حرکتی).

۲- تغییر در پارامترهای داخلی دوربین (بزرگنمایی).

۳- حرکت اشیاء موجود در صحنه.

۴- تغییرات روشنایی صحنه و تغییرات ساختار یا اندازه و شکل ظاهری اشیا.

ما علاقه‌مند به تشخیص اشیاء متحرک موجود در صحنه می‌باشیم، فرض می‌کنیم که اشیاء صلب باشند و تغییراتی در روشنایی صحنه و همچنین تغییری در پارامترهای داخلی دوربین نداشته باشیم، از این‌رو حرکتی که عامل آن موارد چهارم و دوم باشد، در دنباله تصاویر وجود ندارد. بنابراین تا به اینجا تنها دو عامل برای حرکتی که در دنباله تصاویر مشاهده می‌کنیم وجود دارد و در صورت تشخیص حرکت ناشی از حرکت دوربین (خود حرکتی) و جبران سازی این حرکت مابقی حرکت موجود در دنباله تصاویر می‌بایست مربوط به اشیاء متحرک باشد، و با تشخیص این حرکت باقی مانده می‌توانیم اجسام متحرک را شناسایی کنیم.

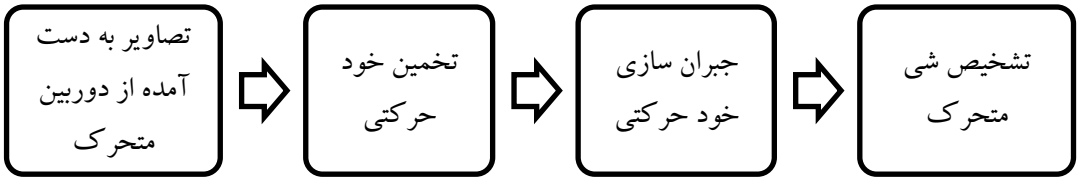
---

[1] Ego-motion

۱-۱-۲ فرآیند تشخیص شی متحرک در تصاویر دوربین متحرک

فرآیند تشخیص شی متحرک در تصاویر دوربین متحرک در شکل ۱-۱ نشان داده شده است. در حین دریافت تصاویر ابتدا خود حرکتی موجود در دنباله تصاویر تخمین زده می‌شود سپس خود حرکتی از دنباله تصاویر حذف می‌شود و در نهایت با استفاده از روش‌های تشخیص شی متحرک که برای تصاویر دوربین ثابت توسعه داده شده است، اشیاء متحرک را شناسایی می‌کنیم.

روش‌های تشخیص شی متحرک در تصاویر دوربین ثابت در سه دسته تقسیم بندی می‌شوند : ۱. تفاضل زمانی ۲. تفریق پس زمینه ۳. شار نوری [۵]. در روش تفاضل زمانی، تفاضل سطح روشنایی و یا مقدار رنگ دو تصویر متوالی و یا دو تصویر که فاصله زمانی کوتاهی با یکدیگر دارند محاسبه می‌شود. با فرض اینکه دوربین ثابت است، مقدار تفاضل در تصویر تفاضلی در مکان‌هایی که اشیاء متحرک حضور دارند، قابل توجه خواهد بود و به این صورت اشیاء متحرک را شناسایی می‌کنند. در روش تفریق پس زمینه، یک مدل از پس زمینه تهیه می‌شود و برای تشخیص شی متحرک در یک تصویر، آن را با مدل پس زمینه مقایسه می‌نمایند. ناحیه‌هایی از تصویر که مطابق با مدل پس زمینه نمی‌باشد به عنوان ناحیه‌هایی که اشیاء متحرک در آنجا حضور دارند تشخیص داده می‌شوند. شار نوری یک میدان سرعت است که مشخص می‌کند یک نقطه از تصویر در دو تصویر متوالی در چه جهتی و با چه سرعتی جابجا شده است. در این روش، شار نوری را محاسبه می‌نمایند، با فرض ثابت بودن دوربین، نقاطی که شار نوری آن‌ها در حدود صفر است به عنوان پس زمینه شناخته می‌شوند و نقاطی که شار نوریشان قابل توجه است به عنوان نقطه متعلق به جسم متحرک شناخته می‌شوند.



شکل۱-۱ فرآیند تشخیص شی متحرک در تصاویر دوربین متحرک

هنگامی که دوربین متحرک است، به دلیل حرکت ظاهری پس زمینه، روش‌های تشخیص شی متحرک در تصاویر دوربین ثابت، قادر به شناسایی اجسام متحرک نمی‌باشند. همان طور که اشاره شده، حرکت ظاهری پس زمینه در اثر حرکت دوربین را خود حرکتی می‌نامند. در حالتی که دوربین متحرک است با حذف خود حرکتی در دنباله تصاویر می‌توانیم با استفاده از تفریق پس زمینه و یا تفاضل زمانی، اشیاء متحرک را شناسایی کنیم برای این کار در روش تفاضل زمانی باید بتوانیم دو تصویر متوالی را بر روی یکدیگر تطبیق[1] دهیم، همچنین در روش تفریق پس زمینه باید تصویر جاری

[1] Registration

را بر روی مدل پس زمینه تطبیق دهیم. برای استفاده از روش شار نوری در تشخیص شی متحرک در تصاویر دوربین متحرک، باید بتوانیم شار نوری ایجاد شده در اثر حرکت دوربین را تشخیص دهیم. نقاطی که شار نوری آن‌ها با شار نوری ایجاد شده در اثر حرکت دوربین، متفاوت می‌باشد به عنوان شی متحرک شناسایی می‌شوند در غیر این صورت این نقاط متعلق به پس زمینه خواهند بود.

کیم و همکارانش [۵] بر اساس همین ایده‌ها برای تشخیص شی متحرک در تصاویر دوربین متحرک به دسته بندی روش‌های مختلف پرداختند. آن‌ها چهار رویکرد را به عنوان رویکردهای اصلی معرفی نمودند: رویکرد تطبیق ویژگی، رویکرد موزائیک پس زمینه، رویکرد شار نوری و رویکرد مدل هندسی دوربین. رویکرد موزائیک پس زمینه، در ابتدا یک موزائیک از پس زمینه تهیه می‌کند. یعنی هر تصویر از دنباله تصاویر را به یک سیستم مختصات منحرف[1] می‌کنند و این تصاویر را به یکدیگر می‌دوزند[2]. با این کار از کل تصاویر صحنه که در موقعیت‌های مختلف دوربین به دست می‌آیند، یک تصویر ایجاد می‌کند این تصویر را تصویر موزائیک می‌گویند. برای تشخیص شی متحرک در تصویر جاری این تصویر را با قسمت متناظرش در تصویر موزائیک تطبیق می‌دهند و با مقایسه این دو تصویر اشیاء متحرک را شناسایی می‌کنند. رویکرد شار نوری، سعی می‌کند شار نوری ایجاد شده بر اثر حرکت دوربین را تشخیص دهد، مرحله کلیدی دسته بندی شار نوری محاسبه شده می‌باشد. برای مثال فرض می‌کنند اشیاء متحرک درصد کوچکی از تصویر را اشغال می‌کنند، در نتیجه با یافتن شار نوری غالب در شار نوری محاسبه شده در دو تصویر متوالی، شار نوری ایجاد شده در اثر حرکت دوربین را مشخص می‌کنند و با مقایسه شار نوری نقاط تصویر با شار نوری ایجاد شده در اثر حرکت دوربین، نقاط را در دو دسته پس زمینه و شی متحرک، دسته بندی می‌نمایند. رویکرد مدل هندسی دوربین، با اندازه گیری پارامترهای داخلی دوربین (مانند فاصله کانونی) و پارامترهای حرکتی دوربین به وسیله سنسورها، مکان یک پیکسل از تصویر فعلی را در تصویر بعدی محاسبه می‌کند. به این شکل خود حرکتی را تخمین می‌زنند و با حذف آن و استفاده از روش‌های تشخیص شی متحرک در تصاویر دوربین ثابت، اشیاء متحرک را شناسایی می‌کنند. در رویکرد تطبیق ویژگی، یک مدل ریاضی برای خود حرکتی فرض می‌شود. این رویکرد سعی می‌کند با یافتن ویژگی‌های متناظر در دو تصویر متوالی، پارامترهای مدل ریاضی فرض شده را تخمین بزند و به این صورت با تخمین خود حرکتی و حذف آن از دنباله تصاویر، می‌توانند با روش‌های تشخیص شی متحرک در تصاویر دوربین ثابت، اشیاء متحرک را شناسایی کنند.

رویکرد شار نوری هزینه زمانی و محاسباتی بالایی دارد. رویکرد مدل هندسی دوربین، وابسته به سخت افزار می‌باشد و نیاز به محاسبه دقیق پارامترهای داخلی و خارجی دوربین دارد و برای استفاده در دوربین‌های مختلف، نیاز به کالیبره شدن دارد. رویکرد موزائیک پس زمینه نیاز دارد ابتدا تصویر موزائیک پس زمینه را ایجاد کند و سپس به فرآیند

[1] Warp
[2] Image stitching

تطبیق فریم جاری بر روی تصویر موزائیک پس زمینه نیازمند می‌باشد. رویکرد تطبیق ویژگی نسبت به تمامی این روش‌ها هزینه زمانی و محاسباتی پایین تری دارد. از این جهت رویکرد تطبیق ویژگی مورد توجه ما قرار گرفت که در فصل سوم به طور خاص به بررسی این رویکرد خواهیم پرداخت.

همان طور که اشاره شد در رویکرد تطبیق ویژگی، برای تخمین خود حرکتی در ابتدا باید بتوانیم مدل ریاضی خود حرکتی را به دست بیاوریم و به تخمین پارامترهای این مدل بپردازیم. لونگوهیگنز و پاردزانی [۷] با فرض مدل سوراخ سوزن با افکنش پرسپکتیو به عنوان مدل دوربین، هنگامی که دوربین حرکت انتقالی و دورانی دارد، یک مدل ریاضی برای خود حرکتی بیان نمودند. این مدل، بردار سرعت هر نقطه تصویر در صفحه تصویر را مشخص می‌کند و تابعی از مختصات نقطه، عمق نقطه و پارامترهای حرکتی دوربین می‌باشد. از این‌رو برای تخمین خود حرکتی ناگزیر می‌باشیم که عمق نقاط صحنه و پارامترهای حرکتی دوربین را تخمین بزنیم. تخمین این پارامترها در همه صحنه‌ها امکان پذیر نیست و ذاتاً مسئله‌ای است که دارای ابهام می‌باشد [۶]، که در فصل دوم این مسائل را تا حدودی توضیح می‌دهیم.

ایرانی و آناندان [۶] با توجه به اینکه مدل ریاضی خود حرکتی، تابعی از عمق نقاط صحنه و پارامترهای حرکتی دوربین می‌باشد به بررسی مسئله تشخیص شی متحرک در صحنه‌های مختلف برای حرکت‌های متفاوت دوربین پرداختند و تلاش نمودند روشی برای تشخیص شی متحرک ارائه دهند که در انواع صحنه‌ها قابل استفاده باشد. آن‌ها به این نتیجه رسیدند که در بعضی از موارد حرکت دوربین و نوع صحنه می‌توان مدل ریاضی خود حرکتی را در قالب یک تبدیل دو بعدی پارامتری بیان نمود و برای تخمین خود حرکتی در این تصاویر کافی است پارامترهای این تبدیل دو بعدی را تخمین زد. با این کار مسئله تخمین خود حرکتی آسان می‌شود. تصاویر هوایی و تصاویر به دست آمده از دوربین PTZ در دسته مواردی قرار می‌گیرند که می‌توان خود حرکتی موجود در آن‌ها در قالب یک تبدیل دو بعدی پارامتری بیان نمود.

تشخیص شی متحرک در صحنه‌ها و حرکت‌هایی از دوربین که خود حرکتیشان قابل بیان در قالب یک تبدیل دو بعدی می‌باشد، مورد توجه قرار گرفت که در فصل چهارم به بررسی خود حرکتی در تصاویر هوایی و دوربین PTZ خواهیم پرداخت. روش‌های مختلف توسعه داده شده در رویکرد تطبیق ویژگی، از ویژگی‌های متفاوتی بهره برده‌اند. که ویژگی‌های متفاوت به لحاظ دقت و زمان محاسباتی با هم متفاوت می‌باشند، علاوه بر این برای صحنه‌های متفاوت به لحاظ شلوغی صحنه و یا یکنواختی صحنه از ویژگی‌های متفاوتی استفاده می‌شود. گوشه، خط و هیستوگرام رنگ از جمله این ویژگی‌ها هستند.

شور و همکارانش [۸] به تخمین و حذف خود حرکتی در تصاویر دوربین PTZ پرداختند. آن‌ها تبدیل مقیاس و انتقال را برای خود حرکتی موجود در این تصاویر در حالتی که نرخ تصویر برداری ۳۰ فریم بر ثانیه باشد، فرض کردند و ثابت کردند که تبدیل فرض شده آن‌ها دقت بالایی دارد. آن‌ها از کمینه‌ها و بیشینه‌های محلی پروفایل افکنش افقی و