



دانشکده برق و کامپیوتر

پایان نامه کارشناسی ارشد مهندسی مکترونیک

شناسایی اشیاء در تصاویر ویدئویی با استفاده از شبکه فازی-عصبی

دانشجو: صادق مهدی زاده

استاد راهنما: دکتر ساسان آزادی

مهر ۱۳۹۱

سپاسگزاری

باتقدیر و شکر شایسته از استاد فرهیخته و فرزانه جناب آقای دکتر آزادی که با نکته های دلاویز و گفته های بلند، صحیفه های سخن را علم پرور نمود و همواره راهنما و راه‌کنشای نگارنده در اتمام و تکمال پایان نامه بوده است.

معلمت ز عرش برتر باد همیشه توست اندیشه ات مظفر باد

باتقدیر و درود فراوان خدمت پدر و مادر بسیار عزیز، دلسوز و خداکارم که پیوسته جرعه نوش جام تعلیم و تربیت، فضیلت و انسانیت آنها بوده ام و همواره چراغ وجودشان روشنگر راه من در سختی ها و مشکلات بوده است.

و باشکر از همسر مهربانم که حس تعهد و مسؤلیت را در زندگی مان تملاتی خدایی داده است.

چکیده

تشخیص اشیاء متحرک در دنباله‌ای از تصاویر ویدئویی، یکی از موارد مورد علاقه‌ی پژوهشگران علم بینایی ماشین در سال‌های اخیر بوده است. این مسئله را می‌توان یکی از شاخه‌های تشخیص الگو دانست. در سال‌های اخیر، علاقه به شناسایی کننده‌های عصبی رو به گسترش بوده و تحقیقات زیادی در این زمینه صورت گرفته است.

در طرح حاضر، تشخیص سه کلاس از اشیاء متحرک با استفاده از یک شناسایی کننده‌ی فازی-عصبی انجام گرفته و نتایج حاصل، با دیگر روش‌های مرسوم از جمله روش عصبی مقایسه گردید. سه کلاس در نظر گرفته شده در این طرح عبارتند از: اتومبیل (کلاس ۱)، حیوان خانگی (کلاس ۲)، و انسان (کلاس ۳). برای شناسایی و تشخیص این سه کلاس، چهار ویژگی خاص که صرفاً از پارامترهای هندسی شکل استخراج می‌شوند، ارائه گردید. این چهار ویژگی عبارتند از: خروج از مرکز بیضی معادل، ضریب پیچیدگی شکلی، جهت گیری و ضریب سختی. در ادامه، مفید و کافی بودن این چهار ویژگی با استفاده از نتایج، اثبات گردید.

برای شناسایی اشیاء، از یک سیستم فازی نوع سوگنو، بهینه سازی شده با الگوریتم آموزش هیبرید در یک شبکه‌ی پس انتشار، استفاده شد. نتایج حاصل از تصاویر تست، تصاویر نویز دار و تصاویر دارای انسداد، با نتایج مشابه پنج شناسایی کننده‌ی دیگر مقایسه گردید. شناسایی کننده‌های مورد مقایسه شامل شناسایی کننده‌های نزدیک‌ترین همسایگی، کمترین فاصله‌ی میانگین، شناسایی کننده‌ی بیزی، شناسایی کننده‌ی عصبی و شناسایی کننده‌ی ساخته شده بر مبنای درخت تصمیم بوده‌اند. نتایج مقایسه‌ها حکایت از مطلوب و قابل اعتماد بودن نتایج شناسایی کننده‌ی فازی-عصبی دارد.

کلمات کلیدی: بینایی ماشین، تشخیص الگو، سیستم‌های فازی، دسته بندی، تشخیص شیء متحرک، شناسایی کننده‌ی فازی-عصبی تطبیقی

۱	فصل اول: مقدمه.....
۱-۱	۱-۱-۱ مروری بر تشخیص شیء در تصاویر.....
۲-۱	۲-۱-۱ روش‌های تشخیص پس زمینه.....
۳-۱	۳-۱-۱ روش‌های استخراج ویژگی.....
۴-۱	۴-۱-۱ روش‌های ایجاد شناسایی کننده.....
۵-۱	۵-۱-۱ جمع بندی فصل اول.....
۱۵	فصل دوم: مبانی تئوری.....
۱-۲	۱-۲-۱ سیستم‌های فازی.....
۱-۱-۲	۱-۱-۲-۱ مجموعه‌های فازی.....
۲-۱-۲	۲-۱-۲-۱ توابع عضویت.....
۳-۱-۲	۳-۱-۲-۱ قوانین فازی.....
۴-۱-۲	۴-۱-۲-۱ سیستم‌های استنتاج فازی.....
۲-۲	۲-۲-۱ شبکه‌های عصبی.....
۱-۲-۲	۱-۲-۲-۱ شبکه‌های عصبی و منطق فازی.....
۳-۲	۳-۲-۱ شبکه‌ی فازی-عصبی.....
۱-۳-۲	۱-۳-۲-۱ مدل سیستم سوگنو.....
۲-۳-۲	۲-۳-۲-۱ الگوریتم بهینه سازی.....
۳-۳-۲	۳-۳-۲-۱ فرا انطباق.....
۴-۲	۴-۲-۱ روش نزدیک‌ترین همسایگی‌ها.....
۱-۴-۲	۱-۴-۲-۱ الگوریتم نزدیک‌ترین همسایگی.....
۵-۲	۵-۲-۱ روش شناسایی کننده‌ی بیزی.....
۱-۵-۲	۱-۵-۲-۱ مدل احتمالی بیز.....
۲-۵-۲	۲-۵-۲-۱ تخمین پارامترها.....
۳-۵-۲	۳-۵-۲-۱ قانون تصمیم‌گیری شناسایی کننده‌ی بیزی.....

۳۵	۲-۶ جمع بندی فصل
۳۶	فصل سوم: طراحی و پیاده سازی
۳۷	۳-۱ تصاویر انتخابی
۳۸	۳-۱-۱ تصاویر آموزشی
۳۹	۳-۱-۲ تصاویر آزمون
۳۹	۳-۱-۳ تصاویر نوپرداز
۴۰	۳-۱-۳-۱ عملیات پیش پردازش
۴۲	۳-۱-۴ تصاویر دارای انسداد
۴۲	۳-۱-۴-۱ انسداد مثبت
۴۳	۳-۱-۴-۲ انسداد منفی
۴۴	۳-۲ انتخاب ویژگی‌ها
۴۴	۳-۲-۱ در دسر ابعادی
۴۵	۳-۲-۲ بردار ویژگی انتخاب شده
۴۹	۳-۳ شناسایی کننده‌ی نزدیک‌ترین همسایگی
۵۲	۳-۴ شناسایی کننده‌ی کمترین فاصله‌ی میانگین
۵۴	۳-۵ شناسایی کننده‌ی بیز
۵۹	۳-۶ شناسایی کننده‌ی عصبی
۶۵	۳-۷ شناسایی کننده‌ی فازی عصبی
۶۶	۳-۷-۱ تکنیک فازی-عصبی
۶۹	۳-۷-۲ آموزش شبکه فازی-عصبی
۷۳	۳-۷-۳ نتایج عملی
۷۷	۳-۸ شناسایی کننده‌ی درخت تصمیم
۸۱	۳-۹ جمع بندی فصل سوم
۸۳	فصل چهارم: نتیجه گیری و بحث
۸۴	۴-۱ بررسی ویژگی‌ها

۸۵.....	۲-۴ دقت کلی شناسایی کننده‌ها
۸۶.....	۳-۴ مقاومت در مقابل نويز
۸۷.....	۴-۴ مقاومت در مقابل انسداد
۸۸.....	۵-۴ زمان آموزش
۸۹.....	۶-۴ زمان اجرا
۸۹.....	۷-۴ مقایسه‌ی کلی
۹۰.....	۸-۴ پیشنهادات
۹۱.....	مراجع
۹۵.....	واژه نامه

فهرست شکل‌ها

- شکل ۱-۱ توزیع گاوسین دو بعدی با میانگین $(0,0)$ و $\sigma = 1$ ۶
- شکل ۲-۱ تخمین گسسته‌ی تابع گاوسی با $\sigma=1$ ۶
- شکل ۳-۱ موارد اشیاء مورد استفاده با یک درجه آزادی و در چهار کلاس مختلف ۸
- شکل ۴-۱ محاسبه نسبت طول به عرض شیء ۱۲
- شکل ۵-۱ تعیین مرکز ثقل شیء و محاسبه فاصله هر نقطه از کانتور تا مرکز ثقل ۱۳
- شکل ۶-۱ (الف) فاصله هر نقطه از کانتور تا مرکز، (ب) فاصله های هموار شده با میانگین گیری ۱۳
- شکل ۱-۲ مجموعه‌ی فازی افراد بلند قد ۱۸
- شکل ۲-۲ یک سیستم فازی نوع ممدانی ۲۱
- شکل ۳-۲ سیستم فازی نوع سوگنو ۲۱
- شکل ۴-۲ شمای کلی شبکه‌ی عصبی ۲۲
- شکل ۵-۲ قوانین سیستم فازی سوگنو مرتبه ۱ ۲۶
- شکل ۶-۲ ساختار سیستم فازی سوگنو ۲۶
- شکل ۷-۲ اعتباریابی خوب با استفاده از داده‌های بررسی ۲۹
- شکل ۸-۲ اعتباریابی نامناسب ۳۰
- شکل ۹-۲ مثالی از روش kNN ، توضیح در متن ۳۱
- شکل ۱-۳ چند نمونه از تصاویر استفاده شده جهت آموزش ۳۹
- شکل ۲-۳ تصویر مربوط به یک اتومبیل ۴۰
- شکل ۳-۳ تأثیر عملیات پردازش در حذف نویز و تغییرات به وجود آمده در ظاهر شیء ۴۱
- شکل ۴-۳ انسداد مثبت ۴۳
- شکل ۵-۳ انسداد منفی ۴۳
- شکل ۶-۳ کاهش عملکرد در اثر افزایش ابعاد بردار ویژگی ۴۴
- شکل ۷-۳ بیضی معادل شیء ۴۵
- شکل ۸-۳ تغییرات میزان خروج از مرکز در مقایسه با تغییر شکل ۴۶
- شکل ۹-۳ مراحل به دست آوردن ضریب پیچیدگی ۴۷
- شکل ۱۰-۳ استفاده از بیضی در به دست آوردن جهت شیء ۴۷
- شکل ۱۱-۳ (الف) تصویر اصلی، (ب) تصویر محدب، (ج) تصویر اصلی به همراه چند ضلعی محدب ۴۸

شکل ۳-۱۲ نمودار مقایسه دقت روش نزدیکترین همسایگی در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۵۱

شکل ۳-۱۳ نمودار مقایسه دقت روش کمترین فاصله‌ی میانگین در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۵۴

شکل ۳-۱۴ نمودار مقایسه دقت روش بیزی در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۵۷

شکل ۳-۱۵ معماری نهایی شبکه‌ی عصبی طراحی شده.....۶۱

شکل ۳-۱۶ نمودار خطای داده‌های آموزشی در حین آموزش شبکه.....۶۱

شکل ۳-۱۷ نمودار عملکرد شبکه در قبال داده‌های آموزشی.....۶۲

شکل ۳-۱۸ نمودار مقایسه دقت شبکه‌ی عصبی در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۶۴

شکل ۳-۱۹ تبدیل دانش خبره به سیستم‌های فازی.....۶۵

شکل ۳-۲۰ توابع عضویت قابل استفاده برای متغیرهای ورودی.....۷۰

شکل ۳-۲۱ نمودار خطای آموزش و چک با شصت بار آموزش.....۷۱

شکل ۳-۲۲ نمودار خطای داده‌های آموزش و داده‌های چک تا قبل از رسیدن به فراانطباق.....۷۲

شکل ۳-۲۳ نمودار مقایسه دقت شبکه‌ی فازی-عصبی در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۷۴

شکل ۳-۲۴ اجرای روش فازی-عصبی با حضور دو شیء متحرک.....۷۵

شکل ۳-۲۵ نتیجه‌ی روش فازی-عصبی در مورد یک نمونه از کلاس ۲.....۷۶

شکل ۳-۲۶ نتیجه‌ی روش فازی-عصبی در مورد یک نمونه از کلاس ۱.....۷۷

شکل ۳-۲۷ نمودار درخت تصمیم طراحی شده با الگوریتم C5.0.....۷۹

شکل ۳-۲۸ نمودار مقایسه دقت روش درخت تصمیم در مواجهه با داده‌های نویز دار به تفکیک کلاس‌های مختلف.....۸۱

شکل ۴-۱ توزیع عددی ویژگی‌های انتخاب شده در کلاس‌های مختلف برای داده‌های آموزشی.....۸۵

فهرست جدول‌ها

جدول ۱-۲ مقایسه‌ی شبکه‌های عصبی و سیستم‌های فازی.....	۲۵
جدول ۲-۲ الگوریتم بهینه سازی ترکیبی برای anfis.....	۲۸
جدول ۱-۳ مشخصات نوع و ویژگی های نویزهای اضافه شده به تصاویر.....	۴۰
جدول ۲-۳ نتایج روش نزدیک ترین همسایگی در مورد دویست داده‌ی آزمایشی.....	۴۹
جدول ۳-۳ نتایج روش نزدیک ترین همسایگی در مواجهه با داده‌های نویز دار.....	۵۰
جدول ۴-۳ عملکرد روش نزدیک ترین همسایگی در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....	۵۱
جدول ۵-۳ نتایج روش کمترین فاصله میانگین در مورد داده های آزمایشی.....	۵۳
جدول ۶-۳ نتایج روش کمترین فاصله میانگین در مواجهه با داده‌های نویز دار.....	۵۳
جدول ۷-۳ عملکرد روش کمترین فاصله‌ی میانگین در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....	۵۴
جدول ۸-۳ میانگین و انحراف معیار داده‌های آموزشی در کلاس‌های مختلف به تفکیک ویژگی.....	۵۵
جدول ۹-۳ نتایج روش بیزی در مورد دویست داده‌ی آزمایشی.....	۵۶
جدول ۱۰-۳ نتایج روش بیزی در مواجهه با داده‌های نویز دار.....	۵۷
جدول ۱۱-۳ کد متلب استفاده از شناسایی کننده‌ی بیزی.....	۵۸
جدول ۱۲-۳ عملکرد روش بیزی در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....	۵۹
جدول ۱۳-۳ نتایج روش استفاده از شبکه‌ی عصبی در مورد دویست داده‌ی آزمایشی.....	۶۳
جدول ۱۴-۳ نتایج روش شبکه‌ی عصبی در مواجهه با داده‌های نویز دار.....	۶۳
جدول ۱۵-۳ عملکرد شبکه عصبی در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....	۶۴
جدول ۱۶-۳ خطای آموزش، تست و بررسی بر حسب توابع عضویت فازی مختلف و دو روش بهینه سازی.....	۷۰
جدول ۱۷-۳ فرمان آموزش سیستم فازی- عصبی تا قبل از رسیدن به فرا انطباق در مرحله‌ی دهم.....	۷۱
جدول ۱۸-۳ عملکرد دو سیستم فازی- عصبی در قبال دویست داده آزمایشی به تفکیک تعداد دفعات آموزش.....	۷۲
جدول ۱۹-۳ دقت سیستم فازی عصبی در مواجهه با نمونه های جدید، بر حسب نوع توابع عضویت فازی.....	۷۳

- جدول ۳-۲۰ نتایج روش فازی-عصبی در مواجهه با داده‌های نويز دار.....۷۴
- جدول ۳-۲۱ عملکرد روش فازی-عصبی در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....۷۵
- جدول ۳-۲۲ خروجی نرم افزار See5 و طراحی درخت تصمیم بر مبنای داده‌های آموزشی.....۷۸
- جدول ۳-۲۳ نتایج روش درخت تصمیم در مورد دویست داده‌ی آزمایشی.....۸۰
- جدول ۳-۲۴ نتایج روش درخت تصمیم در مواجهه با داده‌های نويز دار.....۸۰
- جدول ۳-۲۵ عملکرد روش درخت تصمیم در مواجهه با انسداد به تفکیک کلاس و نوع انسداد.....۸۱
- جدول ۴-۱ مقایسه دقت کلی شناسایی کننده‌های مختلف به تفکیک نوع کلاس.....۸۶
- جدول ۴-۲ دقت شناسایی کننده‌ها در مقابله با داده‌های نويزی در کلاس ۱.....۸۷
- جدول ۴-۳ مقایسه دقت شناسایی کننده‌های مختلف در مواجهه با تصاویر دارای انسداد مثبت به تفکیک نوع کلاس.....۸۷
- جدول ۴-۴ مقایسه دقت شناسایی کننده‌های مختلف در مواجهه با تصاویر دارای انسداد منفی به تفکیک نوع کلاس.....۸۸
- جدول ۴-۵ مقایسه کلی ویژگی‌های شناسایی کننده‌های بحث شده.....۹۰

فصل اول

مقدمه

مقدمه

در دو دهه اخیر، فعالیت های زیادی در زمینه مسائل مربوط به بینایی کامپیوتری انجام شده است. دامنه این فعالیت ها حوزه هایی همچون نظارت و دیده بانی، تشخیص چهره، هدایت ربات و ... را پوشش می دهد. یکی از اهداف اصلی در بینایی کامپیوتری طراحی ماشین هایی بوده است که هوشمندانه تر و یا به صورت انسان نما عمل کنند. یک ماشین برای اینکه هوشمند و مفید باشد، باید قادر باشد محیطی را که در آن جاسازی شده است، درک کند. اخیراً تکنیک های بینایی ماشین که بر شبکه های عصبی و یا سیستم های فازی بنا شده اند، پیشنهاد شده اند.

۱-۱ مروری بر تشخیص شیء در تصاویر

شناسایی اجسام متحرک شامل سه مرحله اصلی است. اولین مرحله، جداسازی سوژه از پس زمینه است. مرحله دوم استخراج بردار ویژگی شیء مورد نظر است. و مرحله سوم شناسایی جسم است. تکنیک های شناسایی جسم معمولاً از شکل یک جسم به عنوان بردار ویژگی استفاده می کنند. استفاده از روش های دنباله گیری خطوط خارجی^۱ و کدگذاری زنجیره ای^۲ به عنوان ابزارهای اولیه برای ضبط موقعیت نقاط لبه خارجی برای استخراج بردار ویژگی مطرح است. در بعضی از مطالعات،

¹ -Contour Following

² -Chain Coding

تبدیل فوریه گسسته (DFT) و ضرایب آن برای نقاط لبه استفاده شده اند که این مسئله به دلیل نامتغیر بودن این ضرایب با وجود چرخش، تغییر مکان و سایز است. علاوه بر ضرایب فوریه، ضرایب ویولت^۱ نیز به عنوان ویژگی استفاده می شود. استخوان بندی ستاره ای^۲ نیز به عنوان ویژگی شناخته شده است. در بعضی منابع و مطالعات یک ضریب فشردگی به صورت نسبت مساحت به مجذور محیط تعریف شده است و برای کلاسه بندی انسان و وسیله نقلیه مورد استفاده قرار گرفته است. [۱] ثابت های ممانی^۳ که نسبت به انتقال و چرخش، نامتغیر هستند نیز برای شناسایی اجسام با اشکال مختلف مورد استفاده قرار گرفته اند. در بعضی مراجع هفت مورد از ثوابت ممانی استاندارد شده به عنوان ویژگی های شکلی انسان و اتومبیل مورد استفاده قرار گرفته اند. [۲]

عمل شناسایی جسم در سیستم های بینایی به خاطر مسائلی همچون انسداد^۴، تغییرات در ژست، نورپردازی ضعیف، تغییر در ابعاد اجسام و چرخش در صفحه و خارج از آن با مشکلاتی مواجه می شود. [۳] یک سیستم تشخیص ممکن است در طی فرآیند تشخیص با قسمتی و یا تمام موارد ذکر شده، مواجه شود. فاکتورهای مختلفی ممکن است در ایجاد مسائل ذکر شده دخیل باشد. مشکل ممکن است در اثر ضعف دقت سنسور، شرایط محدود و یا طبیعت مورفولوژیک جسم به وجود آید. به عنوان مثال تشخیص یک شیء غیرصلب همچون عابر پیاده، به دلیل داشتن درجات آزادی بالا با سختی همراه است. طبیعت بندبند و مفصل دار بدن انسان، عمل تشخیص عابر پیاده را به یکی از مشکل ترین عملیات های تشخیص شیء تبدیل کرده است. از طرف دیگر بدنه های صلب همچون وسایل نقلیه با توجه به کاربرد قطعی و مشخص شان با سختی کمتری شناسایی می شوند. این مسئله می تواند ناشی از این واقعیت باشد که وسایل نقلیه بدنه صلب دارند، قابلیت تغییر شکل ندارند و مسئله خود انسدادی^۵ در اغلب موارد در آن ها وجود ندارد. مسائل دیگر همچون تغییرات در نورپردازی، تغییر ژست و انسداد همچنان به عنوان یک مشکل حتی در اجسام صلب باقی می ماند. فارغ از اینکه یک سیستم تشخیص شیء، چه نوع جسمی را باید شناسایی کند، یک پاسخ قطعی بله یا خیر در ابتدای امر، در پروسه تشخیص شیء، می تواند گمراه کننده باشد. یک الگوریتم تشخیص شیء قدرتمند باید عدم قطعیت را در پروسه تشخیص تا هنگامی که آخرین مرحله تشخیص کامل شود، حفظ کند. [۳] به عنوان مثال اگر یک سیستم بینایی ماشین بر مبنای تشخیص لبه، قطعه بندی، مطابقت شکل^۶ و زیرسیستم های دیگر بنا شده باشد، باید اجازه ظهور تعدادی عدم قطعیت را

¹ -wavelet

² -Star Skeletonization

³ -Moment Invariants

⁴ -Occlusion

⁵ -Self-Occlusion

⁶ -Shape Matching

در هر یک از این زیرسیستم ها بدهد و جواب قطعی بله/خیر را تا زمانی که پاسخ تمام زیرسیستم ها آماده شود، به تأخیر اندازد. تصمیم نهایی باید بر مبنای خروجی های تمامی زیرسیستم ها به طور همزمان اتخاذ شود.

۲-۱ روش های تشخیص پس زمینه

جدا کردن اشیاء موجود در پس زمینه منوط به تشخیص وجود حرکت در دنباله‌ی تصاویر ویدئویی و بالاخص در فریم حاضر است. برای تشخیص حرکت، الگوریتم‌ها و الگوهای متعددی ارائه شده است که به بعضی از آنها اشاره می‌شود.

به طور کلی می‌توان روش‌های تشخیص حرکت را به چهار گروه کلی تقسیم کرد: تفریق پیش-زمینه، دیدگاه‌های آماری، دیدگاه‌های زمانی و دیدگاه‌های مکانی و فلوی نوری. از آن جا که تمرکز این پایان‌نامه بر روی مرحله‌ی دوم و سوم فرآیند دسته‌بندی است (یعنی استخراج بردار ویژگی و شناسایی شیء)، تنها به تعداد معدودی از روش‌های پیشنهادشده برای مدل‌سازی پس‌زمینه اشاره‌ای گذرا می‌شود.

۱. الگوریتم Heikkila and Silver

این روش از جمله روش‌های مبتنی بر تفریق پیش‌زمینه است به طور خلاصه می‌توان گفت که در این روش، مدل پس‌زمینه از رابطه‌ی زیر تبعیت می‌کند:

$$B_{t+1} = \alpha I_t + (1 - \alpha) B_t \quad (1-1)$$

که در آن α ثابت نرخ تغییرات است که برای جلوگیری از دنباله‌های مصنوعی ایجاد شده پشت سر اشیاء کم تحرک قرار داده می‌شود. همچنین تشخیص حرکت با استفاده از رابطه‌ی زیر صورت می‌گیرد:

$$|I_t(x, y) - B_t(x, y)| > Th \quad (2-1)$$

چنانچه در یک پیکسل $P(x, y)$ ، اختلاف مقدار فریم حاضر با مدل پس‌زمینه از حد آستانه‌ی بخصوصی بیشتر باشد، آن پیکسل جزء تصویر پیش‌نما در نظر گرفته می‌شود. در این روش تصحیحاتی هم برای جبران سازی تغییرات روشنایی ناگهانی و ظهور اشیاء ثابت جدید و همچنین جبران سازی نوسانات روشنایی نظیر نوسانات شاخه‌ها در نظر گرفته شده است. [۴]

۲. روش Pfunder

در این روش پیکسل‌های پیش‌زمینه به وسیله‌ی یک مقدار واحد به روزرسانی می‌گردند.

$$B_t = \alpha I_t + (1 - \alpha) B_{t-1} \quad (3-1)$$

پیکسل‌های پس‌زمینه به وسیله‌ی میانگین و کوواریانس مدل شده‌اند. این روش برای شروع احتیاج به یک منظره‌ی خالی دارند. [۵]

۳. روش cutler

در این روش پیش‌نما تخمین میانگین زمانی N فریم آخر است. اگر پیکسل در فریم حاضر در شرط زیر صدق کند، به عنوان پیش‌نما و بخشی از ناحیه‌ی متحرک در نظر گرفته می‌شود:

$$\sum_{C \in R, G, B} |I_t(C) - B_t(C)| > k\sigma \quad (4-1)$$

σ تخمین ایجاد شده آفلاین از انحراف استاندارد نویز و k یک ثابت از قبل انتخاب شده می‌باشد.

[۶]

۴. مدل میانه

در این روش به جای میانگین مقادیر و پیکسل، از میانه برای مدل‌سازی پیش‌زمینه استفاده شده است. ادعا شده است که استفاده از میانه‌ی مقادیر پیکسل دارای میزان دقت تشخیص حرکت بهتری نسبت به مدل میانگین بوده و دارای پیچیدگی زمانی کمتری نسبت به مدل میانگین و همچنین مدل ترکیب گاوسی می‌باشد. [۷]

۵. مدل ترکیب تطبیقی گاوسی (این روش از جمله روش‌های آماری می‌باشد).

در این روش هر پیکسل از تصویر به طور جداگانه به وسیله‌ی k گاوسین مدل می‌شود:

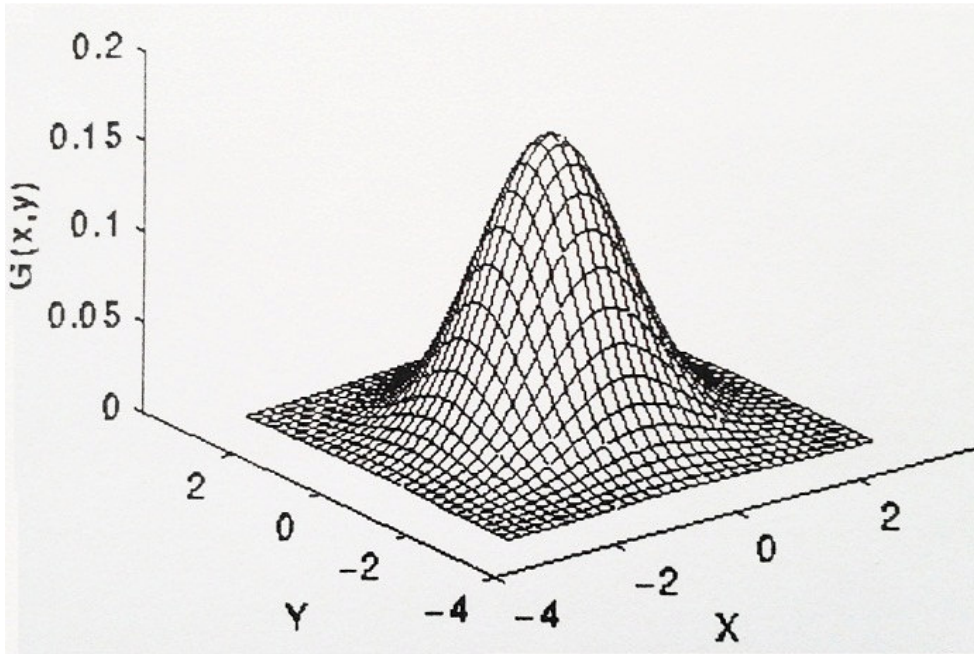
$$P(I_t) = \sum_{i=1}^k \omega_{i,t} \eta(I_t; \mu_{i,t}, \Sigma_{i,t}) \quad (5-1)$$

تعداد k از قبل مشخص و بین ۳ تا ۵ می‌باشد. همچنین $\Sigma_{i,t} = \sigma_{i,t}^2 I$ می‌باشد. یک تابع

گاوسی در فضای دو بعدی به صورت زیر تعریف می‌شود:

$$\eta(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (6-1)$$

نمایش تصویری این توزیع در شکل ۱-۱ دیده می‌شود.



شکل ۱-۱ توزیع گاوسین دو بعدی با میانگین (۰ و ۰) و $\sigma = 1$

به صورت تئوری، توزیع گاوسی در هر جایی از صفحه غیر صفر می‌باشد اما در عمل بهتر است که انحراف استاندارد بیشتر از سه برابر میانگین را صفر قرار دهیم. شکل ۲-۱ ماسک گاوسی با σ برابر ۱ را نشان می‌دهد. [۸]

$$\frac{1}{273}$$

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

شکل ۲-۱ تخمین گسسته‌ی تابع گاوسی با $\sigma=1$

۶. اختلاف دو فریم متوالی

این روش به طور بسیار ساده‌ای بر مبنای اختلاف هر فریم با فریم قبلی عمل می‌کند. طبق فرمول زیر:

$$|frame_i - frame_{i-1}| > Th \quad (7-1)$$

اگر در فریم فعلی اختلاف یک پیکسل با پیکسل متناظر در فریم قبلی از یک حد آستانه فراتر رود، پیکسل مذکور، جزئی از ناحیه‌ی متحرک در نظر گرفته می‌شود. این روش معایبی دارد که در زیر ذکر می‌شود:

- پس‌زمینه‌ی تخمینی فقط فریم قبلی می‌باشد.
- فقط در شرایط خاص سرعت شیء و نرخ فریم کار می‌کند.
- خیلی به آستانه Th حساس می‌باشد.

برای رفع ضعف این روش، یک روش هیبریدی شامل اختلاف سه فریم متوالی با یک مدل تفریق پیش‌زمینه تطبیقی توسعه داده شد. الگوریتم هیبریدی نواحی متحرک را در ویدئو بدون معایب اختلاف زمانی و تفریق پیش‌زمینه، قطعه‌بندی می‌نماید. در این روش از اختلاف سه فریم متوالی برای تعیین پیش‌نما بودن هر پیکسل استفاده می‌شود. [۹] این الگوریتم به صورت زیر است:

$$|I_t - I_{t-1}| + |I_t - I_{t-2}| > Th \quad (8-1)$$

که I_t فریم حاضر است.

در پایان‌نامه‌ی حاضر با توجه به ثابت بودن منظره و مشخص بودن مدل پس‌زمینه، احتیاج به استفاده از روش‌های فوق نبوده است. شایان ذکر است که ترکیب روش پیشنهاد شده با یک روش تشخیص پس‌زمینه می‌تواند به عنوان ادامه‌ی این کار تلقی شود که در قسمت پیشنهادات ارائه شده است.

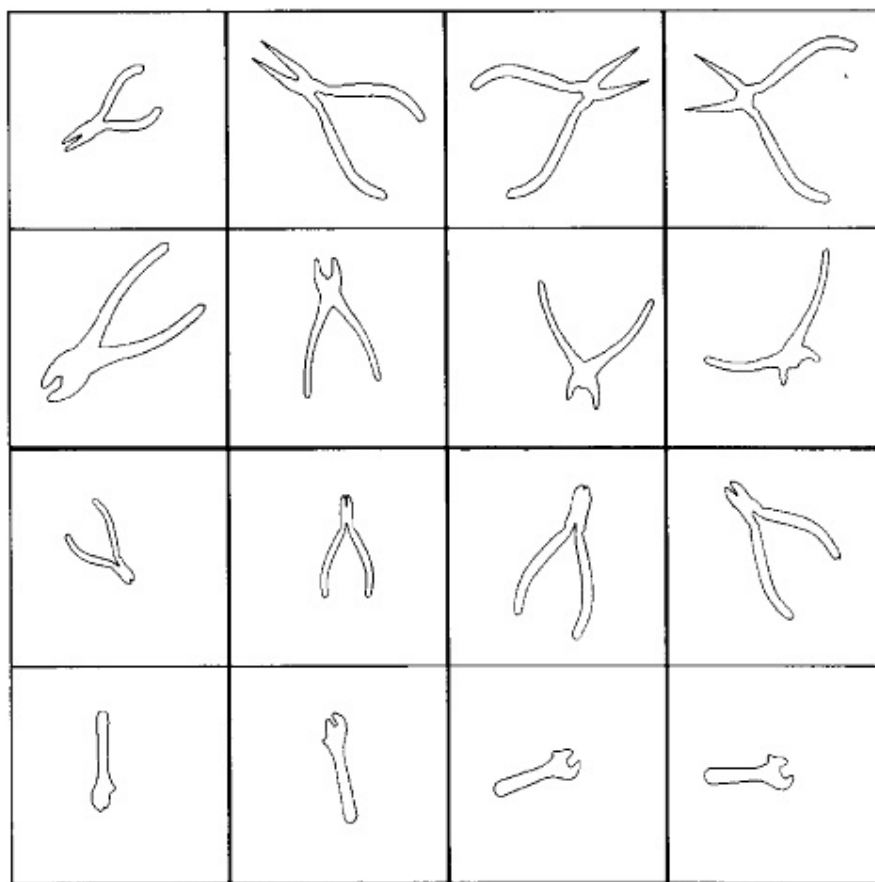
۳-۱ روش‌های استخراج ویژگی

در این قسمت به تعدادی از روش‌های استخراج ویژگی که در مقالات و پژوهش‌های گذشته مورد استفاده قرار گرفته‌اند، اشاره می‌شود. چگونگی ایجاد بردار ویژگی از یک تصویر گرافیکی داده شده، موضوع مهمی در تشخیص الگو می‌باشد. اطلاعات اصلی درباره‌ی یک شیء در مرزهای خارجی^۱ آن یافت می‌شود. به همین جهت تقریباً در تمامی موارد، از اطلاعات مرز اشیاء استفاده می‌شود.

¹ - boundary

در مرجع [۱۰] از توصیف‌گر فوریه جهت استخراج ویژگی استفاده شده است چرا که این توصیف‌گر، مقدار یکتایی را بدون توجه به چرخش، جابجایی و یا تغییر مقیاس، برای یک شیء ارائه می‌دهد. مزیت دیگری که برای این ابزار ارائه شده است، که می‌توان با نادیده گرفتن مولفه‌های درجه بالا، حجم داده را فشرده‌سازی و کوچک نمود.

در این مقاله روشی برای دسته‌بندی ابزارهایی با یک درجه آزادی ارائه شده است. به عنوان نمونه در شکل ۱-۳ تصاویر چند نمونه ابزار دیده می‌شود که زاویه‌ی فک آن‌ها متغیر است و به عنوان وجه تفاوت اصلی آن‌ها عمل می‌کند. چهار نوع کلاس خروجی در این پژوهش در نظر گرفته شده و در نهایت با تعدادی ابزار خارج از کلاس و تعدادی تصویر دارای نویز مورد آزمایش قرار گرفته است.



شکل ۱-۳ موارد اشیاء مورد استفاده با یک درجه آزادی و در چهار کلاس مختلف [۱۰]

در این پژوهش تصویر یک شیء با استفاده از یک دوربین CCD قرار گرفته در بالا، برداشته شده است. اشیاء مورد استفاده، ابزارهای صنعتی (انبردست) بوده‌اند. همچنین در این روش، ابزارها به صورت تک تک در مقابل دوربین قرار گرفته‌اند و به همین جهت امکان انسداد نیز وجود نداشته است.

اطلاعات تصویر به صورت یک ماتریس 256×256 ذخیره شده‌اند. با استفاده از عملگر سوبل^۱ لبه‌های تصویر به دست آمده و با استفاده از تصویر لبه‌ها، اطلاعات داده‌های موقعیت مرزی به دست آمده است. این اطلاعات به صورت $(x[m], y[m])$ برای $1 < m < L$ ذخیره شده است (L تعداد نقاط داده می‌باشد). سپس، با استفاده از تبدیل فوریه^۲ی گسسته^۲ی نقاط داده، توصیف‌گرهای فوریه^۳ به ازای $0 < k \leq L - 1$ با استفاده از روابط زیر به دست آمده است:

$$a[k] = \frac{1}{L} \sum_{m=1}^L x[m] e^{-jk \left(\frac{2\pi}{L}\right) m} \quad (9-1)$$

$$b[k] = \frac{1}{L} \sum_{m=1}^L y[m] e^{-jk \left(\frac{2\pi}{L}\right) m} \quad (10-1)$$

از آن‌جا که مؤلفه‌های dc در روابط فوق (یعنی $b[0]$ و $a[0]$) فقط اطلاعاتی را درباره‌ی مرکز شیء ارائه می‌دهند، این دو مؤلفه کنار گذاشته می‌شوند. در ادامه روابط زیر را داریم:

$$r[k] = \sqrt{|a[k]|^2 + |b[k]|^2} \quad (11-1)$$

$$s[k] = \frac{r[k]}{r[1]} \quad 1 < k \leq L - 1 \quad (12-1)$$

که $|a[k]|$ و $|b[k]|$ مقدار قدر مطلق اعداد مختلف $a[k]$ و $b[k]$ را نشان می‌دهند. در این‌جا به راحتی ملاحظه می‌شود که $r[k]$ نسبت به چرخش و انتقال و $s[k]$ نسبت به تغییر مقیاس، نامتغیر هستند. توصیف‌گر نامتغیر $s[k]$ به عنوان توصیف‌گر فوریه شناخته می‌شود.

از میان توصیف‌گرهای فوریه، چند مقدار اولیه تحت عنوان قسمت‌های فرکانس پایین شناخته می‌شوند و بقیه، قسمت‌های فرکانس بالا را تشکیل می‌دهند. قسمت‌های فرکانس پایین $s[k]$ با استفاده از شکل کلی یک شیء تعیین می‌شوند. در حالی که قسمت‌های فرکانس بالا، در واقع جزئیات شکل را مشخص می‌کنند. از آن‌جا که قسمت فرکانس بالا به راحتی و به شدت تحت تأثیر نویز قرار می‌گیرد و در مقایسه با قسمت فرکانس پایین، نقش کمتری در شناسایی شیء دارد، معمولاً کنار گذاشته می‌شود. در این مقاله از ۱۶ ضریب ابتدایی $s[k]$ استفاده شده است.

علاوه بر توصیف‌گرهای فوریه، ضرایب موجک مربوط به شکل یک شیء نیز می‌توانند در شناسایی

مورد استفاده قرار گیرند. [۱۱]

¹ -Sobel Operator

² - Discrete Fourier Transform

³ - Fourier Descriptors