

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

بازیابی تصویر مبتنی بر محتوا با استفاده از سطوح بالای معنایی

نگارنده: حامد غضنفری

استاد راهنما:

دکتر حمید حسن پور

استاد مشاور:

دکتر علیرضا احمدی فرد

بهمن ۱۳۹۵



دانشگاه تبریز

باسمه تعالی

شماره:

تاریخ:

ویرایش:

مدیریت تحصیلات تکمیلی

دانشکده :

گروه :

پایان نامه کارشناسی ارشد آقای / خانم به شماره دانشجویی:

تحت عنوان:

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

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 :

سپاس گزاری

خدای را بسی شاکرم که از روی کرم، پدر و مادری فداکار نصیبم ساخته تا در سایه
درخت پر بار وجودشان بیسایم و از ریشه آن ها شاخ و برگ گیرم و از سایه وجودشان
در راه کسب علم و دانش تلاش نمایم. والدینی که بودنشان تاج افتخاری است بر
سرم و نشان دلیلی است بر بودنم، چرا که این دو وجود، پس از پروردگار، مایه
هستی ام بوده اندستم را گرفتند و راه رفتن را در این وادی زندگی پر از فراز و نشیب
آموختند. آموزگارانمی که برایم زندگی بودند و انسان بودن را معنا کردند.

تقدیم به وجود بارز نشان...

تشکر و قدردانی

سپاس و ستایش مخصوص خالق بی‌همتایی است که الطاف بی‌کرانش محط زندگی ام را دربر گرفته و فضل بی‌قیاس در مراحل دشوار زندگی شامل حالم گشته که اگر عنایت بی‌منت او نبود، رساندن این بار به سر منزل مقصود غیر ممکن می‌نمود. اکنون که این پژوهش به زیور چاپ آراسته می‌گردد بر خود لازم می‌دانم تا مطابق سنت حسنه سپاسگزاری، والاترین مراتب سپاس خویش را به محضر استادان فرزانه‌ای که در طول تحصیل همواره افتخار ساگردی آنان را داشته‌ام و مرا مرهون الطاف خالصانه خود قرار داده اند تقدیم نمایم.

از استاد گرامیم جناب آقای دکتر حمید حسن پور بسیار سپاسگزارم چرا که بدون راهنمایی‌های ایشان تأمین این پایان‌نامه بسیار مشکل می‌نمود.

همچنین از جناب آقای دکتر علیرضا احمدی فرد به دلیل یاری‌ها و راهنمایی‌های ایشان، سپاس گزارم.

تعهد نامه

اینجانب **حامد غضنفری** دانشجوی دوره کارشناسی ارشد رشته هوش مصنوعی دانشکده کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شاهرود نویسنده پایان نامه **بازیابی تصویر مبتنی بر محتوا با استفاده از سطوح بالای معنایی تحت راهنمایی آقای دکتر حمید حسن پور** 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 مورداستفاده 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 است 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 افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود است. 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.

چکیده:

بازیابی تصویر مبتنی بر محتوا (CBIR)¹ یکی از مهم‌ترین زمینه‌های تحقیقاتی در علم بینایی ماشین و پردازش تصویر است. سیستم CBIR، تصویر جستجو را بر اساس محتوای بصری آن از پایگاه داده‌ی تصاویر، بازیابی می‌کند. به دلیل افزایش سریع مجموعه‌های تصاویر با حجم بالا و پاسخگو نبودن سیستم‌های بازیابی تصویر مبتنی بر متن، وجود سیستم‌های بازیابی تصویر مبتنی بر محتوا ضروری است. سیستم‌های بازیابی تصاویر، اکثراً از ویژگی‌های سطح پایین (مانند رنگ، شکل و بافت) به تنهایی استفاده می‌کنند. ویژگی‌های سطح پایین، محتوای تصاویر ساده را می‌توانند به خوبی توصیف کنند، اما تصاویر حاوی مفاهیم سطح بالا را نمی‌توانند توصیف کنند. منظور از مفاهیم سطح بالا، مفاهیم ادراکی انسان از تصویر است که با استفاده از ویژگی‌های سطح پایین به تنهایی قابل بیان نیست. این مشکل، یکی از چالش‌های اصلی سیستم‌های CBIR است و "شکاف معنایی"² نامیده می‌شود.

در این پایان‌نامه دو روش سریع و کارآمد جهت کاهش "شکاف معنایی" معرفی شده است. این دو روش به دنبال اطلاعات معنایی از تصویر، با استفاده از ویژگی‌های سطح پایین بصورت ترکیبی هستند. بر اساس مطالعات انجام‌شده، سیستم بینایی انسان ویژگی‌های لبه و رنگ را به صورت قابل توجهی در ارزیابی تصویر ملاک قرار می‌دهد. لذا این دو مجموعه ویژگی اطلاعات معنایی زیادی از تصویر را پوشش داده و در سیستم‌های بازیابی تصویر مبتنی بر محتوا نقش بسزایی دارند. یکی از روش‌های مطرح در این حوزه، غالباً از ویژگی‌های لبه و اطلاعات رنگ در فضای رنگ $L^*a^*b^*$ استفاده می‌کند. بردار ویژگی حاصل از اختلاف رنگ مبتنی بر این ویژگی‌ها، "هیستوگرام اختلاف رنگ" نامیده می‌شود. در این پایان‌نامه، ابتدا با اعمال دو معیار آنتروپی و همبستگی بین مولفه‌های این بردار ویژگی در تصاویر مختلف، ویژگی‌های موثر انتخاب می‌شوند. اگرچه ویژگی‌های حاصل، تا حدی بیان‌گر محتوای معنایی تصویر است، با افزودن ویژگی‌های جدید به مجموعه ویژگی حاصل، نرخ

¹ Content Based Image Retrieval

² Semantic Gap

بازیابی تصویر را می‌توان افزایش داد. اولین ویژگی پیشنهادی، هیستوگرام رنگ در فضای رنگ HSV است. دو فضای رنگ HSV و $L^*a^*b^*$ ، نسبت به سایر فضاهای رنگ به سیستم بینایی انسان نزدیک‌تر هستند. فضای رنگ $L^*a^*b^*$ در روش هیستوگرام اختلاف رنگ استفاده شده است. لذا در روش پیشنهادی اول، جهت استفاده از اطلاعات معنایی هر دو فضای رنگ، هیستوگرام رنگ در فضای رنگ HSV به بردار ویژگی اضافه می‌شود. این بردار ویژگی از نظر تعداد ویژگی، سرعت و دقت بازیابی تصویر، نسبت به جدیدترین روش‌های بازیابی تصویر اخیر، بهبود قابل توجهی را نشان می‌دهد. روش پیشنهادی دوم، به دنبال استفاده از اطلاعات معنایی هر دو نوع لبه، شامل لبه‌ی حاشیه‌ای اشیاء و لبه‌ی داخلی اشیاء، ویژگی جدیدی معرفی می‌شود. لبه حاشیه‌ای اشیاء در روش هیستوگرام اختلاف رنگ استفاده شده است. بافت تصویر شامل لبه‌های داخلی اشیاء در تصویر است. لذا هیستوگرام جهت لبه با استفاده از گرادیان حاصل از بافت تصویر به مجموعه ویژگی روش پیشنهادی اول اضافه می‌شود. در نهایت در معیار شباهت تصاویر، به هر گروه ویژگی از مجموعه‌های ویژگی‌های معنایی پیشنهادی، وزنی اختصاص می‌یابد. با انتساب این مقادیر وزنی، سرعت و دقت در سیستم‌های بازیابی پیشنهادی، افزایش می‌یابد. ویژگی‌های معنایی استخراج‌شده به همراه وزن‌دهی آن‌ها، به عنوان سطوح بالای معنایی شناخته می‌شود، چراکه منجر به درک مفاهیم سطح بالای تصویر می‌شوند. در این پایان‌نامه، از دو پایگاه داده استاندارد Corel 5k و Corel 10k استفاده شده است. این افزایش سرعت و دقت بازیابی، حاکی از آن است که سطوح بالای معنایی پیشنهادی، در عین تعداد کم ویژگی، بدون استفاده از قطعه‌بندی تصویر و فرآیندهای یادگیری و خوشه‌بندی، بخوبی فضای دو بعدی تصویر را توصیف کرده و مفاهیم معنایی را از آن استخراج می‌کنند.

کلمات کلیدی: بازیابی تصویر مبتنی بر محتوا، هیستوگرام رنگ، هیستوگرام جهت لبه، آنتروپی،

همبستگی، سطوح بالای معنایی، الگوهای دودویی محلی

لیست مقالات مستخرج از پایان نامه

- غضنفری ح و حسن پور ح، "بازیابی تصویر مبتنی بر محتوا با استفاده از هیستوگرام رنگ"، کنفرانس بین المللی پردازش سیگنال و سیستم‌های هوشمند، دانشگاه صنعتی امیر کبیر، تهران، ص ۸۲، (آذر ۱۳۹۵).
- غضنفری حامد و حسن پور حمید، ۱۳۹۵، "بازیابی تصویر مبتنی بر محتوا با استفاده از تلفیق هیستوگرام رنگ و جهت لبه"، مجله تخصصی انجمن ماشین بینایی و پردازش تصویر ایران، (ارسال شده).

فهرست مطالب

عنوان

صفحه

فصل اول: مقدمه

- ۱-۱- بازیابی تصویر..... ۲
- ۲-۱- تاریخچه بازیابی تصویر..... ۳
- ۳-۱- کاربردهای بازیابی تصویر..... ۳
- ۴-۱- تعریف مساله..... ۶
- ۵-۱- هدف و نوآوری پایان نامه..... ۹
- ۶-۱- ساختار پایان نامه..... ۱۰

فصل دوم: کارهای انجام شده

- ۱-۲- مقدمه..... ۱۲
- ۲-۲- ویژگی های سطح پایین مورد استفاده در سیستم CBIR..... ۱۳
- ۲-۲-۱-۱- فضای رنگ RGB..... ۱۳
- ۲-۲-۱-۲- فضای رنگ HSV..... ۱۴
- ۲-۲-۱-۳- فضای رنگی $L^*a^*b^*$ ۱۷
- ۲-۲-۱-۴- هیستوگرام رنگ..... ۱۷
- ۲-۲-۲- ویژگی بافت..... ۱۸
- ۲-۲-۲-۱- ماتریس هم رخدادی سطح خاکستری (GLCM)..... ۱۸
- ۲-۲-۲-۲- تبدیل گابور..... ۲۰
- ۲-۲-۲-۳- الگوهای دودویی محلی (LBP)..... ۲۰
- ۲-۲-۲-۴- ویژگی های تامورا..... ۲۱
- ۲-۲-۳- ویژگی شکل..... ۲۲
- ۲-۲-۳-۱- هیستوگرام جهت لبه..... ۲۲
- ۲-۲-۳- سیستم های CBIR مبتنی بر یادگیری و هزینه بردار..... ۲۴

۲۴ ۲-۳-۱- تکنیک‌های یادگیری ماشین
۳۲ ۲-۴- سیستم‌های CBIR مبتنی بر ترکیب ویژگی‌های سطح پایین
۴۱ ۲-۵- معیار شباهت
۴۲ ۲-۵-۱- فاصله مینکوفسکی
۴۲ ۲-۵-۲- فاصله اقلیدسی
۴۲ ۲-۵-۳- فاصله کانبرا
۴۳ ۲-۵-۴- فاصله کسینوسی
۴۳ ۲-۵-۵- فاصله χ^2
۴۳ ۲-۵-۶- کانبرای بهبود یافته
۴۴ ۲-۶- جمع‌بندی

فصل سوم: روش پیشنهادی

۴۶ ۳-۱- مقدمه
۴۷ ۳-۲- روش پیشنهادی اول
۴۹ ۳-۲-۱- استخراج ویژگی
۴۹ ۳-۲-۱-۱- هیستوگرام اختلاف رنگ
۵۳ ۳-۲-۱-۲- هیستوگرام رنگی در فضای رنگ HSV
۵۴ ۳-۲-۲- انتخاب ویژگی
۵۴ ۳-۲-۲-۱- آنتروپی شانون و حذف ویژگی‌های ناکارآمد
۵۶ ۳-۲-۲-۲- همبستگی بین بردارهای ویژگی
۵۹ ۳-۲-۳- معیار شباهت
۶۱ ۳-۳- روش پیشنهادی دوم
۶۲ ۳-۳-۱- استخراج ویژگی
۶۲ ۳-۳-۱-۱- جهت لبه از بافت تصویر
۶۴ ۳-۳-۲- انتخاب ویژگی
۶۴ ۳-۳-۲-۱- آنتروپی و حذف ویژگی‌های ناکارآمد
۶۵ ۳-۳-۲-۲- همبستگی بین بردارهای ویژگی
۶۵ ۳-۳-۳- معیار شباهت
۶۷ ۳-۴- جمع‌بندی

فصل چهارم: نتایج و آزمایش‌ها

- ۱-۴- مقدمه..... ۷۰
- ۲-۴- پایگاه داده تصویر..... ۷۰
- ۳-۴- بحث روی نرخ بازیابی ویژگی‌های مختلف..... ۷۱
- ۴-۴- ارزیابی نرخ بازیابی روش پیشنهادی اول..... ۷۴
- ۵-۴- ارزیابی نرخ بازیابی روش پیشنهادی دوم..... ۷۷
- ۵-۴- ارزیابی خروجی سیستم بازیابی پیشنهادی..... ۷۹
- ۶-۴- جمع‌بندی..... ۸۶

فصل پنجم: نتیجه‌گیری و کارهای آینده

- ۱-۵- بحث و نتیجه‌گیری..... ۹۰
- ۲-۵- کارهای آینده..... ۹۱

فهرست امثال

- شکل ۱-۱-۱- بلوک دیاگرام بازیابی تصویر مبتنی بر محتوا. ۶
- شکل ۱-۲- مفهوم شکاف معنایی بین درک سیستم و درک انسان از تصویر. ۷
- شکل ۱-۳- نتایج بازیابی تصویر با استفاده از روش‌های مختلف. ۸
- شکل ۲-۱- فضای رنگی RGB. ۱۴
- شکل ۲-۲- فضای رنگی HSV. ۱۵
- جدول ۲-۱- اطلاعات مربوط به برخی رنگ‌های خاص و معادل RGB و HSV آنها. ۱۶
- شکل ۲-۳- نمونه‌های از تشکیل ماتریس هم‌رخدادی GLCM [۲۰]. ۱۹
- شکل ۲-۴- روش کار الگوریتم بافتی الگوهای باینری محلی (LBP). ۲۱
- شکل ۲-۵- (الف) ماسک رابرتز (ب) ماسک سوبل. ۲۳
- شکل ۲-۶- نمونه‌های از کارکرد شبکه‌ی HOG برای تصویر [۷]. ۲۸
- شکل ۲-۷- آموزش حالت‌های مختلف قرارگیری شیء انسان در مقابل دوربین، با استفاده از DPM [۷]. ۲۹
- شکل ۲-۸- دو نمونه از شناسایی مفهوم بصری. اشیاء با چهارگوش‌های بسته‌ی قرمز و سبز، مفاهیم بصری با چهارگوش‌های بسته‌ی آبی رنگ مشخص شده‌اند. تصویر سمت چپ شامل مفهوم "اسب سواری" و تصویر سمت راست شامل مفهوم "نشستن افراد دور میز ناهار خوری" هستند [۷]. ۲۹
- شکل ۲-۹- نتیجه بخش‌بندی ابتدایی با استفاده از الگوریتم Mean Shift [۹]. ۳۰
- شکل ۲-۱۰- تصویر سمت چپ، ابتدا کاربر، اشیاء را مشخص و سپس الگوریتم قطعه‌بندی اجرا شده، تصویر سمت راست ادغام نواحی معنایی [۹]. ۳۱
- شکل ۲-۱۱- استفاده از واسط کاربری و سپس اعمال الگوریتم بخش‌بندی روی تصویر [۹]. ۳۲
- شکل ۲-۱۲- جداسازی تصویر، (ج) تصویر اصلی، (ب) تصویر ساختار و (الف) تصویر بافت [۲۰]. ۳۳
- شکل ۲-۱۳- مثالی از تشخیص ساختار نوار-شکل در پنجره 3×3 . (a) پنجره‌های از تصویر. (b) و (c) فرآیند تشخیص ساختار نوار-شکل. (d) ساختار نوار-شکل تشخیص داده شده. ۳۵
- شکل ۲-۱۴- تشخیص ساختار برجستگی. (a) چهار الگوی ساختار نوار-شکل، زوایه بین ساختار نوار-شکل و جهت افقی به ترتیب ۰، ۴۵، ۹۰ و ۱۳۵ درجه است. (b) چهار فیلتر گابور جهت‌دار که در آن زوایه مورد نظر به ترتیب ۰، ۴۵، ۹۰ و ۱۳۵ درجه است. (c) فرآیند تشخیص ساختار برجستگی. علامت \oplus به معنای تطبیق ساختار نوار-شکل و فیلتر گابور است. (d) چهار الگوی ساختار برجستگی. ۳۶
- شکل ۲-۱۵- پیکربندی پیکسل در پنجره ناحیه 2×2 و پنجره‌های جابجاشده مربوط به آنها در چهار جهت، که جهت محاسبه (a) $0, -1$ ، (b) $0, 1$ ، (c) $1, 0$ و (d) $1, 0$ مورد نیاز است [۱۵]. ۳۹
- شکل ۳-۱- دیاگرام روش پیشنهادی اول جهت بازیابی تصویر مبتنی بر محتوا. ۴۸

- شکل ۳-۲- تصویر a تصویر تمبر و تصویر b تصویر میوه. نمودارهای هیستوگرام اختلاف رنگ دو تصویر مشاهده میشوند. ویژگیهای ۱-۹۰ و ویژگیهای مربوط به رنگ و ویژگیهای ۱۰۸-۹۱ و ویژگیهای مربوط به جهت لبه هستند [۱۹]. ۵۲.....
- شکل ۳-۳- فضای استوانه‌ای فضای رنگی HSV. ۵۳.....
- شکل ۳-۴- آنتروپی ۱۰۸ ویژگی هیستوگرام اختلاف رنگ یکنواخت ادراکی ۵۵.....
- شکل ۳-۵- نمودار نرخ بازیابی-حذف ویژگی‌های کمتر از آنتروپی. ۵۶.....
- شکل ۳-۶- همبستگی بردار ویژگی ۱۴۵م از هیستوگرام اختلاف رنگ نسبت به سایر بردارهای ویژگی و همچنین بردار خروجی مطلوب. ۵۷.....
- شکل ۳-۷- همبستگی بردار ویژگی ۱۱م از هیستوگرام رنگ HSV نسبت به سایر بردارهای ویژگی و همچنین بردار خروجی مطلوب. ۵۸.....
- شکل ۳-۸- نمودار نرخ بازیابی-وزن گروه‌های ویژگی. ۶۰.....
- شکل ۳-۹- دیاگرام کلی روش پیشنهادی دوم جهت بازیابی تصویر مبتنی بر محتوا. ۶۱.....
- شکل ۳-۱۰- روش کار الگوریتم الگوهای دودویی محلی جهت استخراج ویژگی بافت از تصویر. ۶۳.....
- شکل ۳-۱۱- (الف) ماسک رابرتز (ب) ماسک سوئل. ۶۳.....
- شکل ۳-۱۲- آنتروپی ۱۲۸ ویژگی جهت لبه حاصل از بافت تصویر. ۶۴.....
- شکل ۳-۱۳- نمودار نرخ بازیابی-وزن گروه‌های ویژگی. ۶۶.....
- شکل ۴-۱- نمونه‌هایی از تصاویر پایگاه داده تصویر Corel5k. ۷۰.....
- شکل ۴-۲- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ورزش دفاع شخصی. ۸۱.....
- شکل ۴-۳- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ورزش دفاع شخصی. ۸۱.....
- شکل ۴-۴- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ورزش اتومبیل‌رانی. ۸۲.....
- شکل ۴-۵- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ورزش اتومبیل‌رانی. ۸۲.....
- شکل ۴-۶- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ریزه‌سنگ‌های دریایی. ۸۳.....
- شکل ۴-۷- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ریزه‌سنگ‌های دریایی. ۸۳.....
- شکل ۴-۸- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: خودرو. ۸۴.....
- شکل ۴-۹- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: خودرو. ۸۴.....
- شکل ۴-۱۰- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: وسیله نقلیه قدیمی. ۸۵.....
- شکل ۴-۱۱- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: وسیله نقلیه قدیمی. ۸۶.....

فهرست جداول

- جدول ۱-۲-۱- اطلاعات مربوط به برخی رنگ‌های خاص و معادل RGB و HSV آنها..... ۱۶
- جدول ۱-۴-۱- مقایسه نرخ بازیابی روش هیستوگرام اختلاف رنگ [۱۹] با بهبود یافته‌های آن..... ۷۲
- جدول ۱-۴-۱- مقایسه نرخ بازیابی روش هیستوگرام رنگ نسبت به دو فضای رنگ HSV و $L^*a^*b^*$ ۷۳
- جدول ۳-۴-۳- مقایسه نرخ بازیابی روش هیستوگرام رنگ HSV با بهبود یافته‌های آن..... ۷۳
- جدول ۴-۴-۴- مقایسه نرخ بازیابی روش هیستوگرام جهت لبه با بهبود یافته‌های آن..... ۷۴
- جدول ۴-۴-۵- مقایسه نرخ بازیابی روش پیشنهادی اول با روش‌های اخیر..... ۷۵
- جدول ۴-۴-۶- مقایسه روش پیشنهادی اول با روش‌های اخیر از نظر زمان بازیابی تصویر..... ۷۶
- جدول ۴-۴-۷- مقایسه نرخ بازیابی روش پیشنهادی دوم با روش‌های اخیر..... ۷۸
- جدول ۴-۴-۸- مقایسه روش‌های پیشنهادی با چهار روش از جدیدترین روش‌ها در حوزه بازیابی تصویر..... ۷۹

فصل اول:

مقدمه

۱-۱- باز یابی تصویر

گسترش فناوری‌های دیجیتال و در دسترس بودن وسایل دریافت و ذخیره عکس، باعث ایجاد انبوهی از تصاویر دیجیتال در حوزه‌های مختلف شده است. با توجه به رشد سریع اینترنت و در کنار آن فراگیر شدن موتورهای جستجو، کاربران انتظار دارند که بتوانند تصاویر مورد نظر خود را همانند اطلاعات دیگر به راحتی جستجو و دریافت نمایند. به همین دلیل نیاز به سیستمی است که بتواند جستجو و باز یابی تصاویر را بر اساس خواسته‌های کاربر انجام دهد [۱,۲].

در این راستا سیستم‌های زیادی برای باز یابی تصاویر پیشنهاد شده است. این سیستم‌ها را می‌توان بر اساس ساختار و نوع عملکرد به سه روش مبتنی بر متن، مبتنی بر محتوا و ترکیبی از متن و محتوا تقسیم نمود [۳]. در روش مبتنی بر متن، تصاویر به صورت دستی و غیر خودکار توسط انسان در قالب متن توصیف و برچسب گذاری می‌شوند. سپس این برچسب‌ها با کمک یک سیستم مدیریت پایگاه داده ذخیره شده و در هنگام جستجو مورد استفاده قرار می‌گیرند. این روش دارای دو عیب عمده است، نخست آن که برای برچسب گذاری همه تصاویر به حجم عظیمی از نیروی انسانی نیاز است. عیب دیگر این روش، دقیق نبودن برچسب گذاری است، زیرا این کار به صورت ذهنی و سلیقه‌ای انجام می‌شود، که از فردی به فرد دیگر متفاوت است.

به منظور رفع مشکلات سیستم‌های باز یابی تصویر مبتنی بر متن، باز یابی تصویر مبتنی بر محتوا (CBIR) مطرح می‌شود. در سیستم‌های CBIR، محتوای تصویر مورد جستجوی کاربر با محتوای تصاویر پایگاه داده مقایسه می‌شود، تا تصاویر مشابه با تصویر مورد جستجو، باز یابی شوند. محتوای همه تصاویر با استخراج ویژگی‌های سطح پایین مانند رنگ، بافت و شکل، بدست می‌آیند. شباهت ویژگی‌های تصاویر پایگاه داده با ویژگی‌های تصویر مورد جستجو توسط معیار شباهت، اندازه گیری و تصاویر نزدیک تر باز یابی می‌شوند [۳].

در روش سوم، از ترکیب متن و محتوای تصویر جهت بازیابی استفاده می‌شود. در این روش محتوای تصویر با استفاده از ویژگی‌های سطح پایین استخراج شده و با کمک متن مورد جستجو، تصویر در پایگاه داده تصویری، جستجو می‌شود [۴،۵].

۱-۲- تاریخچه بازیابی تصویر

بازیابی تصاویر از دهه ۱۹۷۰ تاکنون یک مقوله فعال تحقیقاتی می‌باشد. بازیابی تصویر مبتنی بر متن تصاویر، از اوایل دهه ۱۹۷۱ آغاز شد. همانطور که ذکر شد، دو مشکل عمده در این نوع بازیابی وجود دارد، خصوصاً زمانی که حجم مجموعه پایگاه داده بزرگ باشد [۱-۳].

۱. مشکلات و زمان زیادی صرف یادداشت‌گذاری تصاویر می‌شود.

۲. مشکلاتی که بدلیل محتوای زیاد تصاویر و درک متفاوت انسان‌ها از تصاویر حاصل می‌شود.

در اوایل سال ۱۹۹۲ بدلیل افزایش سریع مجموعه‌های تصاویر با حجم بالا و پاسخگو نبودن سیستم‌های مبتنی بر متن، سیستم‌های بازیابی مبتنی بر محتوای تصاویر معرفی شدند. در این سیستم‌ها بجای یادداشت‌گذاری دستی، تصاویر با استفاده از محتوای بصری خود جستجو می‌شوند. این روش، هنوز هم از مهم‌ترین زمینه‌های تحقیقاتی در علم پردازش تصویر است [۳].

۱-۳- کاربردهای بازیابی تصویر

تصویر از زمان پیدایش وب جهان گستر^۳ از لحاظ تعداد و دامنه صفحات، رشد زیادی داشته است. به طوری که امروزه تعداد آن چندین میلیارد برآورد می‌شود. بنابراین بازیابی درست و سریع تصاویر از اهمیت بالایی برخوردار است. بازیابی تصویر کاربردهای زیادی دارد، چند نمونه از مهم‌ترین این کاربردها تشریح می‌شوند [۶].

³ World Wide Web

- **حقوق معنوی:**

یکی از کاربردهای بازیابی تصاویر می‌تواند برای حمایت از حق صاحب امتیاز تصویر باشد. به علت گسترش سریع تصاویر، حقوق مالک در وب، حیاتی است. یکی از روش‌های موجود استفاده از نشانه‌گذاری دیجیتال است، که با اضافه کردن برچسب یا توصیف‌گر بر روی تصویر، مانع از استفاده غیرقانونی آن می‌شوند.

- **فیلتر کردن مواد بالغین:**

امروزه تصاویر غیر اخلاقی فراوانی در وب وجود دارد که دسترسی به این نوع تصاویر برای کودکان و حتی بالغین در برخی فرهنگ‌ها، دغدغه اصلی افراد است. در سال‌های اخیر، پژوهشگران به استفاده از روش‌های تصویر مجازی در رایانه، روی آورده‌اند. این روش به طور خودکار تصاویر مربوط به مسائل جنسی را تشخیص می‌دهد.

- **اجرای قانون:**

بازیابی تصاویر مبتنی بر محتوا کاربردهای گوناگونی در اجرای قانون و جلوگیری از وقوع جرایم دارد، مثل تشخیص اثر انگشت، تشخیص چهره، تطبیق DNA و شناسایی رد پا. نظام‌های نظارتی بسیاری از اینترنت برای فروش و به نمایش گذاشتن کالاهای غیر قانونی خود مثل مواد مخدر، قاچاق اسلحه و غیره استفاده می‌کنند. چون این گونه وبگاه‌ها دارای اطلاعات دیداری زیادتری نسبت به اطلاعات متنی هستند، تکیه بر روش‌های بازیابی متنی در این مورد نمی‌تواند کافی باشد. استفاده از نظام‌های مبتنی بر محتوا می‌تواند کمک شایانی به شناسایی این گونه وبگاه‌ها نمایند.

- **مسافرت و گردشگری:**

مردم قبل از مسافرت به مکانی خاص، خواهان آن هستند که اطلاعات بیشتری راجع به آن مکان به دست آورند. اطلاعات ممکن است شامل دیدن نقشه یک کشور یا شهر باشد که اطلاعات گردشگری راجع به آن محل را ارائه می‌دهد. با توجه به اینکه اینترنت دارای تصاویر زیادی از کشورهای جهان است، هیچ شیوه‌ای بهتر از جستجو در اینترنت وجود ندارد. اما ابزاری برای بازیابی تصاویر در اینترنت لازم است. به علاوه فهرست تصاویر موجود در اینترنت می‌تواند مسافران را در یافتن اطلاعات موردنظر کمک کند.

- **آموزش:**

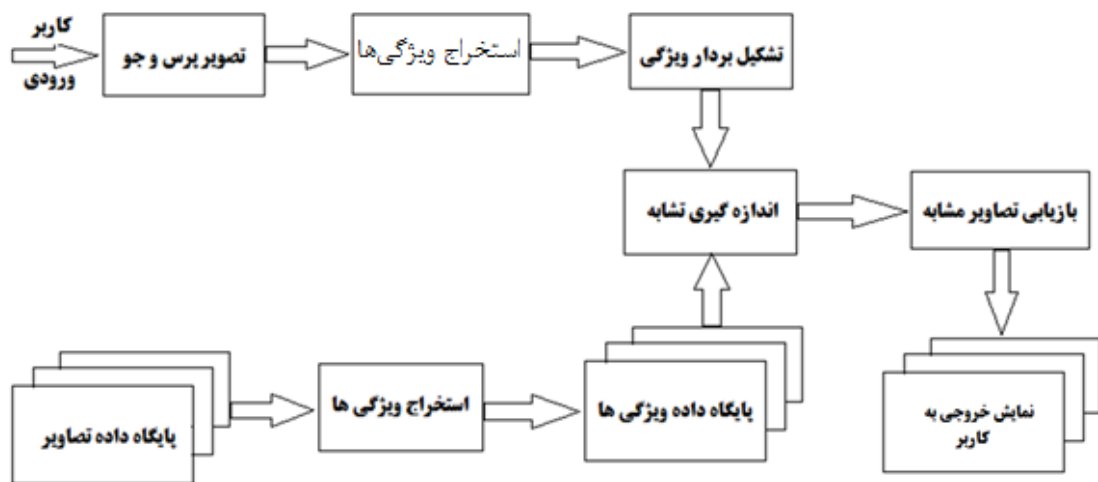
آموزش یکی از کاربردهای مهم تصاویر در اینترنت است. اکثر دانشجویان برای انجام پژوهش به دو دلیل نیازمند تصاویر هستند:

۱. به عنوان منبع اطلاعاتی (به عنوان مثال تعیین قلمرو ایران در دوران قاجاریه).

۲. برای تشریح نظرات خویش با استفاده از تصویر. بسیاری از این گونه تصاویر در اینترنت وجود دارد و دانشجویان می‌توانند آن‌ها را جستجو و بازیابی کنند.

۴-۱- تعریف مساله

اغلب سیستم‌های بازیابی تصویر مبتنی بر محتوا، بصورت دیاگرام شکل (۱-۱) عمل کرده و تصاویر مورد نظر کاربر را بازیابی می‌کنند:



شکل ۱-۱- ۱- بلوک دیاگرام بازیابی تصویر مبتنی بر محتوا.

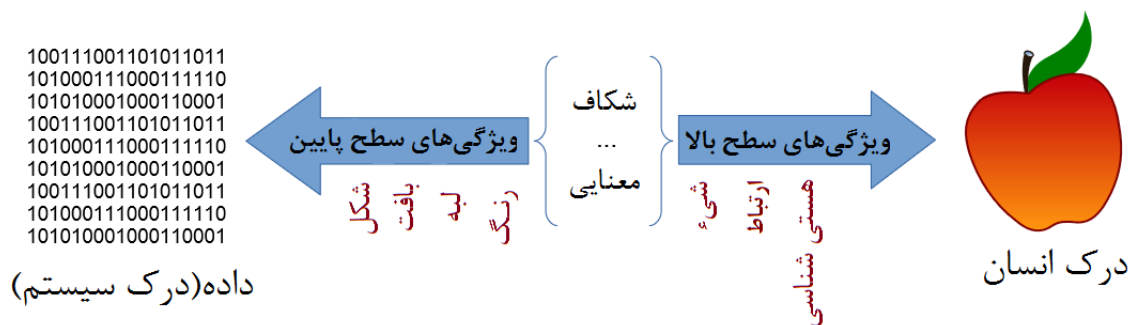
همانطور که از دیاگرام مربوط به بازیابی تصویر مبتنی بر محتوا (شکل ۱-۱) مشخص است، ابتدا تصویر مورد جستجو از کاربر دریافت شده و سپس ویژگی‌های سطح پایین (رنگ، شکل و بافت) از آن استخراج می‌شود. ویژگی‌های استخراجی بردار ویژگی تصویر را می‌سازند. سپس همین ویژگی‌ها از تصاویر پایگاه داده، استخراج شده و بردارهای ویژگی برای این تصاویر ساخته می‌شوند. در مرحله آخر با استفاده از یک معیار شباهت، شباهت بردارهای تصاویر پایگاه داده، با بردار ویژگی تصویر کاربر اندازه‌گیری می‌شود و تصاویر نزدیک‌تر بازیابی و به کاربر نشان داده خواهند شد.

با وجود آنکه روش‌های متعددی برای CBIR ارائه شده است، اما هنوز این روش‌ها نتوانسته‌اند نیازهای کاربران را برآورده نمایند. به همین دلیل مشاهده می‌شود که کارایی کنونی موتورهای جستجوی تصویر با سطح مورد انتظار کاربران تفاوت چشمگیری دارد.

همانطور که ذکر شد، در سیستم‌های بازیابی تصویر مبتنی بر محتوا (CBIR)، یک تصویر جهت جستجو، توسط کاربر داده می‌شود. سپس تصویر مورد جستجو با ویژگی‌های سطح پایین مانند

رنگ، بافت و شکل، توصیف می‌شود. اگرچه ویژگی‌های سطح پایین می‌توانند محتوای تصاویر ساده را توصیف کنند، ولی نمی‌توانند تصاویر پیچیده شامل مفاهیم سطح بالا را نیز بصورت صحیح توصیف کنند [۷]. از نمونه‌هایی از مفاهیم سطح بالا، می‌توان به اشیاء، ارتباط بین آن‌ها و هستی‌شناسی اشاره کرد.

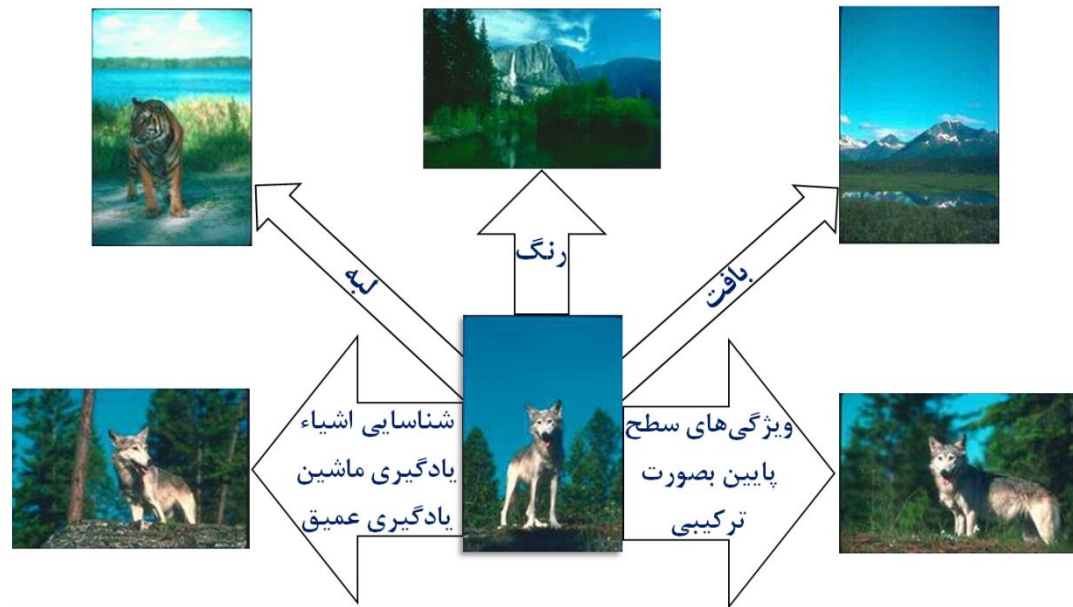
منظور از مفاهیم سطح بالا، مفاهیم ادراکی انسان از تصویر است که با استفاده از یک ویژگی سطح پایین به تنهایی، قابل بیان نیست. برای مثال، تصویری از جنگل و تصویر دیگری با زمینه و بافت سبز را در نظر بگیرید، در تصویر دوم مفهوم جنگل وجود ندارد. در این صورت بر اساس ویژگی سطح پایین (مانند رنگ)، هر دو تصویر به یک کلاس تعلق خواهند داشت. این مشکل، یکی از چالش‌های اصلی سیستم‌های CBIR است و "شکاف معنایی" نامیده می‌شود. مفهوم کلی این چالش در شکل (۱-۲) نشان داده شده است. روش‌های زیادی برای کاهش شکاف معنایی ارائه شده‌اند. اغلب این روش‌ها مبتنی بر یادگیری ماشین، یادگیری عمیق، شناسایی اشیاء، قطعه‌بندی و ترکیب ویژگی‌های سطح پایین هستند.



شکل ۱-۲- مفهوم شکاف معنایی بین درک سیستم و درک انسان از تصویر.

در سال‌های اخیر برای کاهش این شکاف معنایی روش‌های گوناگونی بیان و استفاده شده است. اغلب این روش‌ها، نیاز به پردازش‌های زیاد، آموزش‌های طولانی و در نهایت هزینه‌های زیاد دارند. در این پایان‌نامه به دنبال روشی هستیم که نیاز به پردازش‌های زیاد و آموزش نداشته باشد و در عین حال سریع بوده و تعداد ویژگی کم داشته باشد. این روش به نوعی از ویژگی‌های سطح پایین

به سطوح بالای معنایی تصویر دست پیدا می‌کند و از آن‌ها جهت بازیابی تصاویر از پایگاه داده، استفاده می‌کند. این روش تحت عنوان "استفاده از ویژگی‌های سطح پایین بصورت ترکیبی" است که در کنار دیگر روش‌ها در شکل ۱-۳ آمده است.



شکل ۱-۳- نتایج بازیابی تصویر با استفاده از روش‌های مختلف.

همان‌طور که در شکل ۱-۳ مشاهده می‌شود، استفاده از روش‌های مختلف بازیابی تصویر، ممکن است منجر به نتیجه‌های بازیابی متفاوتی شود. استفاده از تنها یک ویژگی سطح پایین، بدرستی مفهوم تصویر مورد جستجو را بازیابی نمی‌کند. سه تصویر بالا در شکل، این چالش را نشان می‌دهند. در صورت استفاده از ویژگی‌های سطح پایین بصورت ترکیبی، محتوای معنایی تصویر استخراج و مفهوم تصویر مورد جستجو بدرستی بازیابی خواهد شد. همچنین تصاویر بازیابی شده سمت چپ نیز خروجی حاصل از روش‌های مبتنی بر روش‌های یادگیری ماشین، شناسایی اشیاء و یادگیری عمیق در بازیابی تصویر است. این روش‌ها بدرستی مفهوم تصویر مورد جستجو را بازیابی می‌کنند ولی قبل از بازیابی تصویر، به آموزش و پردازش‌های زیادی نیاز دارند.

بنابر مطالب ذکر شده در بالا، سیستم‌های بازیابی تصویر مبتنی بر محتوا علاوه بر مشکل شکاف معنایی از مشکل هزینه آموزش و پردازش‌های زیاد رنج می‌برند. بنابراین در این پایان‌نامه برآنیم که

روشی ارائه دهیم که علاوه به اینکه به آموزش و پردازش‌های زیاد نیازی ندارد، بخوبی مفاهیم معنایی را از تصویر استخراج کند و شکاف معنایی کاهش یابد. این مهم با استفاده از ترکیب ویژگی‌های سطح پایین موثر انجام می‌شود. مهم‌ترین خاصیت این ویژگی‌ها این است که بیشترین تاثیر را بر سیستم بینایی انسان نسبت به سایر ویژگی‌های سطح پایین دارند.

۱-۵- هدف و نوآوری پایان‌نامه

هدف از این پایان‌نامه ارائه روشی جهت بازیابی تصویر مبتنی بر محتوا است. این روش بگونه-ای است که تا حد زیادی شکاف معنایی بین درک انسان و درک سیستم، کاهش یابد. بر اساس مطالعات انجام‌شده، سیستم بینایی انسان ویژگی‌های لبه و رنگ را به صورت قابل توجهی در ارزیابی تصویر ملاک قرار می‌دهد. این دو ویژگی دارای اطلاعات معنایی زیادی از تصویر هستند. لذا این دو مجموعه ویژگی در سیستم‌های بازیابی مبتنی بر محتوا نقش بسزایی دارند. بنابراین با ترکیب این دو ویژگی، محتوای معنایی موجود در تصویر استخراج می‌شود. زمان استخراج این ویژگی‌ها نیز در نظر گرفته می‌شود و سعی بر تسریع سیستم بازیابی در عین دقت بازیابی بالا است. همچنین با انتساب مقادیر وزنی به گروه‌های مختلف ویژگی در معیار شباهت، سرعت و دقت سیستم افزایش می‌یابد. ترکیب ویژگی‌های معنایی تصویر و همچنین وزن‌دهی به گروه‌های مختلف ویژگی، به عنوان سطوح بالای معنایی شناخته می‌شود. چراکه بخوبی محتوای تصویر و مفاهیم معنایی آن را توصیف می‌کنند.

از دستاوردها و نوآوری‌های این پایان‌نامه می‌توان به بهبود روش هیستوگرام اختلاف رنگ اشاره کرد. این روش، یکی از جدیدترین و با دقت‌ترین روش‌ها در حوزه بازیابی تصویر مبتنی بر محتوا است. ابتدا تعداد ویژگی‌های این روش با استفاده از دو معیار آنتروپی و همبستگی، کاهش می‌یابند. این کاهش ویژگی، منجر به بهبود نرخ بازیابی تصویر این روش می‌شود. سپس جهت افزایش نرخ بازیابی، گروه‌های ویژگی معنایی جدیدی به مجموعه ویژگی اضافه می‌شوند. این ویژگی‌های معنایی، بخوبی اطلاعات معنایی تصویر را استخراج می‌کنند. از دستاوردهای مهم دیگر این پایان‌نامه، وزن‌دهی

به گروه‌های مختلف ویژگی در معیار شباهت است. این وزن‌دهی نیز باعث بهبود قابل توجهی در نرخ بازیابی تصویر می‌شود. آزمایش‌ها نشان می‌دهند، سطوح بالای معنایی استخراج شده از تصاویر، بخوبی مفاهیم معنایی را از محتوای آن‌ها استخراج می‌کنند.

۱-۶- ساختار پایان‌نامه

این پایان‌نامه در ۵ فصل تهیه شده است که هر فصل به شرح زیر است. در فصل دوم، چند نمونه از روش‌های انجام‌شده در حوزه بازیابی تصویر مبتنی بر محتوا، شرح داده شده است. در این فصل ابتدا با روش‌های استخراج ویژگی و بازیابی تصویر آشنا شده و سپس چند کار انجام شده اخیر در هر مورد تشریح می‌شود. در فصل سوم، سیستم‌های بازیابی تصویر مبتنی بر محتوای پیشنهادی معرفی خواهند شد. در فصل چهارم، ارزیابی روش‌های پیشنهادی و آزمایش‌ها را خواهیم داشت و در پایان، در فصل پنجم، نتیجه‌گیری و کارهای آینده پیشنهادی معرفی خواهند شد.

فصل دوم:

مروری بر کارهای

انجام شده

تصاویر دیجیتال در حال حاضر بطور گسترده‌ای در پزشکی، معماری، تشخیص چهره، تشخیص اثر انگشت، بیومتریک و غیره استفاده می‌شود. از این رو، جستجو و بازیابی کارآمد و موثر تصویر امری مهم تلقی می‌شود. سیستم‌های بازیابی تصویر مبتنی بر متن، بازیابی را با استفاده از کلمات کلیدی انجام می‌دادند. این کلمات کلیدی توسط نیروی کار انسانی به تصویر داده می‌شد. با توجه به اینکه تعداد تصاویر پایگاه داده افزایش یافته و همچنین برچسب‌گذاری برای تصاویر امری زمان‌بر و هزینه‌بردار است، این نوع بازیابی تصویر منسوخ شده و بازیابی تصویر مبتنی بر محتوا جای آن را گرفته است [۱،۲].

بازیابی تصویر مبتنی بر محتوا ابتدا با استفاده از تنها یک ویژگی سطح پایین انجام می‌شد. استفاده از تنها یک ویژگی سطح پایین باعث بوجود آمدن شکاف معنایی بین درک سیستم و درک انسان از تصویر مورد جستجو می‌شود. لذا جهت کاهش این شکاف معنایی و رسیدن به مفاهیم معنایی تصویر، روش‌های کم‌هزینه و بدون یادگیری و آموزش مثل استفاده از ویژگی‌های سطح پایین بصورت ترکیبی و یا روش‌های با آموزش طولانی و پردازش‌های زیاد مثل روش‌های مبتنی بر یادگیری ماشین و شناسایی اشیاء، بوجود آمدند. در این پایان‌نامه از ترکیب ویژگی‌های سطح پایین استفاده شده است که در عین سرعت و دقت بالا در استخراج محتوای معنایی تصویر، نیازی به آموزش و هزینه‌های زیاد ندارد.

در ادامه بطور کلی مهم‌ترین ویژگی‌های سطح پایین مورد استفاده در بازیابی تصویر، بشکل اجمالی مورد بررسی قرار می‌گیرد (بخش ۲-۲). پس از آن روش‌های کاهش شکاف معنایی بصورت دسته‌بندی شده، مورد بررسی قرار می‌گیرد (بخش ۲-۳). در این بخش، چند کار اخیر انجام شده در حوزه بازیابی تصویر مبتنی بر محتوا بررسی می‌شود.

۲-۲- ویژگی‌های سطح پایین مورد استفاده در سیستم CBIR

ویژگی در سیستم‌های CBIR، مشخصه و خصوصیتی از تصویر است که می‌تواند مشخصه‌ی بصری خاصی از تصویر استخراج کند. این ویژگی بصورت کلی از همه تصویر و یا بصورت محلی از ناحیه‌ها یا اشیاء موجود در تصویر گرفته شود. ویژگی‌های سطح پایینی که معمولاً در CBIR استفاده می‌شود، رنگ، بافت و شکل هستند [۳]. در ادامه مهم‌ترین روش‌های استخراج هریک بصورت مختصر تشریح می‌شود.

۲-۲-۱- ویژگی رنگ

رنگ یکی از متداول‌ترین و تعیین‌کننده‌ترین ویژگی‌های دیداری در زمینه بازیابی تصویر است؛ چرا که نسبت به تغییرات مربوط به اندازه، جهت، دورنمایی و اغتشاش تصویر پایدار است [۸].

استفاده از ویژگی رنگ در سیستم‌های بازیابی تصاویر مبتنی بر محتوا کاربردهای بسیار گسترده‌ای دارد. عمومی‌ترین روش در بازیابی رنگ تصاویر استفاده از هیستوگرام رنگ است، که از مهم‌ترین تکنیک‌های بازیابی تصویر مبتنی بر محتوا است [۸]. یکی از مهم‌ترین بحث‌ها در استخراج ویژگی رنگ، فضای رنگ است که عملیات استخراج رنگ در آن انجام می‌شود. در ادامه مهم‌ترین فضاهای رنگ مورد استفاده در بحث CBIR مختصر توصیف می‌شوند.

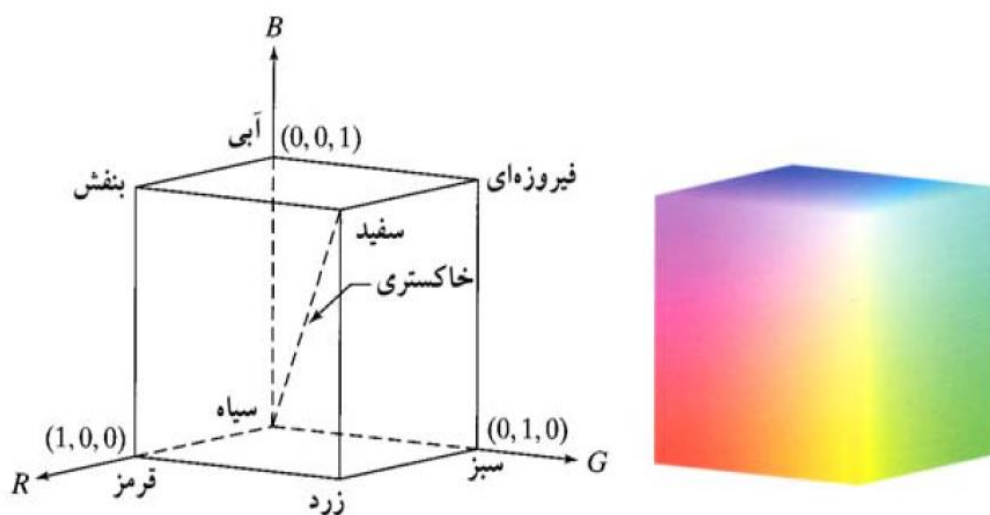
۲-۲-۱-۱- فضای رنگ RGB

برای استخراج ویژگی رنگ از تصویر، ابتدا باید نوع فضای رنگ را تعیین نمود. فضاهای رنگ متفاوتی وجود دارند. مهم‌ترین و ساده‌ترین فضای رنگ، RGB است که برای توصیف رنگ‌ها استفاده می‌شود. این فضای رنگ مرتبط با منابع نور روشن‌کننده می‌باشد. ترکیب تمام رنگ‌ها با هم، رنگ سفید را ایجاد می‌کند [۱].

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. منظور از یکنواختی ادراکی این است که در این فضا تفاوت دو نقطه با توجه به مختصات آن‌ها، متناسب با درک انسان نیست. به عبارت دیگر فاصله اقلیدسی، تفاوت بین دو رنگ را به شکلی که برای انسان معنی‌دار باشد، بیان نمی‌کند [۱].

فضای رنگی RGB با مکعب رنگی RGB نمایش داده می‌شود. شکل ۱-۲ این فضای رنگ را

نشان می‌دهد.

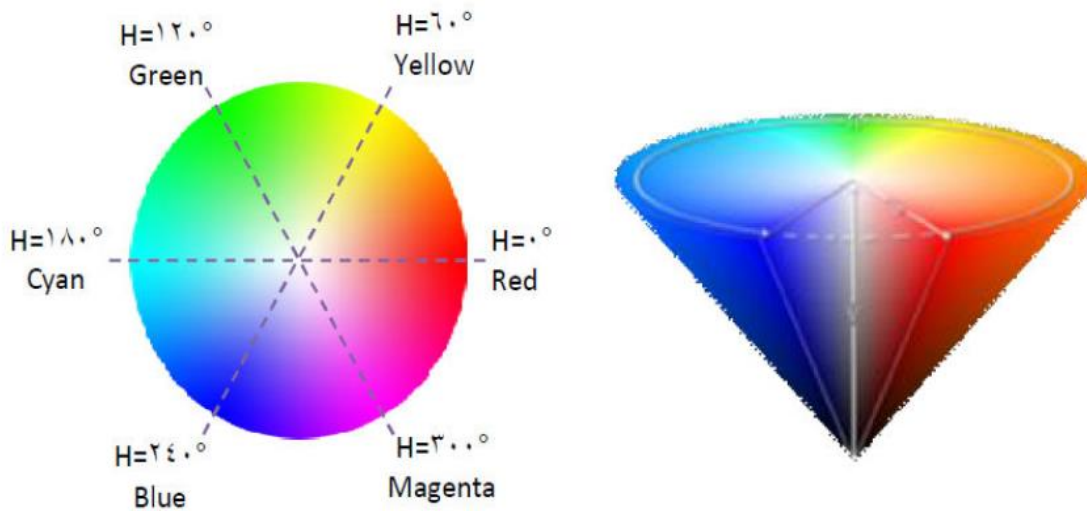


شکل ۱-۲-۱- فضای رنگی RGB

۲-۱-۲-۲- فضای رنگ HSV

فضای رنگ 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 ۲-۲ این فضای رنگ را نشان می‌دهد.



شکل ۲-۲- فضای رنگی 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 [۸-۱۸]. 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 چندی‌سازی می‌شود.

اگر فرض کنیم که $r, g, b \in [0,1]$ به ترتیب معرف محورهای قرمز، سبز و آبی در فضای رنگ RGB باشد و H, S, V معرف مولفه‌های فضای HSV باشند. همچنین $H \in [0,360]$ ، $S \in [0,1]$ و $V \in [0,1]$ باشد. فرض می‌شود که \max معرف بیشینه مقدار r, g, b و \min معرف کمینه آن‌ها باشد، آنگاه برای تبدیل از RGB به HSV از روابط (۲-۱)، (۲-۲) و (۳-۲) استفاده می‌کنیم:









$$h = \begin{cases} 0 & , \text{if } \max = \min \\ \left(60^\circ \frac{g-b}{\max-\min} + 0^\circ\right) \bmod 360^\circ & , \text{if } \max = r \\ 60^\circ \frac{b-r}{\max-\min} + 120^\circ, & , \text{if } \max = g \\ 60^\circ \frac{r-g}{\max-\min} + 240^\circ, & , \text{if } \max = b \end{cases} \quad (1-2)$$

$$S = \begin{cases} 0, & \text{if } \max = 0 \\ 1 - \min/\max & \text{otherwise} \end{cases} \quad (2-2)$$

$$V = \text{Max} \quad (3-2)$$

در جدول ۱-۲ اطلاعات مربوط به برخی رنگ‌ها با نام تجاری و معادل RGB و HSV آن‌ها آمده است.

جدول ۱-۲-۱- اطلاعات مربوط به برخی رنگ‌های خاص و معادل RGB و HSV آن‌ها

HSV	RGB	نام رنگ	رنگ
($0^\circ, 0, 0$)	($0, 0, 0$)	Black	
($0^\circ, 0, 1$)	($255, 255, 255$)	White	
($0^\circ, 1, 1$)	($255, 0, 0$)	Red	
($60^\circ, 1, 1$)	($255, 255, 0$)	Yellow	
($120^\circ, 1, 1$)	($0, 255, 0$)	Green	
($180^\circ, 1, 1$)	($0, 255, 255$)	Cyan	
($240^\circ, 1, 1$)	($0, 0, 255$)	Blue	
($300^\circ, 1, 1$)	($255, 0, 255$)	Magenta	

یکی از مهم‌ترین مزیت‌های فضای رنگ HSV خاصیت یکنواختی ادراکی آن است. این مزیت باعث

شده است که نسبت به دیگر فضاهای رنگ بیشتر مورد استفاده پژوهشگران قرار بگیرد [۱۸-۸].

۲-۲-۱-۳- فضای رنگی $L^*a^*b^*$

فضای رنگ $L^*a^*b^*$ نیز دارای خاصیت یکنواختی ادراکی است. مولفه‌های این فضای رنگی تشکیل شده از L^* که معرف روشنایی است و بین سفید تا سیاه متغیر است، a^* که معرف قرمز-سبز 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، تبدیل سایر فضاهای رنگ به این فضا است زیرا غیرخطی و مستلزم محاسبات بالاست [۱۹].

پس از اینکه مهم‌ترین فضاهای رنگ مورد استفاده در CBIR تشریح شد، روش‌های استخراج ویژگی‌های سطح پایین مختلف از جمله رنگ، بافت و شکل را مختصر توضیح می‌دهیم.

۲-۲-۱-۴- هیستوگرام رنگ

هیستوگرام رنگ یک تصویر، نموداری است که تعداد پیکسل برای هر مقدار سطح خاکستری را نشان می‌دهد. برای یک تصویر خاکستری ۸ بیتی مقادیر شدت پیکسل‌ها در فاصله ۲۵۶ شدت مختلف است. بنابراین، هیستوگرام به شکل گرافیکی مقادیر ۲۵۶ سطح خاکستری را نمایش می‌دهد. این سطوح، نشان‌دهنده‌ی توزیع پیکسل‌ها در میان آن دسته‌ها خواهد بود. هیستوگرام همچنین می‌تواند از تصاویر رنگی گرفته شود. یک هیستوگرام رنگی نمایشی از توزیع رنگ‌ها در تصویر را نشان می‌دهد. در واقع هیستوگرام رنگی، توزیع آماری استاندارد از توزیع رنگ از نظر فرکانس‌های سطوح خاکستری در ناحیه‌های مختلف از یک فضای رنگ است.

برای ایجاد یک هیستوگرام رنگی، فضای رنگی به ناحیه‌هایی تقسیم می‌شود. ۲۴ بیت فضای رنگی RGB، $256 * 256 * 256$ ناحیه رنگی مختلف دارد. یک هیستوگرام شامل این تعداد سطح، برای کار کردن، خیلی بزرگ است. از این رو فضای رنگی به تعدادی بین^۴ چندی‌سازی می‌شود که در آن

⁴ bin

هر بین یک محدوده‌ای از مقادیر رنگی را نشان می‌دهد. تعداد پیکسل‌ها در تصویر برای هر یک از این محدوده‌ها، محاسبه می‌شوند. تعداد بین‌ها براساس کاهش تحمل دقت و نیاز به حافظه، بدست می‌آید. نمودار هیستوگرام رنگ در فضاهای رنگ مختلف ساخته می‌شود [۸].

پژوهش‌های زیادی از هیستوگرام رنگ برای استخراج اطلاعات رنگی تصویر استفاده کرده‌اند.

هیستوگرام رنگ در فضاهای رنگ HSV و $L^*a^*b^*$ نیز استفاده می‌شود [۸-۱۸].

۲-۲-۲- ویژگی بافت

ویژگی‌های بافت با استفاده از ماتریس هم رخدادی سطح خاکستری (GLCM)^۵، تبدیل گابور^۶، الگوهای دودویی محلی^۷ و ویژگی‌های تامورا^۸ استخراج می‌شوند. این روش‌های استخراج ویژگی بافت در ادامه توصیف می‌شوند.

۲-۲-۲-۱- ماتریس هم رخدادی سطح خاکستری (GLCM)

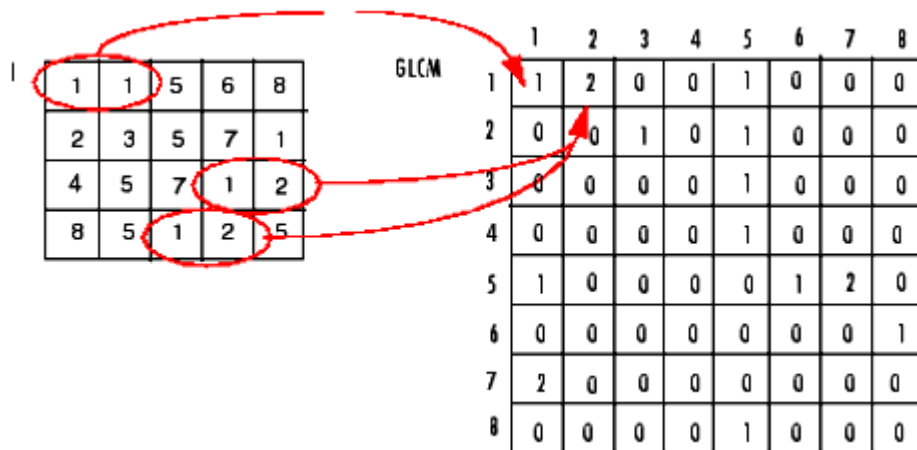
GLCM از یک تصویر در مقیاس خاکستری ایجاد می‌شود. فرض کنید یک پیکسل با سطح خاکستری i بصورت افقی، عمودی و یا مورب با پیکسل‌های مجاور با ارزش z رخ می‌دهد. تعداد این رخدادها توسط GLCM محاسبه می‌شود. این مقادیر توسط فروانی نسبی وقوع دو پیکسل سطح خاکستری i و z داده شده است. فاصله این پیکسل‌ها با d مشخص می‌شود. همچنین جهت محاسبه ماتریس GLCM نیز با θ قابل جدا شدن است. d می‌تواند مقادیر ۱، ۲، ۳، ... و θ می‌تواند مقادیر صفر درجه (افقی)، ۹۰ درجه (عمودی)، ۴۵ درجه و ۱۳۵ درجه (مورب) را داشته باشد. ماتریس GLCM در شکل ۲-۳ نشان داده شده است [۲۰].

⁵ Gray Level Co-occurrence Matrix

⁶ Gabor Transform

⁷ Local binary patterns

⁸ Tamura Feature



شکل ۲-۳- نمونه‌ای از تشکیل ماتریس هم‌رخدادی GLCM [۲۰].

چند خواص بافت آماری مانند کنتراست^۹، همبستگی^{۱۰}، انرژی^{۱۱}، همگن بودن^{۱۲} و آنتروپی^{۱۳}

می‌توانند از GLCM مشتق شوند و فرمول‌های آن‌ها از ۲-۴ تا ۲-۸ ارائه شده است [۲۰].

$$\text{Contrast} = \sum_i \sum_j |i - j|^2 p(i, j) \quad (۴ - ۲)$$

$$\text{Correlation} = \sum_i \sum_j \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (۵ - ۲)$$

$$\text{Energy} = \sum_i \sum_j p(i, j)^2 \quad (۶ - ۲)$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{p(i, j)}{1 + |i - j|} \quad (۷ - ۲)$$

$$\text{Entropy} = \sum_i \sum_j p(i, j) \log p(i, j) \quad (۸ - ۲)$$

[۵] یکی از پژوهش‌هایی است که از این ماتریس برای استخراج ویژگی جهت بازیابی تصاویر استفاده

می‌کند.

^۹ Contrast

^{۱۰} Correlation

^{۱۱} energy

^{۱۲} homogeneity

^{۱۳} entropy

۲-۲-۲-۲- تبدیل گابور^{۱۴}

برای هر تصویر $I(x,y)$ با اندازه‌ی $p*q$ ، تبدیل موجک^{۱۵} گابور گسسته بصورت معادله (۲-۹) بدست می‌آید.

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right] \quad (۲-۹)$$

که در آن W فرکانس نوسان است و پهناى باند فرکانسى فیلتر گابور را مشخص می‌کند. بنابراین پس از انتخاب یک موجک برای فیلتر، موجک در Scale‌های متفاوت و با زوایای متفاوت با تصویر convolve می‌شود. این مقادیر برای استخراج ویژگی استفاده می‌شوند. بطور کلی این روش استخراج بافت، یک نوع فیلتر بالاگذر محسوب شده و از حاصل ضرب یک تابع گوسین در یک تابع مختلط سینوسی ساخته می‌شود و بطور کلی لبه‌های تصویر را استخراج می‌کند [۱۱،۲۰].

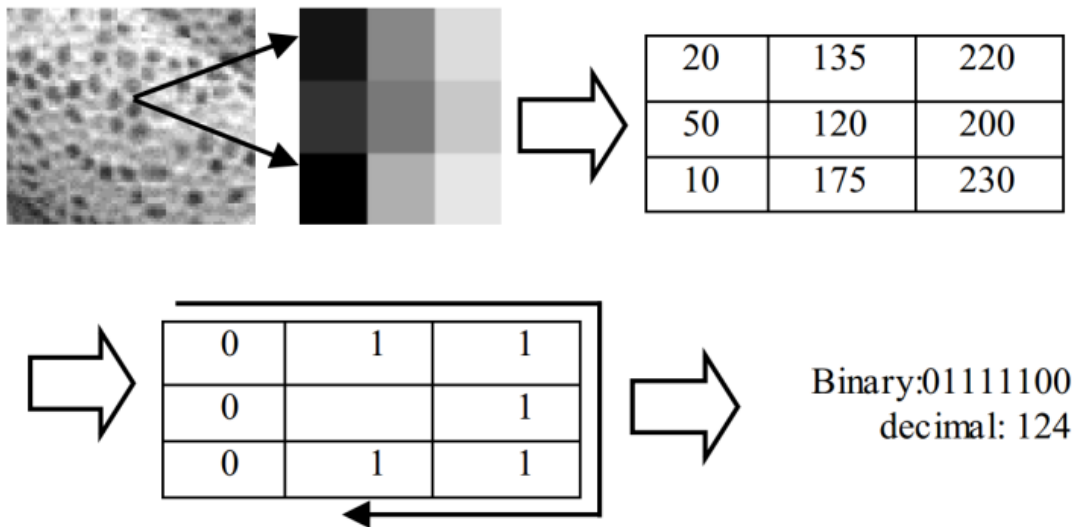
۲-۲-۲-۳- الگوهای دودویی محلی (LBP)

عملگر LBP به عنوان یک توصیف‌گر قدرتمند بافت تصویر در [۲۱] معرفی شده است. این عملگر برای هر پیکسل با توجه به برچسب پیکسل‌های همسایگی 3×3 ، یک عدد دودویی تولید می‌کند. این برچسب‌ها با آستانه‌سازی مقدار پیکسل‌های همسایه با مقدار پیکسل مرکزی به دست می‌آیند. به این صورت که برای پیکسل‌های با مقدار بزرگ‌تر یا مساوی مقدار پیکسل مرکزی برچسب ۱ و برای پیکسل‌های با مقادیر کوچک‌تر از مقدار پیکسل مرکزی برچسب ۰ قرار می‌گیرد. سپس این برچسب‌ها به صورت چرخشی در کنار هم قرار گرفته و یک عدد ۸ بیتی را تشکیل می‌دهند. در نهایت این عدد به صورت دهدهی^{۱۶} جای مقدار پیکسل مرکزی قرار می‌گیرد. نحوه کار این عملگر در شکل ۲-۴ آمده است.

¹⁴ Gabor Transform

¹⁵ Wavelet

¹⁶ Decimal



شکل ۲-۴- روش کار الگوریتم بافتی الگوهای باینری محلی (LBP)

۲-۲-۴- ویژگی‌های تامورا

این ویژگی‌های بافتی جزو روش‌های آماری استخراج بافت هستند. این ویژگی‌ها بر اساس مطالعات روانشناسی بر روی ادراک انسان از تصویر نسبت به بافت، طراحی شده است و شامل چند مشخصه از تصویر می‌شود. این مشخصه‌ها عبارتند از: زبری (دانه دانه شدن بافت)^{۱۷}، کنتراست، جهت‌دار بودن^{۱۸}، منظم بودن^{۱۹}، شباهت خطی^{۲۰} و سختی^{۲۱} بافت [۱۵].

هر یک از این ویژگی‌ها توسط رابطه‌ای جداگانه به دست می‌آید. سه ویژگی اول در سیستم‌های بازیابی نتایج بسیار موفق داشته‌اند. برای محاسبه «درشتی و زبری» بافت، میانگین همسایگی تمام پیکسل‌های تصویر در همسایگی $2^K \times 2^K$ به دست می‌آید. این ویژگی از طریق رابطه (۲-۱۰) به دست می‌آید [۲۷] [۳۰].

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{y+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i, j) / 2^{2k} \quad (2-10)$$

¹⁷ Coarseness
¹⁸ directionality
¹⁹ regularity
²⁰ linelikeness
²¹ Roughness

در هر نقطه، اندازه مناسب برابر با بیشترین مقدار به دست آمده از میانگین های موجود است. میزان درشتی و ریزی بافت در هر نقطه برابر مقدار همسایگی به دست آمده است. اولی از محاسبه انحراف استاندارد مقادیر پیکسل های تصویر خاکستری شده به دست می آید و دومی برابر نقطه اوج تصویر است. بنابراین کنتراست از رابطه (۲-۱۱) به دست می آید.

$$F_{con} = \frac{\sigma}{(a_f)^n}, a_f = \mu_f / \sigma^f \quad (2-11)$$

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

۲-۲-۳- ویژگی شکل

ویژگی شکل با استفاده از روش های مختلفی استخراج می شود که مهم ترین آن ها هیستوگرام جهت لبه است که در ادامه مورد بررسی قرار می گیرد.

۲-۲-۳-۱- هیستوگرام جهت لبه

هیستوگرام جهت لبه، اطلاعات اصلی مربوط به شکل در تصویر را دریافت می کند. اطلاعات لبه موجود در تصویر با استفاده از الگوریتم های تشخیص لبه مانند سوبل^{۲۲}، رابرتز^{۲۳}، کنی و غیره بدست می آید. هر یک از الگوریتم های استخراج لبه، نوعی از لبه های تصویر را استخراج می کنند. برای مثال، الگوریتم سوبل لبه های افقی و عمودی و روش رابرتز لبه های قطری تصویر را استخراج می کنند. این مطلب در شکل ۲-۵ نشان داده شده است. سپس با محاسبه زاویه لبه های بدست آمده، جهت لبه در هر پیکسل محاسبه می شود. جهت های لبه محاسبه شده به تعدادی بین چندی سازی می شود. فراوانی این مقادیر،

²² sobel

²³ roberts

-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 ۲-۵- الف) ماسک رابرتز ب) ماسک سوئل

هیستوگرام جهت لبه را می‌سازند. نوعی از این تشخیص لبه (Histogram of Oriented Gradient) در پژوهشی استفاده شده که در ادامه توضیح داده می‌شود [۱۷].

با توجه به مطالعات انجام شده در حوزه CBIR می‌توان گفت، بطور کلی روش‌های کاهش شکاف معنایی، به دو دسته تقسیم خواهند شد. دسته اول روش‌هایی هستند که نیازمند پردازش‌های سنگین و هزینه‌های زیادی هستند. همچنین به زمان و تصاویر زیادی جهت آموزش در این نوع روش‌ها نیاز است. این دسته با نام "سیستم‌های CBIR مبتنی بر یادگیری و هزینه‌بردار" شناخته می‌شوند. در سال ۲۰۰۹ پروژه ImageNet یک پایگاه داده از ۱۵ میلیون عکس در وسعت ۲۲/۰۰۰ کلاس از شی‌ها آماده شد، که از لحاظ کیفیت و کمیت این مقیاس بی‌سابقه می‌باشد. بعنوان مثال در مورد گربه، بیش از ۶۲۰۰۰ (تصویر) گربه در انواع شکل‌ها و فرم بدن و در تمام گونه‌های اهلی و وحشی موجود می‌باشد. دسته دوم روش‌هایی هستند که در مقابل روش‌های اول قرار گرفته و در عین کاهش شکاف معنایی، به هزینه و پردازش‌های سنگین، کم‌تر نیاز دارند. این دسته با نام "سیستم-های CBIR مبتنی بر ترکیب ویژگی‌های سطح پایین" شناخته می‌شوند. این روش‌ها، طوری از ویژگی‌های سطح پایین جهت بازیابی استفاده می‌کنند که به نوعی مفاهیم معنایی را از تصویر استخراج می‌کنند. پس از استخراج این مفاهیم معنایی، از آن‌ها جهت مقایسه تصاویر و بازیابی تصاویر نزدیک‌تر، استفاده می‌کنند. بنابراین ما نیز در این پایان‌نامه، روشی از نوع دسته دوم ارائه خواهیم کرد. در ادامه از دو دسته مذکور، چند کار انجام شده اخیر بهتر را تشریح خواهیم کرد.

۲-۳- سیستم‌های CBIR مبتنی بر یادگیری و هزینه‌بردار

همان‌طور که ذکر شد، ویژگی‌های سطح پایین به تنهایی نمی‌توانند کارایی لازم را برای بازیابی مناسب تصاویر داشته باشند. بنابراین محققان به دنبال کاهش شکاف معنایی موجود بین ویژگی‌های سطح پایین و مفاهیم ادراکی موجود در تصاویر، روش‌هایی ارائه دادند. تکنیک‌های بسیاری برای کاهش شکاف معنایی بین ویژگی‌های سطح پایین و معناشناسی سطح بالا، ایجاد و تکامل یافته‌اند. یکی از این روش‌ها استفاده از الگوریتم‌های یادگیری ماشین، شناسایی اشیاء و غیره است. یکی از معایبی که این نوع سیستم CBIR دارد، این است که در عین کارایی خوب، نیازمند آموزش طولانی، پردازش‌های زیاد و هزینه‌های سرسام‌آور است. در ادامه مهم‌ترین روش‌هایی که در این سیستم‌ها استفاده می‌شوند، مختصر توصیف می‌شوند.

۲-۳-۱- تکنیک‌های یادگیری ماشین

مفاهیم ادراکی تصویر با استفاده از تکنیک‌های یادگیری ماشین قابل استخراج هستند. این تکنیک‌ها بطور کلی به سه نوع بامربی^{۲۴}، بدون مربی^{۲۵} و یادگیری عمیق^{۲۶} تقسیم می‌شوند.

۲-۳-۱-۱- تکنیک‌های یادگیری ماشین بامربی

شبکه‌های عصبی^{۲۷} [۲۲،۲۳]، درخت‌های تصمیم^{۲۸} [۲۴] و ماشین‌های بردار پشتیبان (SVMها)^{۲۹} [۱۶،۲۵] برخی از روش‌های یادگیری ماشین بامربی هستند که مفاهیم ادراکی را از تصاویر استخراج و یاد می‌گیرند. مفاهیم ادراکی مانند اشیاء، ارتباط بین آن‌ها و غیره است. تکنیک‌های یادگیری ماشین بامربی، فرآیند کلاس‌بندی را با کمک داده‌ی آموزشی که دسته‌بندی اولیه شده، انجام می‌-

²⁴ Supervised learning

²⁵ Unsupervised learning

²⁶ Deep Learning

²⁷ Neural Networks

²⁸ Decision Trees

²⁹ Support Vector Machines

دهند [۲۶،۲۷]. برای داده‌های آموزشی، ورودی (مفاهیم ادراکی) و خروجی مطلوب (تصاویر مربوط به دسته همان تصویر مورد جستجو) هستند. در ابتدا هر دو (داده ورودی و خروجی) شناخته شده هستند. هنگامی که الگوریتم‌های یادگیری با مربی با داده‌های آموزشی شناخته شده آموزش می‌بینند، آن می‌تواند به داده‌های جدید دیده نشده، تعمیم پیدا کند. از این رو با توجه به یک تصویر مورد جستجو، مفاهیم ادراکی از آن استخراج شده و به عنوان ورودی به هریک از الگوریتم‌های یادگیری ماشین داده می‌شود. سپس با داده‌های آموزشی، آموزش اولیه داده می‌شود. این آموزش تحت تعداد تصاویر بسیار زیادی است و به پردازش‌های زیاد و سیستم‌های قدرتمند و نهایتاً هزینه و زمان زیادی نیاز دارد.

الگوریتم یادگیری ماشین، مفهوم ادراکی تصویر مورد جستجو را استخراج و دسته تصاویر مربوط به این مفهوم ادراکی را پیدا می‌کند. از این رو، بجای پیدا کردن شباهت بین تصویر مورد جستجو و سایر تصاویر پایگاه داده، این سیستم، شباهت را فقط بین تصویر مورد جستجو و تصاویری که متعلق به دسته آن تصویر است، محاسبه و تصاویر نزدیک‌تر را بازیابی می‌کند. زمانی که تکنیک‌های یادگیری ماشین استفاده شود، پیش‌بینی از دسته تصویر مورد جستجو، انجام می‌شود. بنابراین نتایج بازیابی شامل فقط تصاویر مربوط به دسته تصویر مورد جستجو است.

۲-۱-۳-۲- تکنیک‌های یادگیری ماشین بدون مربی

روش آموزش بدون مربی، مشکل تلاش برای پیدا کردن ساختار نهان در داده‌های بدون برچسب را ترجیح می‌دهد. این روش هیچ اندازه‌گیری خروجی برای هدایت فرآیند یادگیری ندارد. خوشه‌بندی تصویر نوعی از تکنیک‌های یادگیری بدون مربی است. این روش طوری تصاویر را گروه‌بندی می‌کند که شباهت در بین تصاویر یک خوشه بیشترین باشد و شباهت با دیگر دسته‌ها به حداقل برسد.

در واقع الگوریتم‌های یادگیری بدون مربی اطلاعات تصویر دسته‌بندی نشده را به عنوان ورودی می‌گیرند. سپس این داده‌ها را به تعدادی دسته با پیدا کردن ساختار پنهان داده‌هایی که برچسب ندارند، دسته‌بندی می‌کنند. الگوریتم‌های خوشه‌بندی، داده‌ی داده شده را به n خوشه و

مرکزهای خوشه‌بندی تقسیم‌بندی می‌کنند. هنگامی که یک تصویر مورد جستجو داده می‌شود، الگوریتم خوشه‌بندی فاصله بین تصویر مورد جستجو با سایر مراکز خوشه‌ها را پیدا می‌کند. تصویر مورد جستجو متعلق به خوشه‌ای است که فاصله حداقل را دارد. بنابراین در اینجا نیز فاصله شباهت، بین تصویر مورد جستجو و تصاویر متعلق به خوشه‌ی پیش‌بینی شده، پیدا می‌شود.

از این روش‌ها می‌توان K-Means را نام برد [۲۷]. خوشه‌بند K-Means قصد دارد n مشاهده‌ی داده شده را به k خوشه، تقسیم کند. میانگین هر خوشه مشخص می‌شود. سپس تصویری که میانگین آن دارای کمترین فاصله اقلیدسی با بردار ویژگی تصویر است، در یک خوشه قرار می‌گیرد. با توجه به توزیع پیچیده‌ی داده‌های تصویر، خوشه‌بند K-Means اغلب نمی‌تواند تصاویر با مفاهیم ادراکی مختلف را به اندازه کافی از هم جدا کند.

از دیگر روش‌ها در حوزه‌ی روش‌های بدون مربی، می‌توان به مدل‌های خوشه‌بندی فازی^{۳۰} [۶، ۲۸]، KNN^{۳۱} [۲۷] و یادگیری انتقال برای خوشه‌بندی [۲۹] نام برد. در [۳۰] از یک معیار بیزین جهت امتیازدهی به کلاس تصاویر استفاده می‌شود. در ادامه پژوهش‌هایی از این نوع روش‌ها بیان خواهد شد.

۲-۳-۱-۳-۲- یادگیری عمیق

از دیگر روش‌هایی که به گروه روش‌های یادگیری ماشین تعلق دارند، سیستم‌هایی هستند که از یادگیری عمیق استفاده می‌کنند. این سیستم‌ها، در طی چندین مرحله، یک تصویر را جستجو و مفاهیم موجود در آن را یاد می‌گیرند. طوری که برای آموزش یک مفهوم ادراکی به این نوع سیستم‌ها، تعداد خیلی زیادی نمونه تصویری لازم است. پس از اینکه، سیستم یک مفهوم ادراکی را درک کرد، این نوع سیستم، با درصد بالایی تصاویر دارای آن مفهوم را شناسایی می‌کنند. از معایب این روش،

³⁰ Fuzzy clustering models

³¹ K-nearest neighbor

می‌توان به تعداد زیاد تصاویر مورد نیاز، پردازش‌های زیاد و در پی آن، هزینه‌های سرسام‌آور، اشاره کرد. از پژوهش‌های انجام شده در این حوزه می‌توان به [۳۱,۳۲] اشاره نمود.

۲-۳-۲- شناسایی اشیاء

روش‌های دیگری که در مجموعه روش‌های مبتنی بر یادگیری و هزینه‌بردار قرار می‌گیرد، می‌توان به روش‌های موجود جهت شناسایی اشیاء و مفاهیم ادراکی اشاره کرد. یکی از این روش‌ها قطعه‌بندی تصویر است. یعنی ابتدا باید تصویر با یکسری پیش‌پردازش‌ها، قطعه‌بندی شده و سپس از قطعات بدست آمده، اشیاء شناسایی می‌شوند. از کارهای انجام شده در این حوزه می‌توان به قطعه‌بندی معنایی با استفاده از شبکه‌های پیچش کامل^{۳۲} [۳۲] اشاره کرد. در این مقاله، در طی چندین مرحله و با پیش‌پردازش‌های شبکه‌های کانولوشن، ابتدا تصویر قطعه‌بندی شده و سپس اشیاء موجود در آن شناسایی می‌شوند. از دیگر روش‌ها در این حوزه می‌توان به [۳۳] نیز اشاره کرد.

در ادامه چند نمونه از روش‌های مبتنی بر یادگیری و هزینه‌بردار را بیشتر توضیح می‌دهیم.

در [۷] زارچی و همکاران مدلی مفهومی جهت CBIR ارائه داده‌اند. در این مقاله، به منظور رسیدگی به این مشکل بحرانی شکاف معنایی در توصیف مفاهیم سطح بالا، یک مدل جدید مبتنی بر محتوا معرفی شده است. مدل پیشنهادی تصاویر را در دو مرحله مفهومی بازیابی می‌کند. در مرحله اول، اشیاء با استفاده از مدل مبتنی بر قطعه جداکننده (DPM^{۳۳}) شناسایی می‌شوند. این مرحله، لایه شیء نام دارد. در مرحله دوم، با توجه به رابطه بین اشیاء موجود در تصویر (مثلاً موقعیت قرارگیری فرد نسبت به اسب در یک عکس مربوط به مسابقه اسب‌سواری)، یک مفهوم ادراکی (مفهوم مسابقه اسب-سواری)، از تصویر استخراج می‌شود.

³² Fully Convolutional Networks

³³ Discriminative part-based

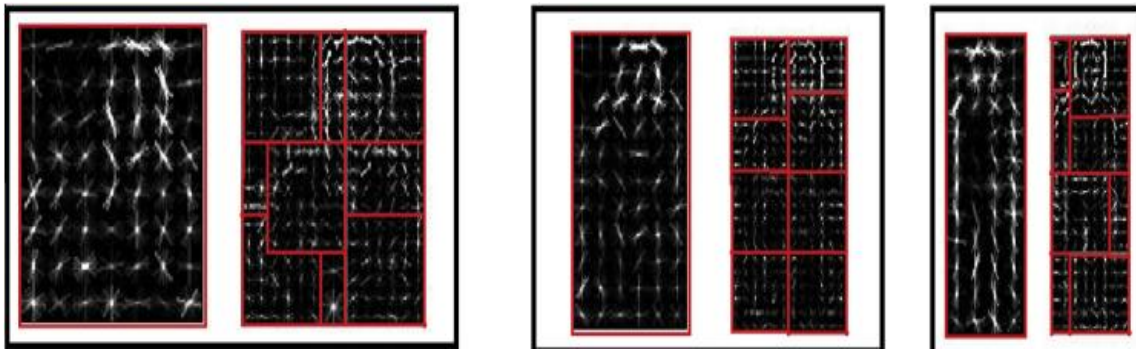
در لایه اول، قبل از عمل شناسایی اشیاء، باید تصویر به شبکه HOG^{۳۴} نگاشت شود. ایده‌ی اصلی در روش HOG این است که توزیع گرادیان‌های محلی یا جهت‌های لبه می‌تواند بخوبی شکل را توصیف کند، حتی اگر اطلاعاتی از موقعیت دقیق گرادیان و یا لبه‌ی متناظر موجود نباشد. این ویژگی، جهت گرادیان‌های تصویر را در یک همسایگی محلی نشان می‌دهد. بسته به اینکه گرادیان بدون علامت یا علامت‌دار در نظر گرفته شود، فاصله‌ی بین ۰-۱۸۰ درجه یا ۰-۳۶۰ درجه به n قسمت مساوی تقسیم می‌شود. این تعداد n، جهت گرادیان یا همان بین‌های هیستوگرام هستند و هر کدام از این فاصله‌ها یک کانال هیستوگرام را تشکیل می‌دهد. برای هر پنجره، هیستوگرام جهت‌های گرادیان برای پیکسل‌های داخل پنجره محاسبه می‌شود و مقدار ماکزیمم در هر کانال فضای رنگی برای هر پیکسل به عنوان مقدار گرادیان آن پیکسل، به خروجی داده می‌شود. نمونه‌ای از این هیستوگرام در شکل (۲-۶) مشاهده می‌شود.



شکل ۲-۶- نمونه‌ای از کارکرد شبکه‌ی HOG برای تصویر [۷].

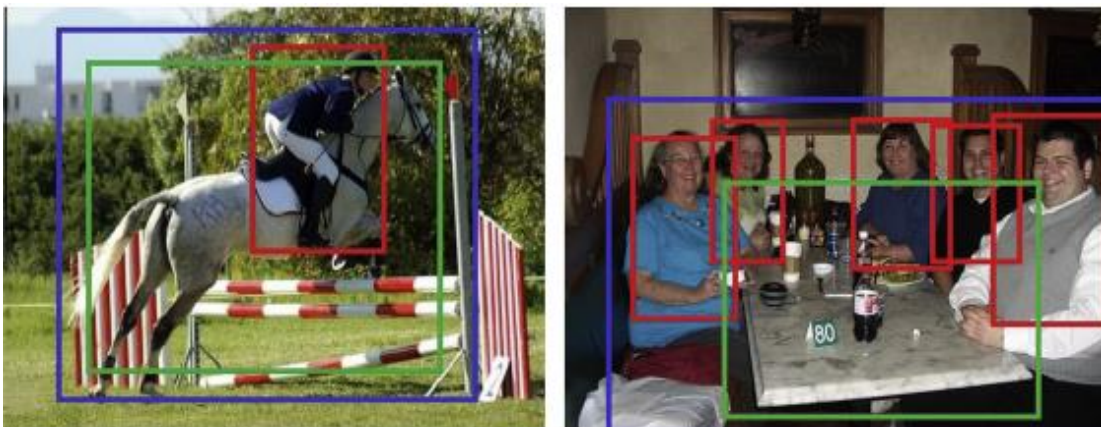
پس از این کار برای آموزش هر شیء به سیستم، از روش شناسایی قطعه جدا کننده استفاده می‌شود. بدین صورت که ابتدا شیء اصلی در تصویر شناسایی و سپس، حالت‌های مختلف قرارگیری آن در پایگاه داده، با استفاده از فیلترهای قطعه، به سیستم CBIR آموزش داده می‌شود (شکل ۲-۷).

³⁴ Histogram of Oriented Gradien



شکل ۲-۷- آموزش حالت‌های مختلف قرارگیری شیء انسان در مقابل دوربین، با استفاده از DPM [۷].

پس از شناسایی اشیاء موجود در تصویر، رابطه‌ی بین آن‌ها (به عنوان مفهوم ادراکی بصری) در تصویر نیز استخراج می‌شود. این رابطه بسته به موقعیت قرارگیری اشیاء نسبت به هم استخراج می‌شود. سپس از اشیاء و روابط شناسایی شده، به عنوان ویژگی جهت بازیابی تصویر، استفاده می‌شود. به عنوان مثال شکل ۲-۸ نمونه‌ای از شناسایی اشیاء و مفهوم بصری در تصویر را نشان می‌دهد.



شکل ۲-۸- دو نمونه از شناسایی مفهوم بصری. اشیاء با چهارگوش‌های بسته‌ی قرمز و سبز، مفاهیم بصری با چهارگوش‌های بسته‌ی آبی رنگ مشخص شده‌اند. تصویر سمت چپ شامل مفهوم "اسب سواری" و تصویر سمت راست شامل مفهوم "نشستن افراد دور میز ناهار خوری" هستند [۷].

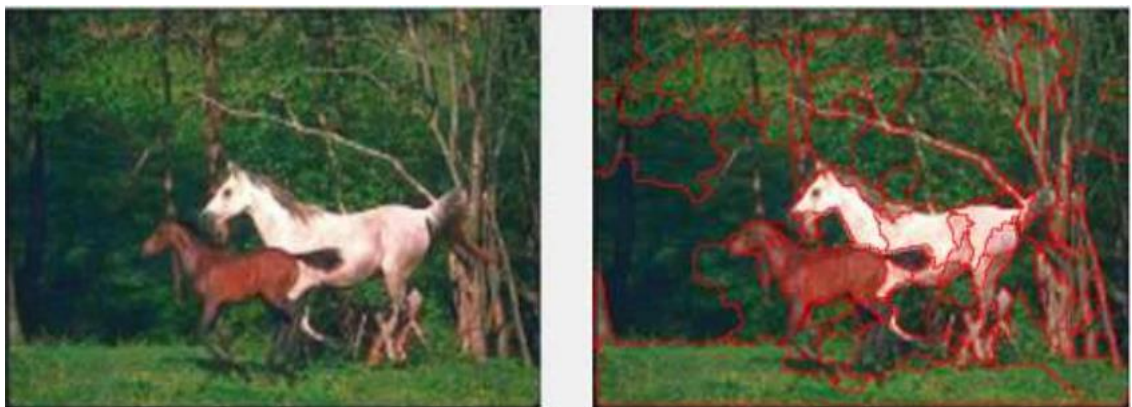
در [۳۰] روشی بین دو روش یادگیری با مربی و بدون مربی (نیمه نظارتی^{۳۵}) با استفاده از معیار بیزین روی الگوریتم پیشنهادی آن، معرفی شده است. این پژوهش در مدل توزیع ویژگی‌های رنگ و بافت در مجموعه‌ای از تصاویر مرتبط، یک چهاچوب بیزین ارائه می‌دهد. با توجه به سوال متنی مشخص کاربر (مثلاً پنگوئن) سیستم ابتدا مجموعه‌ای از تصاویر را از یک مجموعه برچسب‌گذاری شده،

³⁵ Semi-supervised learning

مطابق با آن سوال، استخراج می‌کند. توزیعی روی سراسر ویژگی‌های این تصاویر استفاده می‌شود تا یک امتیاز بیزین را برای هر تصویر در مجموعه بدون برچسب، محاسبه کند. تصاویر بدون برچسب با استفاده از این امتیاز، رتبه‌بندی شده و تصاویر با امتیاز بالاتر، بازیابی می‌شوند. در این پژوهش بردار ویژگی تصاویر با استفاده از دو نوع از ویژگی‌های بافت و یک نوع ویژگی رنگی نمایش داده شده است. ۴۸ ویژگی از نوع بافت گابور، ۲۷ ویژگی از نوع بافت تامورا و ۱۶۵ ویژگی از هسیتوگرام رنگ، مجموعه بردار ویژگی این پژوهش هستند.

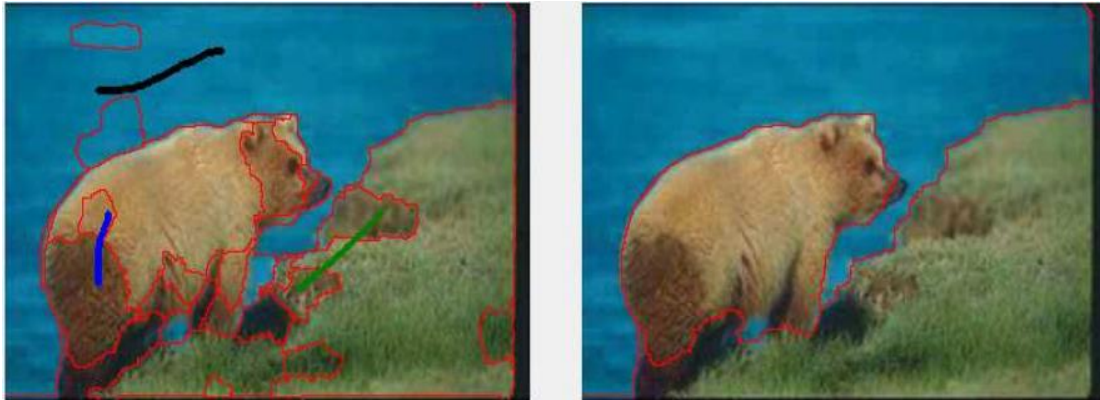
در [۹] به شکلی متفاوت بازیابی تصویر مبتنی بر محتوا ارائه شده است. این پژوهش با استفاده از سه مرحله عمل بازیابی را انجام می‌دهد. در مرحله اول تصویر را بخش‌بندی کرده و در ادامه به منظور بدست آوردن نواحی معنادار از تصویر، آن‌ها را ادغام می‌کند. در مرحله سوم از یک شبکه عصبی سه لایه جهت آموزش و آزمون نواحی معنایی بدست آمده، استفاده می‌کند.

در مرحله اول، ابتدا با استفاده از الگوریتم Mean Shift تصویر بخش‌بندی می‌شود. این الگوریتم به این شکل عمل می‌کند که پنجره‌ای روی تصویر حرکت داده و از پیکسل‌های قرار گرفته داخل پنجره میانگین گرفته و مقدار آن را در پیکسل مرکز پنجره قرار می‌دهد. بدین ترتیب بعد از چندین مرحله تکرار الگوریتم، تصویر به صورت شکل ۲-۹ بخش‌بندی می‌شود.



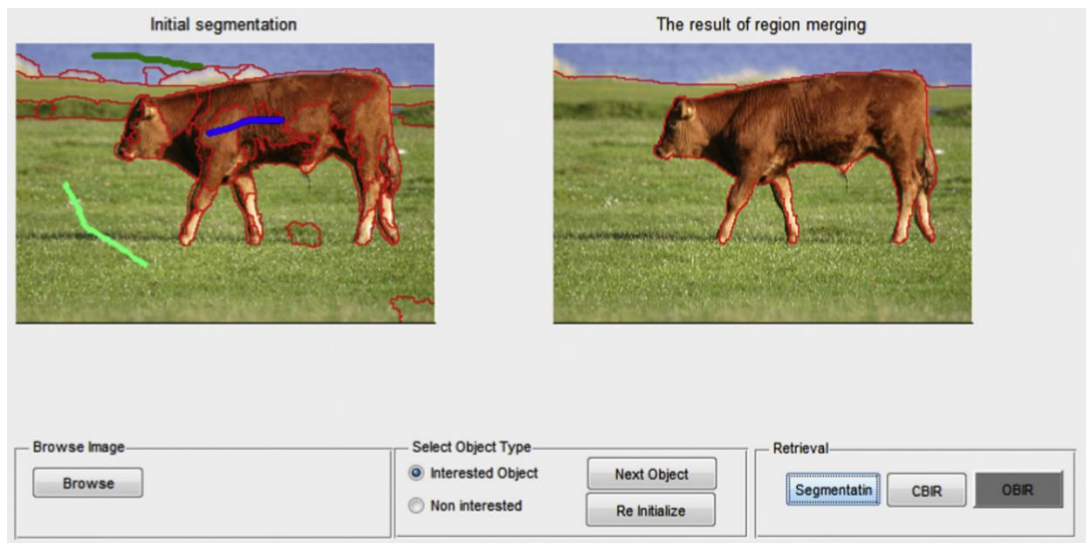
شکل ۲-۹- نتیجه بخش‌بندی ابتدایی با استفاده از الگوریتم Mean Shift [۹].

همانطور که مشاهده می‌شود، این روش بخش‌بندی، مشکل "بخش‌بندی بیش از حد" دارد. برای حل این مشکل، پژوهشگر از واسط کاربری استفاده کرده است. در این واسط کاربری، از کاربر درخواست می‌شود که اشیاء مورد نظر و پس‌زمینه‌ی تصویر را علامت‌گذاری کند. پس از آن، دوباره الگوریتم بخش‌بندی را روی تصویر اعمال کرده، نتیجه به صورت شکل (۲-۱۰) می‌شود.



شکل ۲-۱۰- تصویر سمت چپ، ابتدا کاربر، اشیاء را مشخص و سپس الگوریتم قطعه‌بندی اجرا شده، تصویر سمت راست ادغام نواحی معنایی [۹].

پس از اعمال الگوریتم روی تصویر، گاهی اوقات برخی نواحی از تصویر، بدرستی بخش‌بندی نمی‌شوند. برای ادغام این نواحی از ویژگی‌های سطح پایین استفاده می‌شود. از بین ویژگی‌های سطح پایین، چون هیستوگرام رنگ ارزان‌تر و آسان‌تر است، برای هر ناحیه گرفته می‌شود. سپس نواحی که هیستوگرام رنگ شبیه هم دارند، به عنوان یک ناحیه معنایی باهم ادغام می‌شوند. در نهایت، تصویر بصورت شکل ۲-۱۱ بخش‌بندی می‌شود و از نواحی معنایی برای ساخت بردار ویژگی استفاده خواهد شد.



شکل ۲-۱۱- استفاده از واسط کاربری و سپس اعمال الگوریتم بخش‌بندی روی تصویر [۹].

پس از استخراج نواحی معنایی از تصاویر، از این نواحی جهت آموزش شبکه عصبی استفاده می‌شود. پس از آموزش شبکه عصبی، تصاویر آزمون جهت ارزیابی سیستم بازیابی، مورد جستجو قرار می‌گیرند.

۲-۴- سیستم‌های CBIR مبتنی بر ترکیب ویژگی‌های سطح پایین

این سیستم‌ها با مزیت مهم کم‌هزینه و پردازش کم، شکاف معنایی بین درک سیستم و درک انسان را کاهش می‌دهند. در این گونه سیستم‌ها از هیچ‌گونه قطعه‌بندی، شناسایی اشیاء و هرگونه روش یادگیری و پردازش‌های سنگین، استفاده نمی‌شود. این سیستم‌ها با وجود استفاده از ویژگی‌های سطح پایین، بخوبی فضای دو بعدی تصویر را توصیف و اطلاعات معنایی را از تصویر استخراج می‌کنند. در ادامه این نوع سیستم‌ها تشریح و کارهای انجام شده اخیر بصورت مختصر توضیح داده خواهند شد. همانطور که ذکر شد، از دیگر روش‌ها جهت کاهش شکاف معنایی، ترکیب ویژگی‌های سطح پایین است. این کار باعث می‌شود به محتوای تصویر و مفاهیم موجود در آن، دست پیدا کرد. از کارهای انجام شده در این دسته می‌توان به [۲۰-۱۱] [۸, ۱۰, ۱۵] اشاره کرد که در ادامه چند مورد آن مختصر توصیف می‌شوند.

در [۲۰] روشی برای استخراج ویژگی‌های سطح پایین و نمایه‌سازی تصویر بر مبنای جداسازی تصویر به مؤلفه‌های بافت و ساختار ارائه شده است. ایده اصلی این تحقیق جداسازی تصاویر به مؤلفه‌های بافت و ساختار برای کاهش اثر مخرب وجود بافت و ساختار به صورت همزمان در تصویر در مرحله استخراج ویژگی است. در شکل ۲-۱۲ جداسازی این دو مؤلفه نشان داده شده است.



شکل ۲-۱۲- جداسازی تصویر، (ج) تصویر اصلی، (ب) تصویر ساختار و (الف) تصویر بافت [۲۰].

همچنین نشان داده شده است که با ترکیب بردار ویژگی استخراج شده از مؤلفه‌های بافت و ساختار، دقت در سیستم بازیابی تصویر به مقدار قابل توجهی افزایش می‌یابد.

جداسازی تصاویر به بافت و ساختار و استخراج ویژگی از هر بخش برای ویژگی‌های هیستوگرام لبه، هیستوگرام رنگ و گابور مورد بررسی قرار گرفته است. پس از اینکه مؤلفه‌های بافت و ساختار تصویر جدا شدند، استخراج ویژگی‌های رنگ و لبه و بافت موثرتری از تصویر استخراج خواهند شد. استخراج ویژگی، اساسی‌ترین قدم در ایجاد یک سامانه بازیابی تصویر بر اساس محتوا است و نقش بسیار تعیین‌کننده‌ای در دقت سیستم بازیابی دارد. روش‌های مختلفی برای استخراج بردارهای ویژگی از تصاویر موجود است. از جمله فیلترهای گابور و هیستوگرام راستای لبه برای استخراج اطلاعات فرکانسی و هیستوگرام رنگ برای استخراج اطلاعات مربوط به رنگ در تصاویر رنگی [۲۰].

برای رفع مشکل شکاف معنایی به نوعی باید سیستم بازیابی تصویر را به سیستم بینایی انسان، نزدیک کرد. یکی از روش‌ها در حل این موضوع، روشی است که در پژوهش [۱۱] ذکر شده

است. در این پژوهش از روشی برای شبیه‌سازی مکانیزم‌های دقت بصری برای سیستم بازیابی تصویر مبتنی بر محتوا استفاده شده است. در واقع یک مدل محاسباتی دقت بصری^{۳۶} جدید، یعنی مدل ساختار برجستگی برای بازیابی تصویر مبتنی بر محتوا ارائه شده است. ابتدا، یک نشانه بصری جدید بنام حجم رنگ، با اطلاعات لبه باهم معرفی می‌شوند تا بجای استفاده از ویژگی‌های بصری اولیه (به عنوان مثال رنگ، شدت و جهت‌گیری)، مناطق برجسته تصویر را تشخیص دهند. سپس یک روش نمایش تصویر جدید، بنام هیستوگرام ساختار برجسته پیشنهاد می‌شود تا نمایش تصویر با چهارچوب CBIR را شبیه‌سازی کند.

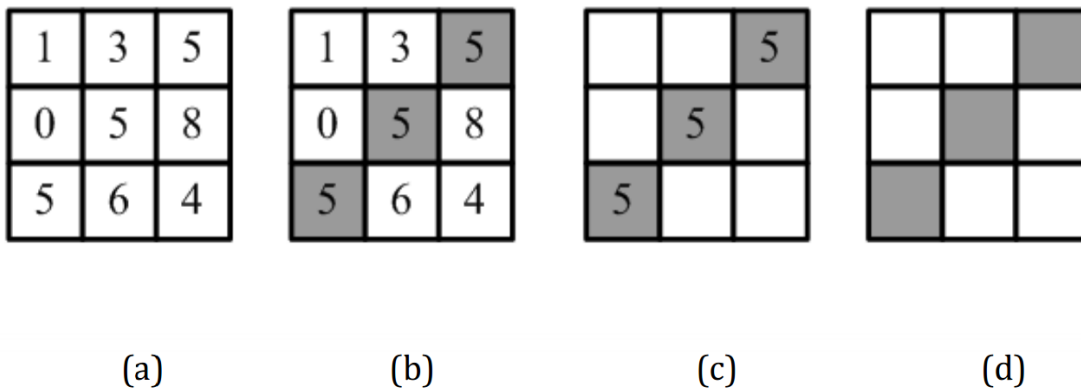
هیستوگرام ساختار برجسته^{۳۷}، ساختار برجسته و فیلترهای گابور جهت‌دار را با یک نوع بسیار خاص در نظر می‌گیرد که در آن اطلاعات رنگ و جهت لبه به هیستوگرام‌ها نگاشت می‌شوند. در الگوریتم این پژوهش، فیلترهای گابور جهت‌دار می‌توانند به ساختارهای میله‌ای اضافه شوند و مشخصه لگاریتم انرژی گابور به عنوان ارزش‌های نمودار هیستوگرام استفاده می‌شود. بنابراین، این روش دارای قدرت تمییز خوبی از ویژگی‌های رنگ، بافت و لبه است [۱۱].

این پژوهش از سه ویژگی سطح پایین در کنار ساختار برجسته استفاده می‌کند. این سه ویژگی، هیستوگرام رنگ در فضای رنگ HSV، شدت رنگ در فضای رنگ HSV و جهت‌گیری لبه هستند. در هیستوگرام رنگ، سه کانال H، S و V به ترتیب به ۶، ۳ و ۳ مقدار چندی‌سازی می‌شود. بنابراین این بردار ۵۴ ویژگی دارد. شدت رنگ در فضای رنگی HSV، مقدار کانال V در نظر گرفته شده است. مقادیر این کانال به ۱۶ مقدار چندی‌سازی می‌شود. جهت لبه با استفاده از عملگر سوبل محاسبه شده است. مقادیر این ویژگی به ۶۰ مقدار چندی‌سازی شده است. در نهایت، بردار ویژگی حاصل در این مقاله، دارای ۱۳۰ بعد است.

³⁶ computational visual attention model

³⁷ saliency structures

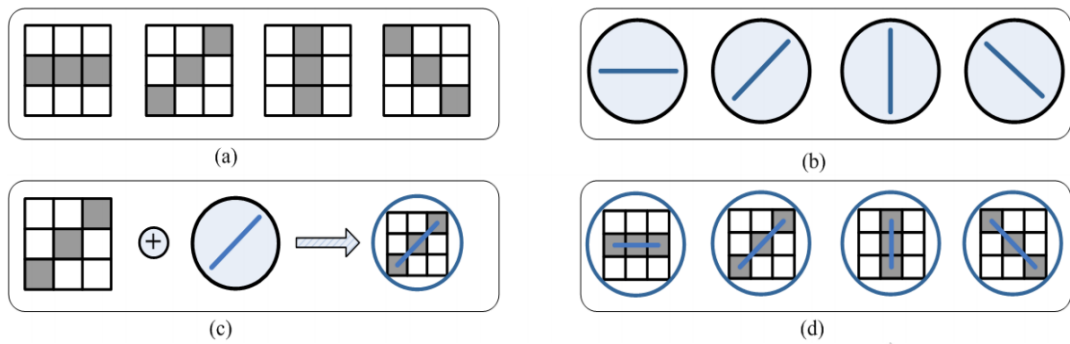
برای توصیف هیستوگرام ساختار برجسته ابتدا ساختارهای نوار-شکل^{۳۸} توصیف می‌شوند. به منظور شبیه‌سازی درست مکانیزم بینایی چشم انسان، ساختارهای نوار-شکل برای تجزیه محتوای تصویر تعریف می‌شوند. این ساختارها کاملاً متفاوت با روش‌های تشخیص لبه و جهت‌گیری لبه است. ساختارهای نوار-شکل، سه پیکسل همجوار متوالی در یک پنجره 3×3 با مقادیر پیکسلی یکسان هستند. نمونه‌ای از این ساختار با زاویه 45° درجه در شکل ۲-۱۳ مشاهده می‌شود.



شکل ۲-۱۳- مثالی از تشخیص ساختار نوار-شکل در پنجره 3×3 . (a) پنجره‌ای از تصویر. (b) و (c) فرآیند تشخیص ساختار نوار-شکل. (d) ساختار نوار-شکل تشخیص داده شده.

ساختارهای نوار-شکل با چهار زاویه 0° ، 45° ، 90° و 135° درجه تعریف می‌شوند. این ۴ ساختار با ۴ حالت فیلتر گابور، جهت تشخیص ساختار برجستگی، تطبیق داده می‌شود. شکل ۲-۱۴ این تطبیق را نشان می‌دهد.

³⁸ bar-shaped structure



شکل ۲-۱۴- تشخیص ساختار برجستگی. (a) چهار الگوی ساختار نوار-شکل، زاویه بین ساختار نوار-شکل و جهت افقی به ترتیب ۰، ۴۵، ۹۰ و ۱۳۵ درجه است. (b) چهار فیلتر گابور جهت دار که در آن زاویه مورد نظر به ترتیب ۰، ۴۵، ۹۰ و ۱۳۵ درجه است. (c) فرآیند تشخیص ساختار برجستگی. علامت \oplus به معنای تطبیق ساختار نوار-شکل و فیلتر گابور است. (d) چهار الگوی ساختار برجستگی.

پس از معرفی الگوهای ساختار برجستگی، ویژگی‌های مورد استفاده در فیلترهای گابور باید مشخص شوند. فیلتر گابور مذکور با چهار زاویه مختلف روی ویژگی‌های چندی‌سازی شده‌ی هیستوگرام رنگ، شدت رنگ و جهت‌گیری لبه اعمال می‌شود. توابع فیلتر گابور برای این ویژگی‌ها به ترتیب با $ES(x, y, C, \theta)$ ، $ES(x, y, I, \theta)$ و $ES(x, y, O, \theta)$ نشان داده می‌شوند. (x, y) یک پیکسل از تصویر، θ زاویه ساختار برجستگی، ES تابع فیلتر گابور، C مقادیر رنگ، I مقادیر شدت رنگ و O مقادیر جهت‌گیری لبه هستند. مطالعات پزشکی نشان می‌دهد روشنایی دریافتی سیستم بینایی انسان، یک تابع لگاریتمی از شدت نور روشنایی به چشم انسان است. بنابراین از ویژگی لگاریتمی انرژی گابور جهت نمایش ویژگی تصاویر استفاده می‌شود. در نهایت هیستوگرام ساختار برجستگی با SSH ^{۳۹} نشان داده می‌شود. این هیستوگرام بصورت رابطه ۲-۱۲ است.

$$SSH = [[H_c], [H_o], [H_l]] \quad (2-12)$$

که در آن:

$$H_c(M_c(x, y)) = \begin{cases} \log \left\{ \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} ES(x, y, C, \theta) \right\}, \theta = \{0, 45, 90, 135\} \\ 0 \leq M_c(x, y) \leq N_c - 1 \end{cases} \quad (2-13)$$

³⁹ saliency structure histogram

$$H_0(M_0(x, y)) = \begin{cases} \log \left\{ \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} ES(x, y, O, \theta) \right\}, \theta = \{0, 45, 90, 135\} \\ 0 \leq M_0(x, y) \leq N_0 - 1 \end{cases} \quad (14-2)$$

$$H_1(M_1(x, y)) = \begin{cases} \log \left\{ \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} ES(x, y, I, \theta) \right\}, \theta = \{0, 45, 90, 135\} \\ 0 \leq M_1(x, y) \leq N_1 - 1 \end{cases} \quad (15-2)$$

در روابط ۲-۱۳ تا ۲-۱۵، $M_c(x, y)$ مقادیر رنگ تصویر بصورت چندی‌سازی شده، $M_0(x, y)$ مقادیر جهت لبه تصویر بصورت چندی‌سازی شده، $M_1(x, y)$ مقادیر شدت رنگ تصویر بصورت چندی‌سازی شده، w عرض تصویر، h طول تصویر، N_c تعداد رنگ‌های چندی‌سازی شده (۵۴)، N_0 تعداد جهت لبه چندی‌سازی شده (۶۰) و N_1 تعداد شدت لبه چندی‌سازی شده (۱۶) هستند.

رابطه ۲-۱۳ بدین معناست که برای کل مقادیر رنگ چندی‌سازی شده، لگاریتم انرژی گابور حاصل از آن مقادیر در چهار جهت ساختار برجستگی، در کل تصویر محاسبه می‌شود و در مقدار رنگ چندی‌سازی شده $M_0(x, y)$ قرار می‌گیرد. روابط ۲-۱۴ و ۲-۱۵ نیز به همین شکل برای به ترتیب جهت‌گیری لبه و شدت رنگ تصویر هستند. در نهایت هیستوگرام ساختار برجستگی، دارای ۱۳۰ ویژگی است. این ۱۳۰ ویژگی به همراه ۱۳۰ ویژگی از ویژگی‌های سطح پایین (که قبلاً بیان شد)، بردار ۲۶۰ تایی این مقاله را تشکیل می‌دهند. این روش، یکی از روش‌های مورد مقایسه با روش‌های پیشنهادی این پایان‌نامه است.

در [۱۵] یک سیستم بازیابی تصویر مبتنی بر محتوای سریع و کارآمد بر اساس ترکیب رنگ و بافت ارائه شده است. ویژگی‌های بافتی شامل اختلاف احتمال معکوس بلوک (BDIP^{۴۰}) و اختلاف

⁴⁰ block difference of inverse probabilities

ضرایب همبستگی محلی بلوک (BVLC)⁴¹ و ویژگی رنگ شامل هیستوگرام رنگ در فضای رنگ HSV است. ابتدا هر یک از ویژگی‌های بافتی و سپس ویژگی رنگی تشریح می‌شود.

• BDIP

این ویژگی بافتی اختلاف روشنایی محلی را در تصویر اندازه‌گیری می‌کند. در واقع پنجره‌ای با اندازه ۲*۲ (بدون اشتراک) بصورت رابطه ۲-۱۶ روی هر یک از کانال‌های تصویر اعمال می‌شود.

$$BDIP(x, y) = \frac{\frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 [I_{max}(x, y) - I(2x + i, 2y + j)]}{I_{max}(x, y)}, x = 0, 1, \dots, \frac{M}{2} - 1, y = 0, 1, \dots, \frac{N}{2} - 1 \quad (16-2)$$

که در آن $I_{max}(x, y) = \text{Max}(I(2x + i, 2y + j))$ برای $i, j = 0, 1$ یا $I_{max}(x, y)$ برابر بیشترین مقدار چهار پیکسل موجود در پنجره ۲*۲ است. بنابراین محدوده مقدار این ویژگی [۰, ۱] است. در واقع این ویژگی حداکثر روشنایی همسایگی ۲*۲ را نشان می‌دهد. با اعمال این پنجره‌گذاری روی کانال‌های تصویر، لبه‌ها و مرزهای ناحیه‌های تصویر بدست می‌آید [۱۵].

• BVLC

این ویژگی بافتی میزان همواری بافت محلی را در تصویر اندازه‌گیری می‌کند. در واقع پنجره‌ای با اندازه ۲*۲ (بدون اشتراک) بصورت رابطه ۲-۱۷ روی هر یک از کانال‌های تصویر اعمال می‌شود.

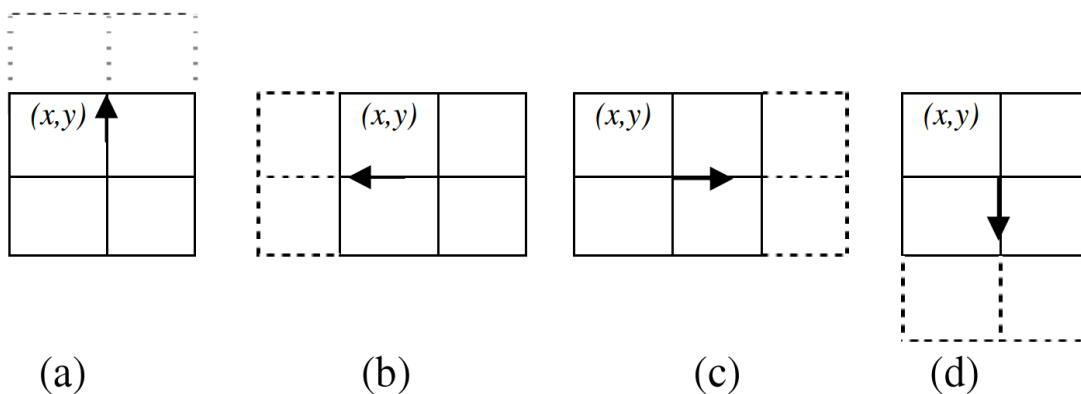
$$\rho(p, q) = \frac{1}{\sigma(x, y)\sigma(p, q)} \times \left[\frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 [I(2x + i, 2y + j)I(p, q) - \mu(x, y)\mu(p, q)] \right] \quad (17-2)$$

⁴¹ block variation of local correlation coefficients

که در آن $(p, q) \in \{(-1,0), (0,-1), (1,0), (0,1)\}$ و نشان‌دهنده‌ی شیفت‌های افقی و عمودی با چهار جهت هستند که در شکل ۲-۱۵ مشاهده می‌شود. $\mu(x, y)$ و $\sigma(x, y)$ به ترتیب میانگین و انحراف معیار استاندارد ناحیه محلی $2*2$ هستند. $I(p, q)$ ، $\mu(p, q)$ و $\sigma(p, q)$ به ترتیب شدت روشنایی، میانگین و انحراف معیار استاندارد روشنایی هستند، وقتی که پنجره محلی با (p, q) در چهار جهت جابجا می‌شود. در نهایت مقدار ویژگی BVLC با رابطه ۲-۱۸ بدست می‌آید.

$$BVLC(x, y) = \max_{(p,q)} [\rho(p, q)] - \min_{(p,q)} [\rho(p, q)] \quad (2-18)$$

بنابراین محدوده مقدار این ویژگی $[0, 1]$ است. در واقع BVLC نشان‌دهنده‌ی اختلاف بین بیشترین و کمترین مقدار ضرایب همبستگی محلی در چهار جهت است. بافت در کل تصویر از جمله مناطق با تفاوت شدت روشنایی کم، توسط این ویژگی، استخراج می‌شود.



شکل ۲-۱۵- پیکربندی پیکسل در پنجره ناحیه $2*2$ و پنجره‌های جابجاشده مربوط به آن‌ها در چهار جهت، که جهت محاسبه (a) $(-1,0)$ ، (b) $(0,-1)$ ، (c) $(0,1)$ و (d) $(1,0)$ مورد نیاز است [۱۵].

این دو بافت محاسبه شده، بصورت دو بعدی هستند. مقادیر هر کدام از این ویژگی‌ها ابتدا مرتب می‌شوند، سپس به یک بعد تبدیل می‌شوند. میانگین بردار یک بعدی گرفته شده و بردار به دو کلاس تبدیل می‌شود. به شکل مشابه، هر کلاس به دو کلاس جدید تبدیل می‌شود تا اینکه کل بردار به ۸ کلاس تبدیل شود. از این ۸ کلاس مختلف، ویژگی‌های میانگین و انحراف معیار استاندارد، به عنوان ویژگی، گرفته می‌شود. این کار برای هر سه کانال تصویر انجام می‌شود. بنابراین هر بافت دارای دارای ۴۸ ویژگی و جمعا بردار ویژگی بافت برای سه کانال تصویر، ۹۶ ویژگی خواهند داشت.

با توجه به اینکه، بافت تصویر بیشتر در مولفه روشنایی تصویر وجود دارد، فقط یک مولفه روشنایی برای استخراج ویژگی بافتی در نظر گرفته شده است. تعداد ویژگی‌های بافتی برای یک کانال روشنایی، ۳۲ ویژگی است.

- هیستوگرام رنگ

در هیستوگرام رنگ، از مقادیر چندی‌سازی شده هر یک از سه کانال فضای رنگ *HSV* استفاده می‌شود. این سه کانال در این مقاله، هریک به ۴ بین چندی‌سازی می‌شود. پس تعداد ویژگی‌های رنگی این مقاله، ۶۴ ویژگی است.

در نهایت ۳۲ ویژگی بافتی در کنار ۶۴ ویژگی از هیستوگرام رنگ، جمعا بردار ۹۶ بعدی روش مقاله [۱۵] را می‌سازند. این روش یکی از جدیدترین، سریع‌ترین و با دقت‌ترین روش‌های بازیابی تصویر مبتنی بر محتواست، که به عنوان یکی دیگر از روش‌های مورد مقایسه با روش‌های پیشنهادی این پایان‌نامه در نظر گرفته می‌شود.

بر اساس مطالعات انجام‌شده، سیستم بینایی انسان در ارزیابی محتوای تصویر، ویژگی‌های لبه و رنگ را بیشتر از سایر ویژگی‌ها مورد توجه قرار می‌دهد. لذا این دو مجموعه ویژگی در سیستم‌های بازیابی مبتنی بر محتوا نقش بسزایی دارند. روشی مطرح در این حوزه، غالبا از لبه‌یاب سوبل برای استخراج ویژگی‌های لبه و از فضای رنگ $L^*a^*b^*$ برای استخراج اطلاعات رنگ در تصویر استفاده می‌کنند. این مجموعه ویژگی، "هیستوگرام اختلاف رنگ (CDH^42)" نام‌گذاری شده است [۱۹].

هیستوگرام اختلاف رنگ با هیستوگرام‌های معمولی تفاوت چشم‌گیری دارد. در هیستوگرام‌های معمولی، فراوانی سطوح خاکستری یا رنگی مختلف اندازه‌گیری می‌شود. اما در هیستوگرام اختلاف رنگ، اختلاف رنگ پیکسل‌های همسایه در راستای رنگ و جهت لبه برابر، محاسبه می‌شود.

⁴² color difference histogram

این کار منجر به استخراج مفاهیم و اطلاعات معنایی از تصویر می‌شود. فضای رنگ مورد استفاده در این پژوهش، $L^*a^*b^*$ است که با توجه به مطالعات پزشکی روی سیستم بینایی انسان، پزشکان برآنند که چشم انسان روی این سه کانال بیشترین حساسیت را دارد. علت دیگر انتخاب این فضای رنگ، خاصیت یکنواختی آن است. ابتدا جهت‌های لبه و مقادیر رنگی در فضای $L^*a^*b^*$ بصورت مجزا به تعداد کلاس‌های مشخص چندی‌سازی می‌شوند. سپس اختلاف رنگ، تنها بین پیسکل‌های هم کلاس در دو حالت کلاس رنگی برابر یا جهت لبه برابر، محاسبه می‌شود. این ویژگی‌های هیستوگرام اختلاف رنگ، منجر به استخراج مفاهیم معنایی از تصویر می‌شود [۱۹]. این روش در پایان‌نامه بهبود یافته است. همچنین یکی از روش‌های دیگر مورد مقایسه با روش‌های پیشنهادی است. لذا جهت عدم تکرار، این روش در فصل سوم بصورت کامل تشریح می‌شود.

۲-۵- معیار شباهت^{۴۳}

شباهت دو تصویر، با فاصله بین ویژگی‌های استخراج شده از آن دو تصویر یا فاصله بین بردارهای تشکیل شده از ویژگی‌های آنها تعریف می‌شود. به عبارت دیگر، تابع شباهت، دو بردار ویژگی را به یک مقدار حقیقی مثبت تصویر می‌کند. استفاده از این مقدار حقیقی مثبت، مقایسه‌ی بین ویژگی‌های دیداری بین دو تصویر را امکان پذیر می‌کند [۳۵].

انتخاب معیاری که عدم شباهت بین دو تصویر را بر پایه‌ی ادراک انسان ارائه کند؛ به میزان قابل توجهی شکاف معنایی را کاهش و رسیدن به هدف بازیابی بر اساس محتوا را افزایش خواهد داد. برای تعیین میزان عدم شباهت دو تصویر، فاصله میان بردارهای ویژگی آن‌ها از یکدیگر محاسبه می‌شود. استفاده از معیارهای متفاوت در تعیین میزان عدم شباهت دو تصویر، جواب‌های متفاوتی خواهد داشت. یعنی این امکان وجود دارد که با این که ویژگی‌های سطح پایین دو تصویر به یکدیگر نزدیک هستند، ویژگی‌های معنایی آن‌ها کاملاً متفاوت باشند.

⁴³ Similarity measure

یک معیار شباهت به عنوان یک متریک فاصله تعریف می شود. در این بخش معیارهای شباهت مختلف با ذکر جزئیات توضیح داده می شوند [۳۵].

۲-۵-۱- فاصله مینکوفسکی

فاصله مینکوفسکی بر اساس نرم L_p به صورت زیر تعریف می شود:

$$d_p(Q, T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^p \right)^{\frac{1}{p}} \quad (19-2)$$

که در آن $T = \{T_1, \dots, T_N\}$ و $Q = \{Q_1, \dots, Q_N\}$ که در واقع بردارهای ویژگی تصویر مورد جستجو می باشند. زمانی که $p=1$ باشد، $d_1(Q, T)$ فاصله بلوک یا فاصله مانهاتان (L_1) هستند.

$$d_1(Q, T) = \sum_{i=1}^{N-1} |Q_i - T_i| \quad (20-2)$$

۲-۵-۲- فاصله اقلیدسی

زمانی که در رابطه ۱۹-۲، $p=2$ باشد، $d_2(Q, T)$ فاصله اقلیدسی L_2 است.

$$d_2(Q, T) = \left(\sum_{i=1}^{N-1} (Q_i - T_i)^2 \right)^{\frac{1}{2}} \quad (21-2)$$

۲-۵-۳- فاصله کانبرا^{۴۴}

فاصله کانبرا با رابطه ۲۲-۲ تعریف می شود. این فاصله فقط به بردارهایی اعمال می شود که عناصر غیر منفی دارند.

$$D(F, G) = \sum_{i=1}^n \frac{|f_i - g_i|}{|f_i + g_i|} \quad (22-2)$$

که در آن، f و g بردارهای ویژگی هستند.

⁴⁴ Canberra Distance

۲-۵-۴- فاصله کسینوسی

فاصله کسینوسی تفاوت را با صرف نظر از طول بردار، در جهت محاسبه می‌کند. تفاوت داده شده در واقع زاویه بین بردارها می‌باشد. با استفاده از قانون ضرب نقطه‌ای داریم:

$$d_{\infty}(Q, T) = 1 - \cos \theta = 1 - \frac{Q \cdot T}{|Q| \cdot |T|} \quad (23-2)$$

که در آن T و Q بردارهای ویژگی هستند.

۲-۵-۵- فاصله χ^2

χ^2 به صورت زیر تعریف می‌شود.

$$d_{\chi^2}(Q, T) = \sum_{i=0}^{N-1} \frac{(Q_i - m_i)^2}{m_i} \quad (24-2)$$

که در آن $m_i = \frac{Q_i + T_i}{2}$ و T و Q بردارهای ویژگی و N طول بردارها است.

۲-۵-۶- کانبرای بهبود یافته

در [۱۹] نیز با آزمایش روی معیارهای مختلف، معیار کانبرا را جهت بازیابی موثر بهبود داده است. این معیار به صورت رابطه (۲۵-۲) است:

$$D(T, Q) = \sum_{i=1}^M \frac{|T_i - Q_i|}{|T_i + u_T| + |Q_i + u_Q|} \quad (25-2)$$

که در آن D فاصله شباهت، T و Q بردارهای مورد مقایسه، M طول بردار ویژگی و u_T و u_Q به ترتیب میانگین بردار ویژگی T و Q است.

۲-۶- جمع‌بندی

در این فصل بطور کلی روش‌های بازیابی تصویر مورد استفاده قدیمی تا روش‌های اخیر مورد بررسی قرار گرفت. ابتدا ویژگی‌های سطح پایین در CBIR و روش‌های استخراج آن‌ها توضیح داده شد. سپس به دنبال کاهش شکاف معنایی بین درک سیستم و درک انسان از تصویر، روش‌های کاهش این چالش بیان شد. روش‌های مبتنی بر یادگیری و هزینه‌بردار توضیح داده شد و کارهای جدیدی که در این حوزه انجام شده بصورت مختصر تشریح شدند. از معایب بزرگ این روش‌ها نیاز به تصاویر زیاد جهت آموزش، شناسایی اشیاء، و نهایتاً هزینه‌ها و پردازش‌های زیاد است. پس از آن به دنبال روش‌هایی که در عین کم‌هزینه و پردازش کم داشتن، کارایی خوبی در بازیابی تصویر داشته باشند، روش‌هایی معرفی شد. این روش‌ها بدون نیاز به هزینه‌ها و پردازش زیاد، بخوبی فضای دو بعدی تصویر را توصیف و اطلاعات معنایی را از تصویر استخراج می‌کنند.

برای روش‌های گوناگون، پژوهش‌های جدید چند سال اخیر مورد بررسی قرار گرفت. در این پایان‌نامه نیز جهت کاهش شکاف معنایی، از روش‌های مبتنی بر ترکیب ویژگی‌های سطح پایین استفاده شده است. این کار به گونه‌ای انجام شده که سطوح بالای معنایی را از تصویر استخراج و از آن‌ها جهت بازیابی استفاده می‌کند. این روش در عین سرعت و دقت بالا در بازیابی تصویر، معاف از هرگونه هزینه‌ها و پردازش‌های زیاد است.

فصل سوم:

روش پیمایشی

پس از اینکه در فصل اول بازیابی تصویر مبتنی بر محتوا، هدف و چالش پیش‌روی آن تشریح شد. در فصل دوم کارهای انجام شده اخیر جهت رفع مشکل این مسئله نیز بیان شد. یکی از مهم‌ترین چالش‌ها در حوزه بازیابی تصویر مبتنی بر محتوا، شکاف معنایی بین درک سیستم و درک انسان از تصویر است. راه‌های گوناگونی جهت کاهش این شکاف معنایی وجود دارند که به اختصار در فصل دوم بیان شد. یکی از مهم‌ترین و جدیدترین این روش‌ها، استفاده از ویژگی‌های سطح پایین بصورت ترکیبی است. در این روش، طوری از ویژگی‌های سطح پایین استفاده می‌شود که می‌توان به اطلاعات معنایی تصویر، دست یافت. به این ترتیب، شکاف معنایی ذکر شده، بسیار کاهش یافته و مفاهیم ادراکی تصویر از آن استخراج می‌شوند.

همان‌طور که ذکر شد، ما در این پایان‌نامه به دنبال روشی جهت کاهش شکاف معنایی هستیم که خصوصیات زیر را داشته باشد:

- روش پیشنهادی تا حد ممکن، پرهزینه و زمان‌بر نباشد. در عین حال بتواند بخوبی اطلاعات معنایی را از تصویر استخراج کند.
- روش پیشنهادی باید از ویژگی‌های سطح پایین استفاده کند. این ویژگی‌ها باید تا حد ممکن به سیستم بینایی انسان نزدیک باشند. همه ویژگی‌های سطح پایین این خاصیت را ندارند.
- مطلب مهم دیگر در ویژگی‌های سطح پایین، فضای رنگ مورد استفاده است. همه فضاهای رنگ به سیستم بینایی انسان نزدیک نیستند. از بین این فضاهای رنگ، نزدیک‌ترین به سیستم بینایی انسان انتخاب می‌شود.
- در فضای رنگ مناسب، استخراج ویژگی‌های سطح پایین موثر از تصویر، منجر به استخراج مفاهیم معنایی از تصاویر می‌شوند.

- جهت استخراج بهتر اطلاعات معنایی، تخصیص وزن به هر گروه ویژگی در معیار شباهت، نیز کارا است. ویژگی‌های مستخرج و وزندهی آن‌ها، به عنوان سطوح بالای معنایی، شناخته می‌شود.

در این فصل روش‌های پیشنهادی جهت بازیابی تصویر مبتنی بر محتوا ارائه می‌شوند. دو روش جهت بازیابی تصویر مبتنی بر محتوا پیشنهاد می‌شود. روش دوم شکل بهبود یافته‌ی روش اول است. روش‌های پیشنهادی مبتنی بر روش هیستوگرام اختلاف رنگ [۱۹] هستند. از دستاوردهای اصلی روش‌های پیشنهادی، بهبود روش هیستوگرام اختلاف رنگ و افزودن ویژگی جدید به آن جهت افزایش نرخ بازیابی است. روش هیستوگرام اختلاف رنگ یکنواخت از جدیدترین روش‌ها در حوزه CBIR است. در ادامه، هر کدام از دو روش پیشنهادی بصورت مجزا توضیح داده می‌شوند.

۲-۳- روش پیشنهادی اول

روش پیشنهادی جهت بازیابی تصویر مبتنی بر محتوا، دارای سه مرحله است.

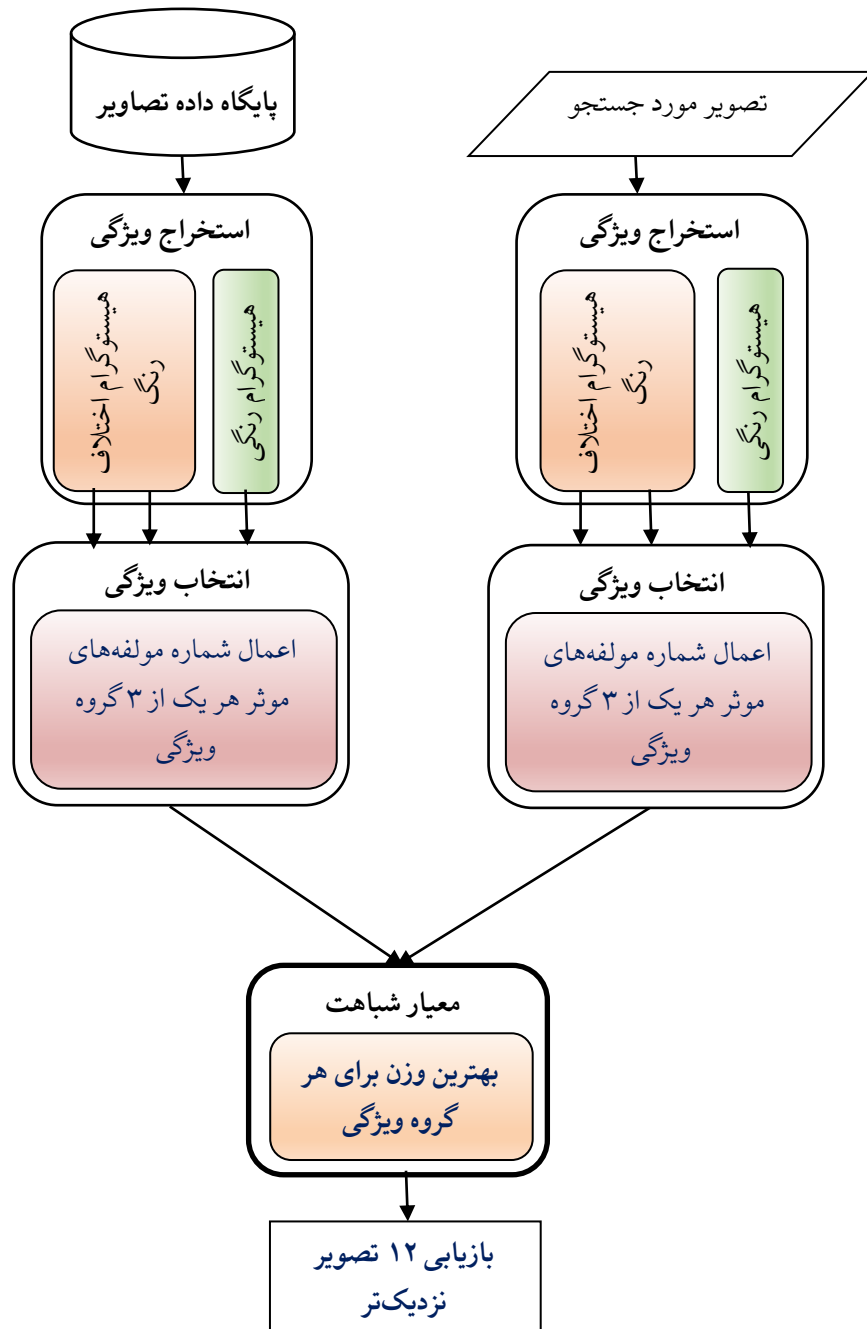
- استخراج ویژگی
- انتخاب ویژگی
- معیار شباهت

در مرحله استخراج ویژگی، ویژگی‌های هیستوگرام اختلاف رنگ و هیستوگرام رنگ در فضای رنگی HSV از تمامی تصاویر پایگاه داده و تصویر مورد جستجو، استخراج می‌شوند.

در مرحله دوم، با اعمال معیارهای آنتروپی و همبستگی بین ویژگی‌ها، ویژگی‌های موثر از بین سایر ویژگی‌ها انتخاب می‌شوند. با اعمال آنتروپی به هر یک از بردارهای اولیه ویژگی، یکسری ویژگی‌های ناکارآمد حذف می‌شوند. سپس معیار همبستگی بین ویژگی‌های باقی‌مانده، محاسبه می‌شود و ویژگی‌های ناکارآمد دیگری نیز حذف می‌شوند.

در نهایت، جهت اندازه‌گیری شباهت تصاویر پایگاه داده با تصویر مورد جستجو، ویژگی‌های بدست آمده با هم ترکیب و فاصله اندازه‌گیری می‌شود. در این مرحله به هر گروه ویژگی، وزنی اختصاص می‌یابد. ابتدا شباهت تصاویر با استفاده از هر نوع ویژگی محاسبه می‌شود، سپس به هر عدد بدست آمده ضریبی بین ۰ و ۱ (با شرط مجموع ضرایب برابر ۱) اختصاص داده می‌شود.

روش پیشنهادی اول جهت بازیابی تصویر مبتنی بر محتوا، به صورت دیاگرام شکل (۳-۱) است.



شکل ۳-۱- دیاگرام روش پیشنهادی اول جهت بازیابی تصویر مبتنی بر محتوا

در این پایان‌نامه، از ویژگی هیستوگرام اختلاف رنگ مطرح شده در [۱۹] استفاده می‌شود. این مجموعه ویژگی شامل ۱۰۸ ویژگی مربوط به رنگ و لبه تصویر است. در روش پیشنهادی اول، ابتدا با محاسبه آنتروپی هر یک از ویژگی‌ها و همبستگی بین بردارهای ویژگی، ویژگی‌های موثر انتخاب می‌شوند. انتخاب ویژگی‌های برتر و کاهش تعداد ویژگی‌ها در این روش، باعث افزایش نرخ بازیابی و بهبود زمان بازیابی می‌شود. اگرچه ویژگی‌های حاصل از "هیستوگرام اختلاف رنگ"، تا حدی بیان‌گر محتوای معنایی تصویر هستند، با افزودن ویژگی‌های جدید به این مجموعه، نرخ بازیابی تصویر را می‌توان افزایش داد. به دنبال افزودن ویژگی‌های برتر رنگی، هیستوگرام رنگی در فضای رنگ HSV پیشنهاد می‌شود. علت انتخاب فضای رنگی HSV خاصیت یکنواختی ادراکی آن است. این خاصیت باعث می‌شود، ویژگی‌های رنگی مستخرج از این روش به درک انسان از تصویر نزدیک‌تر باشد. با ترکیب این ویژگی‌ها با ویژگی‌های موثر از روش هیستوگرام اختلاف رنگ، اطلاعات معنایی از تصویر استخراج می‌شود.

۳-۲-۱- استخراج ویژگی

۳-۲-۱-۱- هیستوگرام اختلاف رنگ

مطالعه‌های فیزیولوژیکی و روان‌شناسی انسان نشان می‌دهد که سیستم بینایی انسان نسبت به رنگ و جهت لبه‌ها در تصویر بسیار حساس است. بنابراین اختلاف رنگ بین رنگ‌ها و جهت‌گیری لبه‌ها در تصویر، اطلاعات بصری زیادی را پوشش می‌دهد و نقش مهمی در تحلیل و فهم محتوای تصویر دارد [۱۹].

برای محاسبه این اختلاف رنگ باید بدنبال فضای رنگی بود که دارای خاصیت یکنواخت ادراکی باشد. منظور از خاصیت یکنواخت ادراکی این است که درک سیستم در فضای رنگی به درک بصری انسان نزدیک و تا حد ممکن، یکنواخت باشد. از بین فضاها رنگ موجود برای تصویر، دو

فضای رنگی $L^*a^*b^*$ و HSV دارای این خاصیت هستند [۷]. با توجه به یکنواختی بالای فضای رنگی $L^*a^*b^*$ ، انتخاب بسیار مناسبی جهت تعیین اختلاف بین رنگ‌ها بین دو نقطه است [۱۹].

در مرحله بعد باید جهت لبه را در این فضا تشخیص داد. جهت لبه نفوذ قوی روی درک انسان نسبت به تصویر دارد و می‌تواند مرزهای جسم و ساختار بافت آن و در نهایت اطلاعات معنایی را نشان دهد. جهت تشخیص موثر لبه‌های ناشی از تغییرات رنگی در فضای رنگ یکنواخت، باید بردارهای واحد L ، a و b را در امتداد محورهای L^* ، a^* و b^* در این فضا، قرار داد. سپس رابطه‌های (۱-۳ و ۲-۳) برای تصویر تمام رنگی $f(x,y)$ محاسبه می‌شوند:

$$u = \frac{\partial L^*}{\partial x} l + \frac{\partial a^*}{\partial x} a + \frac{\partial b^*}{\partial x} b \quad (1-3)$$

$$v = \frac{\partial L^*}{\partial y} l + \frac{\partial a^*}{\partial y} a + \frac{\partial b^*}{\partial y} b \quad (2-3)$$

در روابط زیر g_{xx} ، g_{yy} ، g_{xy} به عنوان حاصل ضرب‌های داخلی این بردارها تعریف می‌شوند:

$$g_{xx} = u^T u = \left| \frac{\partial L^*}{\partial x} \right|^2 + \left| \frac{\partial a^*}{\partial x} \right|^2 + \left| \frac{\partial b^*}{\partial x} \right|^2 \quad (3-3)$$

$$g_{yy} = v^T v = \left| \frac{\partial L^*}{\partial y} \right|^2 + \left| \frac{\partial a^*}{\partial y} \right|^2 + \left| \frac{\partial b^*}{\partial y} \right|^2 \quad (4-3)$$

$$g_{xy} = u^T v = \left| \frac{\partial L^*}{\partial x} \right| \left| \frac{\partial L^*}{\partial y} \right| + \left| \frac{\partial a^*}{\partial x} \right| \left| \frac{\partial a^*}{\partial y} \right| + \left| \frac{\partial b^*}{\partial x} \right| \left| \frac{\partial b^*}{\partial y} \right| \quad (5-3)$$

مشتقات جزئی موردنیاز برای پیاده‌سازی رابطه‌های ۳-۳ تا ۵-۳ می‌توانند با استفاده از

عملگرهای سوبل محاسبه شوند. اگر $I(x,y)$ را در فضای رنگی $L^*a^*b^*$ قرار دهیم، با استفاده از نمادهای بالا، جهت نرخ تغییرات $I(x,y)$ بصورت رابطه ۶-۳ است:

$$\varphi(x,y) = \frac{1}{2} \tan^{-1} \frac{2g_{xy}}{g_{xx} - g_{yy}} \quad (6-3)$$

پس از اینکه جهت‌گیری لبه $\varphi(x, y)$ برای هر پیکسل محاسبه شد، جهت‌ها بطور یکسان به m مقدار چندی‌سازی می‌شوند. پس از آزمایش روی بهترین مقدار برای m ، مقدار ۱۸ برای آن در نظر گرفته شده است [۱۹].

پس از محاسبه لبه‌ها و جهت آن‌ها برای هر پیکسل، نوبت رنگ هر پیکسل است. همانطور که ذکر شد، رنگ دارای مفاهیم سطوح بالای معنایی بیشتری برای چشم و درک انسان است. بنابراین فضای رنگی $L^*a^*b^*$ بخاطر خاصیت یکنواختی ادراکی، انتخاب شد. در این فضا بطور یکنواخت کانال L^* به ۱۰ مقدار و a^* و b^* هر کدام به ۳ مقدار چندی‌سازی می‌شوند. بنابراین فضای رنگی $L^*a^*b^*$ ، به ۹۰ مقدار رنگی، چندی‌سازی می‌شود. مقایسه و انتخاب فضای رنگی $L^*a^*b^*$ و همچنین چندی‌سازی آن به ۹۰ رنگ، طی آزمایش، در پژوهش [۱۹] بدست آمده است.

پس از اینکه، رنگ و جهت لبه برای هر پیکسل، چندی‌سازی و محاسبه شد، روش استفاده از این ویژگی‌ها بیان می‌شود. مقادیر چندی‌سازی شده‌ی رنگ تصویر $C(x, y)$ بصورت $w = 0, 1, \dots, w - 1$ مشخص می‌شوند. موقعیت پیکسل‌های همسایه با (x, y) و (x', y') و مقادیر رنگ آن‌ها با $C(x, y) = w_1$ و $C(x', y') = w_2$ مشخص می‌شوند. زاویه‌ها نیز در (x, y) و (x', y') با $\theta(x, y) = v_1$ و $\theta(x', y') = v_2$ نشان داده می‌شوند. D فاصله بین پیکسل‌های همسایه، W و V به ترتیب تعداد چندی‌سازی برای رنگ و جهت‌گیری لبه هستند. نهایتاً فرمول اختلاف رنگ بصورت رابطه ۷-۳ و ۸-۳ است.

$$H_{color}(C(x, y)) = \begin{cases} \sum \sum \sqrt{(\Delta l)^2 + (\Delta a)^2 + (\Delta b)^2} \\ \text{where } \theta(x, y) = \theta(x', y'); \\ \max(|x - x'|, |y - y'|) = D \end{cases} \quad (7-3)$$

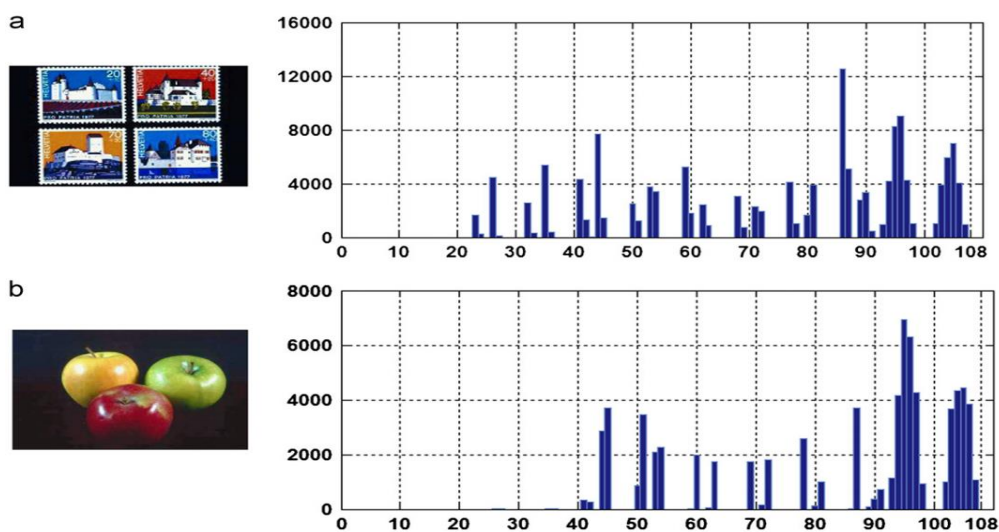
$$H_{ori}(\theta(x, y)) = \begin{cases} \sum \sum \sqrt{(\Delta l)^2 + (\Delta a)^2 + (\Delta b)^2} \\ \text{where } C(x, y) = C(x', y'); \\ \max(|x - x'|, |y - y'|) = D \end{cases} \quad (8-3)$$

که در آن ΔL و Δa و Δb به ترتیب اختلاف رنگ بین دو پیکسل همسایه در کانال‌های L^* ، a^* و b^* است. رابطه (۷-۳) بدین معناست که اگر دو پیکسل همسایه با فاصله D ، جهت لبه (θ) یکسان داشتند، اختلاف رنگ آن‌ها محاسبه شده و در مقدار رنگ پیکسل (x, y) یعنی $C(x, y)$ قرار می‌گیرد. پس H_{color} برداری با ۹۰ ویژگی است.

رابطه (۸-۳) بدین معناست که برای هر مقدار جهت لبه، اگر دو پیکسل همسایه با فاصله D ، رنگ چندی‌سازی شده (C) یکسانی داشتند، اختلاف رنگ آن‌ها محاسبه شده و در مقدار جهت لبه پیکسل (x, y) یعنی $\theta(x, y)$ قرار می‌گیرد. پس H_{ori} برداری با ۱۸ ویژگی است. در نهایت، بردارهای $H_{color}(C(x, y))$ و $H_{ori}(\theta(x, y))$ باهم ترکیب شده و بردار ویژگی نهایی H_{CDH} را با $(90+18=108)$ ویژگی می‌سازند (رابطه ۹-۳)

$$H_{CDH} = \begin{bmatrix} H_{color}(0), H_{color}(1), \dots, H_{color}(W-1), \\ H_{ori}(0), H_{ori}(1), \dots, H_{ori}(V-1) \end{bmatrix} \quad (9-3)$$

با آزمایش در پژوهش [۱۹] روی مقادیر متغیرها جهت بازیابی تصویر، مقدار $D=1$ و $W=90$ و $V=18$ انتخاب شد. نمونه‌ای از بردار ویژگی این روش برای تصاویر، در شکل ۳-۲ آمده است.



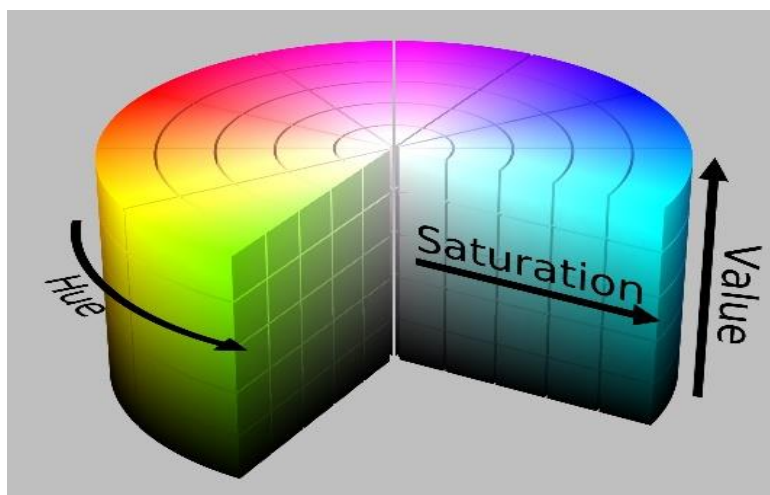
شکل ۳-۲- تصویر a تصویر تمبر و تصویر b تصویر میوه. نمودارهای هیستوگرام اختلاف رنگ دو تصویر مشاهده می‌شوند. ویژگی‌های ۹۰-۱ و ویژگی‌های مربوط به رنگ و ویژگی‌های ۱۰۸-۹۱ ویژگی‌های مربوط به جهت لبه هستند [۱۹].

در این پایان‌نامه روش هیستوگرام اختلاف رنگ یکنواخت پیاده‌سازی با استفاده از آنتروپی و همبستگی بین ویژگی‌ها، ویژگی‌های برتر آن انتخاب می‌شوند. این انتخاب ویژگی‌های برتر علاوه بر کاهش بردار ویژگی روش مذکور، باعث بهبود نرخ بازیابی تصویر می‌شود.

۲-۱-۲-۳- هیستوگرام رنگی در فضای رنگ HSV

همانطور که ذکر شد، رنگ یکی از موثرترین عوامل در تصویر برای درک انسان از تصویر است. فضای رنگی HSV نیز مانند فضای رنگی $L^*a^*b^*$ دارای خاصیت یکنواختی ادراکی است. پس این دو فضا برای ویژگی رنگ، مناسب‌تر از سایر فضاهای رنگ هستند. پس از آزمایش روی ویژگی‌های این دو فضای رنگ، فضای رنگ HSV انتخاب می‌شود. این آزمایش در فصل چهارم تشریح می‌شود.

ابعاد تشکیل‌دهنده فضای HSV، 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. هر برش دایره‌ای از ۰ تا ۳۶۰ درجه علامت‌گذاری می‌شود که هر درجه معرف یک طول موج مشخص رنگ است. شکل (۳-۳) این فضا را نشان می‌دهد.



شکل ۳-۳- فضای استوانه‌ای فضای رنگی HSV.

برای تشکیل هیستوگرام رنگ HSV، تصویر از فضای رنگ RGB به این فضا منتقل می‌شود. سپس این فضا به طور خطی چندی‌سازی می‌شود. در این پایان‌نامه نیز، سه کانال H، S و V به ترتیب

به ۱۸، ۳ و ۳ مقدار چندی‌سازی می‌شود. در نهایت با ۱۶۲ رنگ مختلف، یک بردار $162(R)$ ویژگی برای هر تصویر داریم (رابطه ۳-۱۰).

$$H_{HSV} = [H_{hsv}(0), H_{hsv}(1), \dots, H_{hsv}(R-1)] \quad (3-10)$$

پس از اینکه این بردار ۱۶۲ ویژگی استخراج شد، با استفاده از آنتروپی و همبستگی ویژگی‌های برتر این بردار انتخاب می‌شوند. با کنارهم قرار دادن ویژگی‌های جدید هیستوگرام اختلاف رنگ و هیستوگرام رنگ در فضای رنگ HSV، بردار ویژگی جدید ساخته می‌شود. این بردار ویژگی در عین کوچک بودن، مفاهیم معنایی تصویر را بخوبی استخراج می‌کند.

۳-۲-۲- انتخاب ویژگی

همانطور که ذکر شد، جهت انتخاب ویژگی‌های کارآمد، از دو معیار آنتروپی و همبستگی استفاده می‌شود. این معیارها روی ۱۰۰۰ تصویر تصادفی از پایگاه داده، اعمال می‌شوند. این مجموعه با عنوان داده‌های اعتبارسنجی استفاده می‌شوند. در ادامه دو معیار آنتروپی و همبستگی بصورت جداگانه و مختصر توصیف شده است و روی ویژگی‌های پیشنهادی اعمال می‌شوند.

۳-۲-۲-۱- آنتروپی شانون^{۴۵} و حذف ویژگی‌های ناکارآمد

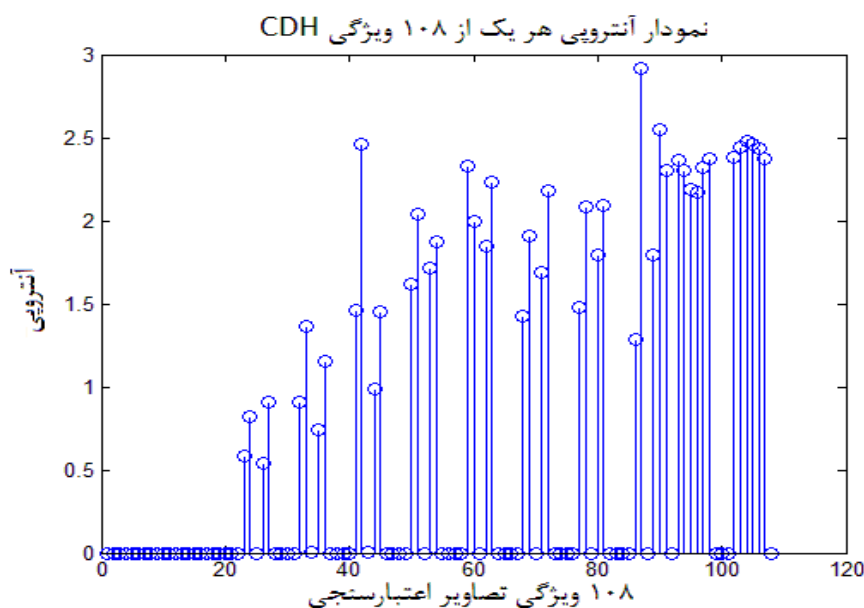
در این قسمت، با استفاده از معیار آنتروپی، ویژگی‌هایی که روی نرخ بازیابی، بی‌تاثیر هستند و یا تاثیر بسیار ناچیزی دارند حذف می‌شوند. در واقع هرچه آنتروپی یک ویژگی بیشتر باشد، آن ویژگی دامنه تغییرات بیشتر و حاوی اطلاعات مهم‌تری است. فرمول آنتروپی برای بردار X بصورت رابطه (۳-۱۱) است.

⁴⁵ Shannon

$$En(X) = - \sum_{i=1}^n P(x_i) \log_2 P(x_i) \quad (3-11)$$

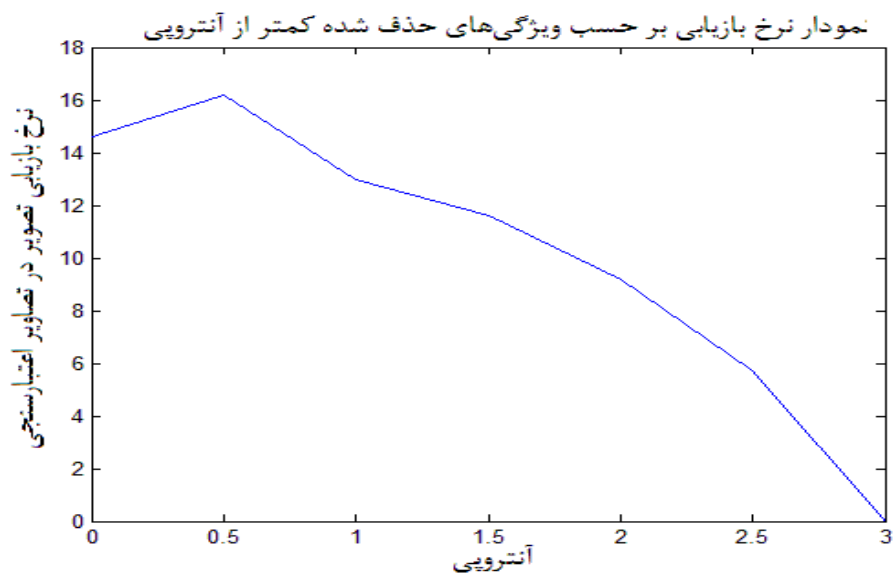
که در آن، $P(x_i)$ مقدار ویژگی نام در بردار X است.

بنابراین آنتروپی را روی ویژگی‌های هیستوگرام اختلاف رنگ در تصاویر اعتبارسنجی اعمال کرده و نمودار آن را در شکل (۳-۴) می‌بینیم.



شکل ۳-۴- آنتروپی ۱۰۸ ویژگی هیستوگرام اختلاف رنگ

همانطور که ملاحظه می‌شود، ویژگی‌های زیادی دارای آنتروپی خیلی کم و حتی صفر هستند. جهت حذف این ویژگی‌ها، روی داده‌های اعتبارسنجی، نرخ بازیابی می‌گیریم. بدین صورت که در هر مرحله ویژگی‌های با آنتروپی صفر تا ۳ را (با گام ۰.۵)، حذف و نرخ بازیابی را محاسبه می‌کنیم. همان‌طور که در شکل (۳-۵) ملاحظه می‌شود، نرخ بازیابی با حذف ویژگی‌های با آنتروپی به ترتیب صفر، ۰.۵ و ۳ محاسبه شده است. با توجه به این نمودار، ویژگی‌هایی که حذف آن‌ها باعث نرخ بازیابی بیشتری می‌شوند (ویژگی‌های با آنتروپی کمتر از ۰.۵)، حذف خواهند شد.



شکل ۳-۵- نمودار نرخ بازیابی-حذف ویژگی‌های کمتر از آن‌روپی

بنابراین، از ۱۰۸ ویژگی، ۶۳ ویژگی ناکارآمد، حذف می‌شوند. از ۴۵ ویژگی باقی‌مانده، ۳۲ ویژگی اول مربوط به ویژگی‌های قسمت رنگ و ۱۳ ویژگی دیگر مربوط به ویژگی‌های قسمت لبه است.

در مرحله بعد، رابطه آن‌روپی روی ۱۶۲ ویژگی هیستوگرام رنگی اعمال می‌شود. عملیات مانند قبل انجام شده و این‌بار ویژگی‌های با آن‌روپی کمتر از ۱.۵ حذف شدند. تعداد این ویژگی‌های حذفی در این نوع ویژگی، ۱۰۸ ویژگی بود. در نهایت بردار ویژگی هیستوگرام رنگ، ۵۴ بعد خواهد داشت.

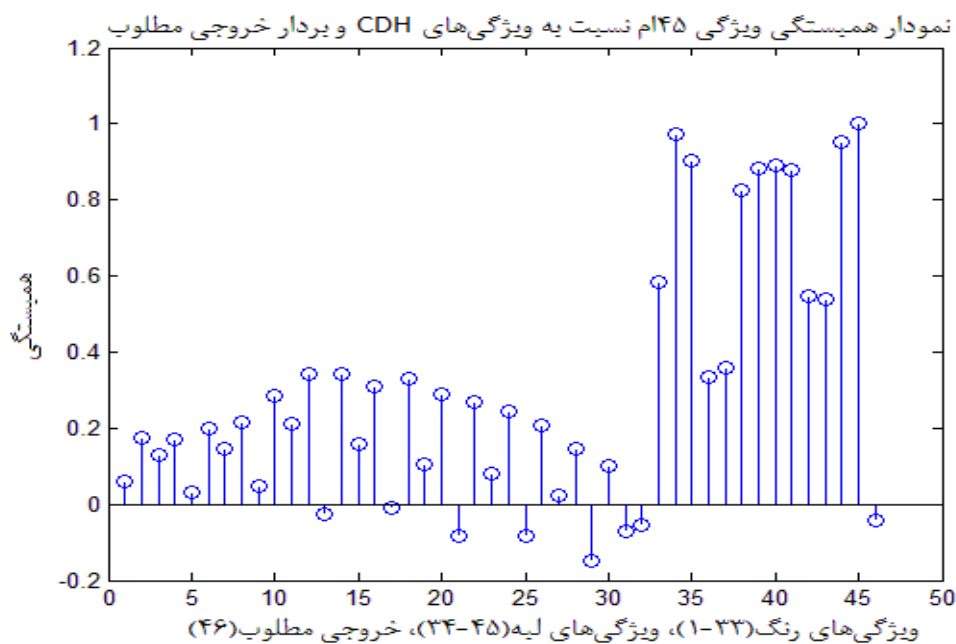
۳-۲-۲-۲- همبستگی بین بردارهای ویژگی

در گام بعدی برای حذف ویژگی‌های ناکارآمد، میزان همبستگی تک تک ویژگی‌ها نسبت به هم و همچنین نسبت به بردار خروجی مطلوب، روی داده‌های اعتبارسنجی، اندازه‌گیری می‌شود. ضریب همبستگی ابزاری آماری برای تعیین نوع (مستقیم یا معکوس) و درجه رابطه، برای دو متغیر کمی

است. این ضریب بین ۱ تا -۱ است و در صورت عدم رابطه بین دو متغیر، برابر صفر است. رابطه این ضریب برای دو بردار X و Y با n ویژگی، بصورت رابطه (۳-۱۲) است.

$$Corr(X, Y) = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}} \quad (3-12)$$

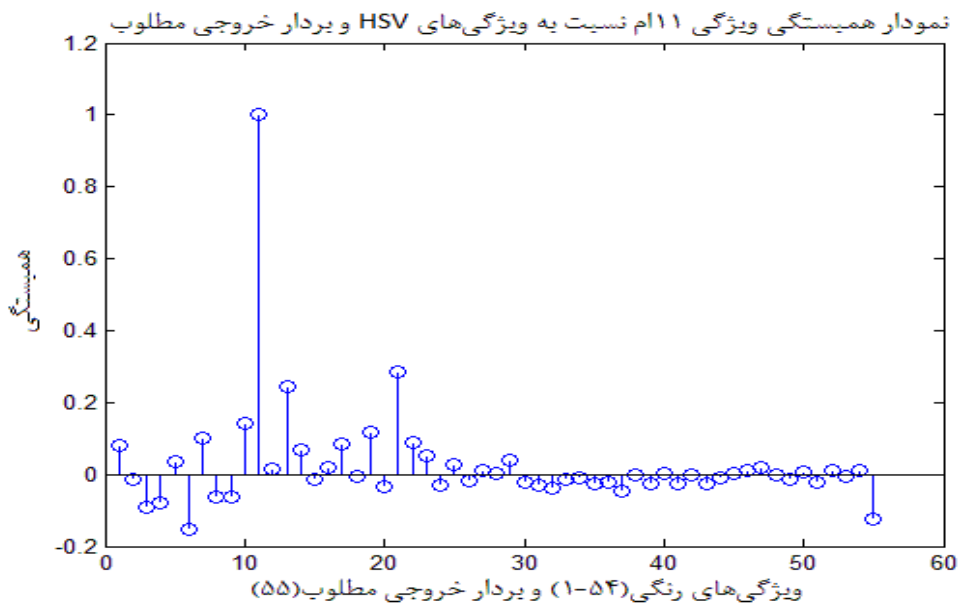
بین دو ویژگی اگر ارتباط وجود داشته باشد، می‌توان از یکی به دیگری رسید. پس اگر یکی از آن‌ها را حذف کنیم، مشکلی بوجود نخواهد آمد. برای هیستوگرام اختلاف رنگ با ۴۵ ویژگی باقیمانده، همبستگی بین هر ویژگی با سایر ویژگی‌ها و همچنین با بردار خروجی مطلوب، اندازه‌گیری می‌شود. منظور از بردار خروجی مطلوب، بردار شماره کلاس‌های تصاویر اعتبارسنجی است. همان‌طور که قبلاً بیان شد، ویژگی‌های ۱ تا ۳۲ مربوط به رنگ و ۳۳ تا ۴۵ مربوط به لبه است. برای نمونه، همبستگی ویژگی ۴۵ از هیستوگرام اختلاف رنگ نشان می‌دهد که این ویژگی نسبت به ویژگی‌های هم‌نوعش (ویژگی‌های مربوط به جهت‌گیری لبه)، همبستگی زیاد و نسبت به بقیه همبستگی کمتر و نسبت به بردار خروجی مطلوب (بردار ۴۶م)، همبستگی منفی دارد (شکل ۳-۶).



شکل ۳-۶- همبستگی بردار مولفه ۴۵م از هیستوگرام اختلاف رنگ نسبت به سایر بردارهای ویژگی و همچنین بردار خروجی مطلوب

همانطور که بیان شد، مقدار همبستگی بین بازه [۱،۱-] تغییر می‌کند. بنابراین در این قسمت، از بین بردارهایی که مقدار همبستگی آن‌ها باهم بیشتر از ۰.۹ باشند، یکی را نگه می‌داریم. جهت انتخاب یک بردار، برداری که با بردار خروجی مطلوب همبستگی مثبت بیشتری دارد، را نگه داشته و دیگری حذف می‌شود. از میان ۴۵ ویژگی هیستوگرام اختلاف رنگ، فقط سه ویژگی با ویژگی‌های هم‌نوع خود، همبستگی زیاد دارند. از بین این ویژگی‌ها، ویژگی ۴۵ با خاطر همبستگی منفی با بردار خروجی مطلوب، حذف خواهد شد. در نهایت هیستوگرام اختلاف رنگ ۴۴ ویژگی دارد.

در نهایت همبستگی بین ویژگی‌های باقیمانده از مرحله آنتروپی مربوط به هیستوگرام رنگ در فضای رنگ HSV بررسی می‌شوند. برای مثال شکل (۳-۷) همبستگی ویژگی ۱۱ام را نشان می‌دهد (ویژگی ۵۵ ام، بردار خروجی مطلوب است). پس از بررسی ویژگی‌های هیستوگرام رنگ، تعداد ۳۱ ویژگی ناکارآمد حذف و بنابراین هیستوگرام رنگ در فضای رنگ HSV تنها دارای (۲۳=۳۱-۵۴) ویژگی می‌شود.



شکل ۳-۷- همبستگی بردار ویژگی ۱۱ام از هیستوگرام رنگ HSV نسبت به سایر بردارهای ویژگی و همچنین بردار خروجی مطلوب

در نهایت ۳۲ ویژگی مربوط به قسمت رنگ از هیستوگرام اختلاف رنگ و ۱۲ ویژگی مربوط به قسمت لبه از هیستوگرام اختلاف رنگ و ۲۳ ویژگی مربوط به هیستوگرام رنگ در فضای رنگ HSV، جمعاً بردار ۶۷ ویژگی پیشنهادی اول را می‌سازند.

۳-۲-۳- معیار شباهت

یکی دیگر از مهمترین معیارها در بازیابی تصویر، بعد از نوع ویژگی‌ها و اندازه بردار ویژگی، معیار شباهتی است که به وسیله آن بردارهای ویژگی باهم مقایسه می‌شوند. اگر ویژگی‌های نزدیک به درک انسان، از تصویر استخراج شوند ولی معیار مناسبی جهت اندازه‌گیری شباهت آن‌ها استفاده نشود، نرخ بازیابی خوبی بدست نمی‌آید.

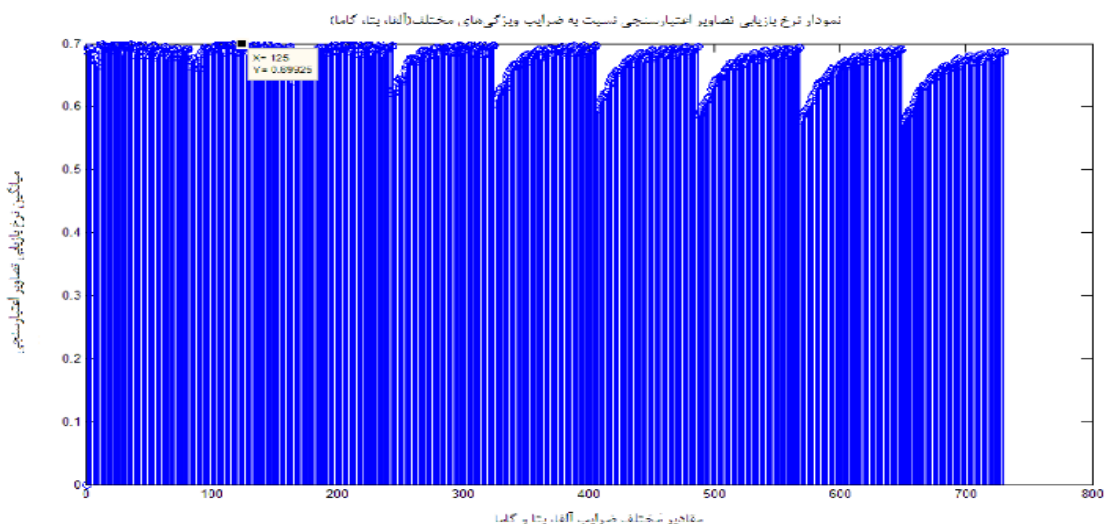
معیارهای شباهت زیادی برای بازیابی تصویر وجود دارند [۳۷]. ما نیز با آزمایش روی معیارهای مختلف، معیار معرفی شده در مقاله [۱۹] را جهت بازیابی استفاده می‌کنیم. همان‌طور که ذکر شد، این معیار بهبود یافته معیار کانبرا و به صورت رابطه (۳-۱۳) است.

$$D(T, Q) = \sum_{i=1}^M \frac{|T_i - Q_i|}{|T_i + u_T| + |Q_i + u_Q|} \quad (3-13)$$

که در آن D فاصله شباهت، T و Q بردارهای مورد مقایسه، M طول بردار ویژگی و u_T و u_Q به ترتیب میانگین بردار ویژگی T و Q است.

از بین ۶۷ ویژگی روش پیشنهادی اول، ویژگی‌های ۱ تا ۲۳ مربوط به هیستوگرام رنگی HSV، ویژگی‌های ۲۴ تا ۵۵ مربوط به اختلاف رنگ در راستای پیکسل‌های همسایه با رنگ (چندی‌سازی- شده) برابر و ویژگی‌های ۵۶ تا ۶۷ مربوط به اختلاف رنگ در راستای پیکسل‌های همسایه با جهت لبه برابر است.

یکی دیگر از عوامل مهم در بحث معیار شباهت و نرخ بازیابی، وزن‌دهی به ویژگی‌های مختلف است. کار دیگری که در این مقاله انجام شده است، وزن‌دهی غیریکنواخت به چهار گروه ویژگی پیشنهادی است. به این صورت که، ابتدا شباهت تصاویر با استفاده از هر نوع ویژگی محاسبه می‌شود، سپس به هر عدد دست آمده ضریبی بین ۰ و ۱ (با شرط مجموع ضرایب برابر ۱) اختصاص داده می‌شود. این کار باعث می‌شود که تاثیر هر گروه ویژگی در معیار فاصله، به میزانی باشد که بهترین استخراج معنایی از تصاویر انجام شود. جهت این کار، پس از محاسبه شباهت توسط هر گروه ویژگی، به هر مقدار بدست آمده، ضریبی بین ۰.۱ تا ۰.۹ اعمال می‌شود. دامنه تغییرات هر کدام از سه ضریب از ۰.۱ تا ۰.۹ است که با گام‌های ۰.۱ تغییر می‌کند. نرخ بازیابی روی داده‌های اعتبارسنجی با استفاده از این ضرایب محاسبه می‌شود. همان‌طور که در نمودار شکل ۳-۸ ملاحظه می‌شود، مرحله ۱۲۵ بیشترین نرخ بازیابی را داشته و معادل ضرایب (با شرط مجموع ضرایب برابر ۱) برای ویژگی‌های هیستوگرام رنگ، اختلاف رنگ در راستای پیکسل‌های همسایه با رنگ برابر و اختلاف رنگ در راستای پیکسل‌های همسایه با جهت لبه برابر به ترتیب، ۰.۱۲، ۰.۲۴ و ۰.۶۴ است.



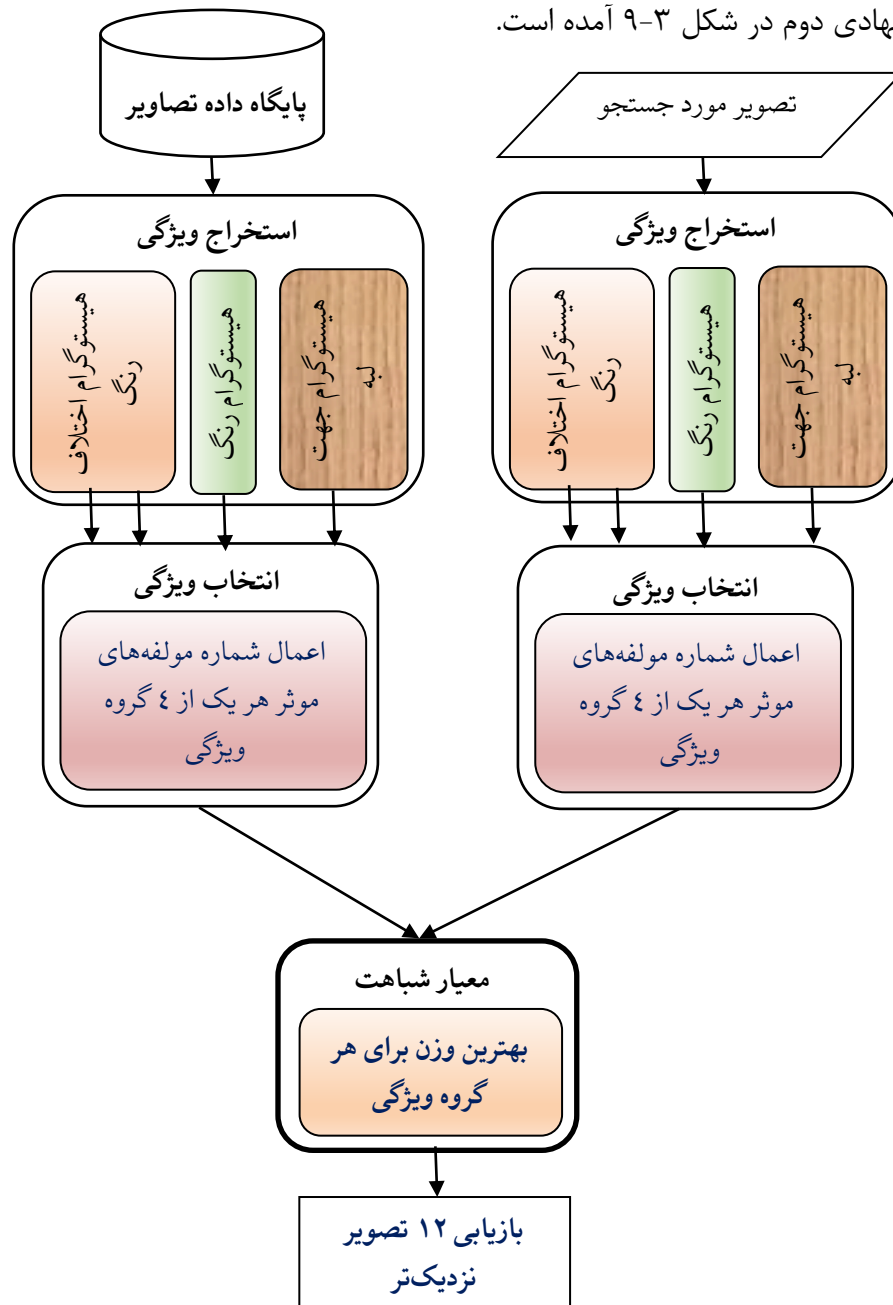
شکل ۳-۸ - نمودار نرخ بازیابی-وزن گروه‌های ویژگی.

این کار منجر به بهبود چشم‌گیری در نرخ بازیابی تصویر می‌شود. در فصل بعد در این مورد بیشتر بحث خواهد شد.

۳-۳- روش پیشنهادی دوم

روش پیشنهادی دوم، در واقع بهبود یافته‌ی روش پیشنهادی اول است. در این روش، به بردار ویژگی روش پیشنهادی اول، ویژگی‌هایی اضافه می‌شوند تا بتوان بهتر سطوح بالای معنایی را از تصویر استخراج کرد. در واقع این روش از حساسیت درک بصری انسان نسبت به جهت لبه، بهره می‌برد و باعث می‌شود اطلاعات معنایی استخراج شده، نرخ بازیابی تصویر را افزایش دهند. دیاگرام کلی روش

پیشنهادی دوم در شکل ۳-۹ آمده است.



شکل ۳-۹- دیاگرام کلی روش پیشنهادی دوم جهت بازیابی تصویر مبتنی بر محتوا.

همانطور که در شکل ۳-۸ دیده می‌شود، ویژگی جدیدی بنام "جهت لبه از بافت تصویر" به مجموعه ویژگی‌ها اضافه شده است. مراحل بهینه‌سازی بردار ویژگی روش پیشنهادی اول نیز روی این ویژگی جدید اعمال می‌شود. در ادامه این مراحل توصیف خواهند شد.

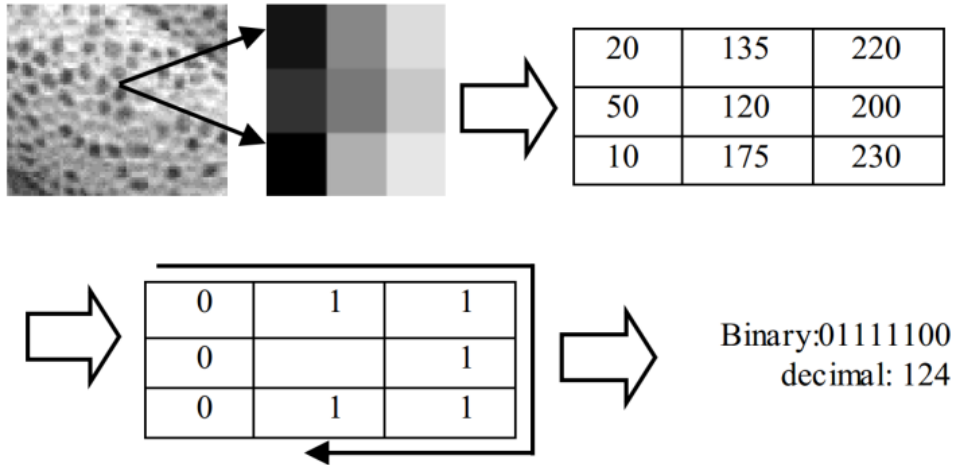
۳-۳-۱- استخراج ویژگی

۳-۳-۱-۱- جهت لبه از بافت تصویر

با توجه با اینکه هدف اصلی این مقاله بهبود روش هیستوگرام اختلاف رنگ یکنواخت ادراکی است و همچنین این روش از دو ویژگی موثر رنگ و لبه استفاده می‌کند، در روش پیشنهادی اول، ویژگی رنگی هیستوگرام رنگ در فضای رنگ HSV به آن اضافه شد. با توجه به این که روش هیستوگرام اختلاف رنگ از لبه‌های حاشیه‌ای اشیاء در تصویر استفاده می‌کند، اکنون جهت افزایش ویژگی‌های مربوط به لبه، ویژگی جدیدی، بیان می‌شود. بنابراین جهت استفاده از همه اطلاعات لبه، استفاده از اطلاعات لبه‌های داخلی اشیاء در کنار لبه‌های حاشیه‌ای، با استخراج اطلاعات لبه از بافت تصویر پیشنهاد می‌شود.

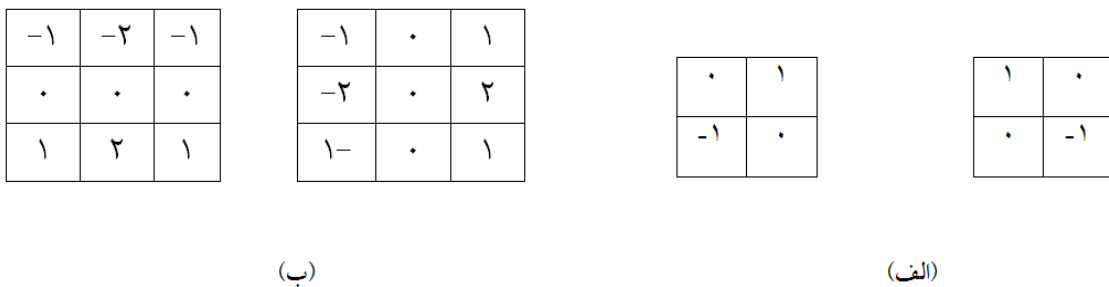
یکی از مهم‌ترین روش‌ها برای استخراج ویژگی بافت تصویر، الگوهای دودویی محلی (LBP) می‌باشد. عملگر LBP به عنوان یک توصیفگر قدرتمند بافت تصویر برای اولین بار توسط Ojala و همکاران معرفی شد [۲۱]. این عملگر برای هر پیکسل با توجه به برچسب پیکسل‌های همسایگی 3×3 یک عدد دودویی تولید می‌کند. این برچسب‌ها با آستانه‌سازی مقدار پیکسل‌های همسایه با مقدار پیکسل مرکزی به دست می‌آیند. به این صورت که برای پیکسل‌های با مقدار بزرگ‌تر یا مساوی مقدار پیکسل مرکزی برچسب ۱ و برای پیکسل‌های با مقادیر کوچک‌تر از مقدار پیکسل مرکزی برچسب ۰ قرار می‌گیرد. سپس این برچسب‌ها به صورت چرخشی در کنار هم قرار گرفته و یک عدد ۸ بیتی را تشکیل می‌دهند. نحوه کار این عملگر در شکل ۳-۱۰ آمده است. بدین ترتیب با جابجایی پنجره 3×3

در سرتاسر تصویر، بافت تصویر بدست می‌آید. پس از استخراج بافت از تصویر، باید به دنبال ویژگی قابل استفاده جهت مقایسه و بازیابی تصاویر بود.



شکل ۳-۱۰- روش کار الگوریتم الگوهای دودویی محلی جهت استخراج ویژگی بافت از تصویر.

در [۳۶] خسروی و همکاران، از ویژگی‌های سوبل و رابرتز جهت بازشناسی متون چاپی فارسی استفاده کرده‌اند. این ویژگی‌ها مشتمل بر دو ماسک سوبل و رابرتز هستند که در عین سادگی و حجم کم، کارایی مناسبی دارند. در شکل ۳-۱۱ ماسک‌های سوبل و رابرتز نشان داده شده‌اند. اپراتور سوبل لبه‌های افقی و عمودی و رابرتز نیز همانطور که در شکل گویاست لبه‌های مورب را مشخص می‌کنند.



شکل ۳-۱۱- (الف) ماسک رابرتز (ب) ماسک سوبل

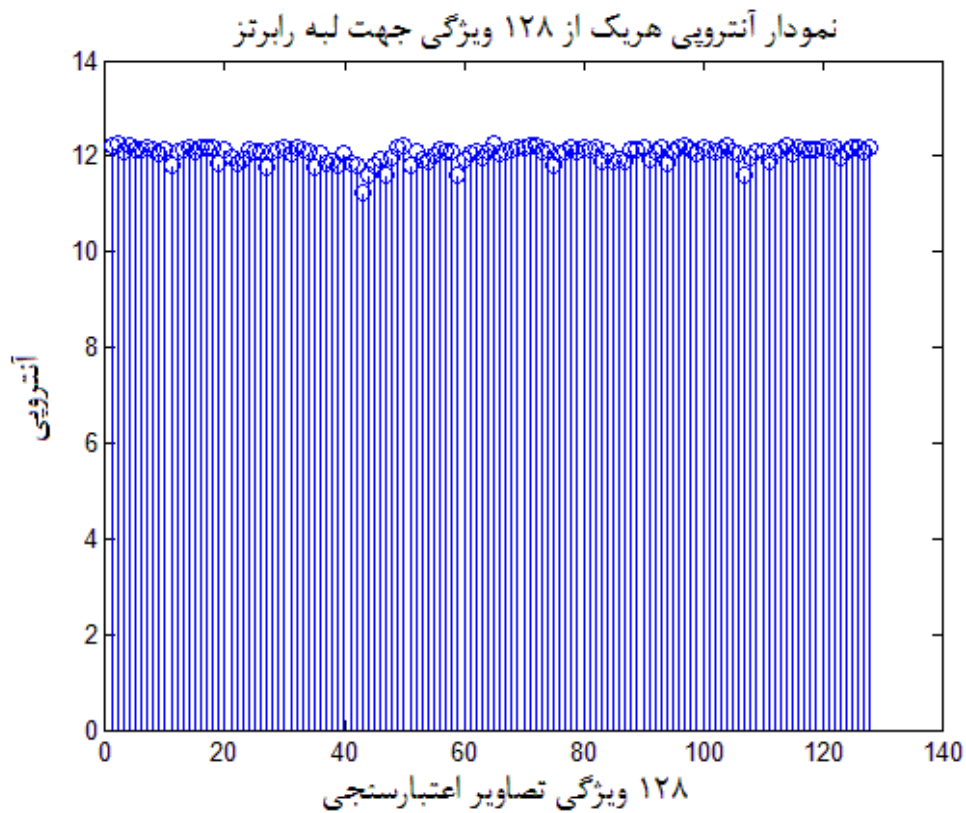
در این پایان‌نامه از ماسک رابرتز که گرادیان‌های قطری را استخراج می‌کند، جهت استخراج لبه از ویژگی‌های بافتی LBP استفاده می‌شود. بدین صورت که عملگر رابرتز را روی بافت حاصل از الگوریتم LBP اعمال کرده و لبه‌های تصویر از آن‌ها استخراج می‌شوند. پس از استخراج لبه‌های

تصویر، جهت لبه در هر پیکسل محاسبه می‌شود. سپس این زاویه‌ها به ۱۲۸ سطح چندی‌سازی شده و از هر تصویر، یک هیستوگرام ۱۲۸ تایی از زوایای لبه بافت آن استخراج می‌شود.

۳-۳-۲- انتخاب ویژگی

۳-۳-۲-۱- آنتروپی و حذف ویژگی‌های ناکارآمد

در روش پیشنهادی دوم، آنتروپی فقط بر روی ویژگی‌های روش جدید (جهت لبه از بافت تصویر) اعمال می‌شود. بنابراین آنتروپی روی ۱۲۸ ویژگی جهت لبه از بافت تصویر اعمال می‌شود. پس از بررسی مقدار آنتروپی این ویژگی‌ها، دیده می‌شود که همه این ویژگی‌ها، آنتروپی‌های بالا و نزدیک هم دارند. شکل (۳-۱۱) بیانگر همین موضوع است. بنابراین تعداد این ویژگی‌ها، در این مرحله، تغییر نمی‌کند.



شکل ۳-۱۲- آنتروپی ۱۲۸ ویژگی جهت لبه حاصل از بافت تصویر

۳-۲-۲- همبستگی بین بردارهای ویژگی

همبستگی بین ۱۲۸ ویژگی جهت لبه حاصل از بافت، در داده‌های اعتبارسنجی، باعث حذف ۲۰ ویژگی از این نوع می‌گردد. از بین همه این ویژگی‌ها، ۳۰ ویژگی باهم همبستگی بالای ۰.۹ دارند که موجب حذف ۲۰ ویژگی از آن‌ها می‌شود. در نهایت تعداد ویژگی‌های این روش به ۱۰۸ ویژگی کاهش می‌یابد.

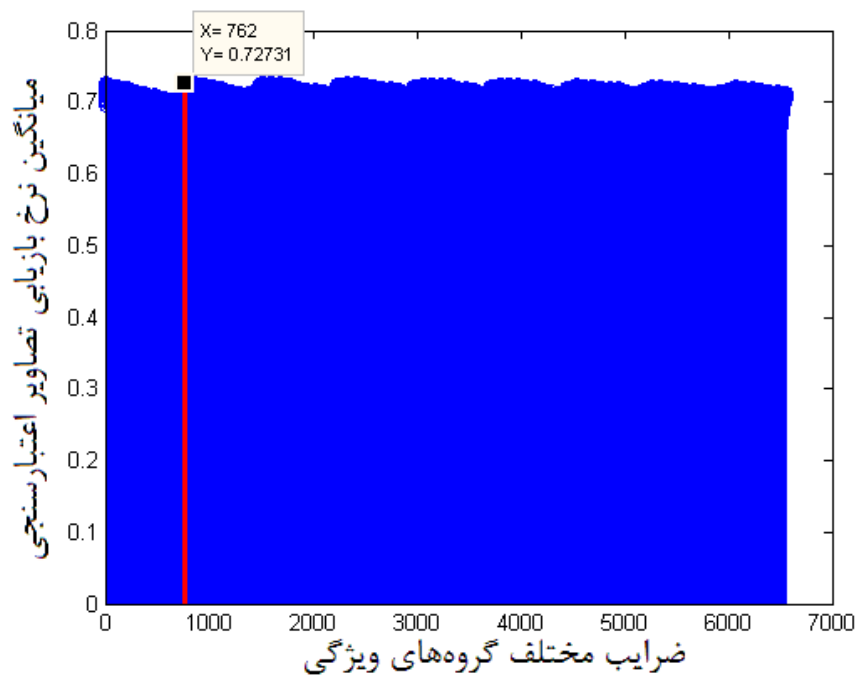
پس از بررسی آنتروپی و همبستگی بین ویژگی‌های روش جهت لبه، بردار ویژگی روش پیشنهادی دوم معرفی می‌شود. با ترکیب ۶۷ ویژگی روش پیشنهادی اول با ۱۰۸ ویژگی برتر هیستوگرام جهت لبه، بردار ۱۷۵ ویژگی روش پیشنهادی دوم ساخته می‌شود. بنابراین از ۱۷۵ ویژگی روش پیشنهادی دوم، ۳۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای رنگ (چندی‌سازی- شده) برابر تصویر، ۱۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای لبه تصویر، ۲۳ ویژگی مربوط به هیستوگرام رنگ در فضای رنگی HSV و ۱۰۸ ویژگی مربوط به هیستوگرام جهت لبه از بافت تصویر خواهد بود.

در ادامه پس از توصیف معیار شباهت، به گروه‌های مختلف ویژگی، وزن‌هایی تخصیص می‌یابد. این وزن‌دهی موجب استخراج بهتر مفاهیم معنایی از تصاویر خواهند شد.

۳-۳-۳- معیار شباهت

همان‌طور که ذکر شد، یکی دیگر از مهمترین معیارها در بازیابی تصویر، بعد از نوع ویژگی‌ها و اندازه بردار ویژگی، معیار شباهتی است که به وسیله آن بردارهای ویژگی باهم مقایسه می‌شوند. در روش پیشنهادی دوم نیز، معیار کانبرای بهبود یافته استفاده شده است.

از ۱۷۵ ویژگی روش پیشنهادی دوم، ۱۰۸ ویژگی مربوط به هیستوگرام جهت لبه از بافت تصویر، ۲۳ ویژگی مربوط به هیستوگرام رنگ در فضای رنگی HSV، ۳۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای رنگ و ۱۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای لبه تصویر خواهد بود. وزن‌دهی به چهار گروه، روی داده‌های اعتبارسنجی اعمال شد. بطوریکه هر ضریب را از مقدار ۰.۱ تا ۰.۹ با گام‌های ۰.۱ تغییر داده و در هر مرحله، نرخ بازیابی محاسبه شد. نمودار شکل (۳-۱۳) نشان می‌دهد که مرحله ۷۶۲، بیشترین نرخ بازیابی را داشته و معادل ضرایب (با شرط مجموع ضرایب برابر ۱) به ترتیب ۰.۲۵، ۰.۱، ۰.۲ و ۰.۴۵ است.



شکل ۳-۱۳- نمودار نرخ بازیابی-وزن گروه‌های ویژگی.

در فصل چهارم با ارزیابی روش‌های پیشنهادی، ملاحظه خواهد شد که بردار ویژگی روش‌های پیشنهادی، به همراه وزن‌دهی به آن‌ها بخوبی مفاهیم معنایی را از تصاویر استخراج می‌کنند. این مجموعه با عنوان سطوح بالای معنایی شناخته می‌شود. چرا که محتوای معنایی تصویر را درک می‌-

کند و وابسته به ویژگی سطح پایین خاصی نیست. استفاده از این سطوح بالای معنایی جهت بازیابی تصویر مبتنی بر محتوا، منجر به بهبود نرخ بازیابی تصویر می‌شود.

۳-۴- جمع‌بندی

در این فصل دو روش جهت بازیابی تصویر مبتنی بر محتوا پیشنهاد شد. سیستم بینایی انسان به دو ویژگی رنگ و لبه حساس‌تر از سایر ویژگی‌ها است. لذا این دو مجموعه ویژگی در سیستم‌های بازیابی مبتنی بر محتوا نقش بسزایی دارند. یکی از روش‌های مطرح در این حوزه، استخراج جهت لبه و اطلاعات رنگ از فضای رنگ $L^*a^*b^*$ است. این مجموعه ویژگی تحت عنوان "هیستوگرام اختلاف رنگ" نام‌گذاری شده است. ابتدا با اعمال دو معیار آنتروپی و همبستگی بین ویژگی‌های روش هیستوگرام اختلاف رنگ، ویژگی‌های کارآمد آن انتخاب شد. این روش دارای ۱۰۸ ویژگی بود که پس از کاهش ویژگی، دارای ۴۴ ویژگی شد. پس از آن در به دنبال افزایش ویژگی رنگ، هیستوگرام رنگ در فضای رنگ HSV پیشنهاد شد. فضای رنگ HSV دارای خاصیت یکنواختی ادراکی است. تعداد ویژگی‌های این هیستوگرام، ۱۶۲ ویژگی بود، که با اعمال آنتروپی و همبستگی بین این ویژگی‌ها، تعداد ۲۳ ویژگی کارآمد انتخاب شد. این ۲۳ ویژگی به همراه ۴۴ ویژگی هیستوگرام اختلاف رنگ، برادر ۶۷ ویژگی پیشنهادی اول را می‌سازند.

در روش پیشنهادی دوم، به دنبال افزایش ویژگی‌های مربوط به لبه، ویژگی هیستوگرام جهت لبه حاصل از بافت تصویر به مجموعه ویژگی پیشنهادی اول اضافه می‌شود. ابتدا بافت تصویر با استفاده از الگوریتم الگوهای دودویی محلی (LBP) استخراج می‌شود. به علت اینکه لبه‌های تصویر در بافت نسبت به تصویر اصلی تیزتر است، لبه‌ها با استفاده از گرادیان رابرتز از بافت تصویر استخراج می‌شود. پس از آن، جهت لبه هر پیکسل در بافت تصویر استخراج و به ۱۲۸ سطح چندی‌سازی می‌شود. سپس معیارهای آنتروپی و همبستگی بین ویژگی‌های این هیستوگرام ۱۲۸ تایی، اعمال می‌شود. این تعداد

ویژگی به ۱۰۸ ویژگی کاهش می‌یابد. ترکیب این ۱۰۸ ویژگی با ۶۷ ویژگی روش پیشنهادی اول، بردار ۱۷۵ ویژگی پیشنهادی دوم را می‌سازد.

این دو بردار ۶۷ ویژگی و ۱۷۵ ویژگی پیشنهادی، بدون نیاز به یادگیری، قطعه‌بندی و شناسایی اشیاء، بخوبی مفاهیم معنایی تصاویر را استخراج می‌کند. جهت استخراج دقیق‌تر اطلاعات معنایی تصاویر، تخصیص وزن به هر گروه ویژگی، کار دیگری است که در این پایان‌نامه انجام شده است. پس از محاسبه شباهت تصاویر با استفاده از هر گروه ویژگی، به مقدار بدست آمده از آن‌ها، وزنی تخصیص داده می‌شود. این وزن‌دهی به همراه بردار ویژگی با عنوان سطوح بالای معنایی شناخته می‌شود. چرا که محتوای معنایی تصویر را درک می‌کند و وابسته به ویژگی سطح پایین خاصی نیست. استفاده از این سطوح بالای معنایی جهت بازیابی تصویر مبتنی بر محتوا، منجر به بهبود نرخ بازیابی تصویر می‌شود.

فصل چهارم:

نتایج و آزمایش‌ها

۴-۱- مقدمه

پس از بیان روش‌های پیشنهادی، در این فصل این روش‌ها مورد ارزیابی و مقایسه قرار می‌گیرند. در ادامه ابتدا پایگاه داده تصویری مورد استفاده مختصر تشریح می‌شود. سپس همانطور که در فصل سوم بیان شد، نرخ بازیابی پس از اعمال آنتروپی و همبستگی روی ویژگی‌های مختلف بصورت جداگانه بحث و تشریح می‌شوند. همچنین پس از اعمال وزندهی به گروه‌های مختلف ویژگی در معیار شباهت، نرخ بازیابی روش‌های پیشنهادی بحث می‌شود. در نهایت نرخ بازیابی روش‌های پیشنهادی با روش‌های اخیر در حوزه CBIR مقایسه می‌شوند.

۴-۲- پایگاه داده تصویر

در این مقاله از دو پایگاه داده استاندارد Corel 10k و Corel 5K به ترتیب با ۵۰۰۰ و ۱۰۰۰۰ تصویر در ۵۰ و ۱۰۰ کلاس مختلف استفاده می‌شود. کلاس‌های این پایگاه داده، شامل تصاویر با پس‌زمینه‌های پیچیده‌ای است. تصاویر در آدرس [۳۷] موجود و قابل دریافت است. نمونه‌هایی از تصاویر این پایگاه داده در شکل ۴-۱ نشان داده شده است.



شکل ۴-۱- نمونه‌هایی از تصاویر پایگاه داده تصویر Corel5k

از ۱۰۰۰۰ تصویر پایگاه داده Corel 10k، ۵۰۰۰ تصویر اولیه مطابق با تصاویر پایگاه داده Corel 5k است. بنابراین از مجموع ۵۰۰۰ تصویر موجود در پایگاه داده Corel 5k، ۱۰۰۰ تصویر بصورت تصادفی از کلاس‌های مختلف، به عنوان داده‌های اعتبارسنجی برای هر دو پایگاه داده استفاده می‌شود

اکثر مقاله‌های موجود در حوزه بازیابی تصویر مبتنی بر محتوا، نرخ بازیابی را بصورت میانگین دقت بازیابی روی ۱۰۰۰ و ۲۰۰۰ تصویر تصادفی به ترتیب از پایگاه‌های داده Corel 5k و Corel 10k، محاسبه می‌کنند. جهت مقایسه روش‌های پیشنهادی با این مقاله‌ها، تعداد تصاویر جهت بازیابی، نسبت گرفته می‌شود. مقاله‌های مذکور ۲۰ درصد تصاویر پایگاه داده را جهت آزمون در نظر می‌گیرند. پس با توجه به این که تعداد تصاویر پایگاه‌های داده ۴۰۰۰ و ۹۰۰۰ تصویر است، به ترتیب ۸۰۰ و ۱۸۰۰ تصویر بصورت تصادفی جهت بازیابی در نظر گرفته می‌شود. پس دقت نرخ بازیابی روش‌های پیشنهادی بصورت میانگین نرخ بازیابی روی این دو زیرمجموعه محاسبه می‌شود. ضمناً چون روش‌های پیشنهادی بدون هرگونه آموزش هستند، بنابراین تصاویری جهت آموزش در نظر گرفته نشده است.

۴-۳- بحث روی نرخ بازیابی ویژگی‌های مختلف

پس از تشریح روش پیشنهادی و معیار شباهت مورد استفاده، نرخ بازیابی محاسبه می‌شود. مقایسه نرخ بازیابی روی ویژگی‌های باقی‌مانده از هر مرحله کاهش ویژگی (آنتروپی و همبستگی) انجام می‌شود.

نرخ بازیابی تصویر روش هیستوگرام اختلاف رنگ [۱۹] با ۱۰۸ ویژگی، روی تصاویر آزمون پایگاه داده Corel 5k و Corel 10k به ترتیب ۵۷.۲۳ و ۴۵.۲۴ درصد است. با حذف ۶۳ ویژگی توسط آنتروپی و همبستگی، نرخ بازیابی حدود ۱ درصد روی هر دو پایگاه داده، افزایش می‌یابد. همچنین با حذف ویژگی ۱۴۵ام این ویژگی‌ها، نرخ بازیابی تغییری نمی‌کند. اگرچه حذف ویژگی در

مواردی ممکن است موجب بهبود نرخ بازیابی نشود ولی موجب بهبود در بار محاسباتی و زمان بازیابی تصویر می‌شود.

جدول ۴-۲- مقایسه نرخ بازیابی روش هیستوگرام اختلاف رنگ [۱۹] با بهبود یافته‌های آن.

الگوریتم	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام اختلاف رنگ (۱۰۸ ویژگی) [۱۹]	۵۷.۲۳ درصد	۴۵.۲۴ درصد
۴۵ ویژگی هیستوگرام اختلاف رنگ	۵۷.۹۲ درصد	۴۶.۱۱ درصد
۴۴ ویژگی هیستوگرام اختلاف رنگ	۵۷.۹۲ درصد	۴۶.۱۱ درصد

در این قسمت ابتدا علت انتخاب فضای رنگ HSV تشریح می‌شود. همان‌طور که قبلاً بیان شد، بین فضاهای رنگ، دو فضای رنگ HSV و $L^*a^*b^*$ دارای خاصیت یکنواخت ادراکی هستند. یک فضای رنگ یکنواخت ادراکی، تضمین می‌کند که تفاوت بین دو رنگ (که توسط چشم انسان درک شده) متناسب با فاصله‌ی اقلیدسی آن‌ها در فضای رنگ است. جهت انتخاب یکی از این دو فضا جهت استخراج ویژگی هیستوگرام رنگ، هر فضای رنگ به ۱۶۲ مقدار چندی‌سازی می‌شود. در فضای رنگ HSV، کانال H به علت اهمیت بیشتر، به ۱۸ بین و دو کانال دیگر هر کدام به ۳ بین چندی‌سازی می‌شوند. در فضای رنگ $L^*a^*b^*$ نیز، کانال L^* که نشان‌دهنده روشنایی تصویر است به ۱۸ بین و دو کانال دیگر هر کدام به ۳ بین چندی‌سازی می‌شوند. نرخ بازیابی حاصل از این دو هیستوگرام در دو فضای رنگ مذکور، محاسبه شده و در جدول ۴-۲ قابل مشاهده است.

جدول ۴-۳- مقایسه نرخ بازیابی روش هیستوگرام رنگ نسبت به دو فضای رنگ HSV و $L^*a^*b^*$.

الگوریتم	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام رنگ در فضای رنگ HSV	۴۴.۲۱ درصد	۳۵.۴۱ درصد
هیستوگرام رنگ در فضای رنگ $L^*a^*b^*$	۴۱.۵۴ درصد	۳۱.۶۷ درصد

همان طور که ملاحظه می‌شود، هیستوگرام رنگ در فضای رنگ HSV نسبت به فضای رنگ $L^*a^*b^*$ قابلیت تمییز بهتری روی تصاویر دارد و درصد نرخ بازیابی آن بیشتر است. بنابراین فضای رنگ HSV جهت استخراج هیستوگرام رنگ، انتخاب شد.

نرخ بازیابی ۱۶۲ ویژگی هیستوگرام رنگی HSV به تنهایی، ۴۴.۲۱ و ۳۵.۴۱ درصد است که پس از حذف ویژگی‌های ناکارآمد توسط آنتروپی و همبستگی، نرخ بازیابی تغییر چندانی نمی‌کند. جدول ۴-۳ این مقادیر را نشان می‌دهند.

جدول ۴-۴- مقایسه نرخ بازیابی روش هیستوگرام رنگ HSV با بهبود یافته‌های آن.

الگوریتم	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام رنگ HSV (۱۶۲ ویژگی)	۴۴.۲۱ درصد	۳۵.۴۱ درصد
۵۴ ویژگی هیستوگرام رنگ HSV	۴۴.۲۳ درصد	۳۵.۴۴ درصد
۲۳ ویژگی هیستوگرام رنگ HSV	۴۴.۲۱ درصد	۳۵.۴۰ درصد

نرخ بازیابی تصویر حاصل از ویژگی هیستوگرام جهت لبه به تنهایی با ۱۲۸ ویژگی، روی تصاویر آزمون به ترتیب ۴۷.۱۶ و ۳۶.۹۴ درصد است. با حذف ۲۰ ویژگی توسط روش همبستگی، نرخ بازیابی تغییر چندانی نمی‌کند (جدول ۴-۴).

جدول ۴-۵- مقایسه نرخ بازیابی روش هیستوگرام جهت لبه با بهبود یافته‌های آن.

الگوریتم	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام جهت لبه (۱۲۸ ویژگی)	۴۷.۱۶ درصد	۳۶.۹۴ درصد
۱۰۸ ویژگی هیستوگرام جهت لبه	۴۷.۱۸ درصد	۳۶.۹۱ درصد

۴-۴- ارزیابی نرخ بازیابی روش پیشنهادی اول

همان‌طور که ذکر شد، تعداد ۲۳ ویژگی مربوط به هیستوگرام رنگی HSV، ۳۲ ویژگی مربوط به قسمت رنگ از هیستوگرام اختلاف رنگ و ۱۲ ویژگی مربوط به قسمت جهت لبه از هیستوگرام اختلاف رنگ است. بنابراین این ویژگی‌ها جمعاً بردار ۶۷ ویژگی پیشنهادی اول را می‌سازند. نرخ بازیابی با استفاده از این ۶۷ ویژگی، بدون وزندهی به گروه‌های مختلف ویژگی، برابر ۵۸.۸۶ و ۴۸.۵۶ درصد می‌شود.

همچنین وزندهی به سه گروه مختلف ویژگی، روی داده‌های اعتبارسنجی اعمال شد. ضرایب بدست آمده برای هر گروه ویژگی به ترتیب، ۰.۱۲، ۰.۲۴ و ۰.۶۴ است. پس از اعمال این ضرایب، نرخ بازیابی به ترتیب به ۶۰.۰۸ و ۴۹.۰۰ می‌رسد.

همان‌طور که گفته شده یکی از روش‌هایی که جهت مقایسه با روش‌های پیشنهادی انتخاب شده، مقاله [۱۵] است. روش مورد استفاده در این مقاله بطور کامل در فصل دوم تشریح شد. بردار

ویژگی این مقاله، ۹۶ بعد دارد. مقایسه این روش با روش پیشنهادی اول بصورت جدول ۴-۵ می‌باشد. همچنین روش پیشنهادی اول با روش هیستوگرام اختلاف رنگ با ۱۰۸ ویژگی مقایسه شده است.

جدول ۴-۶- مقایسه نرخ بازیابی روش پیشنهادی اول با روش‌های اخیر

الگوریتم	تعداد ویژگی	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام اختلاف رنگ [۱۹]	۱۰۸	۵۷.۲۳ درصد	۴۵.۲۴ درصد
روش سریع مبتنی بر بافت و رنگ [۱۵]	۹۶	۵۷.۵۰ درصد	۴۷.۰۱ درصد
روش پیشنهادی اول بدون اعمال وزندهی	۶۷	۵۸.۸۶ درصد	۴۸.۵۶ درصد
روش پیشنهادی اول	۶۷	۶۰.۰۸ درصد	۴۹.۰۰ درصد

همان‌طور که در جدول ۴-۵ ملاحظه می‌شود، روش پیشنهادی اول، با وجود کم‌ترین تعداد ویژگی، بهترین نرخ بازیابی تصویر را دارد. این نشان‌دهنده‌ی این است که این بردار ۶۷ ویژگی، بخوبی فضای دو بعدی تصویر را توصیف کرده و اطلاعات معنایی را از تصویر استخراج می‌کند. این افزایش نرخ حاکی از آن است که ویژگی‌های پیشنهادی، بخوبی مشکل شکاف معنایی بین درک سیستم و درک انسان را کاهش می‌دهند.

روش مقاله [۱۵] یکی از جدیدترین پژوهش‌ها در حوزه بازیابی تصویر از نظر زمان استخراج ویژگی و زمان بازیابی تصویر است. در این قسمت، زمان استخراج ویژگی، کاهش ویژگی و بازیابی تصویر روش پیشنهادی اول با این مقاله مقایسه می‌شود. در جدول ۴-۶ این مقایسه مشاهده می‌شود.

جدول ۴-۷- مقایسه روش پیشنهادی اول با روش‌های اخیر از نظر زمان بازیابی تصویر روی پایگاه داده Corel 5k.

مجموع زمان	زمان بازیابی تصویر		زمان استخراج ویژگی (ثانیه)			تعداد ویژگی	الگوریتم
	خواندن	بازیابی	کاهش ویژگی	ویژگی دوم	ویژگی اول		
۰.۲۳۶	۰.۰۱۵	۰.۱۶۰	۰.۰۶۱			۱۰۸	هیستوگرام اختلاف رنگ [۱۹]
	۰.۰۷۹		۰.۰۷۸				
۰.۱۵۷	۰.۰۷۹		۰.۰۷۸			۹۶	روش سریع مبتنی بر بافت و رنگ [۱۵]
۰.۱۵۵	بازیابی	خواندن	کاهش ویژگی	ویژگی دوم	ویژگی اول	۶۷	روش پیشنهادی اول
	۰.۰۶۰	۰.۰۰۸	۰.۰۰۹	۰.۰۱۷	۰.۰۶۱		

در جدول ۴-۶ منظور از "زمان استخراج ویژگی"، مقدار زمانی است که از یک تصویر، ویژگی‌ها استخراج می‌شوند. "خواندن" به معنای مقدار زمانی است که فایل حاوی ویژگی‌های کلیه تصاویر پایگاه داده Corel 5k فراخوانی می‌شود. "بازیابی" منظور زمان مقایسه تصویر مورد جستجو با سایر تصاویر پایگاه داده است. ویژگی اول منظور ویژگی هیستوگرام اختلاف رنگ و ویژگی دوم، هیستوگرام رنگ است. کاهش ویژگی منظور مقدار زمانی است که طول می‌کشد تا بردار ۶۷ ویژگی پیشنهادی از دو ویژگی ساخته شود. همانطور که ملاحظه می‌شود، روش پیشنهادی اول، از نظر زمان نیز، بسیار بهبود یافته است.

با توجه به دو جدول ۴-۵ و ۴-۶، مشاهده می‌شود بردار ۶۷ ویژگی پیشنهادی اول، از نظر دقت و سرعت، نسبت به سایر روش‌های اخیر بهبود قابل توجهی دارد. باید تاکید کرد، که بردار ویژگی روش پیشنهادی اول، از نظر تعداد ویژگی نیز، از سایر روش‌های اخیر، کم‌تر است.

۴-۵- ارزیابی نرخ بازیابی روش پیشنهادی دوم

همان‌طور که ذکر شد، در روش پیشنهادی دوم، تعداد ۱۰۸ ویژگی مربوط به جهت لبه مبتنی بر بافت LBP تصویر به مجموعه ۶۷ ویژگی روش پیشنهادی اول اضافه می‌شود. بنابراین این ویژگی‌ها جمعاً بردار ۱۷۵ ویژگی پیشنهادی دوم را می‌سازند. نرخ بازیابی با استفاده از این ۱۷۵ ویژگی، بدون وزن-دهی به گروه‌های مختلف ویژگی، روی پایگاه‌های داده Corel 5k و Corel 10k به ترتیب برابر با ۶۲.۳۲ و ۵۲.۵۸ درصد می‌شود.

همچنین وزن‌دهی به چهار گروه مختلف ویژگی، روی داده‌های اعتبارسنجی اعمال شد. ضرایب بدست آمده برای هر گروه ویژگی (۱۰۸ ویژگی مربوط به هیستوگرام جهت لبه، ۲۳ ویژگی مربوط به هیستوگرام رنگ، ۳۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای رنگ و ۱۲ ویژگی مربوط به هیستوگرام اختلاف رنگ در راستای لبه) به ترتیب ۰.۲۵، ۰.۱، ۰.۲ و ۰.۴۵ است. پس از اعمال این ضرایب، نرخ بازیابی به ترتیب به ۶۴.۵۳ و ۵۶.۰۱ می‌رسد.

همان‌طور که در فصل دوم بیان شد، روش دیگری که جهت مقایسه با روش پیشنهادی، انتخاب شده است، مقاله [۱۱] است. علت این که این مقاله با روش پیشنهادی اول مقایسه نشد، تعداد ویژگی‌های آن است. این روش جزو روش‌های با بردار ویژگی بزرگ است. همان‌طور که ملاحظه شد، بردار ویژگی این مقاله، ۲۶۰ بعدی است. بنابراین روش پیشنهادی دوم با ۱۷۵ ویژگی با این روش، مقایسه می‌شود. این مقایسه در جدول ۴-۷ نشان داده شده است.

جدول ۴-۸- مقایسه نرخ بازیابی روش پیشنهادی دوم با روش‌های اخیر.

الگوریتم	تعداد ویژگی	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام اختلاف رنگ [۱۹]	۱۰۸	۵۷.۲۳ درصد	۴۵.۲۴ درصد
روش هیستوگرام برجستگی [۱۱]	۲۶۰	-	۵۴.۸۸ درصد
روش پیشنهادی دوم بدون اعمال وزن‌دهی	۱۷۵	۶۲.۳۲ درصد	۵۲.۵۸ درصد
روش پیشنهادی دوم	۱۷۵	۶۴.۵۳ درصد	۵۶.۰۱ درصد

همان‌طور که در جدول ۴-۷ ملاحظه می‌شود، روش پیشنهادی دوم، بهبود قابل توجهی نسبت به روش هیستوگرام رنگ [۱۹] دارد. همچنین نسبت به روش هیستوگرام برجستگی [۱۱]، با وجود بردار ویژگی کوچکتر، بازهم بهبود نرخ بازیابی تصویر وجود دارد.

همان‌طور که در جدول‌های ۴-۵ و ۴-۷ مشاهده می‌شود، روش‌های پیشنهادی نسبت به جدیدترین روش‌های حوزه بازیابی تصویر، بهبود قابل توجهی در افزایش نرخ بازیابی تصویر دارند. این افزایش نرخ حاکی از آن است که اطلاعات معنایی تصویر بخوبی از تصاویر استخراج می‌شود. از نظر زمان بازیابی نیز، روش پیشنهادی اول، نسبت به جدیدترین و سریع‌ترین روش در حوزه بازیابی تصویر، بهبود یافته است.

یکی دیگر از مهم‌ترین روش‌های پایه جهت مقایسه در حوزه بازیابی تصویر، روش MPEG-7 [۳۸] است. در این روش از ترکیب ویژگی‌های رنگ و بافت استفاده می‌شود. ویژگی‌های رنگ شامل سه هیستوگرام رنگ، شامل رنگ غالب، طرح‌بندی رنگ و رنگ مقیاس‌پذیر است. ویژگی‌های بافتی این روش، شامل هیستوگرام لبه، توصیف‌گر بافت همگن و توصیف‌گر مرور بافت است. در نهایت با ترکیب این دو دسته ویژگی رنگ و بافت، از هر کانال تصویر، ۸۰ ویژگی استخراج می‌شود. جمعاً تعداد ویژگی‌های این روش ۲۴۰ ویژگی است. در جدول ۴-۹ مقایسه کلی هر دو روش پیشنهادی با چهار روش جدید قابل مقایسه، نشان داده شده است.

جدول ۴-۹ - مقایسه روش‌های پیشنهادی با چهار روش از جدیدترین روش‌ها در حوزه بازیابی تصویر.

الگوریتم	تعداد ویژگی	درصد نرخ بازیابی تصویر روی Corel 5k	درصد نرخ بازیابی تصویر روی Corel 10k
هیستوگرام اختلاف رنگ [۱۹]	۱۰۸	۵۷.۲۳ درصد	۴۵.۲۴ درصد
روش هیستوگرام برجستگی [۱۱]	۲۶۰	-	۵۴.۸۸ درصد
روش سریع مبتنی بر بافت و رنگ [۱۵]	۹۶	۵۷.۵۰ درصد	۴۷.۰۱ درصد
روش MPEG-7 [۳۸]	۲۴۰	۳۹.۴۶ درصد	۳۲.۳۱ درصد
روش پیشنهادی اول	۶۷	۶۰.۰۸ درصد	۴۹.۰۰ درصد
روش پیشنهادی دوم	۱۷۵	۶۴.۵۳ درصد	۵۶.۰۱ درصد

جدول ۴-۹ درصدهای نرخ بازیابی تصویر روش‌های پیشنهادی در مقایسه با جدیدترین روش‌ها در حوزه بازیابی تصویر را نشان می‌دهد. همان‌طور که مشاهده می‌شود، روش پیشنهادی اول، در عین تعداد کم ویژگی، روی پایگاه داده corel 5k از نرخ‌های بازیابی ۳۹، ۵۷.۲۳ و ۵۷.۵۰ درصد به نرخ بازیابی ۶۰.۰۸ درصد افزایش یافته است. همچنین روش پیشنهادی دوم نیز، روی پایگاه داده corel 10k از نرخ‌های بازیابی ۳۲، ۴۵، ۵۴ و ۴۷ درصد به نرخ بازیابی ۵۶ درصد افزایش داشته است. بنابراین با توجه به تعداد ویژگی و درصد نرخ بازیابی، مشهود است که روش‌های پیشنهادی بطور قابل توجهی توانسته‌اند محتوای تصویر را توصیف کنند.

۴-۶- ارزیابی خروجی سیستم بازیابی تصویر پیشنهادی

چند نمونه از خروجی‌های سیستم بازیابی تصویر پیشنهادی در ادامه مشاهده می‌شود. بازیابی ۱۲ تصویر مربوط به هر نمونه تصویر مورد جستجو با روش پیشنهادی مقاله [۱۹] مقایسه می‌شود. شکل ۲-۴ یکی از نمونه‌های سیستم بازیابی مقاله [۱۹] را نشان می‌دهد.



شکل ۴-۲- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ورزش دفاع شخصی.

همان طور که در شکل ۴-۲ مشاهده می‌شود، تصویر مورد جستجو دارای مفهوم ورزش دفاع شخصی است (تصویر سمت راست). بازیابی ۱۲ تصویر مشابه بنابر روش پیشنهادی مقاله [۱۹] در سمت چپ شکل آمده است. شماره تصاویر بازیابی شده، به ترتیب فاصله شباهت با تصویر مورد جستجو را نشان می‌دهند. در شکل ۴-۳ بازیابی همین مفهوم با استفاده از سیستم بازیابی تصویر پیشنهادی، مشاهده می‌شود.



شکل ۴-۳- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ورزش دفاع شخصی.

همان‌طور که در شکل ۳-۴ مشاهده می‌شود، سیستم بازیابی تصویر پیشنهادی با حداکثر درصد بازیابی، بخوبی تصاویر مشابه را بازیابی می‌کند. در شکل ۲-۴ سه تصویر بازیابی شده‌ی ۳، ۹ و ۱۲، بطور اشتباه بازیابی شده است. در صورتی که بازیابی همین تصویر با سیستم بازیابی پیشنهادی، همه ۱۲ تصویر، دارای معنای ورزش دفاع شخصی است. در شکل ۴-۴ و ۵-۴ تصویر مورد جستجو، دارای مفهوم مسابقه اتومبیل‌رانی است. بازیابی دو سیستم مذکور مقایسه می‌شوند.



شکل ۴-۴ - سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ورزش اتومبیل‌رانی.



شکل ۴-۵ - سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ورزش اتومبیل‌رانی.

همان‌طور که در دو شکل ۴-۴ و ۴-۵ مشاهده می‌شود، سیستم بازیابی تصویر مقاله [۱۹] دارای ۴ اشتباه است (تصاویر ۴، ۶، ۱۱ و ۱۲) در حالی که سیستم بازیابی پیشنهادی دارای فقط یک اشتباه (تصویر ۵) است. اگرچه این تصویر مربوط به مسابقه اتومبیل‌رانی نیست، ولی تصویر یک خودرو است. این بدین معناست که سیستم پیشنهادی با استفاده از سطوح بالای معنایی بخوبی مفاهیم را از محتوای تصاویر استخراج می‌کند. در ادامه در شکل‌های ۴-۶ و ۴-۷ نمونه‌ی دیگری از بازیابی این دو سیستم را مشاهده می‌کنیم.



شکل ۴-۶ - سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: ریزه‌سنگ‌های دریایی.



شکل ۴-۷ - سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: ریزه‌سنگ‌های دریایی.

همان طور که در شکل های ۴-۶ و ۴-۷ مشاهده می شود، سیستم بازیابی تصویر مقاله [۱۹] ۷ تصویر را بصورت نادرست بازیابی کرده است (تصاویر ۳، ۴، ۶، ۷، ۹، ۱۰ و ۱۲). در حالی که سیستم بازیابی پیشنهادی با حداکثر نرخ، همه تصاویر را بدرستی بازیابی کرده است. این نشان دهنده درک معنایی سیستم پیشنهادی ماست. یعنی ترکیب چهار مجموعه ویژگی و همچنین وزن دهی به آنها در روش پیشنهادی به عنوان یک مجموعه از سطوح بالای معنایی شناخته می شوند. در ادامه در شکل های ۴-۸ و ۴-۹ مقایسه دیگری برای درک بهتر این موضوع ارائه می شود.



شکل ۴-۸ - سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: خودرو.



شکل ۴-۹ - سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: خودرو.

همان‌طور که در شکل‌های ۴-۸ و ۴-۹ مشاهده می‌شود، سیستم بازیابی تصویر مقاله [۱۹] فقط ۴ تصویر را بصورت صحیح بازیابی کرده است (تصاویر ۵، ۶، ۹ و ۱۰). در حالی که سیستم بازیابی پیشنهادی با نرخ بهتری، ۹ تصویر را بدرستی بازیابی کرده است. اگر با دقت بیشتری به شکل ۴-۹ توجه شود، می‌توان به خاصیت ادراکی روش پیشنهادی نسبت به مفهوم تصویر مورد جستجو، پی برد. چرا که تصاویر بازیابی شده، مستقل از ویژگی‌های سطح پایین مثل لبه و رنگ هستند. این یعنی ترکیب چهار مجموعه ویژگی و همچنین وزن‌دهی به آن‌ها طوری معنای تصویر را استخراج می‌کنند که نمی‌توان از آن‌ها، به عنوان ویژگی‌های سطح پایین یاد کرد. مقایسه دیگری در شکل‌های ۴-۱۰ و ۴-۱۱ برای درک بهتر این موضوع ارائه می‌شود.



شکل ۴-۱۰- سیستم بازیابی تصویر مربوط به مقاله [۱۹]. مفهوم تصویر مورد جستجو: وسیله نقلیه قدیمی.



شکل ۴-۱۱- سیستم بازیابی تصویر پیشنهادی. مفهوم تصویر مورد جستجو: وسیله نقلیه قدیمی.

همان‌طور که در شکل‌های ۴-۱۰ و ۴-۱۱ مشاهده می‌شود، سیستم بازیابی تصویر مقاله [۱۹] فقط ۴ تصویر را بصورت صحیح بازیابی کرده است (تصاویر ۱، ۲، ۳ و ۷). در حالی که سیستم بازیابی پیشنهادی با نرخ بهتری، ۹ تصویر را بدرستی بازیابی کرده است. با دقت بیشتر به شکل ۴-۱۱ می‌توان مشاهده کرد که دو تصویر ۸ و ۹ نیز دارای مفهومی مشابه به وسیله نقلیه قدیمی هستند. بدین معنی که سیستم بازیابی پیشنهادی بخوبی مفاهیم (در این‌جا وسیله نقلیه) را درک می‌کند. بنابراین نمی‌توان از روش پیشنهادی مذکور، به عنوان ویژگی‌های سطح پایین یاد کرد.

۴-۷- جمع‌بندی

جهت بازیابی موثر تصویر مبتنی بر محتوا دو روش پیشنهاد شد. این دو روش در عین کم بودن تعداد ویژگی‌های بخوبی اطلاعات معنایی تصویر را استخراج و نرخ بازیابی را افزایش می‌دهند. یکی از دستاوردهای این پایان‌نامه بهبود روش هیستوگرام اختلاف رنگ است. علاوه بر بهبود روش این روش،

در روش پیشنهادی اول، ویژگی رنگی جدیدی پیشنهاد شد. بردار ویژگی روش پیشنهادی اول، از نظر زمان بازیابی، درصد نرخ بازیابی و تعداد ویژگی از جدیدترین روش بازیابی تصویر، بهتر عمل می‌کند.

روش پیشنهادی دوم نیز، با وجود تعداد ویژگی بیشتر نسبت به روش پیشنهادی اول، نسبت به روش جدید دیگری در حوزه بازیابی تصویر، هم تعداد ویژگی کم‌تر و هم درصد نرخ بازیابی بیشتری دارد. یکی دیگر از کارهایی که در این پایان‌نامه انجام شده است، وزن‌دهی به گروه‌های مختلف ویژگی است. این کار باعث بهبود در استخراج مفاهیم معنایی از تصویر شده و نرخ بازیابی را به شکل چشم‌گیری افزایش می‌دهد.

در بخش آخر این فصل، با چندین نمونه بازیابی، خروجی سیستم بازیابی پیشنهادی با یکی از بهترین سیستم‌های بازیابی تصویر اخیر مقایسه شد. در همه‌ی نمونه‌ها، سیستم پیشنهادی نرخ بازیابی تصویر بهتری دارد. همچنین کاملاً مشهود است که روش پیشنهادی درک درستی از مفاهیم داخل تصویر دارد. بدین معنی که مستقل از ویژگی‌های سطح پایین تصویر مورد جستجو مثل لبه و رنگ، است و ترکیب چهار مجموعه ویژگی پیشنهادی به همراه وزن‌دهی به آن‌ها، به عنوان سطوح بالای معنایی در نظر گرفته می‌شود. این سطوح بالای معنایی با وجود تعداد کم ویژگی، بخوبی محتوای تصویر را درک و مفاهیم آن را استخراج می‌کنند.

فصل پنجم:

نتیجه گیری و کارهای آینده

۵-۱- بحث و نتیجه‌گیری

در این پایان‌نامه به منظور کاهش شکاف معنایی بین ویژگی‌های سطح پایین و مفاهیم سطح بالای تصویر، سیستم بازیابی تصویر جدیدی معرفی شد. یکی از روش‌های کاهش شکاف معنایی استفاده از ویژگی‌های سطح پایین بصورت ترکیبی است. زیرا این روش بدون نیاز به یادگیری، شناسایی و قطعه‌بندی عمل بازیابی را انجام می‌دهد. یکی از این روش‌ها هیستوگرام اختلاف رنگ است. بنابر مطالعه‌های فیزیولوژیکی و روان‌شناسی، سیستم بینایی انسان، نسبت به رنگ و جهت لبه‌ها در تصویر بسیار حساس است. این دو ویژگی دارای اطلاعات معنایی زیادی از تصویر هستند. روش هیستوگرام اختلاف رنگ مبتنی بر این دو ویژگی است. پس از اعمال آنروپی و همبستگی روی ویژگی‌های این روش، با افزایش نرخ بازیابی، ۱۰۸ ویژگی این روش به ۴۴ ویژگی کاهش یافت. جهت افزایش ویژگی‌های لبه و رنگ، دو دسته ویژگی جدید به مجموعه ویژگی هیستوگرام اختلاف رنگ، اضافه شدند.

یکی از ویژگی‌های پیشنهادی، هیستوگرام رنگ در فضای رنگ HSV است. دو فضای رنگ HSV و $L^*a^*b^*$ ، نسبت به سایر فضاها رنگی به سیستم بینایی انسان نزدیک‌تر هستند. فضای رنگ $L^*a^*b^*$ در روش هیستوگرام اختلاف رنگ استفاده شده است. لذا در روش پیشنهادی اول، جهت استفاده از اطلاعات معنایی هر دو فضای رنگ، هیستوگرام رنگ در فضای رنگ HSV به بردار ویژگی اضافه شد. این بردار ویژگی از نظر تعداد ویژگی، سرعت و دقت بازیابی تصویر، نسبت به جدیدترین روش‌های بازیابی تصویر، بهبود قابل توجهی را نشان داد. روش پیشنهادی دوم به دنبال استفاده از اطلاعات معنایی هر دو نوع لبه، شامل لبه‌ی حاشیه‌ای اشیاء و لبه‌ی داخلی اشیاء است. لبه حاشیه‌ای اشیاء در روش هیستوگرام اختلاف رنگ استفاده شده است. بافت تصویر شامل لبه‌های داخلی اشیاء در تصویر است. لذا در روش پیشنهادی دوم، هیستوگرام جهت لبه با استفاده از گرادیان حاصل از بافت تصویر به مجموعه ویژگی روش پیشنهادی اول اضافه شد. در نهایت در معیار شباهت تصاویر، به هر گروه ویژگی از مجموعه‌های ویژگی پیشنهادی، وزنی اختصاص یافت. با انتساب این مقادیر وزنی،

سرعت و دقت در سیستم‌های بازیابی پیشنهادی، افزایش یافت. روش پیشنهادی اول نسبت به جدیدترین روش‌های حوزه بازیابی تصویر، از نظر زمان و درصد نرخ بازیابی، بهبود قابل توجهی دارد. روش پیشنهادی دوم نیز، با وجود تعداد ویژگی کم‌تر از روش‌های جدید قابل مقایسه، نرخ بازیابی را بهبود داده است. این افزایش نرخ حاکی از آن است که بردارهای ویژگی مذکور، بدون استفاده از قطعه‌بندی تصویر و فرآیندهای یادگیری و خوشه‌بندی، بخوبی فضای دو بعدی تصویر را توصیف می‌کنند.

ویژگی‌های معنایی استخراج‌شده به همراه وزن‌دهی آن‌ها، به عنوان سطوح بالای معنایی شناخته می‌شود. چراکه منجر به درک مفاهیم سطح بالای تصویر می‌شوند.

۵-۲- کارهای آینده

به عنوان پیشنهادات و کارهای آینده در این زمینه، می‌توان به موارد زیر اشاره کرد:

- استفاده از بازخورد ربط به عنوان واسط بین سیستم بازیابی و کاربر.
- استفاده از PCA جهت کاهش بردار ویژگی‌ها.
- استفاده از ویژگی‌های سطح پایین دیگر جهت استخراج برجستگی تصویر.
- محاسبه اختلاف رنگ پیکسل‌های همسایه در راستای بافت برابر.
- استفاده از روش‌های یادگیری ماشین در کنار ویژگی‌های سطح پایین، به منظور جداسازی بهتر کلاس‌های تصاویر.

- [1] Dharani, T., and I. Laurence Aroquiaraj. "A survey on content based image retrieval." Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on. IEEE, 2013.
- [2] Liu, Ying, et al. "A survey of content-based image retrieval with high-level semantics." *Pattern recognition* 40.1 (2007): 262-282.
- [3] Girija, O. K., and M. Sudheep Elayidom. "Overview of Image Retrieval Techniques." International Journal of Advanced Research in Computer and Communication Engineering 4 (2015) ISSN: 2278-1021.
- [4] Uchihashi, Shingo, and Takeo Kanade. "Content-free image retrieval by combinations of keywords and user feedbacks." *International Conference on Image and Video Retrieval*. Springer Berlin Heidelberg, 2005.
- [5] Mohamadi, Hanieh, Asadollah Shahbahrani, and Javad Akbari. "Image retrieval using the combination of text-based and content-based algorithms." *Journal of AI and Data Mining* 1.1 (2013): 27-34.
- [6] ستاری م، (۱۳۹۱)، پایان نامه ارشد، "بازیابی تصویر مبتنی بر محتوا با استفاده از معیارهای شباهت فازی"، دانشکده فنی و مهندسی، دانشگاه پیام نور.
- [7] Zarchi, Mohsen Sardari, Amirhasan Monadjemi, and Kamal Jamshidi. "A concept-based model for image retrieval systems." *Computers & Electrical Engineering* 46 (2015): 303-313.
- [8] Singha, Manimala, and K. Hemachandran. "Content based image retrieval using color and texture." *Signal & Image Processing* 3.1 (2012): 39.
- [9] Zarchi, Mohsen Sardari, Amirhasan Monadjemi, and Kamal Jamshidi. "A semantic model for general purpose content-based image retrieval systems." *Computers & Electrical Engineering* 40.7 (2014): 2062-2071.
- [10] Chatzichristofis, Savvas A., and Yiannis S. Boutalis. "Fcth: Fuzzy color and texture histogram—a low level feature for accurate image retrieval." *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*. IEEE, 2008.
- [11] Liu, Guang-Hai, Jing-Yu Yang, and ZuoYong Li. "Content-based image retrieval using computational visual attention model." *pattern recognition* 48.8 (2015): 2554-2566.

- [12] Liu, Guang-Hai, et al. "Image retrieval based on micro-structure descriptor." *Pattern Recognition* 44.9 (2011): 2123-2133.
- [13] Liu, Guang-Hai, et al. "Image retrieval based on multi-texton histogram." *Pattern Recognition* 43.7 (2010): 2380-2389.
- [14] Walia, Ekta, and Aman Pal. "Fusion framework for effective color image retrieval." *Journal of Visual Communication and Image Representation* 25.6 (2014): 1335-1348.
- [15] Singh, Chandan, and Kanwal Preet Kaur. "A fast and efficient image retrieval system based on color and texture features." *Journal of Visual Communication and Image Representation* 41 (2016): 225-238.
- [16] Neelima, Nizampatnam, and E. Sreenivasa Reddy. "An Efficient Multi Object Image Retrieval System Using Multiple Features and SVM." *SIRS*. 2015.
- [۱۷] رقیه تبریزی، ابوالفضل لاکدشتی، "ارائه روشی جدید برای بازیابی تصاویر مبتنی بر محتوا با استفاده از تلفیق ویژگی‌های رنگ و شکل"، اولین همایش ملی کاربرد سیستم‌های هوشمند، دانشگاه آزاد واحد قوچان، ۱۳۹۱.
- [۱۸] احمد ذوالفقاری، حسین خسروی، " روشی سریع در بازیابی تصاویر مبتنی بر محتوا با استفاده از ترکیب ویژگی لبه و رنگ"، کنفرانس مهندسی برق و توسعه پایدار- خاوران مشهد ۱۳۹۲.
- [19] Liu, Guang-Hai, and Jing-Yu Yang. "Content-based image retrieval using color difference histogram." *Pattern Recognition* 46.1 (2013): 188-198.
- [۲۰] سعید حیاتی، سعید سریزدی و حسین نظام آبادی پور، " جداسازی تصویر به مؤلفه‌های بافت و ساختار برای بازیابی تصویر بر اساس محتوا"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، ب- مهندسی کامپیوتر، سال ۱۳۹۱، شماره ۱۰.
- [21] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on pattern analysis and machine intelligence* 24.7 (2002): 971-987.
- [22] Xu, Yeong-Yuh. "Multiple-instance learning based decision neural networks for image retrieval and classification." *Neurocomputing* 171 (2016): 826-836.
- [23] Ashraf, Rehan, et al. "Content based image retrieval using embedded neural networks with bandletized regions." *Entropy* 17.6 (2015): 3552-3580.
- [24] Liu, Ying, Dengsheng Zhang, and Guojun Lu. "Region-based image retrieval with high-level semantics using decision tree learning." *Pattern Recognition* 41.8 (2008): 2554-2570.

- [25] Tong, Simon, and Edward Chang. "Support vector machine active learning for image retrieval." *Proceedings of the ninth ACM international conference on Multimedia*. ACM, 2001.
- [26] Carneiro, Gustavo, et al. "Supervised learning of semantic classes for image annotation and retrieval." *IEEE transactions on pattern analysis and machine intelligence* 29.3 (2007).
- [27] Chang, Ray-I., et al. "A novel content based image retrieval system using k-means/knn with feature extraction." *Computer Science and Information Systems* 9.4 (2012): 1645-1661.
- [28] Darwish, Saad M., and Raad A. Ali. "Observations on Using Type-2 Fuzzy Logic for Reducing Semantic Gap in Content-Based Image Retrieval System." *International Journal of Computer Theory and Engineering* 7.1 (2015): 1.
- [29] Shao, Ling, Fan Zhu, and Xuelong Li. "Transfer learning for visual categorization: A survey." *IEEE transactions on neural networks and learning systems* 26.5 (2015): 1019-1034.
- [30] Heller, Katherine A., and Zoubin Ghahramani. "A simple Bayesian framework for content-based image retrieval." *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. Vol. 2. IEEE, 2006.
- [31] Meng, Xin, et al. "Similar image retrieval only using one image." *Optik-International Journal for Light and Electron Optics* 127.1 (2016): 141-144.
- [32] Shelhamer, Evan, Jonathon Long, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *IEEE transactions on pattern analysis and machine intelligence* (2016).
- [33] Bose, Smarajit, et al. "A Hybrid Approach for Improved Content-based Image Retrieval using Segmentation." *arXiv preprint arXiv:1502.03215* (2015).
- [34] Jenni, Kommineni, Satria Mandala, and Mohd Shahrizal Sunar. "Content Based Image Retrieval using colour strings comparison." *Procedia Computer Science* 50 (2015): 374-379.
- [35] Kaur, Sukhdeep, and Deepak Aggarwal. "Image Content Based Retrieval System using Cosine Similarity for Skin Disease Images." *Advances in Computer Science: an International Journal* 2.4 (2013): 89-95.
- [36] Khosravi, Hossein, and Ehsanollah Kabir. "Farsi font recognition based on Sobel-Roberts features." *Pattern Recognition Letters* 31.1 (2010): 75-82.
- [37] <http://www.ci.gxnu.edu.cn/cbir/corel/2016-08-02>. (Accessed Aug15, 2016)

- [38] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, A. Yamada, "*Color and texture descriptors*", IEEE Transactions on Circuit and Systems for Video Technology 11 (6) (2001) 703–715.

Abstract: content based image retrieval(CBIR) is one of the well-known fields of study in computer vision and image processing. A CBIR system searches and retrieves images from an image data base using their visual content. Due to the rapid increase in collections of images with high volume and lack of accountability text-based image retrieval systems, content-based image retrieval systems are essential. Current CBIR systems, mostly use only low-level features(such as color, shape and texture) of images. Low-level features can describe the content of easy images correctly, but can not describe complex images, including high-level concepts. A high-level concepts are concepts of human perceptual in image that are not understood using low-level features. This problem, is one of the main challenges of CBIR systems and called "semantic gap".

In this thesis, two methods are presented to reduce the "semantic gap". These two methods are following semantic information of the image using combination of low-level features. According to studies, the human visual system considers edge and color features in image evaluation. So these two features are covered semantic information and have a significant role in content-based image retrieval systems. One of the methods in this field, often use edge features and color information in the $L^*a^*b^*$ color space. The feature vector include color differences based on these features, is called "histogram color difference". In this thesis, first, by applying two criterias, the entropy and the correlation between feature vector components in different images, effective features are selected. Although these features represent the semantic content of the image partly, by adding new features to the feature set, rate of image retrieval can be increased. One of the proposed features, is the color histogram in the HSV color space. The HSV color space and $L^*a^*b^*$, compared to other color spaces are closer to the human visual system. The $L^*a^*b^*$ color space is used in color difference histogram method. So, in the first proposed method, to use of semantic information in both color spaces, color histogram in the HSV color space is added to the feature vector. This feature vector in terms of features, speed and accuracy of image retrieval, compared to the newest methods, shows a significant improvement. The second proposed method follow semantic information using both edges, include the marginal edges of objects and internal edges of objects. The image texture including the internal edges of objects in the image. So, in the second proposed method, edge orientation histogram using gradient of image texture to add the feature set of first proposed method. Finally in criteria similar, to each group features of the proposed features, the weight is assigned. The assignment of these weights, is increased speed and accuracy in retrieval of proposed systems. The extracted semantic features with these appropriate weights, known as high levels semantic. Because leading to understand the high-level concepts of image. proposed methods improve rate of compared to recent other methods. In this thesis, the two standard databases Corel 5k and Corel 10k are used. This increases the accuracy and speed retrieval, demonstrate that proposed high levels of semantic, despite few features, without using of image segmentation and learning and clustering, describe two-dimensional image space well and semantic concepts are extracted.

Key words: content-based Image retrieval, color histogram, edge orientation histogram, entropy, correlation, high levels semantic, local binary patterns



Faculty of Computer Engineering
M.Sc. Thesis in Artificial Intelligence Engineering

Content-Based Image Retrieval Using High Level Semantics

By:
Hamed Qazanfari

Supervisor:
Prof H.Hassanpour

Advisor:
Dr A.Ahmadifard

Feb 2017