

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده: مهندسی برق و رباتیک

گروه: الکترونیک

ارائه روشی سریع جهت بازیابی تصاویر، مبتنی بر محتوا

دانشجو: احمد ذوالفقاری

استاد راهنما:

دکتر حسین خسروی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

ماه و سال انتشار:

بهمن ۱۳۹۲

دانشگاه صنعتی شاهرود

دانشکده : مهندسی برق

گروه : الکترونیک

پایان نامه کارشناسی ارشد آقای / خانم

تحت عنوان:

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

| امضاء | اساتید مشاور | امضاء | اساتید راهنما |
|-------|----------------------|-------|----------------------|
| | نام و نام خانوادگی : | | نام و نام خانوادگی : |
| | نام و نام خانوادگی : | | نام و نام خانوادگی : |

| امضاء | نماینده تحصیلات تکمیلی | امضاء | اساتید داور |
|-------|------------------------|-------|----------------------|
| | نام و نام خانوادگی : | | نام و نام خانوادگی : |
| | | | نام و نام خانوادگی : |
| | | | نام و نام خانوادگی : |
| | | | نام و نام خانوادگی : |

تقدیم به

پدر و مادر عزیزم

تشر و قدردانی

ابتدا خدای بزرگ را شکر گذارم که همچون تمام لحظات زندگی، در انجام این پروژه یاریم فرمود.
در اینجا لازم می‌دانم به پاس تمام زحمات، از جناب آقای دکتر حسین خسروی کمال تشکر را نمایم.

تعهد نامه

اینجانب دانشجوی دوره کارشناسی ارشد رشته دانشکده

دانشگاه صنعتی شاهرود نویسنده پایان نامه

..... تحت راهنمایی..... متعهد می شوم .

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

* متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد .

چکیده:

در این پایان نامه، روشی جدید در بازیابی تصاویر ارائه شده است که علاوه بر سریع بودن، همزمان دقت بالایی نیز داشته باشد. در سال‌های اخیر، حجم تصاویر موجود در پایگاه‌های داده تصاویر خصوصی و یا عمومی نظیر شبکه جهانی اینترنت رو به افزایش است. با توجه به این موضوع، سیستم‌هایی با قابلیت بازیابی تصاویر بر اساس خصوصیات دیداری در حیطه موضوعات مطرح شده در شاخه پردازش تصویر اهمیت بسزایی یافته است. بازیابی تصاویر کاربردهای بسیاری در زمینه‌های پزشکی، نظامی، صنعتی و محیط چندرسانه‌ای دارد. یک سیستم بازیابی تصویر مبتنی بر محتوا تصاویر را بر اساس مفاهیم بصری-شان مانند رنگ، شکل و بافت بازیابی می‌کند که با بازیابی تصویر مبتنی بر متن متفاوت است. از آنجا سرعت در سیستم‌ها رایانه‌ای تبدیل به مشخصه‌ای مهم شده است؛ علاوه بر دقت، سرعت هم در اولویت قرار گرفته است.

در روش پیشنهادی از دو ویژگی رنگ و لبه استفاده شده است. ترکیب ویژگی هیستوگرام رنگ HSV و ویژگی هیستوگرام لبه (SRF)^۱ بر پایه الگوهای باینری محلی (LBP)^۲، سیستم پیشنهادی را تشکیل می‌دهد. در کنار استفاده از دو ویژگی ذکر شده از معیار تطابق کانبرا، جهت افزایش کارایی نیز بهره گرفته شده است. نحوه بازیابی تصاویر اینگونه است که پس از دریافت تصویر پرس وجو، سیستم تصویر مورد نظر را با تمامی تصاویر موجود مورد ارزیابی قرار می‌دهد و شبیه‌ترین تصاویر را ارائه می‌دهد. جهت ارزیابی روش پیشنهادی، پایگاه داده Wang به کار گرفته شده است که شامل ۱۰۰۰ تصویر در ۱۰ کلاس معنایی می‌باشد. در نهایت با ارزیابی سیستم پیشنهادی با سیستم‌های بروز موجود، مشاهده می‌شود که سیستم ما دقت و سرعت بهتری نسبت به روش‌های موجود دارد.

کلمات کلیدی: بازیابی تصاویر، رنگ، ویژگی لبه، الگوهای باینری محلی، کانبرا

¹ Sobel Roberts feature

² Local binary pattern

لیست مقالات مستخرج از پایان نامه

۱. ذوالفقاری ا، خسروی ح، (۱۳۹۲)، "روشی سریع در بازیابی تصاویر مبتنی بر محتوا با استفاده از ترکیب ویژگی لبه و رنگ"، هشتمین همایش مهندسی برق و توسعه پایدار، ش ۳۸۱، مشهد.

فهرست مطالب

| | |
|----|--|
| ۱ | فصل اول: مقدمه |
| ۲ | ۱-۱ بازیابی تصویر |
| ۵ | ۲-۱ مشکلات سیستمهای بازیابی تصویر |
| ۵ | ۱-۲-۱ فاصله معنایی |
| ۵ | ۲-۲-۱ سلیقههای بودن ادراک انسان |
| ۶ | ۳-۱ هدف پایان نامه |
| ۶ | ۴-۱ ساختار پایان نامه |
| ۷ | فصل دوم: بازیابی تصاویر |
| ۸ | ۱-۲ تاریخچه |
| ۹ | ۲-۲ ساختار |
| ۱۱ | ۳-۲ وظایف یک نمونه از سیستمهای بازیابی تصویر مبتنی بر محتوا (CBIR) |
| ۱۲ | ۴-۲ تحقیقات کاربر |
| ۱۴ | فصل سوم: استخراج ویژگی |
| ۱۵ | ۱-۳ استخراج ویژگی |
| ۱۵ | ۲-۳ استخراج رنگ |
| ۱۶ | ۱-۲-۳ فضاهای رنگ |
| ۱۶ | ۱-۲-۳-۱ فضاهای رنگ RGB |
| ۱۷ | ۱-۲-۳-۲ فضای رنگی HSV |
| ۱۹ | ۱-۲-۳-۳ فضای رنگ $YCbCr$ |
| ۲۰ | ۱-۲-۳-۴ فضای رنگ $L^*a^*b^*$ |
| ۲۰ | ۱-۲-۳-۵ فضای رنگ CMY |
| ۲۰ | ۱-۲-۳-۶ مدل CMYK |

- ۲-۲-۳ روشهای استخراج ویژگی رنگ ۲۱
- ۲-۲-۳ هیستوگرام رنگ ۲۲
- ۲-۲-۳ کورلوگرام رنگ ۲۲
- ۲-۲-۳ هیستوگرام تلفیقی ۲۲
- ۲-۲-۳ هیستوگرام رنگ حلقوی ۲۲
- ۲-۲-۳ ممان رنگ ۲۴
- ۲-۲-۳ بردار ارتباط رنگ ۲۴
- ۲-۲-۳ همبستگی نگار رنگ ۲۴
- ۳-۳ بافت ۲۶
- ۳-۳ استخراج ویژگی بافت ۲۶
- ۱-۳-۳ ویژگیهای فیلتر گابور ۲۷
- ۱-۳-۳ تبدیل موجک ۳۰
- ۳-۱-۳-۳ DT-CWT ۳۱
- ۱-۳-۳ ویژگی های تامورا ۳۲
- ۱-۳-۳ ماتریس مجاورت ۳۲
- ۴-۳ روش های استخراج ویژگی شکل ۳۳
- ۳-۵-۱ روشهای مبتنی بر استخراج ویژگی ۳۵
- ۳-۶-۱ ارزیابی کارایی سیستم های بازیابی تصویر ۳۶
- ۳-۶-۱ آزمون هدف ۳۶
- ۳-۶-۲ نرخ خطا ۳۶
- ۳-۶-۳ راندمان بازیابی ۳۶
- ۳-۷-۱ معیار شباهت ۳۷
- ۳-۸-۱ معیارهای عدم شباهت ۳۹

| | |
|----|--------------------------------------|
| ۳۹ | ۱-۸-۳ فاصله فضای متریک |
| ۴۰ | ۲-۸-۳ فاصله مینکوسکی |
| ۴۰ | ۳-۸-۳ فاصله کسینوسی |
| ۴۰ | ۴-۸-۳ فاصله χ^2 |
| ۴۱ | ۵-۸-۳ فصل مشترک هیستوگرام |
| ۴۱ | ۶-۸-۳ فاصله ماهالانوبیس |
| ۴۲ | ۹-۳ درصد موفقیت وزنی |
| ۴۲ | ۱۰-۳ درصد رتبه بندی شباهت |
| ۴۵ | فصل چهارم : راهکار پیشنهادی |
| ۴۶ | ۱-۴ پیاده سازی برخی از روش های موجود |
| ۴۶ | ۱-۱-۴ ماتریس مجاورت |
| ۴۷ | ۲-۱-۴ ماتریس هم-رخداد رنگ |
| ۴۸ | ۳-۱-۴ تفاضل بین پیکسل های الگوی اسکن |
| ۴۹ | ۴-۱-۴ هیستوگرام جهت لبه های مشابه |
| ۴۹ | ۵-۱-۴ ارزیابی روش های موجود |
| ۵۰ | ۲-۴ معرفی راهکار پیشنهادی |
| ۵۰ | ۱-۲-۴ الگوهای باینری محلی |
| ۵۲ | ۲-۲-۴ هیستوگرام زوایای گرادیان SRF |
| ۵۳ | ۳-۲-۴ هیستوگرام رنگ HSV |
| ۵۵ | ۳-۴ ترکیب ویژگی |
| ۵۶ | ۴-۴ معیار شباهت |
| ۵۶ | ۵-۴ جست و جو و بازیابی |
| ۵۷ | ۶-۴ روش ارزیابی |

| | | |
|-------|--|----|
| ۷-۴ | روش پیشنهادی مبتنی بر ادغام آماری نواحی | ۵۸ |
| ۱-۷-۴ | ادغام آماری نواحی | ۶۱ |
| ۲-۷-۴ | بازیابی تصاویر با استفاده از ادغام آماری نواحی | ۶۱ |
| | فصل پنجم : نتایج | ۶۵ |
| ۱-۵ | پایگاه داده | ۶۶ |
| ۲-۵ | نتایج آزمایش | ۶۸ |
| ۳-۵ | مقایسه سیستم پیشنهادی با کارهای پیشین | ۶۹ |
| ۴-۵ | ارزیابی سرعت سیستم پیشنهادی | ۷۰ |
| ۵-۵ | نمونه ای از نتایج | ۷۱ |
| | منابع: | ۷۵ |

فصل اول

مقدمه

۱.۱ بازتابی تصویر

در سال‌های اخیر، رشد تکنولوژی کامپیوتر، اهمیت فوق‌العاده اطلاعات چندرسانه‌ای، وجود آرشیوهای بزرگ دیجیتال و رشد خیلی سریع شبکه گسترده جهانی، تلاش بسیاری از محققان را برای ایجاد ابزارهای مناسب بازتابی تصویر به خود جلب کرده است. امروزه مردم حجم انبوهی از داده‌های الکترونیکی را در اختیار دارند که بخش عظیمی از آن‌ها را تصاویر تشکیل می‌دهند. ذخیره این تصاویر منجر به تولید پایگاه‌های تصویری حجیمی شده که عمل جست و جو توسط کاربر را مشکل کرده است. بنابراین نیاز به الگوریتم‌های خودکار و کارآمد برای جست و جو پایگاه‌های تصویری و بازتابی تصاویر دلخواه ضروری به نظر می‌رسد.

موضوع بازتابی تصویر از اوایل دهه ۷۰ مطرح شده است. دو رویکرد متفاوت در این زمینه وجود دارد، بازتابی بر اساس متن و بازتابی بر اساس ویژگی‌های دیداری. در رویکرد اول، تصاویر به صورت دستی و با استفاده از کلمات کلیدی، توسط انسان نمایه‌سازی می‌شوند. در رویکرد دوم، برای نمایه‌سازی تصاویر ویژگی‌های دیداری مناسب به‌طور خودکار از آن‌ها استخراج می‌شوند.

روش‌های بازتابی بر اساس متن، از دو مشکل اصلی رنج می‌برند. نمایه‌سازی پایگاه‌های داده تصویری حجیم، برای انسان خسته‌کننده و وقت‌گیر است. همچنین انسان‌ها از یک تصویر خاص برداشت‌های متفاوتی دارند و برچسب‌های متفاوتی برای آن انتخاب می‌کنند.

بعضی از موارد کاربرد بازتابی تصویر عبارت است از: آرشیوهای تصویری، آرشیو موزه‌ها، طراحی مهندسی و معماری، سنجش از راه دور، مدیریت منابع طبیعی و بسیاری کاربردهای دیگر.

در اوایل دهه ۹۰، مشکلات ناشی از نمایه‌سازی متنی که به صورت دستی انجام می‌شد، با بزرگتر شدن پایگاه‌های داده تصویری نمود بیشتری پیدا کرد. برای غلبه بر این مشکلات بود که بازتابی تصویر بر اساس ویژگی‌های دیداری مطرح شد. این روش‌ها ریشه در بازشناسی الگو دارند. در حقیقت تلاش محققان بر این است که روش‌هایی ایجاد کنند که دسترسی به دنیای تصاویر ساکن و ویدیویی را هموار سازند [۱].

در بازتابی تصویر بر اساس محتوا، کاربر تصویر دلخواهش را در قالب ویژگی‌های دیداری توصیف می‌کند و سیستم بازتابی نزدیکترین تصاویر را به آنچه کاربر توصیف کرده است، بازتابی می‌کند. بنابراین بازتابی تصویر بر اساس محتوا، بازتابی بر اساس شباهت است. محتوای یک تصویر ممکن است در سطوح مختلفی

ارائه شود. در پایین ترین سطح، تصویر مجموعه‌ای از نقاط است. از محتوای تصویر در این سطح به ندرت در بازیابی استفاده می‌شود.

اگر تصویر پردازش شود و مشخصات دیداری آن استخراج شود، تصویر در سطح ویژگی ارائه می‌شود. مهمترین ویژگی‌ها برای ارائه تصویر، رنگ، شکل و بافت هستند. ویژگی‌ها بسته به نوع کاربرد ممکن است از تمام تصویر یا بخش‌هایی از آن استخراج شوند. چنانچه ویژگی از تمام تصویر استخراج شود، بیانگر خصوصیات فرامحلی و در غیر اینصورت نشان‌دهنده خصلت‌های محلی تصویر است. عموماً ارائه تصویر در سطح ویژگی نسبت به خود تصویر فضای کمتری نیاز دارد.

سطح دیگری از محتوا، سطح معنا است. ارائه تصویر در سطح معنا، کار بسیار دشواری است و یکی از زمینه‌های فعال تحقیقاتی به حساب می‌آید. در بسیاری از تحقیقات، محتوای معنایی تصاویر از محتوای سطوح پایین‌تر استنتاج می‌شود.

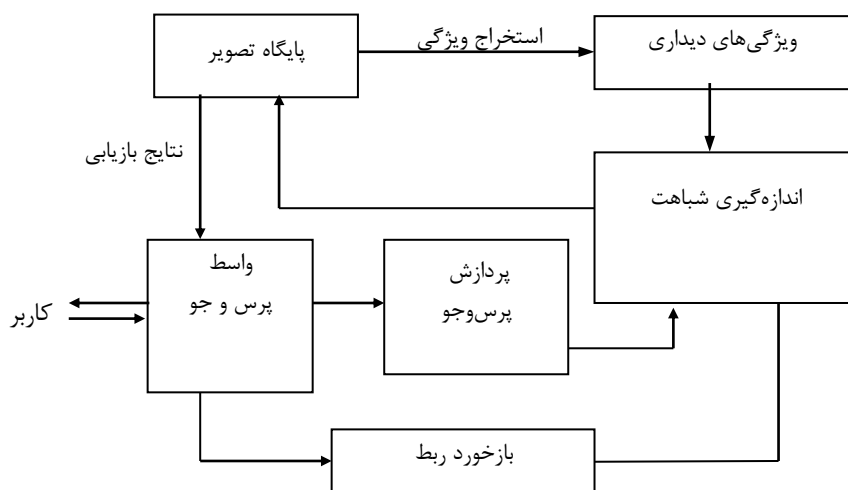
سامانه‌های بازیابی تصویر در دو مرحله عمل می‌کنند. در مرحله اول، ویژگی‌های دیداری تصاویر پایگاه به صورت خودکار استخراج شده و در مرحله دوم، پس از دریافت تصویر پرس و جوی کاربر، ویژگی‌های سطح پایین آن استخراج شده و جستجو در پایگاه ویژگی‌های دیداری برای یافتن نزدیک‌ترین تصاویر به تصویر پرس و جو انجام می‌گیرد.

در سیستم‌های بازیابی تصویر بر اساس محتوا کاربر به دنبال جستجوی معنایی تصویر است در حالیکه سامانه پایگاه داده را بر اساس ویژگی‌های سطح پایین جستجو می‌کند و تصاویری به کاربر ارائه می‌کند که دارای ویژگی‌های سطح پایین نزدیک به تصویر پرس و جوی کاربر هستند، در حالیکه ممکن است دارای معنای دلخواه کاربر نباشند. این موضوع نشان‌دهنده آن است که ویژگی‌های سطح پایین برای ارائه ویژگی‌های معنایی تصاویر کافی نیستند. از سوی دیگر سلیقه افراد مختلف با یکدیگر متفاوت است و ممکن است افراد مختلف معانی متفاوتی از یک تصویر برداشت کنند [۲]. برای رفع این محدودیت‌ها، لازم است عمل بازیابی با تعامل کاربر همراه باشد. در این تعامل سعی می‌شود پرس و جوی دلخواه کاربر مدل شود و با تنظیم ضریب ویژگی‌ها، کاربر به تصاویر دلخواهش برسد.

ساختار یک سامانه بازیابی تصویر در شکل ۱ به تصویر کشیده شده است. در این سامانه پایگاه تصویر شامل تصاویری است که کاربر برای یافتن تصویر دلخواهش آنها را جستجو می‌کند. ویژگی‌های دیداری مناسبی از تصاویر استخراج شده و تصاویر با آنها نمایه‌سازی می‌شوند [۳][۴][۵]. ویژگی‌های استخراج

شده از تصاویر در پایگاه ویژگی‌های دیداری یا سطح پایین نگهداری می‌شوند. این سامانه شامل یک رابط گرافیکی یا واسط پرس و جو است که با استفاده از آن با کاربر ارتباط برقرار کرده و با او در تعامل است.

در زیر سامانه پردازش پرس و جو، ویژگی‌های مناسب را از تصویر پرس و جو استخراج می‌کند. زیر سامانه اندازه‌گیری شباهت، شباهت بین بردار ویژگی تصویر پرس و جو و بردارهای ویژگی تصاویر پایگاه را محاسبه می‌کند. سپس سامانه نزدیکترین تصاویر به تصویر پرس و جو را پیدا می‌کند. این تصاویر از طریق رابط گرافیکی به کاربر ارائه می‌شوند و با دریافت نظر کاربر راجع به آنها، تعامل با کاربر تا رسیدن به تصاویر دلخواه او ادامه می‌یابد.



شکل ۱-۱: ساختار یک سیستم بازیابی تصویر

از این فرآیند که به بازخورد ربط معروف است به منظور یادگیری کوتاه‌مدت استفاده می‌شود.

سیستم‌های بازیابی تصویر، ممکن است از روش‌های مختلفی با کاربر ارتباط برقرار کنند که عبارتند از: پرس و جو اتفافی، ارائه مثال، ارائه طرح ساده تصویری و جست و جو بر اساس متن. در پرس و جو اتفافی، ابتدا چند تصویر از پایگاه داده انتخاب شده و به‌عنوان نمونه به کاربر نمایش داده می‌شود، سپس کاربر تصاویر مرتبط با تصویر یا تصاویر دلخواه خود را انتخاب کرده و به سیستم اعلام می‌کند. سیستم با استفاده از ویژگی‌های این تصویر یک پرس و جو فرضی ساخته و پایگاه ویژگی‌های سطح پایین را برای یافتن نزدیکترین تصاویر جستجو می‌کند. در پرس و جو با ارائه مثال، کاربر یک تصویر از نوع آنچه به

دنبال آن است به سیستم ارائه می‌کند. برای پرس و جوی از نوع ارائه طرح ساده، کافی است کاربر طرحی از آنچه می‌خواهد بکشد و آنرا در پایگاه داده جست و جو کند. پرس و جوی متنی نیز بر پایه توصیف متنی کاربر از تصویر دلخواهش استوار است.

۲.۱ مشکلات سیستم‌های بازیابی تصویر

اینکه کاربر در اولین مرحله کار با سیستم به تصاویر دلخواه خود برسد، دشوار و دور از ذهن است. علت این است که کاربر به دنبال جستجوی معنایی تصویر است یا به عبارت دیگر، کاربر تصویری با یک معنای خاص را جستجو می‌کند، در حالیکه سیستم پایگاه داده را بر اساس ویژگی‌های سطح پایین جستجو می‌کند، چه بسا تصاویری وجود دارند که دارای ویژگی‌های سطح پایین یکسان هستند، در حالیکه معانی با ویژگی‌های سطح بالای متفاوتی دارند، در اوایل توسعه سیستم‌های CBIR، تحقیقات روی روش‌های متفاوت ارائه ویژگی‌ها متمرکز شده بود. در این سیستم‌ها کاربر ویژگی‌های مورد علاقه خود را انتخاب و وزن‌های مربوط به اهمیت هر ویژگی را مشخص می‌کند. دو دلیل عمده برای محدودیت زیاد این سیستم‌ها وجود دارد.

۱.۲.۱ فاصله معنایی

در بازیابی تصاویر، جستجو در پایگاه ویژگی‌های سطح پایین برای یافتن نزدیکترین تصاویر به تصویر پرس و جو انجام می‌شود. اما به طور معمول خواسته کاربر جستجوی معنایی پایگاه است. به‌عنوان مثال ممکن است منظور کاربر، یافتن تصاویری باشد که منظره غروب خورشید را نشان می‌دهد. در چنین شرایطی، ممکن است تصاویری به کاربر ارائه شود که رنگ قرمز در آن‌ها غالب است، ولی دارای معنای غروب خورشید نیستند. در بیشتر موارد ویژگی‌های سطح پایین برای ارائه معانی تصاویر کافی نیستند. به فاصله میان ویژگی‌های معنایی سطح بالا و ویژگی‌های دیداری سطح پایین فاصله معنایی گفته می‌شود.

۲.۲.۱ سلیقه‌ای بودن ادراک انسان

برای افراد مختلف یا حتی یک فرد خاص تحت شرایط متفاوت، ممکن است محتوای دیداری یک تصویر معنای متفاوتی داشته باشد. به‌عنوان مثال در یک تصویر خاص ممکن است یکی شیفته رنگ آن و دیگری شیفته ویژگی دیداری از آن شود. از طرف دیگر در اغلب موارد وقتی که کاربر احساس می‌کند که دو تصویر مشابه‌اند، بدین معناست که این دو تصویر دارای معنای یکسانی هستند نه اینکه ویژگی‌های دیداری یکسانی دارند.

برای رفع این محدودیت‌ها، لازم است عمل بازیابی همراه با تعامل با کاربر باشد. در سیستم‌هایی که اخیراً مطرح شده‌اند، بین سیستم و انسان تا رسیدن به تصاویر دلخواه کاربر تعامل وجود دارد. در این تعامل، سعی می‌شود پرس و جوی دلخواه کاربر مدل شود و با تنظیم وزن مربوط به ویژگی‌های مختلف در ترکیب ویژگی‌ها، کاربر به تصاویر دلخواه خود برسد. عموماً، هنگامی که کاربر جستجوی یک تصویر را شروع می‌کند، به طور خیلی واضح نمی‌داند که دنبال چه می‌گردد و حتی ممکن است ذهنیت او در حین فرایند بازیابی نسبت به آنچه که دنبال می‌کند عوض شود. بنابراین استفاده از تعامل انسان و کامپیوتر در یک حلقه ضروری به نظر می‌رسد. این عمل در قالب بازخورد ارتباط انجام می‌شود.

۱.۳ هدف پایان‌نامه

هدف از تدوین این پایان‌نامه، ارائه روشی سریع در زمینه بازیابی تصویر مبتنی بر محتوا است. سیستم پیشنهادی از خاصیت رنگ و بافت برای بازیابی تصویر استفاده می‌کند.

با در نظر گرفتن دو فاکتور سرعت و دقت، اقدام به تعیین نوع ویژگی می‌کنیم. از تمام تصاویر موجود در پایگاه داده ویژگی اخذ می‌شود تا از ویژگی‌های اخذ شده جهت جستجو تصاویر استفاده شود. در نهایت قصد داریم سیستم پیشنهادی ما در عین داشتن دقت بالا از سرعت خوبی هم برخوردار باشد.

۴.۱ ساختار پایان‌نامه

ساختار پایان‌نامه به شرح زیر است: در فصل دوم مختصری از تاریخچه سیستم‌های بازیابی تصویر و وظایف آن‌ها آورده شده است. در فصل سوم به معرفی برخی از روش‌های متداول استخراج ویژگی رنگ و بافت و شکل پرداخته شده است. فصل چهارم به معرفی راه‌کار پیشنهادی این پایان‌نامه برای یک سیستم بازیابی تصویر پرداخته است. و در فصل پایانی نیز نتایج الگوریتم مورد نظر ارائه و با کارهای پیشین مقایسه می‌گردد.

فصل دوم

بازیابی تصاویر

۱.۲ تاریخچه

درسالهای اخیر با توجه به گسترش روز افزون اطلاعات در اینترنت و افزایش سریع تعداد و اندازه‌ی کلکسیون‌های تصاویر دیجیتالی نیازمند آنیم تا با استفاده از تکنیکهای هوشمند روشی برای دستیابی سریع به این اطلاعات بیابیم. مرور، جستجو و بازیابی در میان حجم وسیعی از داده‌ها، بدون سازماندهی این اطلاعات کاری غیرممکن می‌نماید. حوزه‌ی فعالیتها در زمینه‌ی بازیابی تصاویر از دهه ۱۹۷۰ با گذر از مجموعه تحقیقات وسیعی که عبارت بودند از مدیریت پایگاه داده و بینایی ماشین، بسیار سرعت یافت [۵].

دو رویکرد متفاوت در این زمینه وجود دارد: بازیابی بر اساس متن و بازیابی بر اساس ویژگی‌های دیداری، در بازیابی بر اساس متن، بدین صورت است که مفاهیم موجود در تصویر در قالب کلمات، توسط اپراتور تشخیص داده می‌شود و در پایگاه داده‌ای به‌عنوان پایگاه داده کلمات کلیدی آن تصاویر، ذخیره می‌شود. در این روش قالب پرس و جوی دریافتی از کاربر نیز به صورت کلمات کلیدی می‌باشد و بر این اساس تصاویر با استفاده از معیار شباهت میان کلمات کلیدی کاربر و کلمات منتسب به آن‌ها، بازیابی می‌شوند. به این روش، بازیابی تصاویر مبتنی بر متن (TBIR) گفته می‌شود [۶].

کلمات کلیدی، هنوز یک روش کاملاً رایج برای ارائه اطلاعات درباره‌ی پایگاه داده است، اما برای توصیف تصاویر با جزئیات و به‌طور رضایت‌بخش نیاز به سیستمی بسیار بزرگ و پیچیده خواهد بود. یکی دیگر از نقطه ضعف‌های جدی این رویکرد نیاز به پرسنل آموزش‌دیده برای حاشیه‌نویسی تصاویر می‌باشد که این امر مستلزم صرف وقت و هزینه بسیار است و بستگی مستقیم با درک افراد از تصاویر دارد. این بدان معنی است که ممکن است افراد مختلف از یک تصویر مشابه، محتوای بصری متفاوتی را درک نمایند. درونی بودن ادراک و ابهام حاشیه‌نویسی ممکن است منجر به ناسازگاری در فرآیندهای بازیابی شود. علاوه بر آن سیستم مبتنی بر کلید واژه به سختی قابل تغییر است. بنابراین رویکردهایی جدیدی برای غلبه بر این محدودیت‌ها مورد نیاز هستند [۶].

در ابتدای دهه‌ی ۹۰، با افزایش حجم تصاویر موجود در پایگاه‌های داده نظیر شبکه جهانی اینترنت و نیز برای غلبه بر مشکلات سیستم‌های مبتنی بر متن، توسعه سیستم‌های بازیابی تصویر مبتنی بر محتوا پیشنهاد شد که وظیفه استخراج تصاویر به صورت خودکار و با استفاده از مفاهیم دیداری نظیر رنگ، بافت و طرح‌بندی تصویر را بر عهده داشت [۷]، [۸]، [۹].

با دادن یک تصویر پرس و جو، سیستم بازیابی تصویر مبتنی بر محتوا تصاویر مشابه با تصویر پرس و جو را از پایگاه داده تصویر بازیابی می‌کند. تمام تصاویر پایگاه داده، برای استخراج ویژگی‌های انتخاب شده، که محتوای تصاویر را ارائه می‌دهند، پردازش می‌شوند. این عمل به طور اتوماتیک وقتی که یک تصویر وارد پایگاه داده می‌شود، انجام می‌شود. این فرآیند به هر تصویر توصیفاتی را اختصاص می‌دهد که در یک پایگاه داده ذخیره می‌شوند و بعداً در فاز تطبیق برای بازیابی تصاویر مرتبط مورد استفاده قرار می‌گیرد [۶].

ویژگی‌های تصویر پرس و جو با استفاده از همان فرآیندی که برای استخراج ویژگی‌های پایگاه داده تصویر به کار برده می‌شود، استخراج خواهد شد. سپس بازیابی تصویر توسط یک موتور تطبیق بدین صورت انجام می‌شود که ویژگی‌ها یا همان توصیفات تصویر پرس و جو را با ویژگی‌های تصاویر موجود در پایگاه داده مقایسه می‌کند. مکانیزم تطبیق، مدل بازیابی را با توجه به متریک انتخاب شده یا همان معیار شباهت پیاده‌سازی می‌کند سپس تصاویر موجود در پایگاه داده با توجه به شباهت‌شان با تصویر پرس و جو رتبه-بندی و تصاویر با بالاترین رتبه بازیابی می‌شوند [۶].

۲.۲ ساختار

سیستم‌های بازیابی تصویر از ویژگی‌های تصویری برای دسته‌بندی و بازیابی تصاویر استفاده می‌کنند. در این سیستم‌ها پایگاه داده تصویری وجود دارد که شامل کلیه تصاویر قابل بازیابی برای کاربر است. کاربر برای استخراج تصاویر از این بانک، خصوصیات مورد نظر خود را به سیستم ارائه می‌دهد و سیستم با استفاده از این اطلاعات، بانک تصاویر خود را جست و جو می‌کند تا تصاویر مطلوب را استخراج و به کاربر نمایش می‌دهد. بر همین اساس، یک سیستم بازیابی تصویر از پنج بخش اصلی تشکیل می‌شود که این بخش‌ها در شکل ۱-۲ با شماره مشخص شده‌اند.

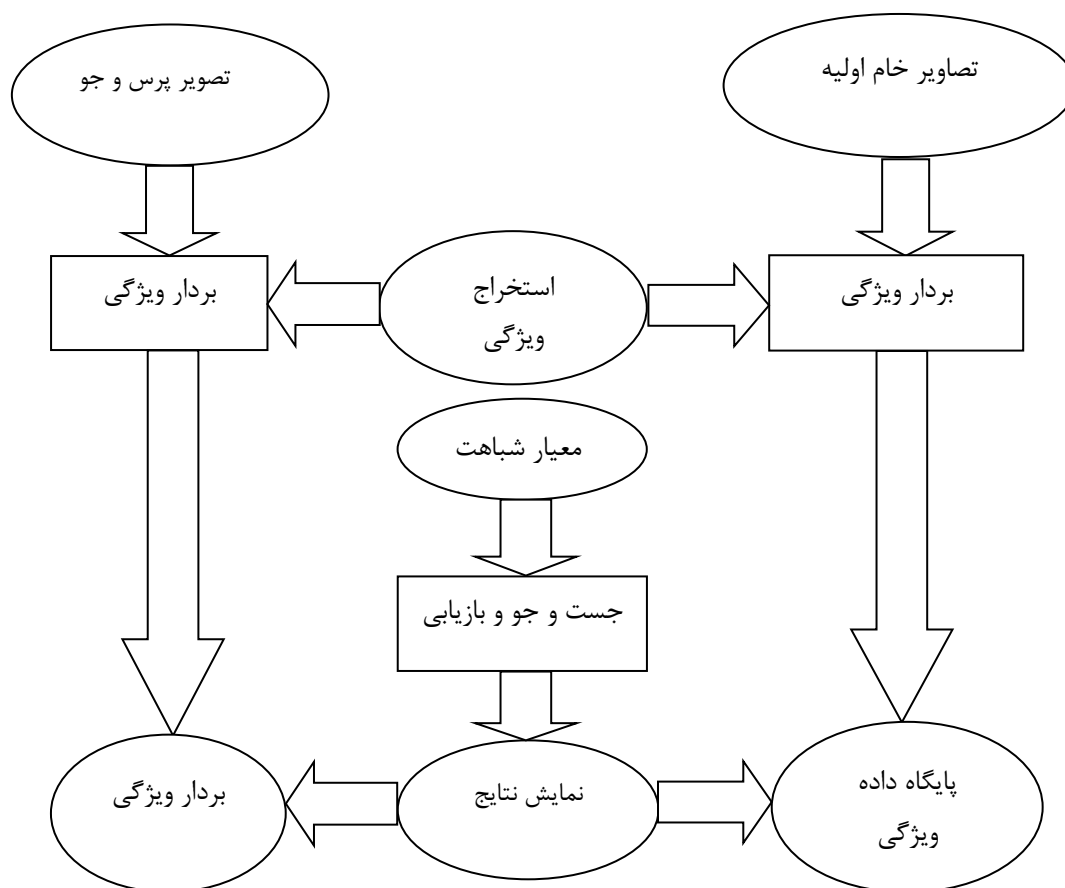
۱. پایگاه داده تصویر که شامل تصاویر خام اولیه است که قابل بازیابی هستند.

۲. واحد استخراج ویژگی که وظیفه استخراج ویژگی‌های تصاویر را بر عهده دارد.

۳. پایگاه داده ویژگی که خصوصیت‌های استخراج شده از تصاویر پایگاه داده در آن ذخیره می‌شوند.

۴. واحد جست و جو و بازیابی تصاویر پایگاه داده را جست و جو می‌کند و بر اساس میزان شباهت آن‌ها به تصویر پرس و جو، شبیه‌ترین تصاویر را بازیابی و به منظور نمایش، به واسطه کاربری می‌فرستد.

۵. واسط کاربری که تصویر پرس و جویی را از کاربر دریافت کرده و تصاویر استخراج شده را برای وی نمایش می‌دهد.



شکل ۲-۱: ساختار کلی یک سیستم بازیابی تصویر (اقتباس از [۱۰])

Gevers در [۱۱] روش‌های بازیابی تصویر را بر اساس کاربردهای مختلف در سه دسته طبقه‌بندی کرده است:

۱. جست و جو بر اساس شباهت تصویر، هدف از این روش به دست آوردن تصاویر مورد نظر کاربر از بین تصاویر موجود در یک گالری تصاویر است که در مراحل پی در پی و با استفاده از بازخورد کاربر صورت گیرد.

۲. جستجو به هدف یافتن تصویر خاص: هدف از این روش به دست آوردن تصاویر مشابه با یک تصویر مشخص از پایگاه داده تصاویر است. منظور از تصویر مشابه، تصویری است که (بخشی از آن) با تصویر مورد نظر کاربر یکسان باشد و یا (بخشی از آن) یک شی در هر دو تصویر یکسان وجود داشته باشد.

۳. جستجو بر اساس دسته‌بندی: هدف از این روش به دست آوردن تصویری است که متعلق به دسته یا کلاس بخصوصی باشد، مانند تصاویر دکوراسیون منزل و یا تصاویر رادیولوژی.

وی همچنین پایگاه داده تصاویر را بر اساس چگونگی میزان تغییر در ویژگی‌ها در دو دسته قرار داده است:

۱. دامنه محدود، شامل تغییرات محدود در ویژگی‌های دیداری تصاویر (مانند پایگاه داده تصاویر مربوط به پزشکی).

۲. دامنه وسیع، شامل تغییرات وسیع و غیر قابل پیش بینی در تصاویر (مانند پایگاه داده تصاویر موجود در اینترنت).

۲-۳- وظایف یک نمونه از سیستم‌های بازیابی تصویر مبتنی بر محتوا^۱

یک نمونه سیستم بازیابی تصویر مبتنی بر محتوا (CBIR) تنها به منابع مختلفی از اطلاعات با قالب‌های متفاوت (برای مثال، تصاویر، متن، فیلم‌های ویدئویی) نمی‌پردازد بلکه شرایط کاربران را نیز مورد توجه قرار می‌دهد. در واقع این هم محتوای منابع اطلاعاتی و هم تحقیقات کاربران را مورد بررسی قرار می‌دهد، و سپس این‌ها را برای بازیابی آیتم‌های مرتبط با آن وفق می‌دهد. وظایف عمده چنین سیستم‌هایی عبارتند از [۶]:

۱. آنالیز محتوای منابع اطلاعاتی و ارائه‌ی محتوای منابع آنالیز شده به روشی که برای همانندسازی تحقیقات کاربران مفید باشد. این مرحله معمولاً خیلی وقت‌گیر است به خاطر اینکه این، فرایندهای پیچیده‌ای را برای همه‌ی منابع اطلاعاتی در پایگاه داده‌ها دارد. با این همه، فقط یکبار انجام می‌گیرد و می‌تواند قطع شود.

¹ Content based image retrieval (CBIR)

۲. آنالیز تحقیقات کاربران و ارائه‌ی آن‌ها به صورتی که برای هماهنگی با منابع پایگاه داده‌ای مناسب باشند. بخشی از این مرحله مشابه مرحله‌ی قبلی است، اما فقط برای تصاویر تحقیقاتی به کار برده می‌شود.

۳. تعریف شیوه‌ای برای هماهنگ‌سازی سوالات مورد جست و جو با اطلاعات در پایگاه داده‌های ذخیره شده، بازیابی اطلاعات که مربوط به هم هستند در روشی موثر انجام می‌گیرد. این مرحله به طور پیوسته انجام می‌گیرد و نیازمند سرعت خیلی بالا است. تکنیک‌های شاخص‌گذاری پیشرفته می‌توانند برای شناسایی مشخصه‌های مکانی مورد استفاده قرار بگیرند تا فرایندهای هماهنگ‌سازی را سرعت بخشند.

۴. ایجاد اصلاحات لازم در سیستم که بر مبنای واکنش‌های کاربران و یا تصاویر بازیابی شده است. واضح است که بحث بالا از یک طرف به سیستم بازیابی تصویر مبتنی بر محتوا مربوط می‌شود، و منابعی از اطلاعات بصری در فرمت‌های مختلف وجود دارد و از طرف دیگر به تحقیقات کاربران مربوط می‌شود. این دو جنبه در سرتاسر مجموعه‌ی وظایف، به هم مرتبط هستند [۶].

۲-۴- تحقیقات کاربر

روش‌های بسیار زیادی وجود دارند که یکی از آن‌ها می‌تواند اعلان تحقیقات بصری باشد. یک روش تحقیقاتی خوب روشی است که برای کاربر عادی باشد و علاوه بر این اطلاعات کافی را از روی نتایج استخراجی هدفمند برای کاربر به دست آورد. روش‌های تحقیقاتی که در زیر آمده‌اند عموماً در تحقیقات در مورد بازیابی تصویر مبتنی بر محتوا مورد استفاده قرار می‌گیرند [۶].

تحقیقات از طریق نمونه‌ها (QBE): در این نوع از تحقیقات، کاربر سیستمی را در نظر می‌گیرد که تصاویر مورد تحقیقات را مطابق با تصاویر پایگاه داده‌ای مشخص می‌سازد که برای تحقیقات و مقایسه‌های دوباره مورد استفاده قرار می‌گیرند. تصاویر تحقیقاتی موردنظر می‌توانند معمولی باشند، و کاربر می‌تواند با استفاده از ابزارهای طراحی خطوط کلی آن‌ها را ترسیم کند. مزیت اصلی این طرح این است که این روشی عادی برای کارشناسان و کاربران همیشگی است تا تصاویر را از پایگاه داده‌ها جست و جو کند [۶].

تحقیقات از طریق ویژگی‌ها (QBF): در سیستم‌های QBF کاربران تحقیقاتشان را از طریق ویژگی‌های مشخص و بارزی مشخص می‌سازند که در تحقیقاتشان به آن‌ها علاقه دارند. برای مثال، یک کاربر ممکن است تحقیقاتش درباره تصاویر پایگاه داده‌ای را از طریق موضوعات مربوط به "بازیابی همه تصاویری انجام

دهد که ربع چپی آن‌ها شامل ۲۵ درصد از پیکسل‌های زرد رنگ باشد". این نوع از تحقیقات با استفاده از ابزارهای تخصصی انجام می‌گیرند. کاربران متخصص سیستم‌های بازیابی تصاویر ممکن است این نوع تحقیقات را عادی بدانند، اما کاربران معمولی چنین نظری را ندارند. QBIC نمونه‌ای از سیستم‌های بازیابی تصویر مبتنی بر محتوا است که از این نوع روش‌های تحقیقاتی استفاده می‌کند [۶].

تحقیقات مبتنی بر خصیصه‌ها: تحقیقات مبتنی بر خصیصه‌ها برای توضیحات متنی مورد استفاده قرار می‌گیرند. تلاش‌های بشر برای کاوش‌های قبلی، راهنمای اولیه برای بازیابی هستند. این نوع از مثال‌ها ایجاب می‌کنند که میزان بالایی از کاوش‌ها به‌سختی و از طریق متدهای کامل به دست آیند به خاطر اینکه تصاویر شامل مقادیر زیادی از اطلاعاتی هستند که به‌سختی و با استفاده از چندین کلمه‌ی کلیدی خلاصه می‌شوند. در حالیکه این روش‌ها معمولاً سریع‌تر هستند و راحت‌تر اجرا می‌شوند فی‌نفسه درجات بالایی از ذهنیت و ابهام در حال حاضر وجود دارد که ما قبلاً آن‌ها را ذکر کرده‌ایم [۶].

فصل سوم

استخراج و پزگی

۳-۱- استخراج ویژگی

استخراج ویژگی کلیدی‌ترین بخش از یک سیستم بازیابی تصویر بر اساس محتوا است. ویژگی‌ها به دو دسته تقسیم می‌شوند.

الف) ویژگی‌های عمومی^۱ (ب) ویژگی‌های خاص^۲

ویژگی‌های عمومی معمولاً شامل رنگ، بافت، شکل و روابط مکانی^۳ هستند. ویژگی‌های خاص، به حوزه مشخصی که در آن مورد استفاده قرار می‌گیرند، وابسته است. ویژگی‌های خاص، در یک سیستم بازیابی عمومی که برای تصاویر گوناگون بکار گرفته می‌شود، کارایی خوبی ندارند. ویژگی‌های عمومی به چهار دسته تقسیم می‌شوند که عبارتند از: رنگ، بافت، شکل و چیدمان رنگ^۴.

۳-۲- استخراج رنگ

رنگ یکی از متداولترین و تعیین‌کننده‌ترین ویژگی دیداری در زمینه بازیابی تصویر است؛ چرا که نسبت به تغییرات مربوط به اندازه، جهت، دورنمایی و اغتشاش تصویر پایدار است [۱۲].

در سیستم‌های بازیابی تصاویر مبتنی بر محتوا، بازیابی از طریق استخراج اتوماتیک ویژگی‌های اولیه تصویر چون رنگ، بافت، شکل و موقعیت مکانی صورت می‌گیرد. استفاده از ویژگی رنگ در سیستم‌های بازیابی تصاویر بر اساس محتوا کاربردهای بسیار گسترده‌ای دارد. عموماً سیستم‌های بازیابی تصویر با تلفیق جست و جوی مبتنی بر رنگ، بافت، شکل و موقعیت مکانی نتایج نهایی را بازیابی می‌کنند. عمومی‌ترین روش در بازیابی رنگ تصاویر استفاده از هیستوگرام رنگ است، که از مهمترین تکنیک‌های بازیابی بر اساس تصویر [۱۳] است. تکنیک‌های دیگری جهت افزودن اطلاعات مکانی به هیستوگرام رنگ در تحقیقات اخیر ارائه شده است مانند مجموعه‌های رنگ [۱۴]، گشتاورهای رنگ [۱۵]، بردار وابستگی رنگ [۱۶]، و همبستگی رنگارنگ [۱۷]، که برخی از آن‌ها با بهبود بردار هیستوگرام رنگ کار می‌کنند و برخی با تقطیع تصویر، اطلاعات مکانی را به ویژگی رنگ استخراج شده می‌افزایند.

1 General features

2 Domain Specific features

3 Spatial relation

4 Color layout

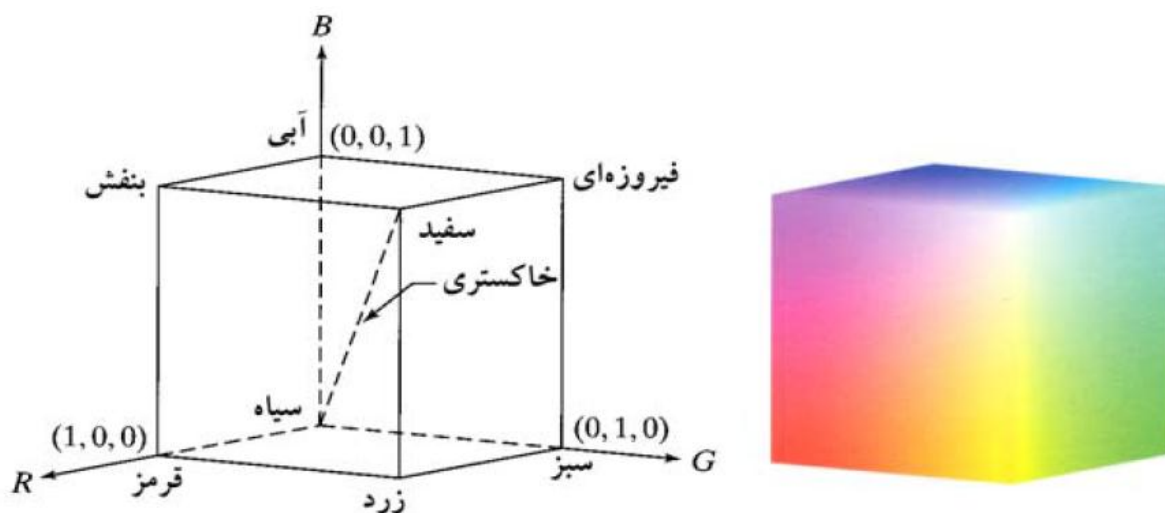
رنگ نه تنها به زیبایی موضوع می‌افزاید بلکه اطلاعات بیشتری را نیز به ما می‌دهد که آن اطلاعات به-عنوان وسیله‌ی قدرتمند در مقدار پایه‌ی زیبایی تصویر به کار می‌رود. در شاخص‌گذاری رنگ یک رنگ مشکوک؛ هدف بازیافتن همه‌ی تصاویری است که ترکیبات رنگی و بافتی آن شبیه آن رنگ موردنظر است. در زیبایی رنگ تصویر چندین روش وجود دارد.

۳-۲-۱- فضاهای رنگ

۳-۲-۱-۱- فضاهای رنگ RGB

برای استخراج ویژگی رنگ از تصویر ابتدا باید نوع فضای رنگ را تعیین نمود. فضاهای رنگی متفاوتی وجود دارند. مهمترین و ساده‌ترین فضای رنگ، RGB است. که برای توصیف رنگها استفاده می‌شود و مرتبط با منابع نور روشن کننده می‌باشد. این رنگها از قانون رنگهای افزایشی مدل RGB پیروی می‌کند. ترکیب تمام رنگها با هم، رنگ سفید را ایجاد می‌کنند [۶].

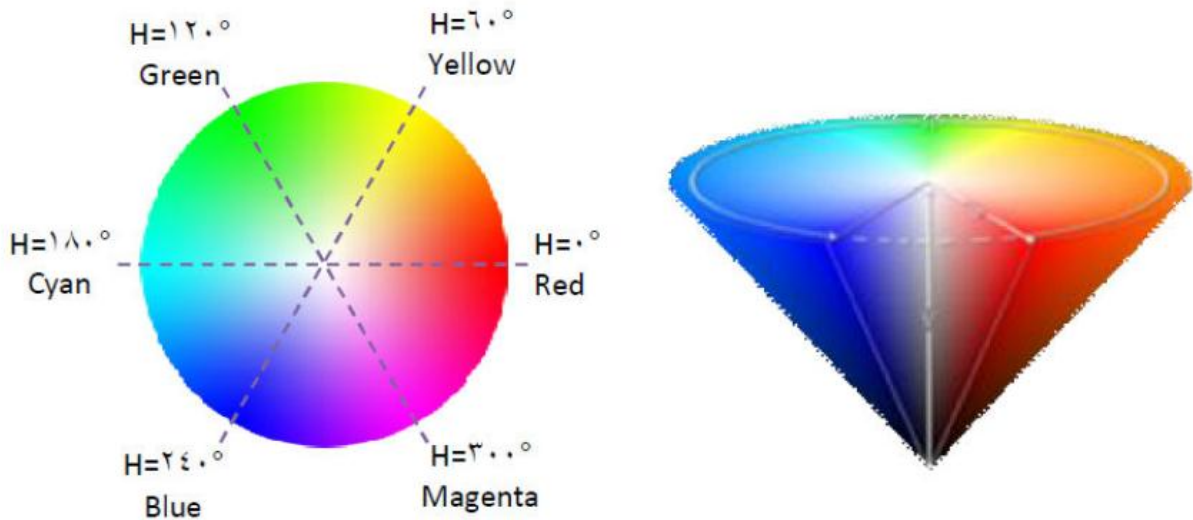
یک تصویر RGB در واقع پشته‌ای از سه تصویر در مقیاس خاکستری می‌باشد. به عبارت ساده‌تر می‌توان گفت هر تصویر RGB از سه تصویر تشکیل می‌شود که به هر یک از آنها تحت عنوان، جزء قرمز، سبز و آبی تصویر ارجاع می‌شود. فضای رنگ RGB به دلیل سادگی به طور گسترده مورد استفاده قرار می‌گیرد اما دارای یکنواختی ادراکی نیست در این فضا تفاوت دو نقطه با توجه به مختصاتشان، متناسب با درک انسان نیست، و فاصله اقلیدسی، تفاوت بین دو رنگ را به شکلی که برای انسان معنی‌دار باشد، بیان نمی‌کند. فضای رنگی RGB با مکعب رنگی RGB نمایش داده می‌شود [۶].



شکل ۳-۱: فضای رنگ RGB به صورت گسترده

۳-۲-۱-۲-۳ فضای رنگی HSV

فضای HSV دارای خاصیت یکنواختی ادراکی است، به همین دلیل از آن در سیستم‌های بازیابی تصویر به طور معمول استفاده می‌شود. ابعاد تشکیل‌دهنده این فضا، رنگ را بر اساس نوع رنگ (H) یا طول موج، درجه اشباع رنگ (S) و درجه روشنایی آن (V) تعریف می‌کنند. فضای مذکور یک فضای مخروطی است.



شکل ۳-۲: فضای رنگ HSV

هر برش دایره ای از ۰ تا ۳۶۰ درجه علامتگذاری می‌شود که هر درجه معرف یک طول موج مشخص رنگ است. از این فضای رنگ در تحقیقات مختلفی استفاده شده است [۱۸-۲۴]. مساله قابل توجه در مورد این فضا این است که، در کنار مزایای استفاده از آن مشکل عمده‌ای وجود دارد و آن محاسبات دشوار در این فضا است. یکی از راه‌های کاهش این محاسبات تبدیل مقادیر از این فضا به RGB و بالعکس است. برای تشکیل هیستوگرام رنگ HSV، تصویر از فضای رنگ RGB به این فضا منتقل می‌شود. سپس این فضا به طور خطی کوانتیزه می‌شود.

اگر فرض کنیم که $r, g, b \in [0,1]$ به صورتی که به ترتیب معرف محورهای قرمز، سبز و آبی فضای RGB باشد و H, S, V معرف مولفه های فضای به صورتی که $b \in [0,360]$ ، $s \in [0,1]$ و $v \in [0,1]$ باشد و فرض کنیم که \max معرف بیشینه مقدار r, g, b و \min معرف کمینه آن ها باشد آن گاه برای تبدیل از RGB به HSV از روابط (۳-۱)، (۳-۲) و (۳-۳) استفاده می‌کنیم:

$$h = \begin{cases} 0 & , if \max = \min \\ \left(60^\circ \frac{g-b}{\max - \min} + 0^\circ\right) \bmod 360^\circ & , if \max = r \\ 60^\circ \frac{b-r}{\max - \min} + 120^\circ, & , if \max = g \\ 60^\circ \frac{r-g}{\max - \min} + 240^\circ, & , if \max = b \end{cases}$$

(۱-۳)

$$s = \begin{cases} 0, & if \max = 0 \\ 1 - \min / \max & otherwise \end{cases}$$

(۲-۳)

$$v = \max$$

(۳-۳)

برای تبدیل از **HSV** به **RGB** از متغیرهای کمکی استفاده می‌کنیم که در روابط (۴-۳) آمده‌اند:

$$f = \frac{h}{60} - \left\lfloor \frac{h}{60} \right\rfloor$$

(۴-۳)

$$p = v_0(1 - s)$$

$$q = v_0(1 - f_0s)$$

$$t = v_0(1 - (1 - f)_0s)$$

به این ترتیب بردار (r, g, b) از رابطه (۵-۳) بدست می‌آید:

$$(v, t, p), \quad if \ h_i = 0$$

(۵-۳)

$$(q, v, p), \quad if \ h_i = ۱$$

$$(r, g, b) = (p, v, t), \quad if \ h_i = ۲$$


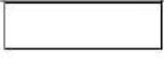






$$(p, q, v), \quad if \ h_i = ۳$$

$$(t, p, v), \quad if \ h_i = 4$$

$$(v, p, q), \quad if \ h_i = ۵$$

در جدول ۱-۳ اطلاعات مربوط به برخی رنگها با نام تجاری و معادل RGB و HSV آن ها آمده است.

جدول ۱-۳: اطلاعات مربوط به برخی رنگها با نام خاص و معادل RGB و HSV آن ها

| رنگ | نام رنگ | RGB | HSV |
|--|---------|-----------------|--------------|
|  | Black | (۰, ۰, ۰) | (۰°, ۰, ۰) |
|  | White | (۲۵۵, ۲۵۵, ۲۵۵) | (۰°, ۰, ۱) |
|  | Red | (۲۵۵, ۰, ۰) | (۰°, ۱, ۱) |
|  | Yellow | (۲۵۵, ۲۵۵, ۰) | (۶۰°, ۱, ۱) |
|  | Green | (۰, ۲۵۵, ۰) | (۱۲۰°, ۱, ۱) |
|  | Cyan | (۰, ۲۵۵, ۲۵۵) | (۱۸۰°, ۱, ۱) |
|  | Blue | (۰, ۰, ۲۵۵) | (۲۴۰°, ۱, ۱) |
|  | Magenta | (۲۵۵, ۰, ۲۵۵) | (۳۰۰°, ۱, ۱) |

۳-۱-۲-۳- فضای رنگ YCbCr

از فضای رنگی YCbCr به صورت گسترده‌ای در تصاویر ویدیویی استفاده می‌شود. در این فرمت اطلاعات از فضای رنگی **luminance** در قالب Y و اطلاعات رنگی در قالب دو جزء C_b و C_r ذخیره می‌شوند. C_b مبین تفاوت بین جزء آبی با مقدار مرجع و C_r مبین تفاوت بین جزء قرمز با مقدار مرجع می‌باشد. رابطه تبدیل فضای رنگی RGB به صورت زیر است:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (۶-۳)$$

۳-۲-۱-۴- فضای رنگ $L^*a^*b^*$

فضای رنگ $L^*a^*b^*$ نیز دارای خاصیت یکنواختی ادراکی است، مولفه‌های این سیستم تشکیل شده از L^* که معرف روشنایی است و بین سفید تا سیاه متغیر است، a^* که معرف قرمز-سبز است و b^* که معادل سبز-آبی است. a^* و b^* مستقل از روشنایی‌اند و بنابراین دارای خواص فضاهای نرمال‌سازی شده هستند. تنها مشکل این فضا این است که تبدیلات سایر فضاهای رنگی به این فضا، غیرخطی و مستلزم محاسبات بالاست.

۳-۲-۱-۵- فضای رنگ CMY

رنگ‌های cyan (آبی مایل به سبز)، Magenta (ارغوانی) و زرد تحت عنوان رنگ‌های ثانویه شناخته می‌شوند که به عنوان جایگزینی برای رنگ‌های اصلی (قرمز، سبز، آبی) مورد استفاده قرار می‌گیرد.

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (۷-۳)$$

البته در این رابطه فرض شده است که کلیه مقادیر در بازه $[0,1]$ نرمال شده باشند.

۳-۲-۱-۶- مدل CMYK

در حالی که مدل RGB برای توصیف رنگ‌های افزایشی به کار می‌رود. مدل CMYK نیز برای توصیف رنگ‌های کاهش‌ی استفاده می‌شود. همانطور که قبلاً توضیح داده شد، این رنگ‌ها به یک منبع نور برای تولید رنگ احتیاج دارند مدل CMYK بر اساس کیفیت جذب جوهر چاپ شده روی کاغذ، پایه ریزی شد. هنگامی که نور سفید به جوهر مات برخورد می‌کند، قسمتی از طیف جذب می‌شود. رنگ‌هایی که جذب نمی‌شوند در چشم منعکس می‌شوند.

در تئوری، فیروزه‌ای و ارغوانی خالص و رنگ‌دانه‌های زرد باید ترکیب شوند تا تمام رنگ را جذب کرده و رنگ سیاه را بوجود آورند. از آنجایی که تمام جوهرهای چاپی ناخالصی دارند، ترکیب این سه رنگ با یکدیگر رنگ قهوه‌ای تیره بوجود می‌آورد که باید با جوهر سیاه مخلوط شود تا رنگ سیاه واقعی را بوجود آورد. CMYK برای پروسه چاپ چهار رنگ بکار می‌رود. در این پروسه، رنگ با ترکیب ۴ جوهر: فیروزه ای (C)، ارغوانی (M)، زرد (Y) و سیاه (K) دوباره تولید می‌شود. حرف K برای جلوگیری از ابهام بکار رفته است زیرا B ممکن است به آبی تعبیر شود.

رنگهای اصلی و فرعی، رنگهای تکمیلی هستند. هر دو رنگ فرعی قادر به ساختن یک رنگ اصلی هستند. مشابه آن ترکیب هر دو رنگ اصلی یک رنگ فرعی را می‌سازد. به عنوان مثال، ترکیب زرد و ارغوانی، قرمز را می‌سازد و قرمز و سبز، زرد را بوجود می‌آورند. این رابطه تکمیلی معمولاً با یک چرخه رنگ توضیح داده می‌شود.

۳-۲-۲- روش‌های استخراج ویژگی رنگ

۳-۲-۲-۱- هیستوگرام رنگ

ویژگی رنگ یکی از ویژگی‌های مبنا در تصویر و قابل درک‌تر و بدیهی در تصویر می‌باشد و استخراج رنگ ساده‌تر می‌باشد. تعداد زیادی از روش‌ها می‌توانند با استفاده از رنگ توصیف شوند. که یکی از این روش‌ها هیستوگرام رنگ است [۲۵].

نمونه واقعی رنگ بوسیله نمودارهای آماری رنگ انجام می‌شود. در ۱۹۹۱ سواین و بالارد یک روش را پیشنهاد دادند و آن را شاخص‌گذاری نامیدند که اشیاء به کار رفته را در نمودار شاخص‌گذاری رنگ نمایش می‌داد. نمودارهای رنگی راهی برای بیان و نمایش توسعه و پخش رنگها در تصاویر هستند که هر جعبه نمودار آماری یک رنگ را در فضای رنگی مناسب نمایش می‌دهد [۲۶].

فاصله‌ی بین نمودار آماری رنگ مورد نظر و نمودار اطلاعات تصویر می‌تواند برای شرح و تطبیق شباهت بین دو انتشار مورد استفاده قرار بگیرد. برای غلبه بر مشکل بوسیله‌ی نمودار آماری در سال ۱۹۹۵ مهتری و همکاران روش‌های دو رنگ نظیر هم را به نام‌های فاصله‌ی روش و منبع لیست روش‌های رنگ را برای بازیابی تصویر پیشنهاد داد.

به‌هنگام استخراج اطلاعات، ستون‌های ارائه شده برای هر یک از مولفه‌های RGB هیستوگرام موردنظر را به وجود می‌آورند. هیستوگرام شامل ۴۸ ستون می‌باشند که هر ستون گستره‌ی کوچکی از مقادیر پیکسل‌ها را تعریف می‌کند. مقادیر ذخیره شده در هر ستون تعداد پیکسل‌ها در تصاویری هستند که در این محدوده وجود دارند. این محدوده‌ها سطوح مختلفی از هر مولفه‌ی RGB را نشان می‌دهند. مقادیر در هر ستون با تقسیم کل پیکسل‌ها در تصاویر بدست می‌آیند [۲۷].

مشکل این روش آن است که ممکن است تصاویری با ظاهری کاملاً متفاوت دارای هیستوگرام رنگ کاملاً مشابه باشند که با افزایش تعداد تصاویر پایگاه داده امکان چنین نقصی بیشتر خواهد شد. برای رفع معایب هیستوگرام رنگ، روش بردار مطرح شد. در این روش سعی شد که محدودیت‌های اضافی که

نشانگر خواص مکانی پیکسل‌ها است بر هیستوگرام متعارف اعمال شود. در اینجا پیکسل‌های موجود در هر کدام از بازه‌های متعلق به هیستوگرام، با توجه به مشخصات محلی خود به دو دسته "یکپارچه" و "غیریکپارچه" تقسیم می‌شوند. یک پیکسل یکپارچه جزئی از یک ناحیه پیوسته و قابل تشخیص است در حالیکه یک پیکسل غیر یکپارچه چنین خاصیتی را ندارد. اشکال این روش آن است که بردارهای ویژگی تصاویر به‌کندی استخراج می‌شوند.

۳-۲-۲-۲ - کورلوگرام رنگ

هانگ و کولای جس (در سال ۱۹۹۷) از کورلوگرام‌های رنگی‌ای استفاده کردند که فضاهای رنگی مرتبط با هم را مورد بررسی قرار می‌دهد [۲۸]. کورلوگرام رنگ نه تنها توزیع کردن رنگ را توصیف می‌کند بلکه همبستگی فضایی در بخش‌هایی از رنگ را توصیف می‌کند. کورلوگرام رنگی از یک تصویر، جدول شاخص-گذاری شده‌ای به‌وسیله‌ی جفت رنگ‌ها است، که در آن K امین ورودی‌ای برای (i, j) است که احتمالاً یک پیکسل رنگی j را در فاصله‌ی k از پیکسل i در تصویر مشخص می‌سازد. به‌خاطر پیچیده بودن این روش، یک autocorrelogram به‌جای گروه‌های متراکم مرتبط با رنگ‌های مشابه مورد استفاده قرار می‌گیرند [۲۹].

۳-۲-۲-۳ - هیستوگرام تلفیقی

هیستوگرام تلفیقی که به هیستوگرام بهبود یافته نیز مشهور است. در تکنیک هیستوگرام تلفیقی برای هر پیکسل علاوه بر اطلاعات رنگ، ویژگی محلی نیز در نظر گرفته می‌شود، تا نقص هیستوگرام را از بین ببرد.

۳-۲-۲-۴ - هیستوگرام رنگ حلقوی

در [۲۵]، خواص آماری و مکانی رنگ در تصویر با یکدیگر تلفیق شده و هیستوگرام مکانی رنگ ایجاد شده است. در این تحقیق، ابتدا فضای رنگ به‌صورت خطی کوانتیزه می‌شود. سپس برای هر یک از M رنگ موجود، یک زیر-تصویر باینری ایجاد می‌شود و مرکز ثقل هر یک از زیر-تصویرهای باینری محاسبه می‌شود. در هر زیر-تصویر، دورترین نقطه تا مرکز ثقل پیدا شده و فاصله آن تا مرکز ثقل در نظر گرفته می‌شود.

شود. این فاصله به N قسمت مساوی تقسیم می‌شود سپس از مرکز ثقل، N دایره رسم می‌شود تعداد پیکسل‌های بین هر دو دایره شمرده شده و یک بردار N بعدی برای هر رنگ ایجاد می‌شود. بنابراین برای هر تصویر یک ماتریس ویژگی $M \times N$ ایجاد می‌شود این ماتریس هیستوگرام رنگ حلقوی نامیده می‌شود.

۳-۲-۲-۵- ممان رنگ^۱

این روش یکی از موفقترین روش‌های استخراج رنگ در سیستم‌های بازیابی است. ممان‌های مرتبه اول، دوم و سوم به ترتیب به نام‌های میانگین^۲، واریانس^۳ و درجه عدم تقارن^۴ [۹]، توزیع رنگ تصویر را به صورت کارآمد و بهینه ارائه می‌دهند. ممان‌های اول تا سوم تصویر بر اساس روابط (۳-۸) تا (۳-۱۰) به دست می‌آیند:

$$\mu_i = \frac{1}{N} \sum_{i=j}^N f_{ij} \quad (۳-۸)$$

$$\delta_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right)^{1/2} \quad (۳-۹)$$

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{1/3} \quad (۳-۱۰)$$

که f_{ij} مقدار آمین مولفه رنگ آمین پیکسل و n تعداد پیکسل‌های تصویر است. ممان‌های رنگ غالباً در سیستم‌های HSV و $L^*a^*b^*$ ارجحیت دارد. استفاده از ممان سوم رنگ در کنار دو ممان اول، بازده کلی

1 Color Moment

2 Mean

3 Variance

4 Skewness

بازیابی را در برخی تصاویر افزایش می‌دهد، اما این ممان به تغییر زاویه دید حساس است و ممکن است خود دلیلی برای کاهش بازده شود.

۳-۲-۲-۶- بردار ارتباط رنگ^۱

این ویژگی اطلاعات مکانی پیکسل‌ها را در هیستوگرام درج می‌کند. در این روش که به طور مخفف به آن CCV می‌گویند، هر مولفه هیستوگرام از دو نوع داده تشکیل می‌شود، بخش مرتبط که شامل تعداد پیکسل‌هایی است که به یک منطقه یکنواخت رنگ متعلق‌اند، اگر α_i معرف پیکسل‌های مرتبط در مولفه i هیستوگرام (مربوط به رنگ i ام) β_i تعداد پیکسل‌های نامرتبط آن مولفه باشد، در آن صورت CCV مربوط به آن تصویر با بردار

$$\langle (\alpha_1, \beta_1) \text{ و } (\alpha_2, \beta_2) \text{ و } \dots \text{ و } (\alpha_N, \beta_N) \rangle$$

معرفی می‌شود. قابل توجه است که

$$\langle (\alpha_1 + \beta_1) \text{ و } (\alpha_2 + \beta_2) \text{ و } \dots \text{ و } (\alpha_N + \beta_N) \rangle$$

همان هیستوگرام تصویر است.

بر اساس نتایج گذشته، CCV بازده بهتری از هیستوگرام رنگ دارد. برای هر دو هیستوگرام و بردار ارتباط رنگ، سیستم HSV بهتر از $L^*a^*b^*$ عمل می‌کند [۲۷][۳۰].

۳-۲-۲-۷- همبستگی نگار رنگ

این ویژگی نه تنها معرف توزیع رنگ پیکسل‌ها است، بلکه همبستگی مکانی هر جفت پیکسل را نیز شامل می‌شود. بعد اول و دوم از سه بعد هیستوگرام رنگ هر جفت پیکسل و بعد سوم معرف فاصله مکانی آنها است. بردار همبستگی نگار رنگ جدولی است که توسط جفت‌های رنگ اندیس‌گذاری شده است، به گونه‌ای که k امین مولفه (i, j) برابر احتمال یافتن یک پیکسل از رنگ j ام در فاصله k از یک پیکسل از رنگ i در تصویر است.

¹ Color Coherence Vector

این بردار در مقابل هیستوگرام و CCV بازده بهتری برای سیستم بازیابی فراهم می‌کند، اما به علت بعد زیاد بار محاسباتی بالایی دارد [۲۷].

۳-۳- بافت

بافت به‌طور گسترده‌ای در پردازش تصویر مورد استفاده قرار می‌گیرد، اما تعریف دقیقی برای انواع گونه‌های جامع ندارد. یکی از تعریف موجود نشان می‌دهد که "یک ناحیه از عکس ساختار ثابتی دارد اگر مجموعه ویژگی‌های محلی این ناحیه ثابت باشد، به‌کندی تغییر کند، یا تقریباً متفاوت باشد".

روش‌های زیادی برای توصیف ساختار وجود دارد: روش‌های آماری اغلب بیشتر فضاها، چارچوب‌های موجود، اکثر لبه‌ها و اندازه‌های قبلی را مورد استفاده قرار می‌دهند. از روی این‌ها بسیاری از مشخصه‌های ساده همانند انرژی، آنتروپی، هموزنتیک، زبری، تفاوت‌ها، بستگی‌ها، گروه‌های متراکم، مراحل مختلف، جهت‌ها، چارچوب‌ها، خطوط، تکرارها به‌دست می‌آیند. این روش‌های توصیف ساختار، ویژگی‌های مختلف بافت‌ها را مورد محاسبه قرار می‌دهند و مناسب برای اندازه‌های قبلی‌ای هستند که با اندازه‌ی پیکسل‌ها قابل مقایسه باشند. روش‌های نحوی و دوگانه روش‌های همانندی دستور زبان‌های زنجیره وار و گرافی هستند که برای ساختارهایی مناسب می‌باشند که می‌توانند به سادگی مشخص شوند و ویژگی‌های‌شان توصیف شود. مقالات بسیاری وجود دارند که این بخش را مورد بررسی قرار می‌دهند. ما به خوانندگان علاقه‌مند توصیه می‌کنیم که برای یافتن اطلاعات بیشتر به آن‌ها مراجعه کنند.

۳-۳-۱- استخراج ویژگی بافت

همانطور که اشاره شد، بافت یکی دیگر از ویژگی‌های مهم تصویر است و از منظر دیداری الگوی یکنواختی است که از حضور بیش از یک رنگ با درجه خاکستری بدست می‌آید. روش‌های استخراج بافت به چهار دسته تقسیم می‌شوند [۱۲].

۱. روش‌های ساختاری

۲. روش‌های آماری^۱

۳. روش‌های تبدیلی^۱

۴. روش‌های مبتنی بر مدل^۲

روش‌های ساختاری شامل اپراتور مورفولوژیک^۳ و گراف همجواری^۴ است که بافت را بر اساس ویژگی‌های ساختاری و قواعد مربوط به نحوه قرار گرفتن پیکسل‌ها و اصطلاحاً بر اساس ریز بافت‌ها تعریف می‌کنند. این روش‌ها برای بافت‌های معمولی مناسب هستند و بیشتر برای ترکیب بافت استفاده می‌شوند تا تحلیل آن، و با وجود اینکه تعریف دقیقی از بافت ارائه می‌دهند، از بازده مناسبی برخوردار نیستند و محاسبات بالایی می‌طلبند. امروزه از روش‌های ساختاری در موارد خاصی استفاده می‌شود.

روش‌های آماری بر اساس هیستوگرام و یا محاسبات مرتبه دوم آماری بر روی درجه خاکستری پیکسل‌های مجاور اعمال می‌شوند و از جمله اولین روش‌های استخراج بافت هستند. L_I در [۳۱] روشی از نوع آماری ارائه کرده که برای استخراج بافت، از هیستوگرام تصویر استفاده شده است. در سایر روش‌های آماری غالباً از ماتریس هم‌رخداد برای استخراج ویژگی‌های بافت استفاده می‌شود.

روش‌های مبتنی بر مدل، به مدل‌سازی بافت می‌پردازند و شامل روش خودبازگشت یا AR، مدل گوسی مارکوف یا RMF، و مدل Gibbs RMF می‌شوند. روش‌های تبدیلی شامل ویژگی‌های تامورا، تبدیل موجک، و غیره است که مبتنی بر اعمال فیلتر بر روی تصویر می‌باشند.

۳-۱-۱-۳-۱- ویژگی‌های فیلتر گابور^۵

ویژگی‌های گابور به علت خاصیت تفکیک‌پذیری خوب و چندگانه در حوزه‌های مکان و فرکانس به‌عنوان ابزاری قوی در تحلیل بافت استفاده می‌شوند. فیلتر گابور می‌تواند برای مدل کردن واکنش سیستم‌های بصری انسان استفاده شود.

1 Model-based

2 Transform

4 Morphological Operator

4 Adjacency Graph

5 Gabor feature

در حوزه مکان، هسته‌های فیلتر گابور از حاصلضرب یک تابع گوسین^۱ با یک تابع سینوسی جهت‌دار به- دست می‌آید. در نتیجه این فیلتر در نقاطی از تصویر که به صورت محلی دارای جهت و فرکانس مکانی معینی هستند پاسخ‌های قوی تولید می‌کند [۳۰]. برای اعمال این فیلتر بهتر است از فضای HSV استفاده شود [۳۲].

خصوصیات بافت موجود در CBIR عمدتاً از موجک‌های گابور، تبدیل‌های موجک گسسته مرسوم (DWT) و فریم‌های موجک گسسته استخراج می‌شوند. نتایج آزمایشات وسیع روی یک مجموعه‌ی بزرگ از تصاویر بافتی نشان داد که اجرای بازیابی با استفاده از فیلترهای گابور، بهتر از استفاده از خصوصیات بر پایه موجک‌های عمودی رایج می‌باشد [۳۳].

خصوصیات بر پایه موجک گابور، نتایج بازیابی را بهتری می‌کند، اما از دو عیب عمده رنج می‌برد:

۱. یک تابع بر پایه گابور متعامد نمی‌باشد، این امر، نیازمندی به حافظه و افزونگی را افزایش می‌دهد.
۲. زمان مورد نیاز برای استخراج خصوصیات بسیار زیاد است که این باعث کاهش سرعت بازیابی می- شود [۳۰].

از طرف دیگر، چون موجکها، هسته تکنولوژی در تولید استانداردهای آینده‌ی کدینگ تصویر، JPEG-2000، هستند انتخاب خصوصیات موجک به‌عنوان خصوصیات تصویر، سیستم‌های CBIR را قادر می‌سازد که بتوانند مستقیماً در حوزه فشرده سازی اجرا شوند.

الف) تابع، یک فیلتر دو بعدی متقارن زوج و تبدیل فوریه آن مطابق روابط (۳-۱۱) و (۳-۱۲) است:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right]\right] \cos(2\pi u_0 x) \quad (3-11)$$

$$F(u, v) = \frac{1}{2} \left[\exp\left[-\frac{1}{2}\left[\frac{(u-u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right]\right] + \exp\left[-\frac{1}{2}\left[\frac{(u+u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right]\right] \right] \quad (3-12)$$

که در آن u_0 ، فرکانس مرکزی موج سینوسی را در راستای محور X ، چنانچه جهت صفر درجه انتخاب شده باشد، مشخص می‌کند. σ_x و σ_y به ترتیب انحراف معیار تابع پوش گوسی در راستای محور

1 Gaussian function

u, v هستند. مجموعه‌ای از فیلترهای گابور از طریق گسترش و چرخش تابع $f(x, y)$ که فیلتر مادر نامیده می‌شود، با تابع مولد رابطه (۳-۱۳)، ایجاد می‌شوند.

$$f_{mn}(x, y) = a^{-m} f(x', y') \quad , a \geq 1 \quad (۳ - ۱۳)$$

$$x' = a^{-m}(x \cos \theta + y \sin \theta), \quad y' = a^{-m}(-x \sin \theta + y \cos \theta)$$

$$\theta = n\pi/k, n = 0, 1, \dots, k-1, \quad m = 0, 1, s-1$$

که در آن f_{mn} مقیاس شده و چرخش یافته‌ای از فیلتر مادر $f(x, y)$ ، ضریب مقیاس، n اشاره مربوط به-جهت فعلی، k تعداد کل جهتها، m اشاره گر مربوط به مقیاس فعلی، s تعداد کل مقیاس‌ها و x' و y' مختصات مقیاس شده و چرخش یافته هستند. ضریب a^{-m} برای مستقل کردن انرژی فیلتر از مقیاس m است. طراحی این فیلتر به گونه‌ای است که محل تقاطع آن‌ها در صفحه نصف توان آن‌ها باشد. به این-ترتیب اطمینان حاصل می‌شود که خروجی فیلترها دارای حداکثر اطلاعات مفید و حداقل اطلاعات زائد هستند. چنانچه u_h و u_l به ترتیب برای نمایش فرکانس مرکزی حداکثر و حداقل بکار روند، باید پارامترها با یک فیلتر با توجه به رابطه (۳-۱۴) محاسبه شوند.

$$a = \left[\frac{u_h}{u_l} \right]^{\frac{1}{s-1}} \quad \sigma_u = \frac{(a-1)u_k}{(a+1)\sqrt{2\ln 2}} \quad \sigma_v = \frac{\tan \left[\frac{\pi}{2k} \right] \left[u_k - \left[\frac{\sigma_u^2}{u_k} \right] (2 \ln 2) \right]}{\sqrt{2\ln 2 - (2 \ln 2)^2 \left[\frac{\sigma_u^2}{u_k^2} \right]}} \quad (۳ - ۱۴)$$

بنابر پیشنهاد مرجع [۳۲]، u_h برابر $128(\sqrt{2})$ سیکل بر پهناي تصوير و u_l برابر $16(\sqrt{2})$ سیکل بر پهناي تصوير انتخاب شده‌اند. بنابراین مقدار a برابر ۲ خواهد بود برای تصویر $I(x, y)$ ، تبدیل موجک گابور با استفاده از رابطه‌ی (۳-۱۵) محاسبه می‌شود.

$$I_{mn}(x, y) = \iint I(x_1, y_1) f^*(x - x_1, y - y_1) dx_1 dy_1 \quad (۳ - ۱۵)$$

از آنجایی که به دست آوردن زیر تصویرهای I_{mn} از طریق رابطه (۳-۱۳) وقت گیر است. این زیر تصویرها با استفاده از رابطه (۳-۱۶) در فضای تبدیل فوریه محاسبه می‌شوند.

$$I_{mn}(x, y) = f^{-1}[I(u, v)F_{mn}(u, v)] \quad (۳ - ۱۶)$$

که در آن $I(u,v)$ تبدیل فوریه تصویر $I(x,y)$ و $F_{mn}(u,v)$ تبدیل فوریه $f_{mn}(x,y)$ است. برای حذف حساسیت فیلترها به میانگین شدت روشنایی تصویر، $F_{mn}(0,0)$ برابر صفر در نظر گرفته شده است.

۳-۱-۲- تبدیل موجک

تبدیل موجک گسسته یک بعدی یا DWT_1 ، یک سیگنال گسسته $f(x)$ را بر حسب مقادیر تابع مقیاس $\Phi(x)$ و تابع شیفت یافته و بسط یافته $\Psi(x)$ تجزیه می کند. رابطه (۳-۱۵) معرف این تبدیل است. که در آن $\varphi_{j_0,i}$ و $\Psi_{j_0,i}$ و به ترتیب توسط روابط (۳-۱۷) و (۳-۱۸) به دست می آیند [۱۲] [۳۰].

$$f(x) = \sum_{L \in \mathbb{Z}} S_{j_0,L} \varphi_{j_0,L}(x) + \sum_{j \geq j_0, L} C_{j,L} \psi_{j,L}(x)$$

$$\varphi_{j,i}(x) = 2^{\frac{j_0}{2}} \varphi(2^{j_0}x - 1) \quad (3-17)$$

$$\psi_{j_0,i}(x) = 2^{\frac{j}{2}} (2^j x - 1) \quad (3-18)$$

تابع $\Psi(x)$ موجک مادر یا به اختصار موجک نامیده می شود. تبدیل موجک یک تصویر (سیگنال دو بعدی)، با پیاده سازی تبدیل موجک یک بعدی در جهت های افقی و عمودی بدست می آید. دلیل کارایی موجک در آنالیز سیستم ها مسائلی از قبیل محمل فشرده و کوتاه، تعامد، تقارن، گشتاور های صفر و غیره است؛ اما تمامی خصوصیات فوق کاملاً در موجک های کلاسیک برآورده نمی شوند. به همین منظور موجک های چند دقتی و تبدیل موجک مختلط که حاصل گسترش موجک کلاسیک هستند ارائه شده اند. ایده موجک چند دقتی حالت کلی تر موجک های کلاسیک می باشد، به این صورت که به جای استفاده از یک تابع مقیاس خاص از چندین تابع مقیاس، موسوم به توابع مقیاس چند دقتی استفاده می شود. در ساخت موجک های چند گانه، برای برآورده شدن همزمان خواص مطلوب موثر در پردازش سیگنال درجه آزادی بیشتری وجود دارد.

در تبدیل موجک مختلط، فیلترها ضرایب مختلط دارند و نمونه های خروجی مختلط تولید می کنند. DWT حقیقی از نظر انتخاب جهت ضعیف و نیز حساس به شیفت است. این دو مسئله توسط تبدیل موجک مختلط قابل حل می باشند. [۳۵]

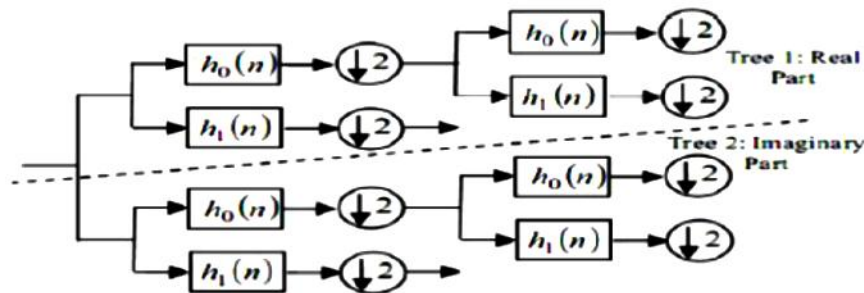
DT-CWT - ۳-۱-۳-۳

تغییر شکل موج مختلط اضافی که توسط کینگسبوی طراحی شده، تغییر شکل موج مختلط درخت دوگان (DT-CWT) نامیده می شود. مزیت اصلی موج مختلط این است که خواص را برحسب حساسیت جابجایی، جهت یابی، و اطلاعات فازی بهبود بخشیده است.

DT-CWT، یک سیگنال را برحسب موج مادر جا بجا کرده و به صورت مختلط تجزیه می کند [۳۶]. DT-CWT، با استفاده از تغییر شکل های جداگانه و با ترکیب سیگنال های زیر گروه به شکلی مناسب اجرا می شود. به خصوص، DT-CWT ۱D با استفاده از ۲ کرانه فیلتری موازی که بر اساس داده هایی یکسان عمل می کنند و میتوانید در شکل ۵ آنرا ببینید، اجرا می شود. یک موج ارزش مختلط رامی توان از رابطه ۳-۱۹ بدست آورد.

$$\psi(t) = \psi_h(t) + j\psi_g(t) \quad (۱۹-۳)$$

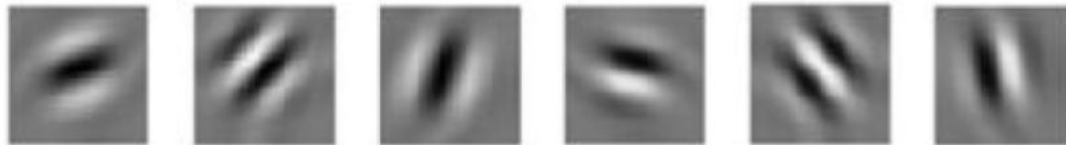
که و، هر دو موج های مقدار واقعی هستند. تا به اینجا، درخت دوگان اصلاً به عنوان تغییر شکل مختلط به نظر نمی رسد. با این حال، زمانی که خروجی های ۲ درخت در شکل ۳-۳ به عنوان بخش های واقعی و خیالی ضرایب مختلط تفسیر شوند، این تغییر به شکلی موثر، مختلط می شود [۳۶].



شکل ۳-۳: تغییر شکل موج مختلط درخت دوگان ۱ D

بسط یک DT-CWT دو بعدی تا ۲ بعدی، با فیلترسازی جداگانه در امتداد ردیف ها و ستون ها انجام می شود. اگر هر دو فیلتر ردیف و ستون، requency های منفی را پنهان کنند، تنها ربع اول طیف سیگنال ۲D واقعی، ضروری هستند، طوری که فیلترسازی، با جفت های مختلط فیلترهای ردیف نیز انجام می شود. این، ۴:۱ سیگنال ۲D تغییر شکل یافته اضافه را ارائه می دهد. چون فیلترهای مختلط می توانند تمام بخش های فضای بسامد ۲ D را جدا کنند، انتخاب جهت دار واقعی را ارائه می دهند. شکل ۳-۴.

واکنش های محرک این ۶ موج مرتبط با تغییر شکل موج مختلط ۲ D را نشان می دهد. این ۶ زیر گروه موج ۲D DT-CWT کاملاً در جهت هستند [۳۶].



شکل ۳-۴: واکنش محرک ۶ موج

۳-۳-۱-۴- ویژگی های تامورا

این ویژگی، ویژگی هایی از بافت را نظیر درشتی و ریزا، تباین، جهت دار بودن ۲، خطی بودن ۳، منظم بودن ۴ و زبری ۵ تعریف می کند و بر اساس مطالعات روانشناسی بر روی ادراک انسان از بافت طراحی شده است. هر یک از این ویژگی ها توسط رابطه ای جداگانه به دست می آید. سه ویژگی اول در سیستم های بازیابی نتایج بسیار موفق داشته و در برخی از سیستم های اولیه مانند QBIC استفاده شده اند.

برای محاسبه «درشتی و زبری» بافت، میانگین همسایگی تمام پیکسل های تصویر در همسایگی $2^k \times 2^k$ به دست می آید. این میان از طریق رابطه ۳-۲۰ به دست می آید [۲۷] [۳۰].

$$A_k(x, y) = \frac{\sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i, j)}{2^{2k}} \quad (۳ - ۲۰)$$

سپس برای هر نقطه اختلاف بین تمام همسایگی های نقاط اطراف آن در هر دو سوی افقی و عمودی محاسبه می شود برای مثال این میانگین برای حالت افقی از رابطه (۳-۲۱) محاسبه می شود.

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \quad (۳ - ۲۱)$$

-
- 1 Coarseness
 - 2 Directionality
 - 3 Line-likeness
 - 4 Regularity
 - 5 Roughness

در هر نقطه، اندازه مناسب برابر است با بیشترین مقدار به دست آمده از میانگین های موجود (E) میزان درشتی و ریزی بافت در هر نقطه برابر مقدار همسایگی به دست آمده است. مفهوم «تقابل»، چگونگی گستردگی سطوح خاکستری را به همراه درجه قطبیت^۱ در تصویر بررسی می کند. اولی از محاسبه انحراف استاندارد مقادیر پیکسل های تصویر خاکستری شده به دست می آید و دومی برابر نقطه اوج^۲ a تصویر است. بنابراین تقابل از رابطه (۳-۲۲) به دست می آید.

$$F_{con} = \frac{\sigma}{(a_f)^n}, a_f = \mu_f / \sigma^4 \quad (22 - 3)$$

جهت دار بودن، به ویژگی کلی یک ناحیه از بافت می پردازد. این ویژگی بیانگر زاویه کلی بافت در یک ناحیه است. برای محاسبه جهت بافت زاویه و طول اجزای ریز بافت محاسبه می شود. سپس هیستوگرام زوایا و طول های به دست آمده از این روش محاسبه شده و به این ترتیب، معرفی برای چگونگی جهت دار بودن بافت به دست خواهد آمد.

۳-۱-۵- ماتریس هم‌رخداد

این ماتریس که به هیستوگرام مرتبه دوم یا GLCM^۳ نیز معروف است، به احتمال تجمعی توزیع جفت پیکسل ها می پردازد و روش بسیار موفقی در تعریف بافت است. GLCM معرف فرکانس حضور هر دو پیکسل در فاصله ای معین در تصویر است، که فاصله آن ها به صورت برداری از پیش تعیین شده در نظر گرفته می شود. پس از محاسبه این ماتریس خصوصیات نظیر انرژی، آنتروپی، تباین، و همگنی از آن استخراج می شود که می توانند، معرفی برای بافت موجود در تصویر باشند. Arvis به همراه سایرین در [۳۷] روش دیگری برای استخراج بافت از این ماتریس ارائه کرده است.

1 Polarization

2 Kurtosis

³ Gray level co-occurrence matrix

۳-۴- روش های استخراج ویژگی شکل

شکل یک ویژگی بصری مهم از یک تصویر است. جست و جوی تصاویر با استفاده از ویژگی شکل مورد توجه زیادی قرار گرفته است. تکنیک های توصیف و نمایش شکل بسیاری وجود دارد. در این بخش ما تکنیک های مهم را دسته بندی و مرور می میکنیم.

یک توصیفگر شکل باید مستقل از کاربرد باشد به جای آنکه تنها برای نوع خاصی از شکل ها به خوبی عمل کند. پیچیدگی محاسباتی کم از ویژگی های مهم یک توصیفگر شکل مطلوب است. تکنیک های توصیف و نمایش شکل به طور عام به دو دسته دسته بندی می شوند روش های مبتنی بر مرز و منتهای مبتنی بر ناحیه. تحت این دسته بندی روش های گوناگون به رویکردهای ساختاری^۱ و رویکردهای سراسری^۲ تقسیم میشوند. این زیرکلاس ها بر اساس این است که آیا شکل به طور کلی نمایش داده می شود یا توسط section/segment ها می باشد. [۳۱]

۳-۴-۱ هیستوگرام جهت لبه ها

از روش های مبتنی بر شکل که بدون استفاده از قطعه بندی تصویر عمل می کنند می توان از روش های مبتنی بر لبه یاد کرد. مطلوبیت استفاده از لبه ها به دو علت است: اولاً حجم لبه های تصویر نسبت به حجم پیکسل های تصویر بسیار کم است، بنابراین به علت کاهش اطلاعات تحت پردازش می توان سیستمی کارا با استفاده از لبه ها ساخت. ثانیاً لبه ها به علت استفاده از همبستگی میان پیکسل ها حجم زیادی از اطلاعات تصویر را در بر دارند و همچنین از آنجایی که از دیدگاه درک انسانی نیز، اشکال به واسطه لبه هایشان از یکدیگر تمیز داده می شوند، پس می توان سیستمی دقیق با استفاده از لبه ها ساخت. بر این اساس جین و وایلا هیستوگرام جهت لبه^۳ (EDH) را ابداع کرده اند که در این حال با دسته بندی لبه های تصویر روی جهت لبه با یک کمیت دهی پنج درجه هیستوگرامی از فراوانی لبه ها

1 Structural

2 Global

3 Edge Direction Histogram

ایجاد می شود و از آن به عنوان بردار ویژگی تصویر استفاده می شود. این روش دقت نسبتاً مطلوبی دارد و نسبت به انتقال، تغییر اندازه و چرخش تصویر مستقل عمل می کند. چون در ساخت EDH با لبه ها به صورت منفرد برخورد شده و از همبستگی میان لبه های مجاور استفاده ای نشده است این روش از حداکثر دقت قابل حصول برخوردار نیست. به طور مثال شکل مبین دو تصویر متفاوت است که EDH آن ها تقریباً یکسان است. [۳۲]



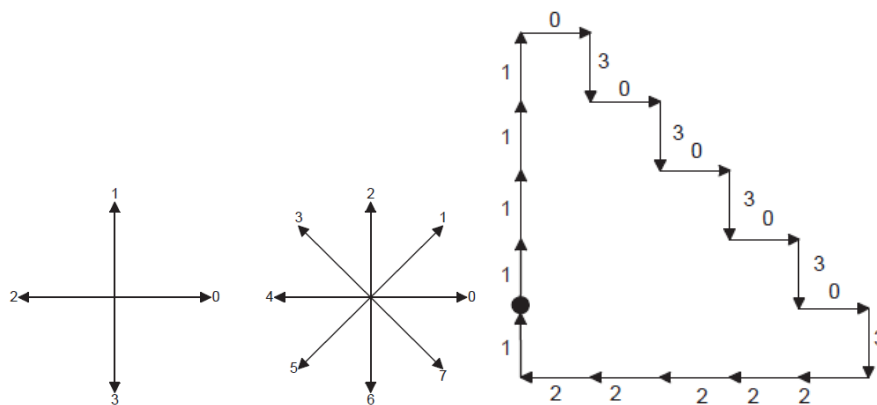
شکل ۳-۶ توصیفگر هیستوگرام جهت لبه ها

۳-۴-۲ کد زنجیره ای

کد زنجیره ای شی را توسط یک دنباله از قطعات با اندازه واحد و با یک جهت داده شده توصیف میکند. این روش در سال ۱۹۶۱ توسط فریمن معرفی شد که پیکربندی های دلخواه هندسی را کدگذاری میکند. در این رویکرد یک منحنی داخواه توسط یک دنباله از بردارهای کوچک با طول واحد و یک مجموعه از جهت های ممکن نشان داده می شود بنابراین متد بردار-واحد نیز نامیده می شود. [۳۱]

در پیاده سازی، مرزهای دیجیتال از یک تصویر روی یک شبکه grid قرار میگیرند نقاط مرزی به نزدیکترین نقطه شبکه تقریب زده می شوند سپس یک تصویر نمونه بدست می آید. از یک نقطه شروع، کد زنجیره ای می تواند در ۴ یا ۸ جهت ایجاد شود.

کد زنجیره ای معمولا ابعاد بالایی دارد و حساس به نویز است و معمولا به عنوان ورودی برای تحلیل های سطح بالاتر استفاده می شود. برای مثال می تواند برای تقریب چندگوشه ای یا یافتن انحنای مرز استفاده شود.



chain code: 11110303030303222221

شکل ۳-۷ کد زنجیره ای

۳-۶- ارزیابی کارایی سیستم های بازیابی تصویر

رایج ترین معیارهای ارزیابی در بازیابی اطلاعات،^۱ معیارهای دقت^۲ و فراخوانی^۳ هستند. این دو معیار به شکل یک گراف که در آن دقت بر حسب فراخوانی رسم شده است ارائه می شوند که به گراف PR مشهور است [۳۹].

1 Information retrieval

2 precision

3 Recall

نرخ فراخوانی و دقت به طور گسترده برای اندازه گیری عملکرد بازیابی مورد استفاده قرار میگیرد. اساسا بر پایه تطبیق قطعی است. در این روش مجموعه داده ای به مجموعه باینری، برطبق مرتبط بودن یا نامرتب بودن با تصویر پرس و جو، تبدیل می شود

نرخ دقت و فراخوانی به صورت زیر تعریف می شود [۴۱ و ۴۴]

$$P = \frac{r}{n_1} = \frac{\text{number of relevant retrieved items}}{\text{number of retrieved images}} \quad (24 - 3)$$

$$R = \frac{r}{n_2} = \frac{\text{number of relevant retrieved items}}{\text{total number of relevant images in DB}} \quad (25 - 3)$$

دقت، اندازه گیری دقت و صحت بازیابی است در حالیکه فراخوانی، اندازه گیری توانایی بازیابی تصاویر مرتبط از پایگاه داده است.

۳-۶-۱- آزمون هدف^۱

در آزمون هدف [۴۱]، کاربر برای رسیدن به تصاویر دلخواه، پرس و جوی خود را از طریق ارائه مثال انجام می دهد. تعداد دفعاتی که کاربر تصویر پرس و جوی خود را تا رسیدن به تصاویر دلخواهش عوض می کند، به عنوان یک عامل برای تعیین کارایی سیستم معرفی شده است.

۳-۶-۲- نرخ خطا^۲

این روش بیشتر در بازشناسی چهره و اشیا به کار می رود. در حقیقت این روش، پارامتر دقت را می دهد.

$$\text{Error rate} = \frac{\text{Number of non-relevant image retried}}{\text{Total No. of image retrieved}} \quad (26 - 3)$$

۳-۶-۳- راندمان بازیابی^۳

اگر تعداد تصاویر بازیابی شده کوچکتر یا مساوی با تعداد کل تصاویر مرتبط با تصویر پرس و جو باشد، این معیار، همان معیار دقت و در غیر اینصورت ارائه دهنده معیار فراخوانی است.

1 Target test

2 Error rate

3 Retrieval efficiency

$$\left[\begin{array}{l} \text{Retrieval} \\ \text{efficiency} \end{array} \right] \left\{ \begin{array}{ll} \frac{\text{No.of relevant image retried}}{\text{Total No.of image retrieved}} & \text{if NO.of retrieved} < \text{NO.of relevant} \\ \frac{\text{No.of relevant image retried}}{\text{Total No.of relevant image}} & \text{otherwise} \quad (27 - 3) \end{array} \right.$$

*گراف دقت بر حسب تعداد تصاویر بازیابی شده

*گراف فراخوانی بر حسب تعداد تصاویر بازیابی شده

علاوه بر آنچه ذکر شد معیارهای دیگری وجود دارد که به ندرت در مقالات ارائه می شود . آنچه که امروز محققان حوزه بازیابی تصویر مبتنی بر محتوا، در مقایسه ی سیستم های مختلف رنج می برند، مشخص نبودن پایگاه داده تصویری استاندارد و معیارهای استاندارد است که با استفاده از آنها بتوان نتایج کار تحقیقات را مقایسه کرد.

۳-۷- معیار شباهت^۱

شباهت دو تصویر، با فاصله بین ویژگی های استخراج شده از آن دو یا فاصله بین بردارهای تشکیل شده از ویژگی های آنها تعریف می شود. به عبارت دیگر ، تابع شباهت دو بردار ویژگی را به یک مقدار حقیقی مثبت تصویر می کند که استفاده از آن، مقایسه ی بین ویژگی های دیداری بین دو تصویر را امکان پذیر می کند.

به عنوان مثال در مورد مقایسه هیستوگرام از توابع شباهتی چون فصل مشترک هیستوگرام^۲ و فاصله اقلیدسی استفاده می شود.

1 Similarity measures

2 Histogram intersection

انتخاب معیار عدم شباهت تصاویر بر اساس هر یک از ویژگی های سطح پایین

انتخاب معیاری که عدم شباهت بین دو تصویر را بر پایه ی ادراک انسان ارائه کند ؛ به میزان قابل توجهی فاصله معنایی را کاهش و رسیدن به هدف بازیابی بر اساس معنا را افزایش خواهد داد . برای تعیین میزان عدم شباهت دو تصویر، فاصله میان بردارهای ویژگی آنها از یکدیگر محاسبه می شود. استفاده از معیارهای متفاوت در تعیین میزان عدم شباهت دو تصویر، جوابهای متفاوتی خواهد داشت. در حالیکه ویژگی های سطح پایین دو تصویر به یکدیگر نزدیک هستند، ویژگی های معنایی آن ها کاملا متفاوت باشند.

۳-۸- معیارهای عدم شباهت

هر معیاری که بتواند عدم شباهت دو بردار را با یک مقدار حقیقی بیان کند، می تواند معیاری برای تعیین فاصله بین دو تصویر در بازیابی استفاده شود. عدم شباهت دو تصویر x, y توسط تابع $d(E, G)$ محاسبه می شود که این تابع فاصله بین بردارهای ویژگی $F=(f_1, f_2, \dots, f_n)$ و $G=(g_1, g_2, \dots, g_n)$ را بیان می کند. F و G به ترتیب بردارهای ویژگی تصاویر x و y هستند. هر چه d کوچکتر باشد، نشان می دهد که دو بردار و به تبع آن دو تصویر به یکدیگر نزدیکتر هستند.

یک معیار شباهت به عنوان یک متریک فاصله تعریف می شود. در این بخش معیارهای شباهت مختلف را با ذکر جزئیات توضیح می دهیم.

۳-۸-۱- فاصله فضای متریک

یک فضای R^N یک متریک [۴۱] است اگر برای هر یک از دو المان X, Y یک عدد حقیقی $D(X, Y)$ وجود داشته باشد که در واقع همان فاصله می باشد، که ویژگی های زیر را باید داشته باشد:

$$۱) d(x, y) \geq 0$$

$$۲) d(x, y) = 0 \quad \text{اگر تنها و اگر } x = y$$

$$۳) d(x, y) = d(y, x)$$

$$۴) d(x, z) \leq d(x, y) + d(y, z)$$

۳-۸-۲- فاصله مینکوسکی

فاصله مینکوسکی [۴۱] بر اساس نرم L_p به صورت زیر تعریف می شود:

$$d_p(Q,T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^p \right)^{\frac{1}{p}} \quad (28-3)$$

که $Q = \{Q_1, \dots, Q_N\}$ و $T = \{T_1, \dots, T_N\}$ در واقع بردارهای ویژگی تصویر هدف و جست و جو می باشند. زمانی که $p=1$ و $d_1(Q,T)$ فاصله بلوک شهری یا فاصله مانهاتان (L_1) هستند.

$$d_1(Q,T) = \sum_{i=1}^{N-1} |Q_i - T_i| \quad (29-3)$$

زمانی که $p=2$ ، $d_2(Q,T)$ فاصله اقلیدسی L_2 است.

$$d_2(Q,T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^2 \right)^{\frac{1}{2}} \quad (30-3)$$

زمانی که $p \rightarrow \infty$ داریم L_∞

$$L_\infty(Q,T) = \max_{0 \leq i \leq 1} \{|Q_i - T_i|\} \quad (31-3)$$

فواصل کانبرا و زکانووسکی به ترتیب با روابط ۳۲-۳ و ۳۳-۳ تعریف می شوند. این فواصل فقط به بردارهایی اعمال می شوند که عناصر غیر منفی دارند [۴۱].

$$D(F,G) = \sum_{i=1}^n \frac{|f_i - g_i|}{|f_i + g_i|} \quad \text{Canberra distance} \quad (32-3)$$

$$D(F,G) = 1 - \frac{2 \sum_{i=1}^n (f_i g_i)}{\sum_{i=1}^n (f_i + g_i)} \quad \text{Czekanowski distance} \quad (33-3)$$

۳-۸-۳- فاصله کسینوسی

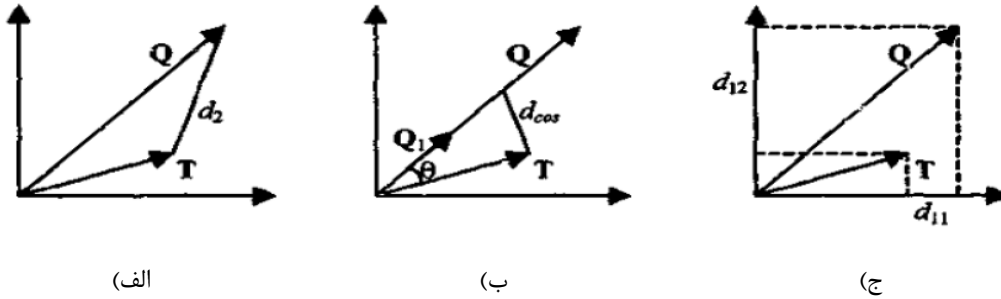
فاصله کسینوسی [۴۱] تفاوت را در جهت، صرف نظر از طول بردار محاسبه می کند. تفاوت داده شده در واقع زاویه بین بردارها می باشد. با استفاده از قانون ضرب نقطه ای داریم:

$$Q \cdot T = |Q| \cdot |T| \cos \theta$$

$$d_\infty(Q,T) = 1 - \cos \theta = 1 - \frac{Q \cdot T}{|Q| \cdot |T|} \quad (34-3)$$

شکل ۳-۵ فاصله کسینوسی و فاصله مینوسکی را در یک فضای دو بعدی تشریح می کند. همانطور که می بینیم فاصله اقلیدسی هم طول و هم زاویه را برای محاسبه فاصله در نظر می گیرد (شکل ۳-۵ الف)،

در حالی که فاصله کسینوسی تنها زاویه را به حساب می آورد. در نتیجه Q_1 و Q فاصله یکسانی تا T دارند، یعنی $d_{\cos}(Q,T)=d_{\cos}(Q_1,T)$ (شکل ۳-۵ (ب)). d_1 فاصله هر بعد از بردارهای ویژگی را محاسبه می کند (شکل ۳-۵ (ج)).



شکل (۳-۵): (الف) فاصله اقلیدسی $d_2(Q,T)$ ؛ (ب) فاصله کسینوسی $d_{\cos}(Q,T)$

؛ (ج) L_1 distance $d_1=d_{11}+d_{12}$

۳-۸-۴ - فاصله χ^2

χ^2 [۴۱] به صورت زیر تعریف می شود.

$$d_{\chi^2}(Q,T) = \sum_{i=0}^{N-1} \frac{(Q_i - m_i)^2}{m_i} \quad (35-3)$$

که در آن $m_1 = \frac{Q_1 + T_1}{2}$

۳-۸-۵ - فصل مشترک هیستوگرام

فصل مشترک هیستوگرام [۴۱] توسط سوین و بالارد ارائه شد. برای تصاویری که از هیستوگرام رنگ استفاده می کنند، که در واقع قادر است تفاوت های جزئی را بررسی کند، زمانی که اندازه ی شی (Q) کوچکتر از اندازه ی تصویر T است. تعریف اصلی از هیستوگرام بدین صورت داده شده است .

$$d_{h_i}(Q,T) = 1 - \frac{\sum_{i=0}^{N-1} \min(Q_i, T_i)}{|Q|} \quad (36-3)$$

که بعدها به صورت یک فاصله متریک گسترش داده شده است [۹۷].

$$d_{h_i}(Q,T) = 1 - \frac{\sum_{i=0}^{N-1} \min(Q_i, T_i)}{\min(|Q|, |T|)} \quad (37-3)$$

۳-۸-۶- فاصله ماهالانوبیس

فاصله ماهالانوبیس [۴۱] یک مورد خاص از فاصله ی درجه دوم است در حالی که ماتریس تبدیل توسط یک ماتریس کواریانس به وجود آمده باشد، که از یک مجموعه ی آموزش دیده از بردارهای ویژگی به دست آمده باشد، که $A = \Sigma^{-1}$. به عبارت دیگر برای به کار گیری فاصله ماهالانوبیس بردارهای ویژگی به صورت متغیرهای تصادفی $X = \{x_0, x_1, \dots, x_{N-1}\}$ در نظر گرفته می شوند. سپس ماتریس همبستگی به صورت $R = [r_{ij}]$ که $r_{ij} = E[x_i x_j]$ و $E\{y\}$ میانگینی از متغیر تصادفی y است. سپس، ماتریس کواریانس توسط Σ ، که $\Sigma = [\sigma_{ij}^2]$ و $\sigma_{ij}^2 = r_{ij} - E\{x_i\}E\{x_j\}$ داده می شود.

فاصله ماهالانوبیس بین بردارهای ویژگی Q, T با قرار دادن $X_Q = Q$ و $X_T = T$ ، به صورت زیر محاسبه می شوند.

$$D_{mah} = [(X_Q - X_T) \Sigma^{-1} (X_Q - X_T)]^{\frac{1}{2}} \quad (38-3)$$

در مورد خاص که X_i به صورت مستقل آماری است، اما واریانس نابرابر دارند، Σ یک ماتریس قطری به صورت زیر می باشد.

$$\begin{bmatrix} \sigma_0^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{n-1}^2 \end{bmatrix}$$

در این مورد فاصله ماهالانوبیس به فرم ساده تری کاهش می یابد.

$$d_{mah}(Q, T) = \sum_{i=0}^{N-1} \frac{(Q_i - T_i)^2}{\sigma_i^2} \quad (39-3)$$

این یک فاصله L_2 وزنی است. آن وزن بیشتر را به بعدی با واریانس کوچکتر می دهد و وزن کمتر را به بعدی با واریانس بیشتر می دهد.

فاصله اقلیدسی برای تعیین بردار ویژگی M بعدی تصویر پرس وجو h_q و تصویر هدف h_t در پایگاه داده مطابق زیر تعریف می شود [۴۰].

$$Euc. d_{q,t} = \sqrt{\sum_{m=0}^{M-1} |h_q(m) - h_t(m)|^2} \quad (40-3)$$

۳-۹- درصد موفقیت وزنی^۱

Pwh [۴۱] مشابه معیار فراخوانی است. آزمون هایی که انجام می شود مشابه است یعنی موارد مرتبط با پرس و جو از مجموعه داده ای انتخاب می شود اگرچه به جای اندازه گیری فراخوانی بر اساس مقدار باینری موارد مرتبط، مقدار وزنی W_i به هر مورد در مجموعه داده ای انتساب داده می شود. W_i وزنی است که نشان می دهد موارد انتساب شده به عنوان مرتبط یا مشابه با پرس و جو هستند.

$$PWH = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^N w_i} \quad (۴۱-۳)$$

که n تعداد موارد بازیابی شده و N تعداد کل موارد در مجموعه داده ای است. در واقع فراخوانی نوع خاصی از pwh است زمانی که W_i مقدار ۰ و ۱ را میگیرد.

۳-۱۰- درصد رتبه بندی شباهت^۲

درصد رتبه بندی شباهت توسط بیمبووپالا [۴۱] ارائه شد. در این اندازه گیری به هر مورد یک رتبه شباهت نسبت داده می شود که بر اساس شباهت آن ها به پرس و جو می باشد، به جای اینکه تنها مرتبط/نامرتب را در rpp و pwh مشخص کنیم.

نتیجه نهایی آزمون به صورت یک ماتریس به شکل $\{Q_j(i,k)\}$ می باشد که $Q_j(i,k)$ رتبه ی آمین مورد در مکان k را برای پرس و جو j نشان می دهد.

سپس میانگین $\bar{p}_j(i)$ و واریانس $\bar{\sigma}_j(i)$ هر سطر محاسبه می شود. $\bar{p}_j(i)$ و $\bar{\sigma}_j(i)$ میانگین رتبه آمین تصویر را برای پرس و جو j و اندازه گیری مقدار نزدیک بودن رتبه به $\bar{p}_j(i)$ را به ترتیب ارائه می دهد.

اگر برای یک پرس و جو j یک الگوریتم بازیابی آیت i را با رتبه $p_j(i)$ برگرداند، توافق (سازش، مطابقت) بین رتبه الگوریتم و رتبه انسان توسط درصد رتبه بندی شباهت $S_j(i)$ اندازه گیری می شود.

$$S_j(i) = \sum_{k=p_j(i) - \frac{\sigma_j(i)}{2}}^{p_j(i) + \frac{\sigma_j(i)}{2}} Q_j(i, k) \quad (۴۲-۳)$$

1 percentage of weighted hits(pwh)

2 percentage of similarity ranking(psr)

طرح $s_j(i)$ به عنوان یک تابع از $p_j(i)$ ، عملکرد بازیابی الگوریتم بازیابی را نشان می دهد $s_j(i)$ بالا، نشان دهنده دقت بازیابی بالای الگوریتم است.

در بالا ۳ معیار کارایی (ارزیابی عملکرد) تشریح شده است.

pwh مطابقت بیشتری با رتبه بندی بشر نسبت به فراخوانی در rpp دارد اگرچه pwh توانایی برگشت (رد) موارد نامرتب را اندازه گیری نمی کند در لیستی که برمی گرداند تعداد زیادی موارد نامرتب وجود دارد. عیب دیگر pwh این است که تعداد ثابت از موارد را بر می گرداند. این مطلوب نیست چون پرس و جویهای متفاوت تعداد موارد مرتبط متفاوتی دارند.

Rpp توانایی اندازه گیری موارد مرتبط و نامرتب را دارد تنها عیب آن این است که رتبه موارد مرتبط با پرس و جو را در نظر نمی گیرد این عیب خیلی مهم نخواهد بود وقتی که مجموعه داده ای کلاس بندی می شود.

rpp معیار ارزیابی بسیار مطلوبی نسبت به pwh و psr است و بسیار مناسب برای ارزیابی کارایی بازیابی در مجموعه داده های بزرگ و کلاس بندی شده است.

فصل چهارم

راه کار پیشهادی

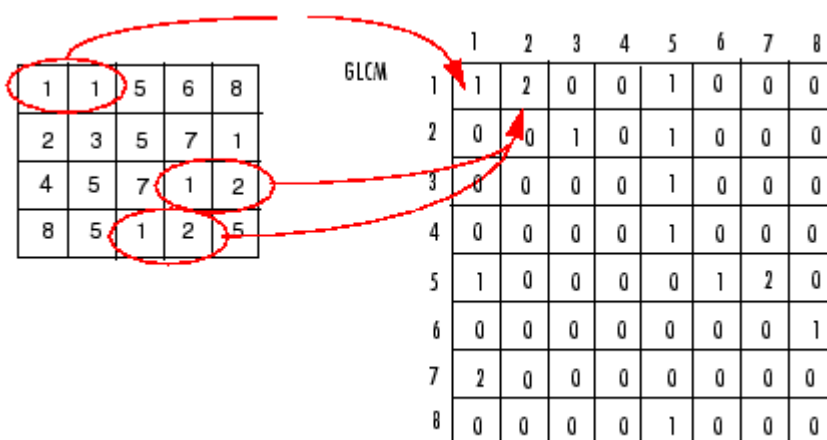
پس از مطالعه و پیاده سازی روش های مختلف با توجه به اهمیت دو مسئله دقت و سرعت، در نهایت روش پیشنهادی خود را ارائه داده ایم. در این فصل معرفی راهکار پیشنهادی خود می پردازیم. روش پیشنهادی ما بر مبنای دو ویژگی لبه و رنگ می باشد. ویژگی رنگ، هیستوگرام فضای رنگ HSV می باشد و در بخش ویژگی لبه از هیستوگرام زوایای گرادیان SRF بر پایه الگوهای باینری محلی بهره برده ایم. بحث معیار شباهت هم مورد اهتمام ما بوده است. فاصله کانبرا در آزمایشات نتایج بهتری را به همراه داشته است.

۴-۱ پیاده سازی برخی از روش های موجود

جهت یافتن ویژگی های مناسب جهت ارائه به عنوان راهکار پیشنهادی، ناگزیر هستیم تا برخی از روش های موجود را پیاده سازی کرده و با علم به نقاط ضعف و قوت آنها، مجموعه ویژگی های مورد نیاز را انتخاب کنیم. جهت ارزیابی روش پیشنهادی آنرا با روش های موجود مقایسه می کنیم.

۴-۱-۱ ماتریس هم رخداد^۱

ماتریس هم رخداد به توزیع پیکسل ها از جهت سطح خاکستری و زاویه مکانی با یکدیگر اشاره دارد. در این روش ابتدا تصویر خاکستری به چند سطح کمی سازی می شود، سپس با توجه به نحوه قرارگیری پیکسل ها در تصویر ماتریس هم رخداد تشکیل می شود. به نوعی این ماتریس تعداد وقوع همجواری سطوح مختلف را در حالت مشخص تعیین می کند. شکل ۴-۱ سطح خاکستری را به ۸ سطح تبدیل کرده و در زاویه صفر با یک همسایگی، ماتریس هم رخداد را تشکیل داده است.



شکل ۴-۱ نحوه تشکیل ماتریس هم رخداد

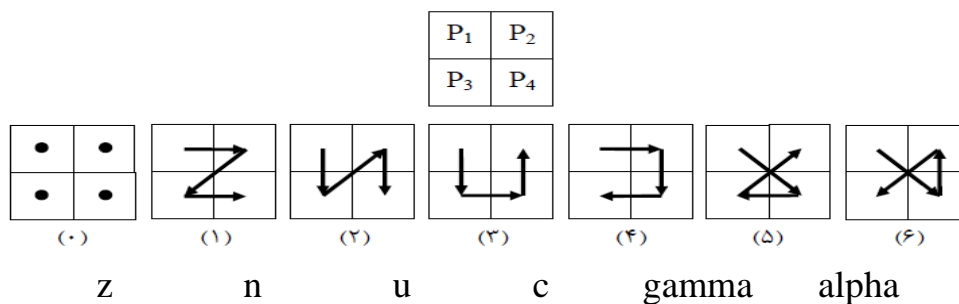
¹ Gray level co-occurrence

از نکات مثبت روش بیان شده، سادگی پیاده سازی و نمایش مناسب تصاویر بافتی منظم می باشد. در مقابل در برابر تصاویری که بافت های پیچیده ای دارند نتایج بالایی نشان نمی دهد.

۲-۱-۴ ماتریس هم‌رخداد رنگ^۱ CCM

ماتریس هم رخداد رنگ در مقاله [۴۸] پیشنهاد شده است. البته چارچوب اصلی آن بر مبنای ویژگی الگوی اسکن (motif) می باشد اما در مقاله ذکر شده در فضای رنگ RGB آزمایش شده به علاوه جزئیات جدیدی به آن اضافه شده است که در ادامه به آن پرداخته می شود.

برای هر ۴ پیکسل که در همسایگی یکدیگر قرار دارند ۷ نوع الگوی اسکن وجود دارد که در شکل ۲-۴ نشان داده شده است. در این حالت فرض شده P1 پیکسل اصلی است و اسکن از این پیکسل آغاز می گردد. جهت اسکن به سمتی است که کمترین اختلاف را با P1 داشته باشد. اگر تمام مقادیر پیکسل ها یکسان باشد یعنی هیچ اختلافی وجود ندارد و الگوی اسکن شماره ۰ می باشد. مثلا اگر مقادیر P1, P2, P3, P4 به ترتیب ۱۰ و ۲۹ و ۴۰ و ۱۰۵ باشد الگوی اسکن شماره ۱ خواهد شد. چرا که میزان اختلاف این پیکسل ها با P1 به ترتیب ۱۹ و ۳۰ و ۹۵ می باشد.



شکل ۲-۴ الگوهای اسکن ماتریس هم رخداد رنگ

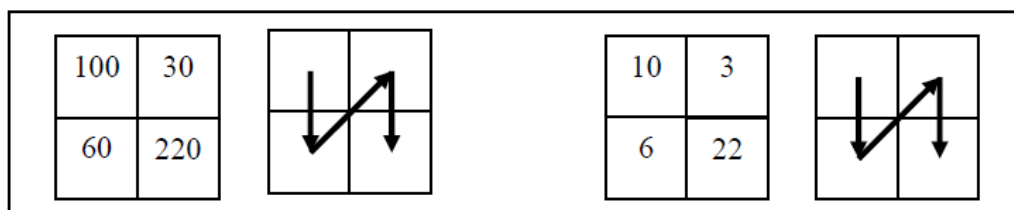
پس از اسکن کامل تصویر، تصویر جدید بدست می آوریم که به ازای مقدار هر پیکسل مقداری را جایگزین می کند که نشان دهنده الگوی اسکن مورد نظر می باشد. تصویر بدست آمده دارای ۷ مقدار از ۰ تا ۶ خواهد بود، از این تصویر همانطور که در ماتریس مجاورت معرفی شد، ماتریس هم رخداد بدست می آید به ازای هر رنگ یک ماتریس به اندازه ۴۹ مقدار خواهد داشت. نتایج آزمایشات دقت متوسط را برای این روش نشان می دهد. از نقاط ضعف این روش، حساسیت زیاد به مقدار پیکسل ها است در نتیجه تصاویری

¹ Color co-occurrence matrix (CCM)

که شباهت بسیار به هم دارد را به آسانی شناسایی می کند اما در طبقه بندی تصاویر که اختلافی در پس زمینه دارند به سختی عمل می کند.

۳-۱-۴ تفاضل بین پیکسل های الگوی اسکن^۱ (DBPSP)

این روش در ترکیب با ماتریس هم‌رخداد رنگ استفاده می شود. هم‌رخداد رنگ (CCM) الگوی اسکن را مشخص می کند اما به تنهایی اطلاعات کافی در زمینه بافت را ارائه نمی دهد. به طور مثال در شکل ۳-۴ دو بخش تصویر الگوی یکسانی از دارند، اما مقدار پیکسل ها با هم تفاوت زیادی دارند.



شکل ۳-۴ دو بخش تصویر با الگوی یکسان و مقادیر متفاوت

ویژگی مذکور، تفاضل بین الگوهای اسکن را طبق روابط ۱-۴ محاسبه می کند. این کار باعث می شود تا مقادیر پیکسل ها نیز در تشکیل ویژگی دخیل شوند.

$$\Delta 1(x, y) = |p1 - p2| + |p2 - p3| + |p3 - p4|$$

$$\Delta 2(x, y) = |p1 - p3| + |p3 - p2| + |p2 - p4|$$

$$\Delta 3(x, y) = |p1 - p3| + |p3 - p4| + |p4 - p2|$$

$$\Delta 4(x, y) = |p1 - p2| + |p2 - p4| + |p4 - p3|$$

$$\Delta 5(x, y) = |p1 - p4| + |p4 - p3| + |p3 - p2|$$

$$\Delta 6(x, y) = |p1 - p4| + |p4 - p2| + |p2 - p4|$$

(۱-۴)

از مشخصه های مهم این روش حجم کم و سرعت آن است. اما به تنهایی نرخ دقت متوسطی داراست و معمولاً به همراه چند ویژگی دیگر از جمله ماتریس هم‌رخداد رنگ به کار گرفته می شود.

¹ Difference between pixels of scan pattern (DBPSP)

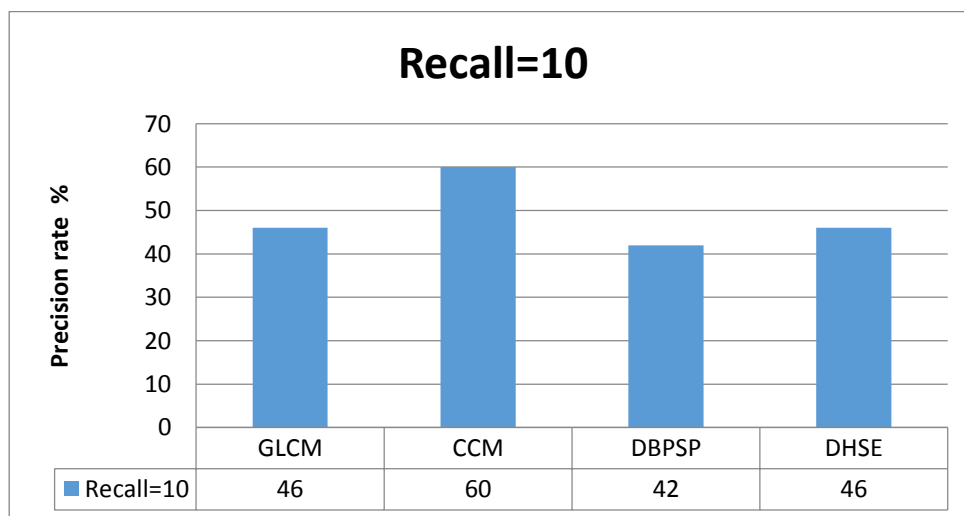
۴-۱-۴ هیستوگرام جهت لبه های مشابه

هیستوگرام لبه های مشابه (DHSE^۱) ابتدا لبه های تصویر را محاسبه می کند. این کار با گرادیان گرفتن از تصویر و بدست آوردن اندازه و فاز هر پیکسل بدست می آید. سپس لبه هایی که از آستانه ای مشخص، بیشتر باشند را به عنوان لبه اصلی تصویر قرار می دهیم. بعد از بدست آوردن لبه های اصلی هر لبه ای که مشابه لبه های اصلی باشد را تعیین می کنیم. لبه های مشابه دو شرط دارند؛ اول اینکه در همسایگی لبه های اصلی قرار داشته باشند، شرط دوم هم به اختلاف اندازه لبه ها بر می گردد که از یک آستانه ای کمتر نباشد. پس از بدست آوردن لبه های اصلی و مشابه آن، گرادیان زوایای تصویر را با فواصل ۵ درجه ای تشکیل می شود.

به علت جستجوی لبه اصلی و مشابه، در نتیجه سرعت الگوریتم پائین می آید. همچنین کمی سازی با فاصله ۵ درجه با توجه به مختصر بودن اطلاعات لبه، باعث توزیع بیش از حد می گردد و کارایی را کاهش می دهد.

۴-۱-۵ ارزیابی روش های موجود

به منظور بررسی دقیق تر روش های بیان شده، نمودار دقت بازیابی تصاویر با تعداد فراخوانی و معیار فاصله یکسان را رسم می نمائیم. در این نمودار درصد دقت بازیابی برحسب تعداد ۱۰ تصویر فراخوانی شده است. معیار فاصله این مجموعه، فاصله اقلیدسی می باشد.



شکل ۴-۴ نمودار دقت روش های موجود

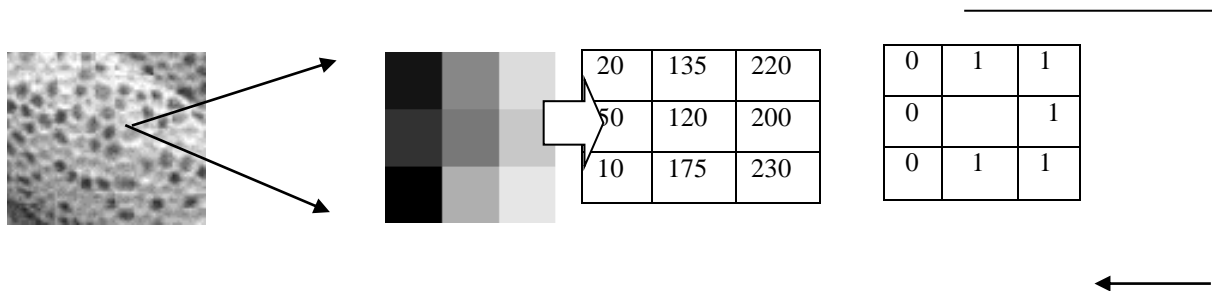
¹ Direction Hidtogram of Similar Edge

۴-۲ معرفی راهکار پیشنهادی

پس از بررسی و پیاده سازی روش های پیشین، ما با دو رویکرد دقت و سرعت سیستم پیشنهادی را ارائه کرده ایم. در روش پیشنهادی ما دو ویژگی رنگ و لبه در نظر گرفته شده است. ویژگی رنگ با توجه به نزدیک بودن به ادراک انسان هیستوگرام HSV می باشد. از آنجا که لبه ها نسبت به کل تصویر حجم کمتری دارند در نتیجه با کاهش اطلاعات، فرآیند پردازش سریعتر خواهد شد، از طرف دیگر لبه ها اطلاعات ارزشمندی را به همراه دارند زیرا در دیدگاه انسان، اشکال بواسطه لبه هایشان تمیز داده می شوند. با توجه به این مسائل، ویژگی لبه سیستم ما هیستوگرام زوایای گرادیان بر پایه الگوهای باینری محلی می باشد. دلیل استفاده از الگوهای باینری محلی دو علت را می توان برشمرد. ابتدا اینکه با اعمال الگوهای باینری محلی لبه های اصلی برجسته می شوند. نکته دیگر، تصویر حاصل شده از این الگوریتم بافتی از تصویر ارائه می دهد به عبارت بهتر ما از زمینه بافت، ویژگی شکل را بدست می آوریم و این به نوعی تلفیق ویژگی به حساب می آید.

۴-۲-۱ الگوهای باینری محلی

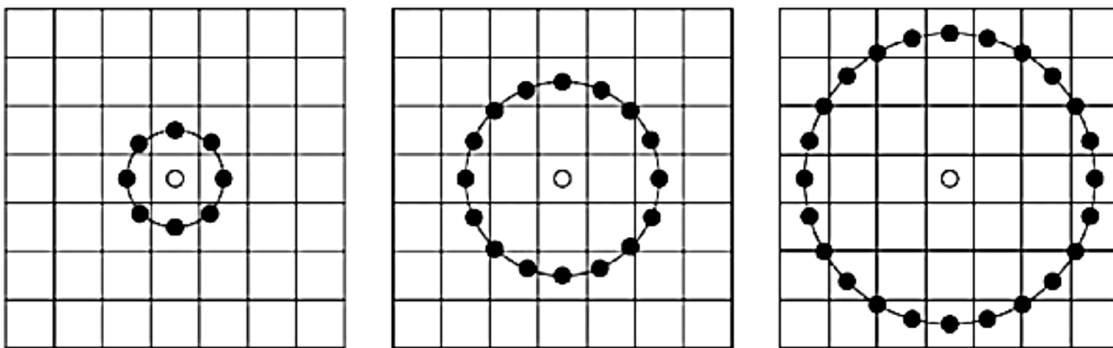
عملگر LBP اصلی به عنوان یک توصیفگر قدرتمند بافت تصویر در [۴۲] معرفی شد. این عملگر برای هر پیکسل با توجه به برچسب پیکسل های همسایگی 3×3 یک عدد دودویی تولید می کند. این برچسب ها با آستانه سازی مقدار پیکسل های همسایه با مقدار پیکسل مرکزی به دست می آیند. به این صورت که برای پیکسل های با مقدار بزرگتر یا مساوی مقدار پیکسل مرکزی برچسب ۱ و برای پیکسل های با مقادیر کوچکتر از مقدار پیکسل مرکزی برچسب ۰ قرار می گیرد. سپس این برچسب ها به صورت چرخشی در کنار هم قرار گرفته و یک عدد ۸ بیتی تشکیل می دهند. نحوه کار این عملگر در شکل ۴-۱ آمده است.



Binary: 01111100 decimal: 124

شکل ۴-۵ نحوه محاسبه عملگر LBP

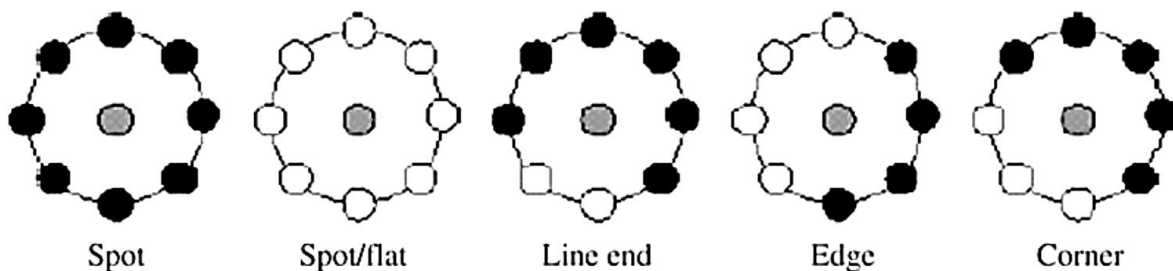
محدودیت عملگر LBP پایه همسایگی کوچک آن (۳×۳) است که باعث می شود بر تصاویر با مقیاس بزرگ تسلط نداشته باشد. بدین منظور بعداً این عملگر با توسعه ای بر اندازه همسایگی در [36] مطرح شد؛ که به صورت $LBP_{P,R}$ نشان داده می شود و حداکثر می تواند 2^P مقدار مختلف، مطابق با 2^P الگوی دودویی تولید شده توسط P پیکسل موجود بر روی شعاع همسایگی R ، تولید کند. برای مثال شکل ۴-۲ نحوه انتخاب پیکسل های همسایه در این نوع الگوی دودویی محلی را به ازای سه شعاع مختلف نشان می دهد.



شکل ۴-۶ عملگر LBP با شعاعها و تعداد همسایگی های مختلف

الگوهای تولید شده توسط LBP معنا و مفهوم خاصی دارند که در شکل ۴-۳ نمونه هایی از آن قابل مشاهده است. طبق این شکل LBP می تواند لبه در جهت های مختلف، نقطه، نواحی صاف، انتهای خط و... را تشخیص دهد. بعد از برچسب گذاری پیکسل ها توسط LBP هیستوگرام این برچسب ها برای ارائه بافت تصویر محاسبه می شود.

به دلیل سادگی در محاسبه و نتایج مناسب، مطالعات زیادی بر روی استفاده از LBP متمرکز شده است و در این راستا نسخه های متفاوت و جدیدی از آن برای استفاده در شرایط مختلف پیشنهاد شده است [۴۳].



شکل ۴-۷ مثال هایی از مفهوم الگوهای تولید شده در LBP، دایره های سفید و مشکی به ترتیب بیانگر عدد یک و صفر هستند.

پس از برچسب گذاری پیکسل ها توسط عملگر LBP، هیستوگرامی از الگوهای تولید شده به صورت رابطه (۲-۴) تعریف می شود.

$$H_i \sum_{x,y} I(f_i(x,y) = i), \quad i = 1, \dots, n - 1 \quad (2-4)$$

که n تعداد الگوهای تولید شده توسط عملگر LBP و تابع I به صورت رابطه (۳.۲) تعریف می شود.

$$I(A) = \begin{cases} 1 & A \text{ is True} \\ 0 & A \text{ is False} \end{cases}$$

در روش پیشنهادی الگوهای باینری محلی با شعاع ۱ و ۸ همسایگی در نظر گرفته شده است.

۲-۲-۴ هیستوگرام زوایای گرادیان SRF

در تصاویر موجود در هر کلاس طبقه بندی وجود لبه های یکسان می تواند به عنوان یک ابزار جهت طبقه بندی به کارآید همانطور که در فصل گذشته اشاره شد؛ استفاده از ویژگی شکل بعنوان یکی از ویژگی های سطح پائین همواره مورد توجه بوده است. در سیستم پیشنهادی ما، با توجه به در نظر گرفتن کارایی روش پیشنهادی و سرعت الگوریتم مورد نظر؛ ویژگی هیستوگرام زوایای گرادیان SRF [۴۴] از آنجا که هر دو مسئله را به خوبی پوشش می دهد انتخاب شد.

ویژگی SRF مشتمل بر دو ماسک سوبل و رابرتز می باشد که در عین سادگی و حجم کم، کارایی مناسبی دارد. اپراتور سوبل اطلاعات ۸ همسایه را بررسی میکند و عملگر رابرتز ۴ همسایه را مورد بررسی قرار می دهد. به جهت اینکه هر کدام از این دو اطلاعات متفاوتی را ارائه می دهند، یکی گرادیانهای افقی و عمودی و دیگری گرادیانهای قطری را محاسبه می کند. به همین دلیل با ترکیب هر دو، ویژگی کاراتری را شاهد خواهیم بود.

نحوه استخراج ویژگی به این شکل است؛ پس از اعمال الگوهای باینری محلی LBP عملگر سوبل را طبق رابطه زیر بدست می آوریم

$$g(x,y)=[g_x, g_y]^T \text{ where}$$

$$g_x(x,y) = f(x+1,y+1) + 2f(x+1,y) + f(x+1,y-1) - f(x-1,y-1) - 2f(x-1,y) - f(x-1,y+1)$$

$$g_y(x,y) = f(x-1,y+1) + 2f(x,y+1) + f(x+1,y+1) - f(x-1,y-1) - 2f(x,y-1) - f(x+1,y-1) \quad (3-4)$$

همچنین عملگر رابرتز را طبق رابطه زیر بدست می آوریم

$$g_x(x,y) = f(x,y) - f(x+1,y+1)$$

$$g_y(x,y) = f(x+1,y) - f(x,y+1) \quad (4-4)$$

در نهایت زاویه گرادیان به طور جداگانه برای هر عملگر محاسبه می شود.

$$P(x,y) = \tan^{-1}(g_y, g_x)$$

$$A(x,y) = \sqrt{gx^2 + gy^2} \quad (5-4)$$

در شکل ۴-۸ ماسک های سوبل و رابرتز که در روابط بالا مورد استفاده قرار گرفته، نشان داده شده است. همانطور که گفته شد اپراتور سوبل لبه های افقی و عمودی و رابرتز نیز همانطور که در شکل گویاست لبه های مورب را مشخص می کنند. با ترکیب دو عملگر طیف کاملتری از لبه ها را خواهیم داشت. زوایای بدست آمده از عملگرهای سوبل و رابرتز را کوانتیزه میکنیم، هر کدام به ۶۴ سطح تقسیم می شوند و یک ماتریس ۱۲۸ تایی ویژگی لبه ما را تشکیل می دهد.

| | | |
|----|----|----|
| -۱ | -۲ | -۱ |
| ۰ | ۰ | ۰ |
| ۱ | ۲ | ۱ |

| | | |
|----|---|---|
| -۱ | ۰ | ۱ |
| -۲ | ۰ | ۲ |
| ۱- | ۰ | ۱ |

| | |
|----|---|
| ۰ | ۱ |
| -۱ | ۰ |

| | |
|---|----|
| ۱ | ۰ |
| ۰ | -۱ |

(ب)

(الف)

شکل ۴-۸ الف) ماسک رابرتز ب) ماسک سوبل

نقطه متمایز کار ما در استفاده از ویژگی SRF این است که بر خلاف رویه معمول در استخراج بافت و شکل که الگوریتم مورد نظر بر تصاویر خاکستری اعمال می شود. در اینجا ما الگوریتم SRF را بر تصویر حاصل از اعمال LBP اجرا می کنیم که به نوعی نوآوری ما به حساب می آید (SRF Based LBP). نتایج حاصل از آزمایشات نشان می دهد این نوآوری، افزایش حدود ۶ درصدی میانگین کلی بازیابی را در پی داشته است.

۳-۲-۴ هیستوگرام رنگ HSV

رنگ یکی از متداولترین و تعیین کننده ترین ویژگی دیداری در زمینه بازیابی تصویر است؛ چرا که نسبت به تغییرات مربوط به اندازه، جهت و دورنمایی و اغتشاش تصویر پایدار است [۱۲]، همچنین حجم کم، پیچیده نبودن محاسبات و سادگی پیاده سازی از دیگر خصوصیات فضای رنگ به حساب می آید. عمومی ترین روش در بازیابی رنگ تصاویر استفاده از هیستوگرام رنگ است، که از مهمترین تکنیک های

بازیابی بر اساس تصویر [۱۳] است. این هیستوگرام بصورت توزیع تجمعی از زیر فاصله های مرتبط با هر پیکسل تشکیل می شود. که هر یک احتمال وقوع پیکسل رنگ مشخص در تصویر را ارائه می دهد.

اولین گام در تشکیل هیستوگرام رنگ، مشخص کردن فضای رنگ است. به صورت عمومی فرمت های مختلف تصاویر همچون BMP، GIF و ... فضای رنگ RGB را در خود دارند. در صورت انتخاب این فضا محاسبات کمتری را خواهیم داشت اما مسئله قابل توجه یکنواخت نبودن از جهت ادراکی است. به همین دلیل ما در روش پیشنهادی فضای RGB را به فضای HSV تبدیل کردیم. از مزایای فضای HSV را می توان اینگونه بیان کرد.

تناسب آن با ادراک رنگ توسط انسان، است و مستقل بودن مولفه های نوع رنگ و روشنایی که قابلیت برتری دادن هریک از مولفه ها را به ما می دهد.

ابعاد تشکیل دهنده این فضا، رنگ را بر اساس رده (H) یا طول موج، درجه اشباع رنگ (S) و درجه روشنایی آن (V) تعریف می کنند. فضای مذکور یک فضای مخروطی است، که در آن:

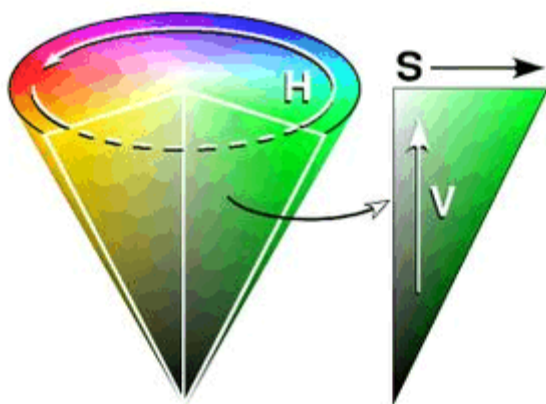
۱. طول موج رنگ برابر زاویه ی رنگ در برش دایره مخروط می باشد و در فاصله $[0, 2\pi]$ تعریف می شود، به صورتی که قرمز در زاویه صفر درجه، سبز در زاویه $\frac{2\pi}{3}$ و آبی در زاویه $\frac{4\pi}{3}$ قرار گرفته که در انتها، دوباره به قرمز در زاویه 2π باز می گردد.

۲. درجه روشنایی رنگ معادل محور مرکزی مخروط است.

۳. درجه اشباع رنگ برابر فاصله نقطه روی دایره تا محور مرکزی است به گونه ای که در نزدیکی محور، میزان غلظت کمتر است و رنگ مزبور به دسته رنگهای خاکستری می پیوندد.

H مولفه اصلی مورد استفاده در سیستم های بازیابی است چرا که نسبت به تغییرات جهت تصویربرداری از شیئی یا منظره پایدارتر است. اما با این حال به تغییرات روشنایی به شدت حساس می باشد. جهت تشکیل هیستوگرام رنگ پس از این فضا به طور خطی کوانتیزه می شود. به دلیل اهمیتی که فضای H نسبت به سایر مولفه ها دارد.

این مولفه به ۱۶ بازه و دو مولفه دیگر هر کدام به ۴ بازه کوانتیزه می شوند. هیستوگرام رنگ تصویر با شمردن نقاطی که در هر بازه قرار می گیرند، محاسبه شده و می گیرند، محاسبه شده و به تعداد کل نقاط تصویر نرمالیزه می شود. بردار حاصل از نمایه سازی هر تصویر با استفاده از ویژگی هیستوگرام رنگ HSV، یک بردار ۲۵۶ بعدی است.



شکل ۴-۹ فضای رنگ HSV

۴-۳ ترکیب ویژگی ها

پس از مرحله استخراج ویژگی باید نحوه کنار هم قرار گرفتن آن ها مشخص شوند. روش های مختلفی برای ترکیب ویژگی وجود دارد که راه حل اولیه کنار هم قرار دادن تمام ویژگی ها در یک ردیف به عنوان یک ماتریس واحد و اندازه گیری فاصله با سایر تصاویر است. این کار دارای چند مشکل است؛ اکثر ویژگی ها اندازه ای متفاوت از هم دارند لذا در صورت کنار هم گذاشتن، ویژگی که اندازه بیشتری دارد تأثیر بیشتری خواهد گذاشت در حالیکه لزوماً اندازه بیشتر به معنای کارایی بالاتر و اهمیت بیشتر نیست، همچنین نمی توانیم تأثیر برخی ویژگی ها را با توجه به عملکرد بهتری که دارد نسبت به دیگر ویژگی ها افزایش دهیم.

راه حل بهتری که به کار می آید؛ وزن دهی و تعیین فاصله مجزا برای هر یک از ویژگی هاست. در اینجا ما دو ویژگی داریم که آنها را با نام های به ترتیب $F1$ و $F2$ نام گذاری می شود. برای هر ویژگی وزن خاصی لحاظ می شود که در اینجا $w1$ و $w2$ نامیده می شود. بدین ترتیب با توجه به ضریبی که $w1,2$ به خود میگیرند تأثیر ویژگی ها تغییر می کنند می توان با تغییرات مختلف بهترین وزن را بدست آورد.

در روش پیشنهادی ما، وزن ویژگی لبه 0.4 و ویژگی رنگ 0.6 انتخاب شده اند.

F =feature vector

Q =Query vector

$$D2 = \text{distance}(F2, Q2)$$

$$D1 = \text{distance}(F1, Q1)$$

$$D_{\text{total}} = (w1 \times D1) + (w2 \times D2)$$

(۴-۶)

۴-۴ معیار شباهت

جهت ارزیابی شباهت بین تصاویر، فاصله بین ویژگی های تصاویر را محاسبه می کنیم. در این مقاله از دو معیار اقلیدسی و کانبرا طبق روابط ۷-۴ و ۸-۴ استفاده شده است که با توجه به آزمایشات صورت گرفته، معیار کانبرا نتایج بهتری را همراه داشته است. دلیل آن هم می توان به این نکات اشاره کرد که معیار کانبرا بر خلاف اقلیدسی در ارزیابی شباهت در بردارهایی که صفرهای زیادی دارند بهتر عمل می کند. علاوه بر این اختلاف بین ویژگی ها در صورت کسر نسبت به اندازه آنها در مخرج نرمالیزه می شود، در نتیجه انتظار می رود نتایج دقیق تری به همراه داشته باشد.

$$d_2(Q,T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^2 \right)^{\frac{1}{2}} \quad (7-4)$$

$$dc(\underline{F}, \underline{G}) = \sum_{i=1}^n \frac{|f_i - g_i|}{|f_i + g_i|} \quad (8-4)$$

۵-۴ جست وجو و بازیابی

در غالب سیستم ها برای جست و جوی تصاویر و پیدا کردن تصاویر مشابه از جست و جوی نزدیکترین همسایه استفاده می شود. الگوریتم طبقه بندی KNN¹ برای هر تصویر پرس و جو بین مجموعه تصاویر آموزشی به دنبال تصاویر مشابه می گردد. اگر سیستم از n ویژگی استفاده شود آنگاه تصاویر را به صورت یک بردار در فضای n بعدی در نظر می گیرد. به طوری که هر تصویر در این فضا معادل یک نقطه است در این فضا K همسایه تصویر پرس و جو از تصاویر آموزشی را جست و جو می کند و در بین این همسایه ها برچسب کلاسی که در اکثریت باشد را به عنوان برچسب دسته ی این تصویر پرس و جو پیش بینی می کند. k می تواند عددی از پیش تعیین شده باشد و یا کاربر آن را مقاردهی کند [۴۹].

اولین کار برای استفاده از k-NN یافتن معیاری برای شباهت یا فاصله بین صفات در داده ها و محاسبه آن است. هنگامی که فاصله بین نمونه های مختلف را توانستیم اندازه گیریم، یکی از مشکلات این روش

¹ K Nearest Neighbour

محاسبه فاصله تصاویر با یکدیگر است اگر فاصله تصاویر به خوبی اندازه گیری نشود الگوریتم کارایی خوبی نخواهد داشت. در سیستم پیشنهادی فاصله بین دو نمونه با استفاده از فاصله اقلیدسی بصورت زیر محاسبه می شود:

$$d_2(Q,T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^2 \right)^{\frac{1}{2}} \quad (9-4)$$

که $Q = \{Q_1, \dots, Q_N\}$ و $T = \{T_1, \dots, T_N\}$ که و در واقع بردارهای ویژگی تصویر هدف و جست و جو می باشند.

۴-۶ روش ارزیابی

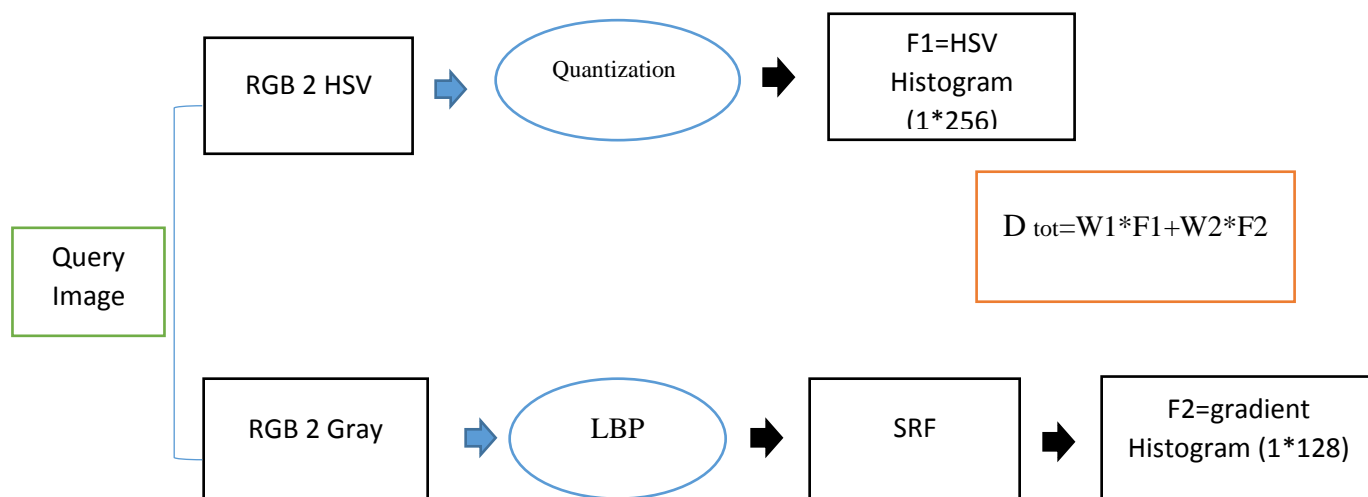
رایج ترین معیارهای ارزیابی در بازیابی تصاویر معیارهای دقت و فراخوانی هستند. این دو معیار به شکل یک گراف که در آن دقت بر حسب فراخوانی رسم شده است ارائه می شوند که به گراف PR مشهور است [۳۹]. نرخ فراخوانی و دقت به طور گسترده برای اندازه گیری عملکرد بازیابی مورد استفاده قرار می گیرد. اساساً بر پایه تطبیق قطعی است. در این روش مجموعه داده ای به مجموعه باینری، برطبق مرتبط بودن یا نامرتب بودن با تصویر پرس و جو، تبدیل می شود.

نرخ دقت (P) و فراخوانی (R) به صورت زیر تعریف می شود.

$$P = \frac{r}{n_1} = \frac{\text{number of relevant retrieved items}}{\text{number of retrieved images}}$$

$$R = \frac{r}{n_2} = \frac{\text{number of relevant retrieved items}}{\text{total number of relevant images in DB}}$$

شکل ۴-۱۰ روند نما راهکار پیشنهادی را نشان می دهد.



شکل ۴-۱۰ روند نما راهکار پیشنهادی

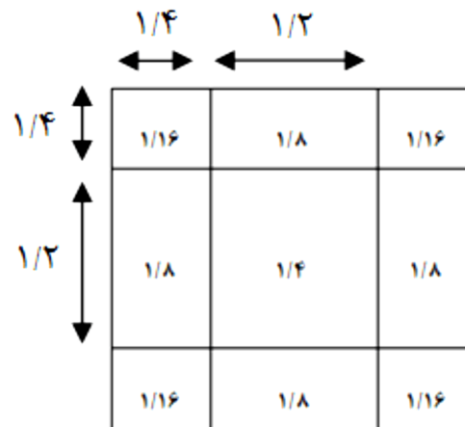
۴-۷ روش پیشنهادی مبتنی بر ادغام آماری نواحی

ناحیه بندی در بازیابی تصاویر، مسئله قابل توجهی است از این نظر که هر تصویر به بخش های با بار معنایی متفاوت تقسیم می شود به طور مثال در شکل ۴-۱۰ تصویر مورد نظر به چهار قسمت تقسیم می شود. هر ناحیه در این شکل مفهوم خاص خود را دارد. ناحیه سمت راست بالا مربوط به بخش ساختمان می شود، سمت راست پائین اسب را در بطن خود دارد، همینطور بقیه نواحی معنا مجزا از هم دارند. در صورتیکه به صورت کلی استخراج ویژگی شوند. مفهوم گرفته شده ترکیبی از همه نواحی خواهد شد و قاعدتا تمام جزئیات را پشتیبانی نمی کند.



شکل ۴-۱۰ تصویر ناحیه بندی شده

به منظور بالا بردن کارائی سیستم های بازیابی تصاویر؛ قسمت اصلی را مرکز تصویر قرار دادند [۵۰]. در شکل ۴-۱۱ نمونه ای از ناحیه بندی مشاهده می شود که اعداد داخل شکل ضرایب اهمیت هرناحیه و اعداد بیرون از شکل نسبت اندازه ناحیه به اندازه تصویر را مشخص می کند.



شکل ۴-۱۱ نحوه ناحیه بندی تصاویر

هر چند روش گفته شده باعث بهبود نسبی عملکرد می شود اما در تمام تصاویر این مرکزیت ذکر شده، صدق نمی کند به عنوان مثال در شکل ۴-۱۲ ناحیه اصلی تصویر به طور کامل در مرکز تصویر نمی باشد.



شکل ۴-۱۲ عدم محوریت مرکز در تصویر

با توجه به مطالب آورده شده در این بخش نیاز به بخش بندی به طور اتوماتیک جهت پیدا کردن قسمت اصلی تصویر به عنوان یک اولویت به ذهن خطور می کند.

۴-۷-۱ ادغام آماری نواحی

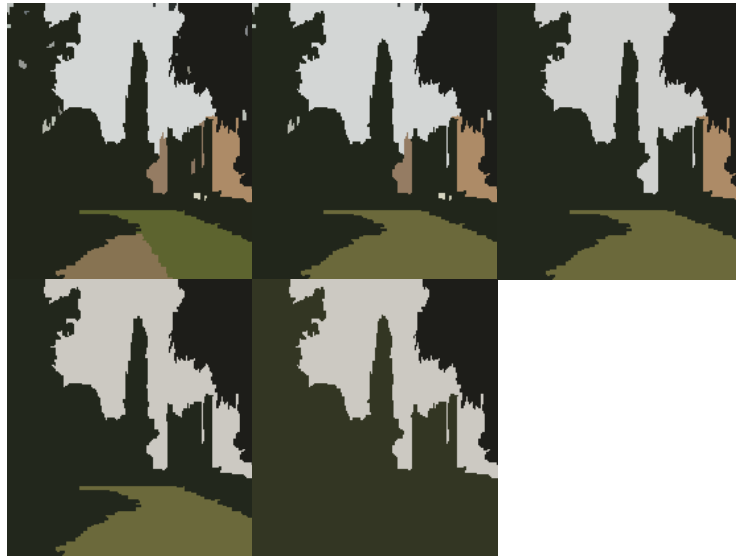
ما برای رسیدن به هدف خود یعنی ناحیه بندی مرتبط با هر تصویر از روش ادغام آماری نواحی SRM^۱ [۴۵] استفاده کرده ایم. الگوریتم SRM به این شکل عمل می کند که بر اساس اطلاعات آماری هر ناحیه، هر تصویر را به بخش های شبیه به هم تقسیم می کند. تعداد نواحی در اختیار کاربر می تواند باشد، در صورتیکه کل تصویر دارای خصوصیات نزدیک به هم باشد بخش بندی صورت نمی گیرد. برای روشن شدن این مطلب به بررسی مثال عملی می پردازیم. در شکل ۴-۱۳ تصویری که قرار است ناحیه بندی بر آن اعمال شود را ملاحظه میکنیم، چند قسمت در این تصویر قابل اشاره است؛ بخش سبز تصویر، قسمت شامل دیوار و آسمان موجود در تصویر. از الگوریتم ناحیه بندی انتظار داریم که نواحی ذکر شده را در خروجی به نمایش گذارد.



شکل ۴-۱۳ تصویر ورودی به الگوریتم SRM

در شکل ۴-۱۴ خروجی حاصل از روش SRM را ملاحظه می کنید. در اینجا الگوریتم را بر مبنای بخش بندی در ۵ سطح برنامه ریزی کرده ایم. بخش های ذکر شده از تصویر اصلی در اینجا مجزا شده اند و این به معنای عملکرد قابل قبول در ناحیه بندی تصاویر است. نکته دیگر این است که در هر سطح از بخش بندی نواحی به دست آمده در سطح قبلی، ادغام شده و به ناحیه با خصوصیات نزدیک به خود می پیوندند. در آخر هم به دو یا سه ناحیه تقسیم می شوند.

¹ Statistical region merging

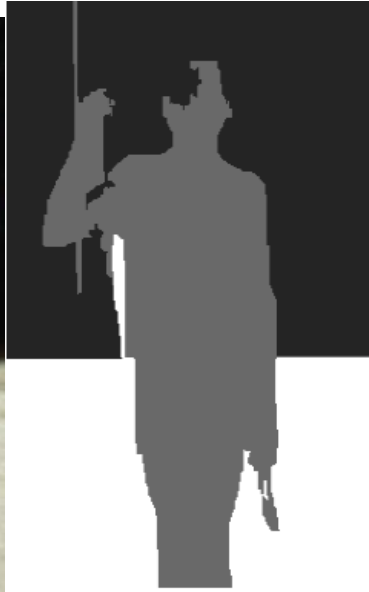


شکل ۴-۱۴ خروجی حاصل از ناحیه بندی به روش SRM

۲-۷-۴ بازیابی تصاویر با استفاده از ادغام آماری نواحی

پس از به دست آمدن نواحی تصویر با استفاده از روش SRM، درصدد استفاده بهتر از ناحیه بندی در بازیابی تصاویر خواهیم بود. هدف ما در این کار جدا کردن یک قسمت از تصویر به عنوان ناحیه اصلی می باشد. جهت برآورده شدن این منظور نیاز به یک مرحله پردازش اطلاعات حاصل شده از نواحی بدست آمده می باشد. در این قسمت باید نواحی که مرکزیت بیشتر و مساحت بیشتری را دارند را در اولویت به عنوان قسمت اصلی قرار داد. در صورتیکه تصویری یافتیم تقریباً یکدست باشد و یک ناحیه تمام تصویر را پوشش دهد، از بخش بندی به دو بخش اصلی و پس زمینه خودداری می کنیم. این کار در مقابل بخش بندی ثابت قرار دارد و باعث افزایش کارایی سیستم خواهد شد.

نمونه های از خروجی حاصل از بخش بندی به عنوان قسمت اصلی تصویر در شکل ۴-۱۱ آورده شده است. در تصاویر حاصل از اعمال الگوریتم، بخش سفید به صورت اتوماتیک به عنوان پس زمینه انتخاب شده است.



رد ۱



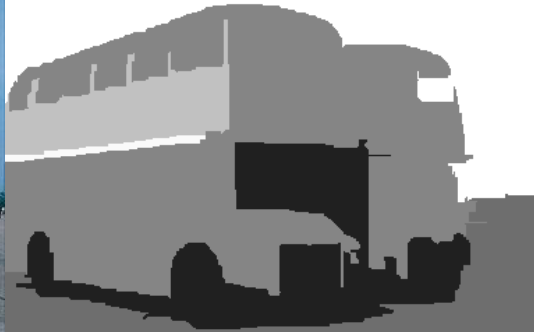
رد ۲



رد ۳



رده ۴



رده ۵



رده ۶

(ب)

(الف)

شکل ۴-۱۳ (الف) تصاویر اصلی (ب) تصاویر حاصل از بخش بندی به روش ادغام آماری نواحی

فصل پنجم

نتیجہ تجربی

۱-۵ پایگاه داده

جهت ارزیابی سیستم های بازیابی تصاویر نیاز به بررسی آنها در چارچوب پایگاه داده تصویری می باشد. اما پایگاه داده مرتبط با بازیابی تصاویر شامل چند ویژگی می باشد؛ به طور مثال در ارزیابی شناسایی چهره نیاز به تصویری از چهره افراد متفاوت می باشد در سیستم های بازیابی هم نیاز به وجود کلاس های معنایی متفاوت داریم که اولاً مشخصه کلاس نامگذاری شده در آن وجود داشته باشد، همچنین همبستگی و تصاویر نزدیک به هم در یک کلاس به اندازه کافی باشد و اختلاف اطلاعات بصری در کلاس های مختلف از حد معمول کمتر نباشد و باعث تداخل شناسایی نشود.

مجموعه تصویری که ما در این پایان نامه با در نظر گرفتن مطالب گفته شده مورد استفاده قرار گرفته است پایگاه داده Wang یکی از مجموعه تصاویر مورد قبول در زمینه بازیابی تصاویر می باشد که اکثر مقالات معتبر بر مبنای آن به ارزیابی سیستم خود می پردازند. این مجموعه برای اولین بار توسط Wang و دیگران در سیستم Simplicity [۴۶] معرفی و بکار گرفته شد.

پایگاه داده Wang مشتمل بر ۱۰۰۰ تصویر در ۱۰ کلاس معنایی متفاوت در اندازه های ۳۸۴*۲۵۶ و ۳۸۴*۲۵۶ می باشد. در جدول ۱-۵ جزئیات آن اشاره شده، همچنین شکل ۱-۵ نمونه از تصاویر آن را نشان می دهد.

جدول ۱-۵ کلاس های معنایی پایگاه داده Wang

| ردیف | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | ۹ | ۱۰ |
|-------------|--------|------|---------|--------|----------|-----|----|-----|---------|-----|
| کلاس معنایی | آفریقا | ساحل | ساختمان | اتوبوس | دایناسور | فیل | گل | اسب | کوهستان | غذا |

از نکات مهم این مجموعه وجود کلاس های متفاوت است. همانطور که در شکل ۱-۵ می بینید در بعضی کلاس ها تصاویر را به لحاظ شناسایی به سختی می توان طبقه بندی کرد مانند کلاس آفریقا، کوهستان، ساختمان. از طرف دیگر تعدادی از کلاس ها شامل تصاویر با جزئیات کمتر و شباهت بیشتر هم وجود دارد. شناسایی این کلاس ها راحت تر از دسته قبلی است از جمله این رده می توان به کلاس های دایناسور، اتوبوس و گل اشاره کرد. در مجموع دلیل این طبقه بندی در نظر گرفتن دقت سیستم مورد آزمایش است. در صورتیکه سیستمی نتواند کلاس های آسانتر را طبقه بندی کند از متوسط پائینتر در نظر گرفته می شود، همچنین روش هایی دقت بالایی دارند که علاوه بر رده های راحت تر در کلاس های پیچیده تر که جزئیات بیشتری دارند کارایی خود را نشان دهد

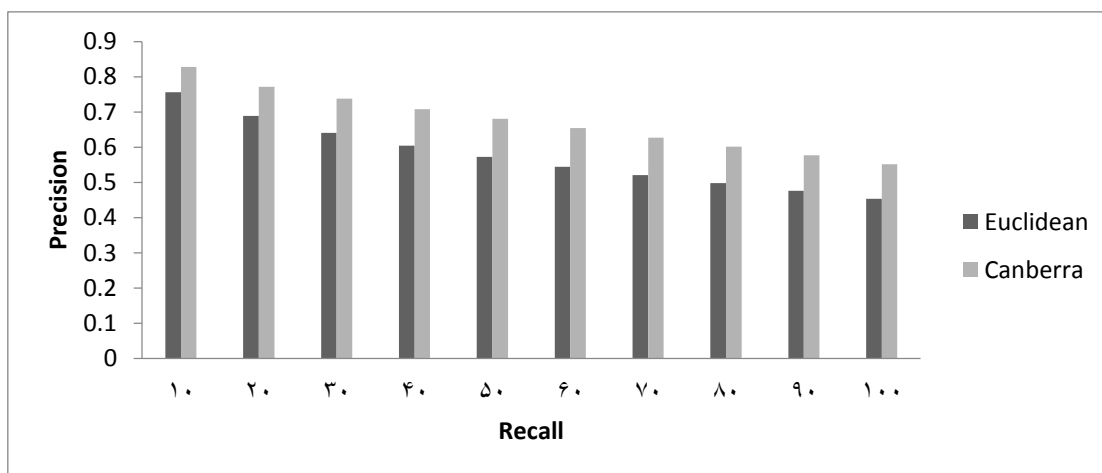


شکل ۵-۱ نمونه از تصاویر پایگاه داده Wang

۵-۲ نتایج آزمایش

همانطور که در فصل چهارم به معرفی سیستم پیشنهادی پرداختیم، روش ما مبتنی بر ترکیب دو ویژگی رنگ و لبه است. ویژگی رنگ از هیستوگرام رنگ HSV تشکیل شده و ویژگی لبه شامل الگوریتم SRF بر پایه الگوهای باینری محلی است. در ترکیب ویژگی ها از روش ضریب دهی استفاده کرده ایم. ضریب ویژگی رنگ ۰.۶۵ و ضریب ویژگی لبه ۰.۳۵ می باشد. ارزیابی شباهت مورد استفاده، معیار تطابق کانبرا است که جهت مقایسه بهتر، فاصله اقلیدسی هم آورده شده است.

همانطور که در فصل ۴ عنوان شد نمودار دقت-فراوانی (Recall-Precision) یکی از مشخصه های مهم در ارزیابی بازیابی تصاویر به شمار می آید. نمودار مورد نظر میزان دقت سیستم را بر اساس تعداد تصاویر فراخوانی شده نشان می دهد. به اینصورت که محور پائین نمودار بیانگر تعداد تصاویر فراخوانی شده و ستون سمت چپ هم مشخص کننده دقت می باشد. در شکل ۵-۲ نمودار دقت-فراوانی با دو معیار آورده شده است. با افزایش تعداد تصاویر مطابق انتظار، نرخ بازیابی کاهش می یابد. این امر در تمام سیستم های بازیابی اتفاق می افتد و دلیل این پدیده، تطابق هر تصویر با تعداد مشخصی از تصاویر هر مجموعه است به طوریکه با افزایش تعداد تصاویر بازیابی شده، نرخ دقت بازیابی کاهش می یابد. در نمودار ارائه شده مشاهده می شود معیار تطابق کانبرا عملکرد بهتری دارد.



شکل ۵-۲ نمودار دقت-فراوانی سیستم پیشنهادی

با توجه به استفاده از ویژگی SRF مبتنی بر الگوهای باینری محلی، نیاز است دقت این ویژگی را با روش های موجود که در فصل ۴ آورده شد، مقایسه کنیم. جدول ۳-۱ مقایسه بین روش ها را نشان می دهد. با بررسی در خواهیم یافت که روش ما به لحاظ دقت و سرعت بهترین نتیجه را در بر دارد.

جدول ۱-۳ مقایسه روش SRF و سایر روش های موجود

| SRF based LBP | BDPSP | DHSE | CCM | کلاس معنایی |
|---------------|-------|-------|-------|-------------|
| ۵۲.۴۰ | ۲۸.۸۰ | ۳۳.۱۰ | ۵۴.۷۰ | آفریقا |
| ۶۰.۱۰ | ۳۴.۷۰ | ۵۳.۱۰ | ۵۲.۱۰ | ساحل |
| ۵۶.۷۰ | ۳۰.۷ | ۵۵.۸۰ | ۴۳.۱۰ | ساختمان |
| ۹۳.۵۰ | ۵۸.۸ | ۸۳ | ۸۳.۲۰ | اتوبوس |
| ۹۴.۳۰ | ۶۰.۹ | ۲۶.۵۰ | ۹۴.۱۰ | دایناسور |
| ۴۶.۰ | ۲۹.۳ | ۵۳.۳۰ | ۳۷.۳۰ | فیل |
| ۸۸.۶۰ | ۸۳.۱۰ | ۱۹.۷۰ | ۸۱.۹۰ | گل |
| ۷۰.۸۰ | ۳۸.۲ | ۷۴.۹۰ | ۶۶.۲۰ | اسب |
| ۳۲.۵۰ | ۲۲.۷ | ۳۲.۷۰ | ۴۰.۷۰ | کوهستان |
| ۵۲.۵۰ | ۳۰.۴ | ۲۹.۲۰ | ۴۲.۳۰ | غذا |
| ۶۴.۴۷ | ۴۱.۷۶ | ۴۶.۲۰ | ۵۹.۵۶ | میانگین |

۳-۵ مقایسه سیستم پیشنهادی با کارهای پیشین

به منظور مقایسه سیستم پیشنهادی از همان پایگاه داده معرفی شده استفاده کرده ایم که هماگونه که اشاره شد پایگاه معمول، در ارزیابی سیستم های بازیابی تصاویر می باشد. همچنین متوسط دقت بازیابی به ازای ۲۰ تصویر فراخوانی، بررسی شده است.

روش پیشنهادی با دو سیستم بازیابی تصاویر مقایسه شده است. در [۴۷] وادیول و دیگران، از ویژگی رنگ و ماتریس هم رخداد خاکستری^۱، جهت استخراج ویژگی بافت استفاده شده است و در [۴۸] لین و دیگران، از ماتریس هم رخداد رنگ^۲ (Ccm) بعنوان توصیف گر اطلاعات مکانی استفاده می کند؛ Ccm به اینصورت عمل می کند، ارتباط مکانی هر پیکسل را با پیکسل های همسایه بررسی می کند و نحوه تغییر مقدار شدت روشنایی پیکسل ها را نسبت به پیکسل اصلی در نظر می گیرد و با ۷ الگوی اصلی

¹ Gray level co-occurrence

² Color co-occurrence matrix

مقایسه می کند و در نهایت ماتریس هم رخداد آن را برای سه رنگ اصلی بدست می آورد. همچنین مقاله مذکور از هیستوگرام رنگ با استفاده از خوشه یابی K-mean و اختلاف بین پیکسل های الگوی اسکن، هم بهره برده است. در جدول ۳-۲ همانطور که ملاحظه می کنید در ۱۰ کلاس معنایی، میانگین دقت بازیابی آورده شده است در آخرین سطر هم میاگین کلی اشاره شده است.

با دقت در جدول ۳-۱ مشاهده می شود که میانگین دقت بازیابی سیستم پیشنهادی از سایر سیستم با ۷۷.۱۵٪ بالاتر است و در ۷ کلاس معنایی نسبت به دو روش دیگر عملکرد بهتری را نشان می دهد. می توان نتیجه گرفت سیستم پیشنهادی بهترین نتایج را در قالب پایگاه داده موردنظر ارائه می دهد.

جدول ۳-۲ مقایسه روش پیشنهادی و سایر سیستم ها

| سیستم پیشنهادی | لین و دیگران [48] | وادبول و دیگران [47] | کلاس معنایی |
|----------------|-------------------|----------------------|-------------|
| ۷۹.۳۵ | ۶۸.۳ | ۷۸.۲۵ | آفریقا |
| ۵۴.۷۰ | ۵۴.۰ | ۴۴.۲۵ | ساحل |
| ۶۹.۷۵ | ۵۶.۲ | ۵۹.۱۰ | ساختمان |
| ۹۴.۲۵ | ۸۸.۸ | ۸۶.۰۵ | اتوبوس |
| ۹۹.۱۵ | ۹۹.۳ | ۹۸.۷۰ | دایناسور |
| ۵۹.۲۵ | ۶۵.۸ | ۵۹.۰۰ | فیل |
| ۹۴.۸۵ | ۸۹.۱ | ۸۵.۳۵ | گل |
| ۹۲.۱۰ | ۸۰.۳ | ۷۴.۹۵ | اسب |
| ۴۵.۵۵ | ۵۲.۲ | ۳۶.۵۵ | کوهستان |
| ۸۲.۵۰ | ۷۳.۳ | ۶۴.۴۰ | غذا |
| ۷۷.۱۵ | ۷۲.۷۰ | ۶۸.۶۶ | میانگین |

۵-۴ ارزیابی سرعت سیستم پیشنهادی

در کنار نرخ دقت بازیابی تصویر، عملکرد سیستم در سرعت جستجو هم، معیار مهمی در ارزیابی به حساب می آید. سیستم پیشنهادی با استفاده از نرم افزار Matlab 2011 و کامپیوتر شخصی با مشخصات پردازنده Intel Core i5 , 2.3 GHz و حافظه 4 GB مورد ارزیابی قرار گرفت و زمان مورد

نیاز برای پرس و جو هر تصویر حدود ۰.۷ ثانیه (0.7 s) بدست آمد، که یکی از سریعترین سیستم های بازیابی تصاویر مبتنی بر محتوا می باشد. مقایسه سرعت روش پیشنهادی با مقاله [۴۸] Lin و دیگران، نتایج قابل توجهی به همراه دارد. در سیستم پیشنهادی [۴۸] با توجه به اینکه ویژگی رنگ از الگوریتم K-mean با تکرار زیاد استفاده می شود و پس از استخراج ویژگی با استفاده از الگوریتم انتخاب ویژگی^۱، بهترین ویژگی ها را انتخاب می کند، در مجموع سرعت پرس و جو یک تصویر کاهش می یابد و به میزان ۸۲ ثانیه می رسد که در مقابل ۰.۷ ثانیه زمان مورد نیاز این مقاله، نشان دهنده سرعت بسیار خوب روش پیشنهادی ما دارد.

۵-۵ نمونه ای از نتایج

از آنجا که ارزیابی عملکرد سیستم های بازیابی تصاویر نیاز به مشاهدات بصری دارد. در این قسمت چند نمونه از خروجی سیستم را نشان می دهیم. دقت شود که اولین تصویر از سمت چپ بعنوان تصویر جستجو (Query) انتخاب شده است و تصاویر دیگر نتایج همان تصویر جستجو خواهند بود و در زیر هر تصویر تعداد تصاویر صحیح آورده شده است.



نمونه از کلاس ۱ - دقت ۱۴ از ۱۴

¹ Feature selection



نمونه از کلاس ۲ - دقت ۱۱ از ۱۴



نمونه از کلاس ۳ - دقت ۱۴ از ۱۴



نمونه از کلاس ۴ - دقت ۱۳ از ۱۴

Query



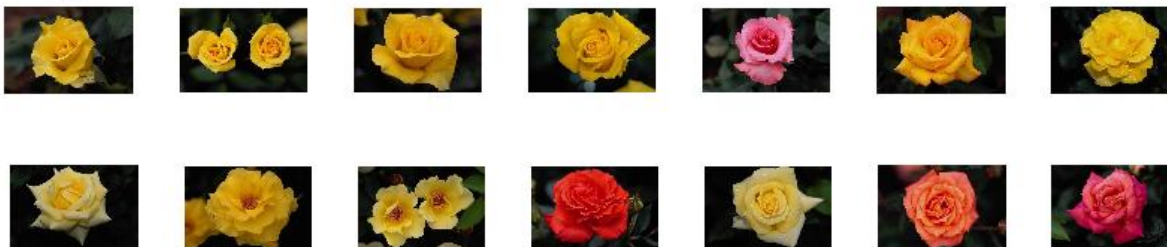
نمونه از کلاس ۵ - دقت ۱۴ از ۱۴

Query



نمونه از کلاس ۶ - دقت ۱۰ از ۱۴

Query



نمونه از کلاس ۷ - دقت ۱۴ از ۱۴

Query



نمونه از کلاس ۸ - دقت ۱۳ از ۱۴

Query



نمونه از کلاس ۹ - دقت ۱۰ از ۱۴

Query



نمونه از کلاس ۱۰ - دقت ۱۴ از ۱۴

منابع:

[1] Y.Liu- D.Zhang- G.Lu and W.Ma- "A survey of content-based image retrieval with high-level semantics-" Pattern Recognition-40- PP. 262-282- 2006.

[۲] سریزدی. سعید و نظام‌آبادی‌پور. حسین، "روش جدیدی برای نمایه‌سازی تصاویر رنگی مبتنی بر فشرده سازی و مورفولوژی ریاضی"، نشریه علمی پژوهشی انجمن کامپیوتر ایران، علوم و مهندسی کامپیوتر، ۱۳۸۴، مجلد ۳، شماره ۳(الف)، صفحات ۸۱-۹۴.

[3] H. Nezamabadi-pour and S. Saryazdi, "Object-based image indexing and retrieval in DCT domain using clustering techniques", World Academy of Science, Engineering and Technology, vol. 3, pp. 207-210, 2005.

[4] H._ Nezamabadi-pour and E._ Kabir- "Image retrieval using block-based color histogram and local distribution of intensity gradient", 2nd Int. Symposium on Telecommunication- IST2003- Isfahan- Iran- PP. 176-180- 2003.

[5] Y.Rui and T.S.Huang and S.F.Change, "Image Retrieval:Current Techniques,Promising Directions ,and Open Issues ",Journal off Visual Communication and Image Representation and Image Representation ,1999.

[6] R.Datta,D.Joshi,J.Li_ and J.Z. Wang, "Image Retrieval:Ideas, Influences, and Trends Of the New Age", ACM Computing Survey, Vol.40,pp.5:1-60,2008.

[7] Lon g F;Zhang H.and Dagan Feng D., "Fundamentals of content-based image retrieval ,in Multimedia I nformation Retrieval and Management Technological Fundamentals and Application ", Springer-verlage,pp.26-1,2003.

[8] Li X,: Chen S.C:M.L.Shyu and Furht B., "Image Retrieval by color,Texture,and Spatial Information", in 8th International Conference on Distributed Multimedia Systems (DMS'2002),San Francisco Bay,California ,USA ,2002,pp.,159-152.

[9] Muwei Jian, Junyu Dong, Ruichun Tang, "Image Combining Color, Texture and Region with Objects of user's Interest for Content-Based Image Retrieval", Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing,IEEE Computer Society, 2007.

[10] Einarsson S.H., "Data structures for intermediate search in the Eff 2image retrieval system" , Reykjavik University, thechincal report, 2004.

[11] Gevers Th. And Smeulders A.W.M., " Image Search Engines, An Overview," The International Society for Optical Engineering(SPIE) ,vol.VIII,pp, 337-327,2003.

[12] Einarsson S.H. ; Gretarsdottir R.Y. ; Jonsson B.P and Amsaleg L.,“ The EFF 2 Image retrieval System Prototype”, in ASTED Intel .Conf . on Databases and Application (DBA), Innsbruck, Austria, 2005.

[13] M.J.Swaim,and D.H.Ballard, “color indexing ”, International Journal of computer vision ,vol,7,no,1,pp.11-32,1991.

[14] J.R.Smith, S.F.Change, “tools and techniques for color image retrieval”, Proce of the SPIE conference on the Storage and Retrieval for Image and Video Database IV,pp.426-437,San Jose ,CA,USA,46 BIBLIOGRAPHY,1996.

[15] M.A.Stricker,and A.Dimai, “color indexing with weak spatial constraints”, Proc,of the SPIE conference on the storage and Retrieval for Image and video Databases IV,pp.29-40,San Diego,CA,USA,February 1996.

[16] G.Pass,R.Zabih,and J.Miller, “companing images using color coherence”, Proc,of the Fourth Acm Multimedia conference .pp.65-74,New York,NY ,USA,November 1996.

[17] J.Huaang, “Color-Spatial Image Indexing and Application”.Ph.D.thesis,comell Univ.,1998.

[18] H.Nezamabadi-pour,and E.Kabir , “Image Retrieval Using Histograms of Unicolor Blocks and Directional Changes in Intensity Gradient”, Accepted for Publication in Pattern Recognition Letters,2004.

[۱۹] ح. نظام آبادی پور، ا. کبیر و س. سریزدی، "بازیابی تصویر با استفاده از اطلاعات رنگ و شکل"، دومین کنفرانس بینایی ماشین و پردازش تصویر ایران. جلد دوم، صفحات ۲۲۶ تا ۲۳۲، ۱۳۸۱.

[۲۰] ح. نظام آبادی پور، ا. کبیر و س. سریزدی، "بازیابی تصویر بر اساس هم‌رخدادی رنگها در بلوکهای لبه"، هشتمین کنفرانس انجمن کامپیوتر ایران. صفحات ۳۵۹ تا ۳۶۴، ۱۳۸۱.

[21] H. W. YOO, S. H. Jung, D. S. Jang and Y. K. Na.”Extraction of Major object Features Using VQ Clustering for Content-based image retrieval ,” Pattern Recognition, vol.35,pp,1115-1126,2002.

[22] M. G. Albanesi, S. Bandelli and M. Ferreti,“Quantitative Assessment of Qualitative Color Perception in Image Database Retrieval ”, IEEE Int. Conf. on Image Analysis and Processing , pp,410-415,2001.

[23] H. W. YOO, , D. S. Jang, S. H. Jung and J. H.Park,“ Visual Information Retrieval System via Content-based Approach, Pattern Recognition” , vol.35,pp,479-769,2002.

[24] J. R. Smit, S. F. Chang,“Tools and Techniques for color image retrieval” , Storage and Retrieval for Image and Video Database IV, Proc. SPIE ,vol.2670,426- 437,1996.

[25] Z. Huang, P. K. Chan, W. Y. NG, S. Yeung, "Content-Based Image Retrieval Using Color Moment and Gabor Texture Feature", Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, v.2, pp. 719 - 724 ,2010.

[26] Brigale L., Kokare M., Doye D., "Colour and Texture Features for Content Based Image Retrieval", IEEE Proceeding of International Conference on Computer Graphics, Imaging and Visualisation, pp. 146-149, 2006.

[۲۷] منصورى.زهرا و جم زاده.منصور، "مرورى بر سيستم هاى بازيابى در كاربردهاى عام"، ۱۳۸۷.

[28] J.Huang, S.R.Kumar, N.Mitra, W.Zhu and R.Zabih, "Image indexing using color correlograms", IEEE Con. Computer Vision and Pattern Recognition, pp.762-768,1997.

[29] Young Deok Chun, Nam Chul Kim, Member, IEEE, and Ick Hoon Jang, Member, IEEE, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 10, NO. 6, OCTOBER 2008.

[30] F Long, H Zhang, DD Feng, "Fundamentals of content-based image retrieval", in Multimedia Information Retrieval and Management - Technological Fundamentals and Applications," Springer-Verlag, pp. 1-26_F2003.

[31] Li X. ; Chen S. ; Shyu M. and Furht B., "An Effective Content-Based Visual Image Retrieval System," in . • th IEEE Computer Society International Computer Software and Applications Conference (COMPSAC), Oxford, .uu., pp. 914-919.

[32] Deselaers Th., "Features for Image Retrieval," 2003.

[33] Manesh Kokare, P. K. Biswas, B. N. Chatterji, "Texture_ Image Retrieval Using New Rotated Complex Wavele_ Filters",_ IEEE_ TRANSACTIONS_ON_SYSTEMS, , VOL_35,_ NO._ 6,_ DECEMBER_2005.

[34] j.A.k.Jain and S. K. Bhattachavjee, "Address B lock Location on Evelopes using Gabor Filters," Pattern Recognition , vol ,no.12,pp.1449-1477,1992.

[35] N. G. Kingsbury, "Image Processing with complex wavelet" , Phil. Trans. Roy. Soc. London A, vol.357,pp.2543-2560,Sep.1999.

[36] Jianhua Wu, Zhaorong Wei, Youli Chang, "Color and Texture Feature For Content Based Image Retrieval," International Journal of Digital Content Technology and its Applications Volume 4, Number 3, June 2010.

[37] Bhagavathy S. ; Tesic J. and Manjunath B. S., "On the Rayleigh Nature of Gabor Filter Outputs," in Intl. Conf. on Image Processing (ICIP), 2003.

[38] Lindsay I Smith, A tutorial on Principal Components Analysis, 2002.

[39] H.Muller, W.Muller, D.M.Square,S.M.Maillent and T.Pue,"Performance evaluation in content-based image retrieval: overview and proposals",Pattern Recognition letters, vol.22,pp.539-601,2001.

[۴۰] لاکدشتی، م.معین، ک. بدیع، "بازیابی مبتنی بر محتوای تصاویر رادیوگرافی بر اساس هیستوگرام جهت الگو"، مجله فیزیک پزشکی ایران، دوره ۵، شماره ۱، پیاپی (۱۸،۱۹)، بهار و تابستان ۸۷.

[41] Y. Rubner, " Perceptual Metrics for Image Database Navigation", PhD thesis, Stanford University, 1999.

[42] T. Ojala, M. Pietikäinen, D. Harwood, "A comparative study of texture measures with classification based on featured distribution", Pattern Recognition, vol. 29, no.1, pp. 51–59, 1996.

[43] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen, "Local binary patterns and its application to facial image analysis: A survey", IEEE Transactions on Systems, Man, and Cybernetics-part C: Applications and reviews, vol. 41, no. 6, 2011.

[44] H.Khosravi: Integrated Approach the recognition of printed Farsi , Ph.D. Dissertation,University of Tarbiat modares 2009.

[45] Richard Nock and Frank Nielsen,"Statistical region merging", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 26, NO. 11, NOVEMBER 2004 .

[46] J. L. J. Z .Wang, and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 23, pp. 963-947, .2001 .

[47] Vadivel A, Sural Shamik, Majumdar AK. An integrated color and intensity co-occurrence matrix. Pattern Recognit Lett 2007;28:974–83.

[48] Lin, C. H., Chen, R. T., & Chan, Y. K. (2009). "A smart content-based image retrieval system based on color and texture feature". *Image and Vision Computing*, 27(6), 658-665.

[49] Zakariya, S. M., Ali, R., & Ahmad, N., "Combining visual features of an image at different precision value of unsupervised content based image retrieval", In *Computational Intelligence and Computing Research (ICIC)*, 2010 IEEE International Conference on (pp. 1-4).

[۵۰] منصورى ز، (۱۳۸۷) ، پایان نامه ارشد، " بازیابی تصاویر رنگی با استفاده از دانش بافت و رنگ در ساختار درختی دودویی"، دانشکده کامپیوتر، دانشگاه شریف.

Abstract

In this thesis, a fast and accurate method is proposed for image retrieval. In the recent years, the number of images increased drastically in social networks in internet and public databases. According to the subject, retrieval systems became important. Image retrieval has many usages in the fields of medical, military, industrial and multimedia environment. A Content based image retrieval system retrieves the image based on its visual contents which is different from text based image retrieval. So the speed of computer systems is becoming a significant feature. Beside acceptable accuracy, the speed is also a priority.

In the proposed method, both color and edge features are used. The combination of color histogram features HSV and edge histogram features (SRF) based on local binary pattern (LBP) form the proposed system. Beside the two mentioned features of Canberra match criteria, it is used for increasing efficiency. How to retrieve image is that the system assess the sample with all images after receiving the query image and provide the most similar image. Wang database has been used to evaluate the proposed method including 1000 images in 10 concept classes. Eventually, according to evaluating proposed system with the up to date and current systems, it is seen that our system have better accuracy and speed than current methods.

Key words: image retrieval, color, edge features, local binary pattern, Canberra



Shahrood University of Technology

Faculty of Electrical and Robotic Engineering

Providing A Fast Method For Content Based Image Retrieval

Ahmad Zolfaghari

Supervisor:

Dr. Hossein Khosravi

2014