

بازیابی تصاویر رنگی با استفاده از دانش بافت و رنگ

در ساختار درختی دودویی

زهره منصوری ۸۵۷۰۲۷۹۴

استاد راهنما: دکتر منصور جمزاد

استاد ممتحن: دکتر حمیدرضا ربیعی

۱۴ بهمن ۱۳۸۶

چکیده

در سال های اخیر روش های بازیابی تصویر در زمینه های گوناگون به خصوص در زمینه بازیابی تصاویر موجود در اینترنت اهمیت بسزایی پیدا کرده است. در این پروژه، ارائه روشی برای بازیابی تصاویر رنگی مبتنی بر اطلاعات بافت و رنگ با استفاده از ایده های یادگیری ماشین در نظر است؛ که در آن ورودی سیستم تصویری به عنوان تصویر مرجع است. سیستم موظف است تا کلیه تصاویر موجود در بانک داده خود را که با تصویر مرجع تا حد مناسبی مشابهت دارند به عنوان خروجی به کاربر تحویل دهد.

کلمات کلیدی

بازیابی تصویر، رنگ، بافت، اندیس گذاری

۱- مقدمه

غلبه بر مشکلات سیستم های مبتنی بر متن، توسعه سیستم های بازیابی تصویر مبتنی بر محتوا^۲ پیشنهاد شد که وظیفه استخراج تصاویر به صورت اتوماتیک و با استفاده از مفاهیم دیداری نظیر رنگ، بافت و طرح بندی^۳ تصویر را بر عهده داشت [۱، ۲]. سیستم های اولیه از این رده، برای بازیابی تصویر از کاربر درخواست می کرد که یک یا چند ویژگی دیداری را انتخاب و محدوده ای برای مقدار آن ها تعیین کند، سپس عمل بازیابی بر اساس این اطلاعات صورت می گرفت. در سیستم های پیشرفته تر، کاربر قادر بود تا درجه اهمیت ویژگی های انتخاب شده را تعیین کند. به دلیل ناتوانی در جلب رضایت کاربر، سیستم های آتی دریافت تصویر به عنوان الگوی بازیابی را در سیستم خود امکان پذیر ساختند. به این شیوه که کاربر در کنار قابلیت های ذکر شده، می توانست تصویر یا شمای تصویر مطلوب خود را نیز به سیستم وارد کند تا سیستم، تصاویر مشابه با آن تصویر را از پایگاه داده خود استخراج نماید. در سیستم های پیشرفته کاربر قادر بود تا میزان اهمیت هر یک از ویژگی ها را خود تعیین نماید، همچنین قابلیت بهبود نتایج جستجو بوسیله تعامل با کاربر و اعمال نظر وی در جستجوهای بعدی نیز به این سیستم ها اضافه شد [۳، ۴].

در این گزارش سعی بر آن است تا راه کارهایی که تاکنون برای توسعه سیستم های بازیابی تصاویر انجام شده اند معرفی گردد، سپس به

در سال های اخیر، به دلیل گسترش نیاز به بازیابی بهینه تصاویر در پایگاه های داده حجیم، تحقیقات گسترده ای در زمینه بازیابی تصاویر صورت گرفته است که منجر به توسعه سیستم های گوناگون دسته بندی تصویر شده است. اولین ایده ها در اوایل ده ۷۰ میلادی مطرح شد که در آنها دسته بندی تصاویر بدون توجه به ویژگی های دیداری آنها و تنها بر اساس حاشیه نویسی^۱ متنی انجام می گرفت؛ به صورتی که مفاهیم موجود در تصویر توسط اپراتور تشخیص داده شده و در پایگاه داده ای به عنوان کلمات کلیدی آن تصویر ذخیره می شد. به این ترتیب کاربران با استفاده از کلمات کلیدی مورد نظر خود به تصاویر مربوطه دسترسی داشتند. به این روش بازیابی تصاویر مبتنی بر متن گفته می شود.

سیستم های فوق الذکر با چند مشکل عمده مواجه بودند. نخست اینکه حاشیه نویسی تصاویر مستلزم وقت و هزینه بسیار و نیز به میزان زیادی به درک اپراتور از تصویر وابسته بود. دوم، از آنجا که مفاهیم موجود در یک تصویر از دید کاربران متفاوت یکسان نیست بنابراین حاشیه های الصاق شده به تصاویر تمامی حیطه پرس و جو را نمی پوشانند؛ و این به این معنا است که پرس و جوهای مبتنی بر متن به میزان کافی کامل و گویا نیستند. در ابتدای دهه ۹۰، با افزایش حجم تصاویر موجود در پایگاه های داده نظیر شبکه جهانی اینترنت و برای

محدود در ویژگی‌های دیداری تصاویر (مانند پایگاه داده تصاویر مربوط به پزشکی)؛ ۲. دامنه وسیع: شامل تغییرات وسیع و غیرقابل پیش‌بینی در تصاویر (مانند پایگاه داده تصاویر موجود در اینترنت).

نمونه‌ای از سیستم‌های بازیابی تصویر که به تازگی توسعه داده شده‌اند می‌توان VisualSEEK [۶]، EFF^2 [۷] و PicToSeek [۸] و Viper [۹] را نام برد که در این مطالعه به برخی از اشاره شده است. همچنین برخی از ویژگی‌های QBIC [۱۰] که توسط شرکت IBM توسعه داده شده است نیز توضیح داده می‌شود. این سیستم جزو مهم‌ترین سیستم‌های بازیابی به شمار می‌رود، چرا که سایر سیستم‌های بازیابی از آن به عنوان مدل اولیه توسعه استفاده کرده‌اند.

۳- استخراج ویژگی‌های تصویر

پیش از شروع کار بازیابی، ویژگی‌های تصاویر بانک داده باید از آن‌ها استخراج شوند. یک تصویر آرایه‌ای دو بعدی از پیکسل‌ها است. برای سهولت دسترسی به ویژگی‌های آن تصویر، به جای استفاده از مقادیر پیکسل‌ها، با استفاده از روش‌هایی خاص ویژگی‌های تصویر را استخراج و در قالبی نمایش می‌دهند که قابل مقایسه باشد. به این روش‌ها «استخراج ویژگی» می‌گویند. خروجی این روش‌ها عدد یا برداری است که یکی از خصوصیات تصویر را معرفی می‌کند، که به آنها «کلاس ویژگی» یا «توصیف‌گر» می‌گویند (مانند هیستوگرام رنگ). به برداری که از کنار هم قراردادن این کلاس‌های ویژگی به دست می‌آید، «برداری ویژگی» تصویر گفته می‌شود [۳، ۶، ۷].

ویژگی‌های تصویر را می‌توان از لحاظ نحوه توصیف تصویر به سه دسته تقسیم کرد: ۱. ویژگی‌های عمومی یا سطح پایین؛ ۲. ویژگی‌های معنایی؛ ۳. ویژگی‌های مختص کاربردهای خاص. ویژگی‌های سطح پایین معرف کلیات تصویرند و توانایی توصیف اشیا یا مفاهیم موجود در تصویر را ندارند. رنگ و بافت دو نوع از این ویژگی‌ها است. ویژگی‌های معنایی به بیان مفاهیم و اشیای موجود در تصویر می‌پردازند و غالباً از روش‌های قسمت بندی^۱ تصویر برای اینکار استفاده می‌کنند. دسته سوم مربوط به ویژگی‌های مربوط به کاربردهای خاص است، نظیر ویژگی‌های توصیفی چهره در سیستم‌های بازیابی تصاویر مربوط به چهره [۸].

به دلیل تمرکز این پروژه بر ویژگی‌های سطح پایین، در بخش‌های بعد به معرفی برخی از مهم‌ترین و پرکاربردترین توصیف‌گرهای رنگ و بافت پرداخته می‌شود.

۳-۱- استخراج رنگ

رنگ یکی از متداول‌ترین و تعیین‌کننده‌ترین ویژگی دیداری در زمینه بازیابی تصویر است؛ چرا که نسبت به تغییرات مربوط به اندازه، جهت و دورنمایی^۱ و اغتشاش تصویر پایدار است [۶].

معرفی و توضیح ایده پیشنهادی برای توسعه یک سیستم بازیابی تصویر بر اساس دو ویژگی رنگ و بافت پرداخته می‌شود.

۲- خصوصیات سیستم‌های بازیابی تصویر

همانطور که ذکر شد سیستم‌های بازیابی تصویر از ویژگی‌های تصویری برای دسته‌بندی و بازیابی تصاویر استفاده می‌کنند. فلوچارت مراحل و اجزای این سیستم‌ها در شکل (۱) (اقتباس شده از [۴، ۵]) آمده است. در این سیستم‌ها پایگاه داده تصویری وجود دارد که شامل کلیه تصاویر قابل بازیابی برای کاربر است. کاربر برای استخراج تصاویر از این بانک، خصوصیات مورد نظر خود را به سیستم ارائه می‌دهد، و سیستم موظف است تا با استفاده از این اطلاعات، بانک تصاویر خود را جستجو کند تا تصاویر مطلوب را استخراج و به کاربر نمایش دهد. برای تساهل و تسریع در امر جستجو، بانک تصویری باید به صورتی اندیس‌گذاری شود که زمان جستجو حداقل شده و حتی‌المکان تصاویر مشابه به سریع‌ترین وجه قابل بازیابی باشند. بهترین راه این است که تصاویر بر اساس ویژگی‌های خود به گونه‌ای کلاسه‌بندی شوند.

سیستم‌های یابی نوعاً از سه بخش تشکیل می‌شوند:

۱. بخش استخراج ویژگی که اطلاعات دیداری تصویر را برای دسته‌بندی آنها استخراج می‌کند؛
۲. بخش اندیس‌گذاری که موظف به دسته‌بندی تصاویر بر اساس ویژگی‌های استخراج شده است؛
۳. بخش بازیابی که پرس‌وجو کاربر را پردازش و واسط کاربری را فراهم می‌کند؛

Gevers در [۴] روش‌های بازیابی تصاویر را بر اساس کاربردهای مختلف در سه دسته طبقه‌بندی کرده است:

۱. جستجو بر اساس شباهت تصاویر^۲ که هدف این روش به دست آوردن تصاویر موردنظر کاربر از بین تصاویر موجود در یک گالری است که در مراحل پی‌درپی و با استفاده از بازخورد کاربر صورت می‌گیرد.
۲. جستجو به هدف یافتن تصویر خاص^۳: هدف این روش به دست آوردن تصاویر مشابه یک تصویر از پایگاه داده تصاویر است. منظور از تصویر مشابه، تصویری است که (بخشی از آن) با تصویر مورد نظر کاربر یکسان باشد و یا (بخشی از) یک شی در هر دو تصویر یکسان وجود داشته باشد.
۳. جستجو بر اساس دسته بندی^۴: هدف از این روش به دست آوردن تصویری است که متعلق به دسته یا کلاس بخصوصی باشد، مانند تصاویر دکوراسیون منزل و یا تصاویر رادیولوژی.

وی همچنین پایگاه داده تصاویر بر اساس چگونگی میزان تغییر در ویژگی‌ها در دو دسته قرار داده است: ۱. دامنه محدود: شامل تغییرات

رنگ و فضاهای رنگی

IQ معرف اطلاعات رنگ است. فضاهای رنگی CIE LAB و CIE LUV نیز دارای خاصیت یکنواختی ادراکی است، اما مشکل آنها دشواری تغییر رنگ از سیستم RGB است [۶].

از آنجا که هر یک از این فضاهای رنگ برای استخراج برخی ویژگی‌ها کاراتر از سایر فضاهاست، بنابراین ممکن است که در یک سیستم بازیابی تصویر از چندین فضای رنگ متفاوت استفاده شود. در بسیاری از روش‌ها جهت سهولت کار، تعداد رنگ‌های مورد استفاده از فضای رنگ به تعداد محدودی کاهش داده می‌شود، که به این روش کوانتیزه کردن فضای رنگی [۶] و به فضای جدید، پالت رنگ گفته می‌شود. برای مثال سیستم Viper [۹] از پالت رنگ ۱۶۶ تایی استفاده کرده که این کار را با کوانتیزه کردن فضای رنگ HSV به ۱۸ سطح در بعد H، ۳ سطح در S و ۳ سطح در V انجام داده است.

ممان رنگ

این روش یکی از موفق‌ترین روش‌های استخراج رنگ در سیستم‌های بازیابی است. ممان‌های مرتبه اول، دوم و سوم به ترتیب به نام‌های میانگین، واریانس و درجه اریبی رنگ [۱۰]، توزیع رنگ تصویر را به صورت کارآمد و بهینه ارائه می‌دهند. ممان‌های اول تا سوم تصویر به ترتیب بر اساس رابطه‌های (۱) تا (۳) به دست می‌آیند:

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

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

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

که f_{ij} مقدار i امین مولفه رنگ i زمین پیکسل و N تعداد پیکسل‌های تصویر است. ممان‌های رنگ غالباً در سیستم‌های HSV و $L^*a^*b^*$ محاسبه می‌شوند که $L^*a^*b^*$ ارجحیت دارد. استفاده از ممان سوم رنگ در کنار دو ممان اول، بازده کلی بازیابی را افزایش می‌دهد، اما این ممان به تغییر زاویه دید حساس است و ممکن است خود دلیلی برای کاهش بازده شود.

تعداد مولفه‌های به دست آمده از این ممان‌ها ۹ تا است (سه ممان برای ۳ مولفه رنگی) بنابراین معرف فشرده‌ای برای تصویر است، ولی قدرت جداسازی کمی دارد. غالباً از ممان رنگ برای فیلتر کردن تصاویر در مراحل اولیه جستجو استفاده می‌شود [۱]. در [۶] ویژگی‌های دیگری نظیر انرژی و آنتروپی آمده است. QBIC از این سه ممان رنگ در سه فضای RGB، YIQ و CIE Lab استفاده کرده است [۱۱].

از دید شخص، رنگ از سه مولفه تشکیل شده است: قرمز، سبز و آبی؛ که این سه یک فضای رنگ^{۱۰} را می‌سازند. بر طبق تعریف یک فضای رنگ، فضایی چندبعدی است که هر بعد آن معرف یکی از اجزای تشکیل‌دهنده رنگ می‌باشد. برای استخراج ویژگی رنگ از تصویر، ابتدا باید نوع فضای رنگ را تعیین نمود. فضای رنگی تعیین شده باید خصوصیات زیر را ارضا کند: ۱. مستقل از سخت‌افزار باشد. ۲. دارای یکنواختی ادراکی^{۱۱} باشد به این معنا که فاصله عددی بین رنگ‌ها در آن فضای رنگی متناسب با تمایزات ادراکی انسان از آن رنگ‌ها باشد. ۳. توسط کاربر قابل تصور و درک باشد [۲].

فضای رنگ RGB به دلیل سادگی استفاده و درک و نیز پشتیبانی سخت‌افزاری به‌طور گسترده مورد استفاده قرار می‌گیرد، اما در مسائل تکنیکی مربوط به سیستم‌های پردازش تصویر دارای کاستی‌هایی است، برای مثال بین سه مولفه آن همبستگی وجود دارد و در صورت تغییر فاکتور روشنایی تصویر، هر سه مولفه RGB تغییر می‌کنند. بنابراین این فضا اغلب در سیستم‌های بازیابی تصویر مورد استفاده قرار نمی‌گیرد. راه‌حل رفع این همبستگی آن است که فضای RGB نرمالسازی شود. در این فضای رنگ جدید که به «فضای رنگ متقابل»^{۱۲} موسوم است، اطلاعات روشنایی از سایر مولفه‌ها جدا شده و بنابراین این فضا در سیستم‌های بازیابی تصویر قابل استفاده است [۴]. مولفه‌های جدید برابرند با R-G و 2B-R-G که نماینده رنگ هستند و R+G+B معرف روشنایی است [۱].

فضای رنگ $L^*a^*b^*$ دارای خاصیت یکنواختی ادراکی است. همچنین در این فضا هر چه فاصله اقلیدسی دو نقطه رنگی کمتر باشد، خواص دیداری آن‌ها نیز شبیه‌تر است. مولفه‌های این سیستم تشکیل شده از L^* که معرف روشنایی است و بین سفید تا سیاه متغیر است، a^* که معرف قرمز-سبز است و b^* که معادل سبز-آبی است. a^* و b^* مستقل از روشنایی هستند و بنابراین دارای خواص فضاهای نرمالسازی شده هستند. تنها مشکل این فضا این است که تبدیلات سایر فضاهای رنگی به این فضا غیرخطی و مستلزم محاسبات بالاست.

فضای HSI^{۱۳} نیز یک فضای ادراکی است، به همین دلیل برای استفاده در سیستم‌های بازیابی مناسب است [۶]. این فضا از سه مولفه I معرف میزان نور، S درجه سفیدی یا اصطلاحاً غلظت رنگ و H معرف طول موج (مانند قرمز، سبز یا آبی) رنگ است. H مستقل از تغییرات جهت تصویر است چرا که نسبت به تغییرات نور و جهت تصویربرداری پایدار و بنابراین برای سیستم‌های بازیابی بسیار مناسب است. با این حال H به تغییرات روشنایی به شدت حساس است [۴].

دو فضای رنگی YIQ و YUV برای نمایش اطلاعات رنگی در سیگنال‌های تلویزیونی استفاده می‌شوند. Y معرف روشنایی و UV یا

هیستوگرام رنگ

پیکسل‌های مرتبط در مولفه i ام هیستوگرام (مربوط به رنگ i ام) و β_i تعداد پیکسل‌های نامرتب آن مولفه باشد، در اینصورت CCV مربوط به آن تصویر با بردار $\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_N, \beta_N) \rangle$ معرفی می‌شود. قابل توجه است که بردار $\langle \alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_N + \beta_N \rangle$ همان هیستوگرام تصویر است.

بر اساس نتایج گذشته، CCV بازده بهتری از هیستوگرام دارد. برای هر دو هیستوگرام و بردار ارتباط رنگ، سیستم HSV بهتر از $L^*a^*b^*$ عمل می‌کند [۱].

همبستگی نگار رنگ^{۲۰}

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

این بردار در مقابل هیستوگرام و CCV بازده بهتری برای سیستم بازبازی فراهم می‌کند، اما به علت بعد زیاد، بار محاسباتی بالایی دارد [۲].

۳-۲- استخراج ویژگی بافت

بافت یکی دیگر از ویژگی‌های مهم تصویر است و از منظر دیداری الگوی یکنواختی است که از حضور بیش از یک رنگ یا درجه خاکستری به دست می‌آید [۱]. روش‌های استخراج بافت به چهار دسته اصلی تقسیم می‌شود: ۱. روش‌های ساختاری؛ ۲. روش‌های آماری؛ ۳. روش‌های مبتنی بر مدل؛ ۴. روش‌های تبدیلی.

روش‌های ساختاری شامل اپراتور مورفولوژیک^{۲۱} و گراف همجواری^{۲۲} است که بافت را بر اساس ویژگی‌های ساختاری و قواعد مربوط به نحوه قرار گرفتن پیکسل‌ها و اصطلاحاً بر اساس ریزبافت‌ها تعریف می‌کنند. این روش‌ها برای بافت‌های معمولی مناسب است و بیشتر برای سنتز بافت استفاده می‌شود تا تحلیل آن، و با وجود اینکه تعریف دقیقی از بافت ارائه می‌دهند، از بازده مناسبی برخوردار نیستند و محاسبات بالایی می‌طلبند. امروزه از روش‌های ساختاری به ندرت استفاده می‌شود [۱، ۱۰].

روش‌های آماری بر اساس هیستوگرام و یا محاسبات مرتبه دوم آماری بر روی درجه خاکستری پیکسل‌های مجاور اعمال می‌شوند و از جمله اولین روش‌های استخراج بافت هستند. Li در [۷] روشی از نوع

هیستوگرام معرف توزیع تجمعی^{۱۵} نقاط در سه کانال رنگ است، به صورتی که هر مولفه هیستوگرام برابر با مجموع تعداد پیکسل‌هایی است که متعلق به یک رنگ یا یک مجموعه رنگ هستند. این روش به دلیل بار محاسباتی پایین در سیستم‌های بازیای تصویر به صورت گسترده استفاده می‌شود، همچنین نسبت به تغییر اندازه کم، تغییر زاویه، دورنمایی و اغتشاش تصویر پایدارتر است [۲، ۶].

به طور کلی هرچه تعداد مولفه‌های هیستوگرام بیشتر باشد قدرت جداسازی آن بیشتر است، اما باعث افزایش بار محاسباتی است و اندیس‌گذاری تصویر را دشوارتر می‌کند. همچنین تعداد زیاد مولفه‌های هیستوگرام لزوماً کارایی بازیایی را افزایش نمی‌دهد. مشکل دیگری که در پایگاه‌های داده تصویر بسیار بزرگ رخ می‌دهد، مساله اشباع شدن هیستوگرام^{۱۶} است، به این معنا که تصاویر بسیار متفاوت ممکن است هیستوگرام مشابهی داشته باشند. در [۱، ۲] راه‌هایی برای حل مشکل اشباع شدن و بهبود کارایی هیستوگرام آمده است.

Gevers به همراه سایرین در [۱۲] راه‌کاری برای بهبود هیستوگرام مطرح کرده است که عمدتاً بر محور حذف نویز از رنگ‌های هیستوگرام استوار است. همچنین QBIC از ۲۵۶ مولفه رنگ برای هیستوگرام تصویر در سیستم RGB استفاده کرده است [۱۱].

هیستوگرام برچسب رنگ^{۱۷}

در برخی از آن سیستم‌ها برای سهولت محاسبه و استفاده از هیستوگرام و بهبود قدرت تفکیک‌دهی، رنگ‌های موجود در فضا یا پالت رنگ بر اساس شباهت دیداری به دسته‌هایی تقسیم می‌شوند که هر دسته برچسب خاصی دارد. سپس هر پیکسل برچسب‌گذاری شده و هیستوگرام این برچسب‌های رنگ محاسبه می‌شود. هیستوگرام حاصله به جای ۳ بعد، از یک بعد تشکیل شده است.

Li به همراه سایرین در [۷] برای سیستم خود، فضای رنگ HSV را به سیزده دسته تقسیم کرده و هیستوگرام برچسب رنگ را با سیزده مولفه به دست آورده است. این سیستم از نظر بازده و کارایی بهتر از هیستوگرام معمولی است. همچنین ابعاد کم هیستوگرام به دست آمده در حل «مشکل ابعاد بالا»^{۱۸} حائز اهمیت است.

بردار ارتباط رنگ^{۱۹}

این ویژگی اطلاعات مکانی پیکسل‌ها را در هیستوگرام درج می‌کند. در این روش که به طور مخفف به آن CCV می‌گویند، هر مولفه هیستوگرام از دو نوع داده تشکیل می‌شود، بخش مرتبط که شامل تعداد پیکسل‌هایی است که به یک منطقه یکنواخت رنگ متعلق‌اند، و بخش نامرتب که پیکسل‌های گسسته را شامل می‌شود. اگر α_i معرف

آماري ارائه کرده که برای استخراج بافت، از هیستوگرام تصویر استفاده شده است. در سایر روش‌های آماری غالباً از ماتریس Co-occurrence برای استخراج ویژگی‌های بافت استفاده می‌شود.

روش‌های مبتنی بر مدل، به مدل‌سازی بافت می‌پردازند و شامل روش خودبازگشت^{۲۳} یا AR، مدل گوسی مارکوف یا RMF و مدل Gibbs RMF می‌شوند.

روش‌های تبدیلی شامل ویژگی‌های Tamura، تجزیه Wold^{۲۴}، تبدیل موجک^{۲۵} و غیره است که مبتنی بر اعمال فیلتر بر روی تصویر می‌باشند. ذیلاً توضیح برخی از مهمترین این روش‌ها آمده است.

ماتریس Co-occurrence

این ماتریس که به هیستوگرام مرتبه دوم [۱۰] یا GLCM^{۲۶} [۱۳] نیز معروف است، به احتمال تجمعی توزیع جفت پیکسل‌ها می‌پردازد و روش بسیار موفقی در تعریف بافت است. GLCM معرف فرکانس حضور هر دو پیکسل در فاصله‌ای معین در تصویر است، که فاصله آنها به صورت برداری از پیش تعیین شده در نظر گرفته می‌شود. پس از محاسبه این ماتریس خصوصیات نظیر انرژی، آنروپی، تباين و همگنی از آن استخراج می‌شود که می‌توانند معرفی برای بافت موجود در تصویر باشند [۱۳]. Arvis به همراه سایرین در [۱۴] روش دیگری برای استخراج بافت از این ماتریس ارائه کرده است.

ویژگی‌های Tamura

این ویژگی و ویژگی‌هایی از بافت را نظیر درشتی^{۲۷}، تباين^{۲۸}، جهت‌دار بودن^{۲۹}، خطی بودن^{۳۰}، نظم^{۳۱} و زبری^{۳۲} تعریف می‌کند و بر اساس مطالعات روانشناسی بر روی ادراک انسان از بافت طراحی شده است. هر یک از این ویژگی‌ها توسط رابطه‌ای جداگانه به دست می‌آید. سه ویژگی اول در سیستم‌های بازیابی نتایج بسیار موفقی داشته و در برخی از سیستم‌های اولیه مانند QBIC استفاده شده‌اند [۱، ۱۱، ۱۳].

ویژگی‌های فیلتر گابور

استفاده از فیلتر گابور یکی از رایج‌ترین روش‌های مبتنی بر فیلتر برای استخراج بافت می‌باشد. فیلتر گابور می‌تواند برای مدل کردن واکنش سیستم‌های بصری انسان استفاده شود.

این فیلتر در هر دو حوزه مکان و فرکانس عمل می‌کند. در حوزه مکان، هسته‌های فیلتر گابور از حاصلضرب یک تابع گوسین با یک تابع سینوسی جهت‌دار به دست می‌آید. در نتیجه این فیلتر در نقاطی از تصویر که به صورت محلی دارای جهت و فرکانس مکانی معینی هستند پاسخ‌های قوی تولید می‌کند [۱]. برای اعمال این فیلتر بهتر است از فضای رنگ HSV استفاده شود [۱۵].

استفاده از فیلتر گابور کارایی بازیابی را افزایش می‌دهد اما دارای دو مشکل عمده است: ۱. توابع پایه گابور متعامد نیست که باعث افزایش حافظه مصرفی می‌شود؛ ۲. زمان موردنیاز برای استخراج این ویژگی‌ها زیاد است که باعث کاهش سرعت بازیابی می‌شود [۱]. Andrysiak در [۱۶] روشی برای استخراج بردار ویژگی بافت از فیلتر گابور آورده است. به گونه‌ای که بردار ویژگی حاصل می‌نیمم طول ممکن را دارا باشد.

ویژگی‌های تبدیل موجک

تبدیل موجک گسسته یک بعدی یا DWT^{۳۳}، یک سیگنال گسسته $f(x)$ را بر حسب مقادیر تابع مقیاس $\varphi(x)$ و تابع شیفت یافته و گسترش یافته $\psi(x)$ تجزیه می‌کند. رابطه (۴) معرف این تبدیل است.

$$f(x) = \sum_{l \in \mathbb{Z}} S_{j_0, l} \varphi_{j_0, l}(x) + \sum_{j \geq j_0, l} C_{j, l} \psi_{j, l}(x) \quad (4)$$

که در آن $\varphi_{j_0, l}$ و $\psi_{j, l}$ و به ترتیب توسط روابط (۵) و (۶) به دست می‌آیند.

$$\varphi_{j, l}(x) = 2^{\frac{j_0}{2}} \varphi(2^{j_0} x - l) \quad (5)$$

$$\psi_{j_0, l}(x) = 2^{\frac{j_0}{2}} \psi(2^{j_0} x - l) \quad (6)$$

تابع $\psi(x)$ موجک مادر یا به اختصار موجک نامیده می‌شود. تبدیل موجک یک تصویر (سیگنال دو بعدی)، با پیاده‌سازی تبدیل موجک یک‌بعدی در جهت‌های افقی و عمودی به دست می‌آید [۱].

دلیل کارایی موجک در آنالیز سیستم‌ها مسائلی از قبیل محمل فشرده^{۳۴} و کوتاه، تعامد، تقارن، گشتاورهای صفر و غیره است؛ اما تمامی خصوصیات فوق کاملاً در موجک‌های کلاسیک برآورده نمی‌شوند. به همین منظور موجک‌های چنددقتی^{۳۵} و تبدیل موجک مختلط که حاصل گسترش موجک کلاسیک هستند ارائه شده‌اند. ایده موجک چنددقتی حالت کلی‌تر موجک‌های کلاسیک می‌باشد، به این صورت که به جای استفاده از یک تابع مقیاس خاص از چندین تابع مقیاس، موسوم به توابع مقیاس چنددقتی استفاده می‌شود. در ساخت موجک‌های چندگانه، برای برآورده شدن همزمان خواص مطلوب موثر در پردازش سیگنال درجه آزادی بیشتری وجود دارد.

در تبدیل موجک مختلط، فیلترها ضرایب مختلط دارند و نمونه‌ها خروجی مختلط تولید می‌کنند. DWT حقیقی از نظر انتخاب جهت ضعیف و نیز حساس به شیفت است. این دو مساله توسط تبدیل موجک مختلط قابل حل می‌باشند.

انواع توابع تبدیل موجک عبارتند از *Mexican Morlet*, *Coiflet*, *Haar Hat* و *Daubechies* که از این میان، فیلتر *Haar* به دلیل سهولت در سیستم‌های بازیابی بیشتر استفاده می‌شود [۱۷].

۳-۳- بلوک‌بندی تصویر

بلوک‌بندی به این معنا است که به جای استفاده از اطلاعات پیکسل-های رنگ، از اطلاعات دسته‌ای از آن‌ها با هم استفاده کرد و به این صورت برای تک‌تک پیکسل‌های موجود در آن مجموعه تصمیم‌گیری نمود. این شیوه برای به دست آوردن برخی ویژگی‌های مبتنی بر بافت لازم است، برای مثال *Andrysiac* و سایرین در [۱۸] برای اعمال فیلتر گابور تصویر را به بلوک‌های مساوی تقسیم کرده‌اند. اما استفاده از آن در سایر ویژگی‌ها مخصوصاً ویژگی‌های مبتنی بر رنگ باعث می‌شود که برخی از اطلاعات تصویر از دست برود. اما این شیوه از مزایایی برخوردار است، برای مثال:

۱. استفاده از این شیوه در پردازش و بازیابی تصاویر باعث می‌شود تا دید وسیع‌تر و کلی‌تری نسبت به تصاویر ایجاد شود.

۲. به دلیل کاهش درجه تفکیک‌پذیری^{۳۶}، حجم تصویر پردازش شده کاهش می‌یابد که این امر باعث می‌شود: الف) پردازش تصاویر موجود در بانک تصاویر و همچنین تصویر پرس‌وجو^{۳۷} ورودی با سرعت بیشتر انجام شود. ب) در اندیس‌گذاری و ذخیره ویژگی‌های بانک تصاویر، فضای حافظه کمتری مصرف شود.

۴- جستجو

برای بازیابی تصویر باید از مطلوبات کاربر مطلع شد. کاربر می‌تواند ویژگی‌ها و درجه اهمیت آنها را برای تصاویر مطلوب خود به سیستم وارد کند، و یا اینکه با استفاده از تصویری موسوم به تصویر پرس‌وجو یا خوراک^{۳۸} [۱۷] مشخصات تصاویر مطلوب خود را اعلام کند. از روش دوم در اکثر سیستم‌های بازیابی تصویر نظیر *PicToSeek* [۱۹]، *EFF*^۲ [۶]، *QBIC* [۱۱]، *VisualSEEK* [۲۰] و *Viper* [۹] استفاده شده است.

در غالب سیستم‌ها برای جستجوی تصاویر و پیدا کردن تصاویر مشابه از جستجوی نزدیک‌ترین همسایه^{۳۹} استفاده می‌شود، که بنابه تعریف *Chiueh* در [۱۷]، این روش جستجو به دنبال نقطه‌ای از فضای ویژگی است که فاصله آن با بردار ویژگی تصویر پرس‌وجو کمینه باشد. روش‌های متنوعی برای به دست آوردن این مجموعه تصاویر عنوان شده است، که اکثر آنها از شیوه‌های اندیس‌گذاری تصاویر و قسمت‌بندی فضای ویژگی استفاده می‌کند. بدین ترتیب k نزدیکترین تصویر^{۴۰} به تصویر مطلوب کاربر استخراج می‌شود، که k می‌تواند عددی از پیش تعیین شده باشد و یا کاربر آن را مقداردهی کند. در *VisualSEEK*

[۱۷] به کاربر این امکان داده می‌شود تا سایر تصاویر استخراج شده که از نظر سیستم از این فاصله همسایگی دورتر است را نیز مشاهده کند^{۴۱}.

Einarsson در سیستم بازیابی خود به نام *Eff*^۲ [۶] برای پیدا کردن این تصاویر از جستجوی پشت‌سرهم^{۴۲} استفاده کرده و آن را به عنوان بهترین روش بازیابی عنوان نموده است. به طوری که بردار ویژگی تصویر پرس‌وجو با تمام تصاویر موجود در پایگاه داده مقایسه شوند، ولی مشکل آن را بار محاسباتی بالا عنوان کرده است.

۴-۱- اندیس‌گذاری

همانطور که ذکر شد، در سیستم‌های بازیابی، هر تصویر با برداری موسوم به بردار ویژگی شناخته می‌شود. مولفه‌های این بردار می‌تواند یک یا چندبعدی باشد. برای مثال، ممان رنگ یک عدد و هیستوگرام یک بردار با طول معین است. برای استفاده از بردارهای ویژگی، این بردارها در ابتدای کار سیستم به صورت برون‌خطی^{۴۳} برای همه تصاویر پایگاه داده محاسبه و در پایگاه داده دیگری به نام پایگاه داده ویژگی [۲] ذخیره خواهند شد.

از دید مدیریت داده‌ای بهتر است بجای اینکه تمام ویژگی‌ها در کنار هم و تداخل یافته ذخیره شوند، هر توصیف‌گر یا کلاس ویژگی در بردار ویژگی‌ها به صورت مجزا جایگذاری شود، چرا که ۱. با جدا کردن آنها امکان دسترسی مجزا به هر ویژگی فراهم می‌آید؛ ۲. کنار هم قرار دادن این بردارها که بعضاً دارای ابعاد زیاد هستند، باعث بروز مشکل ابعاد بالا می‌شود [۲۱].

هدف از اندیس‌گذاری تصاویر، مرتب کردن بردارهای ویژگی و در نتیجه تصاویر ماناظر با آنها است، به گونه‌ای که زمان لازم برای پیدا کردن تصاویر مشابه یک تصویر در این فضای گسترده کمینه و بنابراین بازده سیستم بازیابی بیشینه شود. بنابراین در اندیس‌گذاری باید به دنبال راهکاری برای کلاس‌بندی تصاویر بود. روش‌های متعددی برای اندیس‌گذاری تصاویر استفاده می‌شود که اکثر آن‌ها از ساختارهای درختی استفاده می‌کنند. برای مثال *R*-tree* که در *QBIC* استفاده شده است [۱۱] یا *SS-Tree*. *Gevers* در سیستم *PicToSeek* [۱۹] ذکر کرده که به دلیل بار محاسباتی مربوط به جستجوی این درخت‌ها، از ساختار *SR-tree* استفاده کرده است. همچنین نشان داده که کارایی *SR-tree* از *R*-tree* بیشتر است.

Chiueh و سایرین در [۲۲] از *VP-tree* برای اندیس‌گذاری تصویر استفاده کرده‌اند. روش آنها بر اساس فاصله نسبی بردارهای ویژگی در فضای ویژگی است. برای اینکار از نقطه‌ای به نام نقطه تفوق^{۴۴} به عنوان مرکز ثقل فضای ویژگی استفاده کرده است و محاسبات فاصله بر اساس آن انجام می‌گیرد. در برخی سیستم‌ها از روش‌های کلاستر بندی (*k-means* و یا *SOM*) برای تقسیم‌بندی فضای ویژگی استفاده می‌شود. به این ترتیب برای پیدا کردن تصاویر

که در اینجا θ زاویه بین دو بردار \bar{k} و \bar{l} است. اگر این دو هم جهت باشند، فاصله (۸) برابر یک خواهد شد. Θ می تواند تابعی از \bar{k} و \bar{l} باشد (رابطه (۹)). این معیار برای ویژگی‌هایی با مقادیر حقیقی مناسب است [۴].

فاصله دیگری بر اساس روی هم افتادگی هیستوگرام‌ها تعریف شده است که دو هیستوگرام \bar{k} و \bar{l} با n مولفه را بر اساس مولفه‌های روی هم افتاده آنها مقایسه می‌کند (رابطه (۱۰)). پیش از به دست آوردن این فاصله، دو هیستوگرام باید نرمالسازی شوند. [۴].

$$D_H(\bar{k}, \bar{l}) = 1 - \frac{\sum_{i=1}^n \min(k_i, l_i)}{\sum_{i=1}^n k_i} \quad (10)$$

۴-۳- امتیازدهی به تصاویر

Li در [۲، ۷] برای امتیازدهی به تصاویر، از جمع فواصل به دست آمده برای ویژگی‌های سیستم استفاده کرده است. از آنجا که این فواصل برای در مقیاس‌های متفاوت محاسبه شده‌اند، بنابراین لازم است پیش از جمع‌بندی مقادیر آنها نرمالسازی شوند. روش محاسبه وی برای دو ویژگی رنگ و بافت در فرمول (۱۱) آمده است.

$$D_{sim}(\bar{k}, \bar{l}) = \frac{D_{color}(\bar{k}, \bar{l})}{MAX_{d \in [1, k]} (D_{color}(\bar{k}, \bar{d}))} + \frac{D_{texture}(\bar{k}, \bar{l})}{MAX_{d \in [1, k]} (D_{texture}(\bar{k}, \bar{d}))} \quad (11)$$

که در آن D_{color} معیار فاصله برای ویژگی یا کلاس‌های ویژگی رنگ و $D_{texture}$ معیار فاصله برای ویژگی بافت است. بر گفته Li در رابطه (۱۱)، d مجموعه تصاویری است که از هر دو فیلتر رنگ و بافت عبور کرده‌اند. وی برای استخراج رنگ از بردار هوموگرام رنگ (روشی بر اساس بهینه‌سازی هیستوگرام) و برای استخراج بافت از تبدیل wavelet استفاده کرده است، بنابراین هم اندیس‌گذاری و هم تعیین میزان شباهت در ویژگی رنگ و بافت هر دو بر روی یک بردار ویژگی اعمال شده است.

این روش برای سیستم‌هایی که از چند ویژگی برای استخراج ویژگی رنگ/بافت استفاده می‌کنند و دو بخش اندیس‌گذاری و معیار شباهت بر روی ویژگی‌های متفاوت انجام می‌گیرد، باید تعمیم داده شود. چراکه در این سیستم‌ها D_{color} و $D_{texture}$ خود معرف مجموع چند معیارهای شباهت هستند.

پس از امتیازدهی، تصاویر بر اساس امتیاز خود به صورت نزولی مرتب شده و k تصویر اول برای نمایش به کاربر انتخاب می‌شوند.

۵- پایگاه‌های داده تصاویر

برای تست و استفاده از سیستم‌های بازیابی تصویر، پایگاه داده‌های مخصوصی وجود دارد که در این بخش برخی از پرکاربردترین آنها

مشابه با یک تصویر، فاصله بردار ویژگی با مرکز کلاسترها اندازه‌گیری می‌شود و تصاویر موجود در کلاستری که با مرکز آن بردار مطلوب فاصله کمینه دارند انتخاب می‌شوند. سپس از بین این تصاویر با استفاده از معیارهای شباهت قضاوت می‌شود که کدام تصاویر به عنوان خروجی کاربر نمایش داده شوند.

۴-۲- معیارهای شباهت

فرض می‌شود که به ازای تمامی یا برخی از ویژگی‌های استفاده شده در سیستم بازیابی تصویر، مجموعاً N تصویر نزدیک به تصویر پرس‌وجو به دست آمده است. برای به دست آوردن تصاویر خروجی باید به هر تصویر بر اساس میزان شباهت آن به تصویر پرس‌وجو امتیازی تعلق گیرد و در انتها k بهترین تصویر انتخاب و به عنوان خروجی به کاربر نمایش داده شود. برای این کار از معیارهای شباهت و فاکتور وزن استفاده می‌شود. معیار شباهت تابعی غیرمنفی و صعودی است [۲۳] که میزان شباهت بردارهای ویژگی را مشخص می‌کند، و فاکتور وزن درجه اهمیت این ویژگی‌ها را در تعیین تصاویر نهایی مشخص می‌کند. نحوه انتخاب معیار شباهت برای افزایش دقت نتیجه خروجی مهم است.

غالب معیارهای شباهت از رابطه (۵) تبعیت می‌کنند که به فرم فاصله مینکوفسکی^{۴۵} مشهور است [۴]. در این رابطه \bar{k} معرف بردار ویژگی تصویر پرس‌وجو و \bar{l} معرف بردار ویژگی تصاویر موجود در پایگاه داده است که از فیلتر یا فیلترهایی عبور کرده‌اند.

$$D_M^\rho(\bar{k}, \bar{l}) = \left(\sum_{i=1}^n |k_i - l_i|^\rho \right)^{1/\rho} \quad (7)$$

فاصله مرتبه اول^{۴۶} که برای محاسبه اختلاف هیستوگرام دو تصویر از آن استفاده می‌شود [۲۳، ۷] با قرار دادن $\rho = 1$ در رابطه (۷) به دست می‌آید. استفاده از این معیار در هیستوگرام تصویر باعث بروز منفی نابجا^{۴۷} می‌شود، به این معنا که همه تصاویری که از لحاظ مفهومی باید استخراج شوند، با این روش استخراج نمی‌شوند. برای اصلاح این مساله از روش که در [۲۳] آمده است استفاده می‌شود.

فاصله مرتبه دوم^{۴۸} یا فاصله اقلیدسی با قرار دادن $\rho = 2$ در رابطه (۷) به دست می‌آید. از این معیار نیز به طور عمومی برای سایر کلاس‌های ویژگی استفاده می‌شود، چرا که فرض می‌شود مقادیر موجود در هر کلاس ویژگی از یکدیگر مستقل‌اند [۷].

فاصله کسینوسی^{۴۹} که در رابطه (۸) آمده است، بردارهای ویژگی دو تصویر را مقایسه می‌کند و کسینوس زاویه بین دو بردار را به دست می‌دهد.

$$D_C(\bar{k}, \bar{l}) = 1 - \cos\theta \quad (8)$$

$$\cos\theta = \frac{\bar{k} \cdot \bar{l}}{\|\bar{k}\| \|\bar{l}\|} \quad (9)$$

برای استخراج بافت می‌توان از یکی از دو فیلتر گابور و موجک استفاده کرد.

برای فیلتر گابور از راه‌کار ارائه شده در [۱۸] و برای موجک از [۲۴] استفاده می‌شود. این دو روش راه‌کاری برای استخراج بافت‌های پیچیده-تر از سطح بافت‌های معمولی، با بردار ویژگی مینیمم شده ارائه می‌دهند. هر دو راه‌کار پیاده‌سازی و مقایسه شده و بهترین راه برای توصیف بافت انتخاب خواهد شد.

در مرحله استخراج ویژگی، تصاویر به بلوک‌هایی با اندازه کوچک تقسیم می‌شوند. همانطور که ذکر شد این روش در استخراج ویژگی‌های رنگ باعث کاهش دقت و در کنار آن باعث افزایش سرعت استخراج ویژگی، اندیس‌گذاری و بازیابی تصویر می‌شود. در روش‌های استخراج بافت توسط دو فیلتر گابور و موجک، استفاده از فیلتر مستلزم بلوک-بندی تصویر است.

۶-۲- اندیس‌گذاری و جستجو

اندیس‌گذاری با استفاده از درخت جستجوی دودویی انجام می‌گیرد. برای ممان‌های رنگ و هیستوگرام برچسب رنگ، که قابلیت اندیس‌گذاری را دارند، یک درخت جستجوی دودویی ایجاد خواهد شد. سپس با استفاده از تصاویر حاصل از جستجو در این درختها، معیار شباهت بر روی بردار سایر ویژگی‌های استخراج شده اعمال می‌شود و به این ترتیب k نزدیکترین تصویر به تصویر پرس‌وجو انتخاب و به کاربر نمایش داده می‌شود.

درخت جستجوی دودویی

درخت جستجوی دودویی، بردارهای ویژگی را بر اساس مقادیر هر یک از مولفه‌های آن دسته‌بندی می‌کند، به صورتی که در هر سطح درخت یکی از مولفه‌های بردار ویژگی قضاوت می‌شود. اگر مقدار مولفه بردار ویژگی از مقدار آستانه‌ای کوچکتر باشد، تصاویر مشابه در زیردرخت سمت چپ و اگر بزرگتر بود در زیر درخت سمت راست جستجو می‌شود. به این میزان آستانه، مقدار برش $^5 [۲۵]$ می‌گویند. برای تعیین این مقدار برش از روش‌های گوناگونی استفاده می‌شود. یکی از این روش‌ها که در رابطه (۱۱) آمده است، از میانگین هریک از مولفه-های بردار ویژگی تصاویر در پایگاه داده استفاده می‌شود.

$$A_i = \alpha \left(\frac{1}{N} \sum_{m=1}^N H_{im} \right) \quad (11)$$

که در اینجا A بردار مقادیر برش، N تعداد تصاویر موجود در پایگاه داده تصویر و H_{im} مولفه i ام بردار ویژگی H برای تصویر m ام است. α ضریبی است که میزان دقت بازیابی را تعیین می‌کند. می‌توان برای تعادل درخت $\alpha = 1/2$ و برای بازیابی دقیقتر $\alpha = 2/3$ قرار داد. شایان ذکر است که پیش از محاسبه مقادیر برش، مقادیر بردار ویژگی

اشاره می‌شود. از بین تمامی آنها، تنها مورد آخر برای تست سیستم جاری در دسترس است.

- پایگاه داده مشهور Corel برای بازیابی تصویر به صورت گسترده استفاده می‌شود، اما مشکل آن حجم بالای تصاویر متشکله است. تعداد تصاویر آن بین ۳۰۰۰۰ تا ۲۰۰۰۰۰ در نسخه‌های گوناگون متفاوت است.
- پایگاه داده IRMA از ۳۸۷۹ تصویر رادیولوژی تشکیل شده که برای تست سیستم‌های تک-منظوره با کاربرد پزشکی استفاده می‌شود.
- پایگاه داده دانشگاهی UW شامل ۱۱۰۹ تصویر حاشیه‌نویسی شده است که توسط دانشگاه واشنگتن جمع‌آوری و آماده شده است.
- پایگاه داده ZuBuD شامل ۱۰۰۵ تصویر از ۲۰۱ خانه تشکیل شده که هر خانه از ۵ زاویه مختلف عکس‌برداری شده است.
- پایگاه داده موجود، که شامل ۱۰۰۰ تصویر می‌باشد که در ده کلاس دیداری طبقه‌بندی شده است. نمونه‌ای از تصاویر این پایگاه داده در شکل (۶) آمده است.

۶- بیان مساله و راه‌کار پیشنهادی

هدف از این پروژه ارائه روشی برای بازیابی تصاویر رنگی مبتنی بر اطلاعات بافت و رنگ است. فلوجارت این سیستم در شکل (۱) آمده است. بر این اساس، ابتدا ویژگی تصاویر پایگاه داده استخراج شده و پس از اندیس‌گذاری، در پایگاه داده تصویر ذخیره می‌شوند. پس از این سیستم آماده دریافت پرس‌وجوی کاربر است. کاربر برای جستجوی تصاویر مشابه یک تصویر خاص، آن را به سیستم وارد می‌کند. سیستم بازیابی، ویژگی‌های تصویر پرس‌وجو را با استفاده از روشی مشابه با روش استخراج ویژگی سایر تصاویر پایگاه داده، استخراج می‌کند. پس از آن، پایگاه داده ویژگی توسط بخش بازیابی، جستجو می‌شود تا نزدیک-ترین تصاویر به تصویر پرس‌وجو به دست آید. این تصاویر به کاربر نمایش داده می‌شوند. جزئیات مربوط به هر کدام از بخش‌های این سیستم در قسمت‌های بعد آمده است.

این سیستم بازیابی بر اساس دسته‌بندی Gevers در [۴] جزو سیستم‌های بازیابی برای یافتن تصویر خاص است. عمده کاربرد این سیستم در بازیابی تصاویر اینترنتی است. این سیستم به علت استفاده ویژگی‌های سطح پایین رنگ و بافت برای کاربردهای مفهومی بازیابی تصویر چندان مناسب نیست.

۶-۱- استخراج ویژگی

در استخراج ویژگی رنگ، از ممان‌های اول تا سوم و هیستوگرام برچسب رنگ برای اندیس‌گذاری و فیلتر اولیه تصاویر و از ویژگی‌هایی نظیر بردار همبستگی‌نگار رنگ برای تعیین میزان شباهت تصویر استفاده می‌شود.

شده‌اند. به این ترتیب سایر محاسبات مربوط به استخراج ویژگی بر روی تصویر حاصل شده انجام می‌گیرد.

برای استخراج ویژگی از دو هیستوگرام برچسب رنگ که در اینجا به نام هیستوگرام کلی و جزئی معرفی می‌شوند استفاده شده است. برای محاسبه این دو هیستوگرام باید رنگ‌های موجود در پالت رنگ دسته‌بندی شوند. در اینجا این دسته‌بندی در دو مرحله انجام گرفته است، ابتدا رنگ‌های پالت ۲۱۶ تایی به ۶ دسته کلی تقسیم شده و سپس هر دسته خود به ۶ دسته جزئی‌تر تقسیم می‌شود، بنابراین در دسته‌بندی دوم ۳۶ دسته رنگ به دست خواهد آمد. برای دسته‌بندی رنگ‌ها از الگوریتم *k-means* استفاده شده است. این دسته‌ها در شکل (۳) آمده‌اند. در شکل (۴) پیکسل‌های چند تصویر نمونه براساس شماره کلاس‌ها برچسب خورده‌اند. تصاویر حاصله پس از نرمالسازی هیستوگرام به مقادیر ۱ تا ۲۵۵ نگاشت شده‌اند.

برای به دست آوردن هیستوگرام کلی و جزئی، پیکسل‌های موجود در تصویر یکبار بر اساس شماره دسته‌های کلی (۱ تا ۶) و بار دیگر بر اساس شماره دسته‌های جزئی (۱ تا ۳۶) برچسب خواهند خورد. هیستوگرام کلی و جزئی با محاسبه تعداد پیکسل‌های متعلق به کلاس‌ها به دست خواهند آمد. در شکل (۵) هیستوگرام کلی و جزئی به دست آمده برای یک تصویر نشان داده شده است. هیستوگرام‌های به دست آمده به نرمالسازی شده هستند، به این صورت که مقادیر به دست آمده برای هر مولفه بر تعداد بلوک‌های تصویر تقسیم شده‌اند.

هیستوگرام کلی حاوی اطلاعات کلی درباره رنگ‌های تصویر است و با استفاده از آن، در زمان جستجو می‌توان تصاویری را که با تصویر پرس‌وجو تفاوت عمده‌ای دارند تمیز داد. هیستوگرام جزئی حاوی اطلاعات جزئی‌تر درباره تصاویر است که برای انتخاب *k* نزدیک‌ترین همسایه تصویر پرس‌وجو از میان تصاویر فیلتر شده استفاده خواهد شود. برای این کار از معیار شباهت استفاده می‌شود.

۲-۲- اندیس‌گذاری

برای اندیس‌گذاری از درخت جستجوی دودویی استفاده شده که شکل آن در شکل (۷) آمده است. این درخت، بر اساس بردار هیستوگرام کلی تصاویر را دسته‌بندی می‌کند. مقادیر جداکننده روی گره‌ها با استفاده از رابطه (۱۱) به دست آمده است. در اینجا $N = 18, \alpha = 2/3$ (به تعداد تصاویر پایگاه داده) و H_i معرف بردار هیستوگرام کلی است.

۳-۲- نتایج

در شکل (۶) تصاویر موجود در پایگاه داده تست و دو هیستوگرام منتسب به هر کدام آورده شده است. همانطور که مشهود است هیستوگرام‌های کلی و جزئی تصاویر مشابه به یکدیگر شباهت دارند،

باید نرمالسازی شده باشند. در شکل (۷) درخت جستجوی دودویی با $\alpha = 2/3$ نشان داده شده است.

برای تعیین تصاویر نهایی از میان تصاویر به دست آمده از درخت دودویی، از معیارهای شباهت ذکر شده در بخش‌های قبل استفاده می‌شود. سپس نتایج این معیارها مقایسه شده و برای هر ویژگی کاراترین آنها انتخاب می‌شود.

۳-۳- پایگاه داده

برای تست سیستم از پایگاه داده موجود که شامل ۱۰۰۰ تصویر است، استفاده می‌شود. این امکان وجود دارد که یکی از پایگاه‌های داده ذکر شده برای تست سیستم خریداری شود.

۴-۳- زمان‌بندی

در شکل (۲) روند پیشرفت پروژه تاکنون آمده است. زمان‌بندی تقریبی برای اتمام پروژه در جدول (۱) درج شده است.

جدول (۱): زمان‌بندی تقریبی اتمام پروژه

مرحله	زمان تقریبی
استخراج رنگ	یک ماه
استخراج بافت	دو ماه
جستجو و اندیس‌گذاری	یک ماه
ارزیابی روش‌ها	یک ماه
بهبود بازده بازیابی	دو ماه
جمع‌بندی و تهیه پایان‌نامه	یک ماه
مجموع	هشت ماه

۷- پیاده‌سازی

در اینجا به توضیح بخشی از پیاده‌سازی سیستم بازیابی پرداخته می‌شود. در این مرحله سیستم قادر است تا هیستوگرام تصاویر موجود در پایگاه داده و نیز تصویر پرس‌وجو را استخراج کرده و طی دو مرحله مقایسه کند تا تصاویر مشابه به تصویر پرس‌وجو را به دست آورد.

۷-۱- محاسبه ویژگی

برای محاسبه ویژگی رنگ، فضای رنگ RGB به ۲۱۶ رنگ متشکله مکعب *safe-colors* که حاوی ۲۱۶ رنگ استاندارد می‌باشد، کوانتیزه شده است. به این ترتیب یک پالت رنگ ۲۱۶ تایی به دست خواهد آمد. برای کاهش بار محاسباتی، تصاویر بلوک‌بندی شده‌اند، به این صورت که اندازه هر تصویر به $1/4$ تقلیل یافته است. در این تصویر، رنگ هر پیکسل معرف میانگین رنگ بلوک مرتبط با آن است. پس از این کار، مقادیر پیکسل‌های تصویر به ۲۱۶ رنگ موجود در پالت رنگ نگاشت

ها به جای استفاده از الگوریتم k -means و الگو گرفتن از نحوه دسته بندی Li در [۷]، می توان تضمین بهتری از مشابهت رنگ های موجود در یک دسته ارائه داد. راه کار دیگر افزایش تعداد کلاس های رنگ است که باعث افزایش دقت و البته افزایش بار محاسباتی سیستم شود.

در این سیستم با اضافه کردن اطلاعات مکانی با استفاده از ویژگی های بافت، بردار ارتباط رنگ و یا همبستگی نگار رنگ می توان دقت بالاتری برای استخراج تصویر به دست آورد.

در راه کاری که ارائه داده شد، اندازه بلوک ها وابسته به اندازه تصویر است و تصاویر انتخاب شده تقریباً به یک اندازه هستند. برای بهبود راه کار جاری می توان برای اندازه بلوک های تصویر مقداری ثابت تعیین کرد، به شیوه ای که هیستوگرام های به دست آمده کاملاً مستقل از اندازه تصویر باشد.

مراجع

- [۱] Long F. ; Zhang H. and Dagan Feng D., "Fundamentals of content-based image retrieval, in Multimedia Information Retrieval and Management - Technological Fundamentals and Applications," Springer-Verlag, pp. 1-26, 2003.
- [۲] Li X. ; Chen S.C. ; M.L. Shyu and Furht B., "Image Retrieval by Color, Texture, and Spatial Information," in 8th International Conference on Distributed Multimedia Systems (DMS'2002), San Francisco Bay, California, USA, 2002, pp. 152-159.
- [۳] Einarsson S. H., "Data structures for intermediate search results in the Eff2 image retrieval system," Reykjavík University, technical report 2004.
- [۴] Gevers Th. and Smeulders A.W.M., "Image Search Engines, An Overview," *The International Society for Optical Engineering (SPIE)*, vol. VIII, pp. ۲۲۷-۲۳۷ . ۲۰۰۳.
- [۵] Schettini R. ; Ciocca G. and Zuffi S., "A Survey of Methods for Color Image Indexing and Retrieval in Image Databases".
- [۶] Einarsson S. H. ; Grétarsdóttir R. Ý. ; Jónsson B. Þ. and Amsaleg L., "The EFF² Image retrieval System Prototype," in *ASTED Intl. Conf. on Databases and Applications (DBA)*, Innsbruck, Austria, 2005.
- [۷] Li X. ; Chen S. ; Shyu M. and Furht B., "An Effective Content-Based Visual Image Retrieval System," in *26th IEEE Computer Society International Computer Software and Applications Conference (COMPSAC)*, Oxford, 2002, pp. 914- 919.
- [۸] Rui Y. ; Huang Th. S. and Chang Sh., "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," *Journal of Visual Communication and Image Representation*, vol. 10, pp. ۳۹-۶۲, ۱۹۹۹.
- [۹] Squire D. ; Muller W. and Muller H., "Relevance feedback and term weighting schemes for content-based image retrieval," *Huijmsmans and Smeulders* vol. 5, pp. 549-556, 1998.
- [۱۰] Materka A. and Strzelecki M. , "Texture Analysis Methods - A Review," Technical University of Lodz, Institute of Electronics, Brussels, COST B11 1998.

همچنین هیستوگرام کلی قدرت تفکیک دهی بیشتری از هیستوگرام جزئی دارد. بنابراین از هیستوگرام کلی برای اندیس گذاری تصاویر استفاده شده است.

برای اندیس گذاری، پس از محاسبه، بردار مقادیر برش به صورت $A = \langle 1, 32, 6, 6, 18, 5 \rangle$ به دست آمده است؛ از آنجا که A_1 برابر ۱ است دارای خاصیت جداکنندگی نیست و بنابراین به عنوان مقدار تعیین کننده در سطوح درخت در نظر گرفته نشده است. بنابراین درخت دودویی از ۶ سطح به ۵ سطح کاهش یافته است.

شکل (۵) تصویر مورد پرس و جو است که هیستوگرام کلی و جزئی آن نمایش داده شده است. $H = \langle 0, 76.34, 0.48, 6, 23.11, 0.014, 0 \rangle$ معرف مقادیر عددی هر یک از مولفه های هیستوگرام کلی آن است.

از آنجا که در این مرحله تنها معیار تشخیص شباهت، رنگ تصاویر است و نیز تعداد تصاویر پایگاه داده محدود می باشد، بنابراین انتظار می رود که یک یا حداکثر دو تصویر مشابه برای تصویر پرس و جو از لحاظ دیداری به آن شبیه باشد. شکل (۸) تصاویر استخراج شده از پایگاه داده را نمایش می دهد. شماره تصاویر استخراج شده از درخت دودویی در پایگاه داده برابر (۷ و ۴) است. در جدول (۲) مقادیر شباهت به دست آمده از طریق فاصله های مرتبه اول، دوم (اقلیدسی) و سوم درج شده است. روابط مربوط به این فواصل در بخش های قبل آمده است.

همانطور که انتظار می رود فاصله بین مقادیر به دست آمده از فاصله مرتبه اول بیشتر است و قدرت تفکیک بیشتری در مورد هیستوگرام دارد.

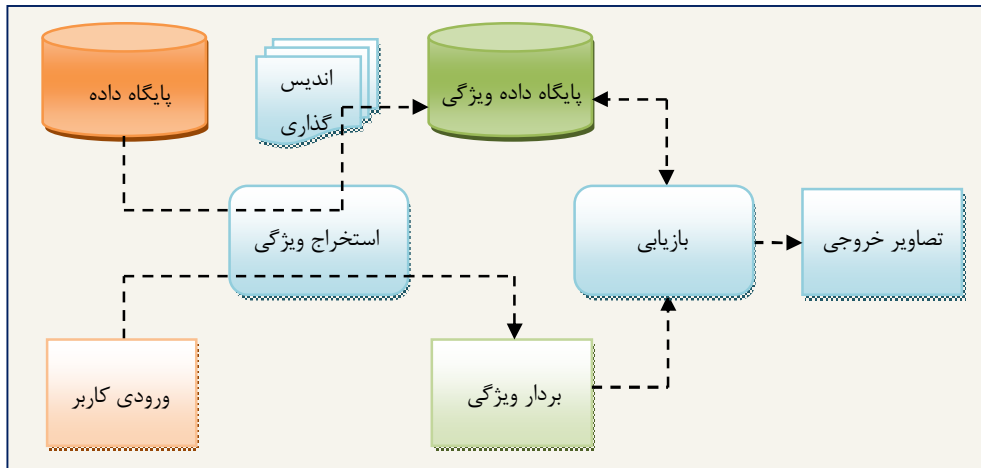
جدول (۲): فاصله محاسبه شده از تصویر پرس و جو برای

تصاویر استخراجی از درخت دودویی			
فاصله	مرتبه اول	مرتبه دوم	مرتبه سوم
تصویر ۷	۷.۱۰۷۲	۱۰.۶۶۷۸	۷.۸۸۶۷
تصویر ۴	۱۲.۱۱۴۸	۱۱.۱۹۹۹	۱۱.۷۹۰۱

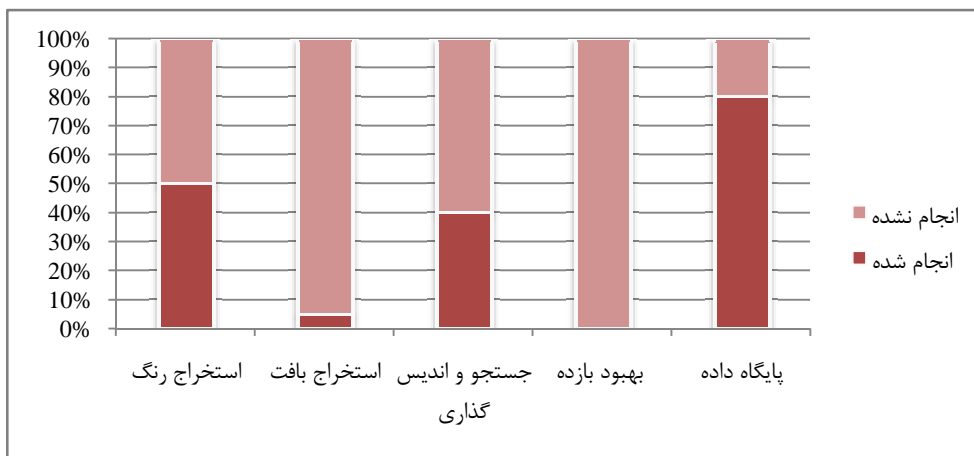
۷-۴- کارهای آینده

در این راه کار به استخراج ویژگی رنگ از تصویر اساره شده است. در این سیستم از فضای رنگ RGB استفاده شده که همانطور که توضیح داده شد، در سیستم های بازیابی تصویر دارای کاستی هایی است. می توان با استفاده از سیستم HSV و $L^*a^*b^*$ (که در مورد هیستوگرام بسیار معمول ترند) نتیجه جستجو را بهبود داد. محاسبه جداگانه هیستوگرام برای بعد روشنایی تصویر نیز می تواند در بهبود نتیجه بازیابی موثر باشد. از آنجا که فضای HSV یک فضای ادراکی است، بنابراین با دسته بندی رنگ در این فضا و استفاده از محاسبات ریاضی و مبتنی بر فاصله رنگ -

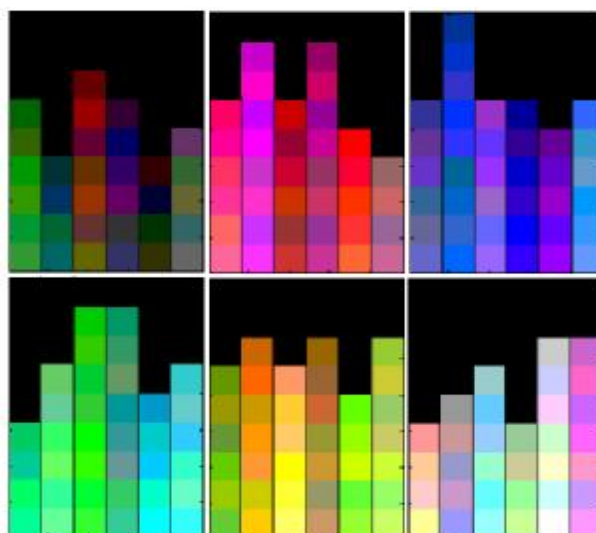
- [١٠] Smith J. R. and Chang S. F., "VisualSEEK: A fully automated content-based image query system," in *ACM Multimedia Conference*. Boston, MA, USA, 1996.
- [١١] Markov I., "VP-tree: Content-Based Image Indexing," in *4th Spring Young Researcher's Colloquium on Database and Information Systems (SYRCODIS'2007)*, Moscow, Russia, 2007.
- [١٢] Chiueh T., "Content-based image indexing," in *Proceedings of VLDB '94*, Santiago ,Chile, 1994, pp. 582-593.
- [١٣] Stricker M. A. and Orengo M., "Similarity of Color Images," in *SPIE*, 1995, pp. 381--392.
- [١٤] S. S. a. P. J. Hiremath P.S., "Wavelet Based Feature for Color Texture Classification with application to CBIR," *Intl. Journal of Computer Science and Network Security (IJCSNS)*, vol. 6, Sep. 2006.
- [١٥] Shi Y. and Liu Y., "Binary Tree-based Clustering Algorithm and Used in Color Image Segmentation," in *4th Intl. Conf. on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2007, pp. 219-22.٢
- [١٦] Veltkamp and Tanase, "Content-Based Image Retrieval Systems: A Survey," Dept. of Computing Science, Utrecht University, Technical Report 2000.
- [١٧] T. Gevers" ,Robus Histogram Construction from Color Invariants," *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 26, pp. 113-118, 2004.
- [١٨] Howarth P. and Ruger S., "Evaluation of Texture Features for Content-Based Image Retrieval," in *Third International Conference, CIVR 2004*, Dublin, Ireland, 2004.
- [١٩] Arvis V. ; Debain C. ; Berducat M. and Benassi A., "Generalization of the Co-occurrence Matrix for Colour Images: Application to Colour Texture Classification " *Image Analysis and Stereology*.٢٠٠٤ ,
- [٢٠] Deselaers Th., "Features for Image Retrieval," 2003.
- [٢١] Bhagavathy S. ; Tesic J. and Manjunath B. S., "On the Rayleigh Nature of Gabor Filter Outputs," in *Intl. Conf. on Image Processing (ICIP)*, 2003.
- [٢٢] Smith J. R. and Chang S., "Tools and Techniques for Color Image Retrieval," in *SPIE*, 1996, pp. 1630-1639.
- [٢٣] Andrysiak T. and Chora's M. , "Image Retrieval Based on Heirarchical Gabor Filter," *Intl. Journal on Applied Mathematics and Computer Science*, vol. 15, pp. 471–480, 2005.
- [٢٤] Gevers Th. and Smeulders A. W. M. , "The PicToSeek WWW Image Search System " in *IEEE ICMCS*, 1999.



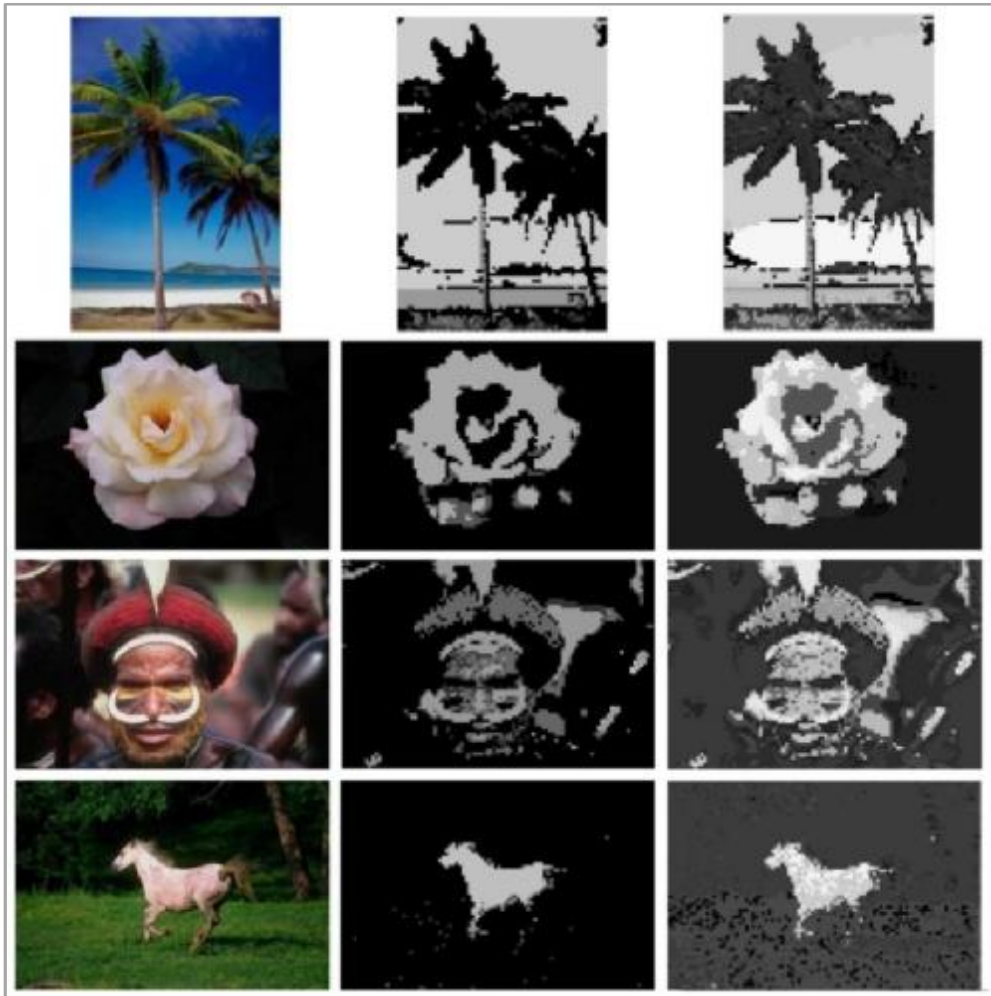
شکل (۱): شمای کلی یک سیستم بازیابی تصویر (اقتباس شده از [۴، ۵])



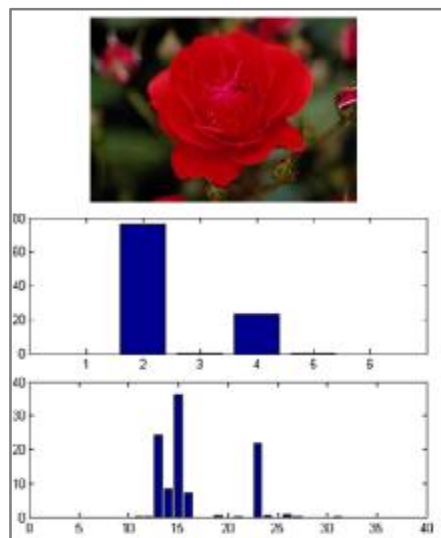
شکل (۲): نمودار میزان پیشرفت کار تاکنون



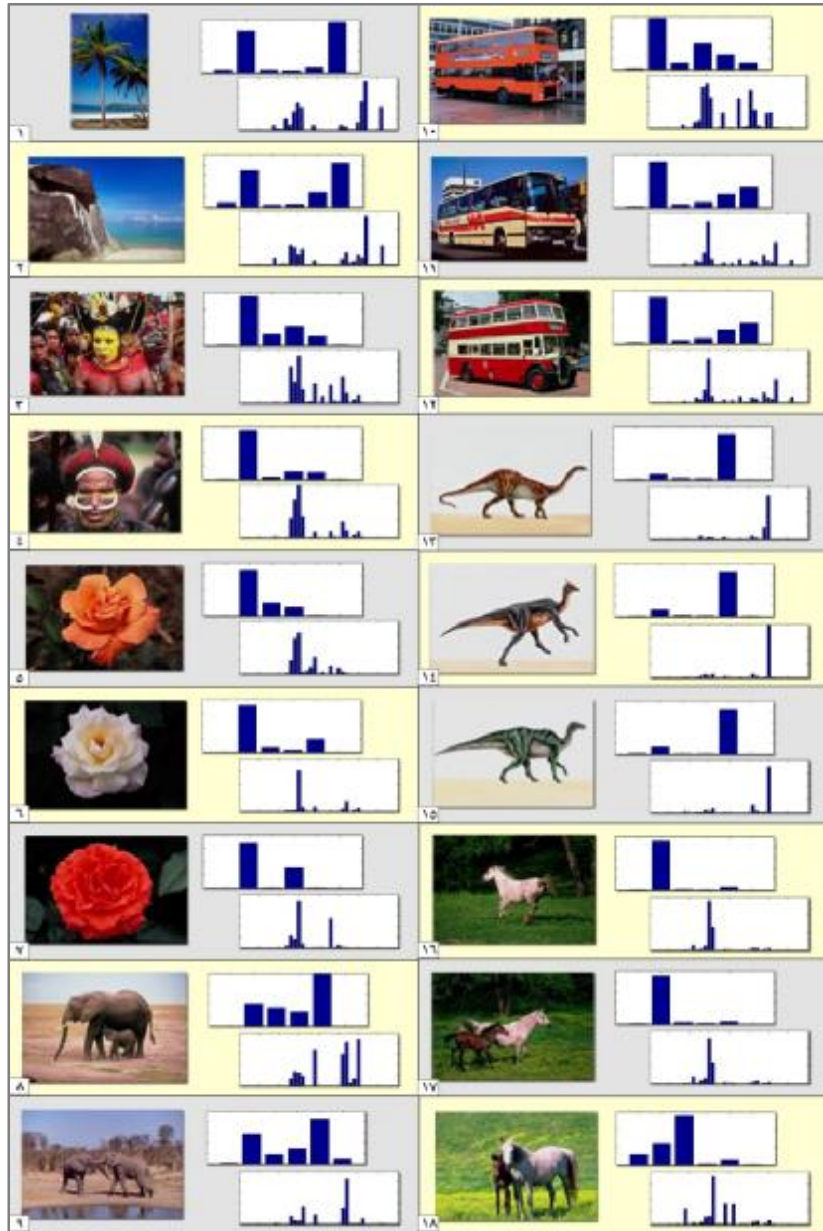
شکل (۳): پالت رنگ سیستم بازیابی تصویر پیشنهادی با ۲۱۶ رنگ



شکل (۴): ستون اول از چپ: تصاویر اصلی، ستون دوم: بلوک‌بندی تصویر با استفاده از کاهش درجه تفکیک پذیری آن و برچسب‌گذاری بلوک‌ها بر اساس شماره دسته‌های اصلی، ستون سوم: برچسب‌گذاری بر اساس شماره دسته‌های فرعی



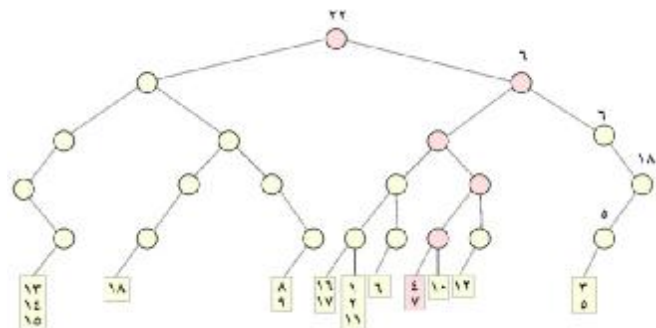
شکل (۵): از بالا به پایین: یک تصویر و هیستوگرام کلی و جزئی معادل آن



شکل (۶): پایگاه داده تصاویر به همراه هیستوگرام کلی (بالا) و جزئی (پایین) هر یک از آن ها



شکل (۸): تصاویر مشابه تصویر شکل (۵)، استخراج شده از درخت دودویی شکل (۷)



شکل (۷): اندیس گذاری تصاویر بر اساس هیستوگرام کلی در درخت دودویی

- 1 Annotation
- 2 Content-based Image Retrieval
- 3 Layout
- 4 Search by Association
- 5 Target Search
- 6 Category Search
- 7 Semantic
- 8 Segmentation
- 9 Perspective
- 10 Color Space
- 11 Perceptual Uniformity
- 12 Opponent color space
- 13 Hue-Saturation-Value
- 14 Skewness
- 15 Joint Distribution
- 16 Histogram Saturation
- 17 Color Label Histogram
- 18 Dimensionality Curse
- 19 Color Coherence Vector
- 20 Color Correlogram
- 21 Morphological Operator
- 22 Adjacency Graph
- 23 Auto-Regressive
- 24 Wold Decomposition
- 25 Wavelet
- 26 Gray-Level Co-occurrence Matrices
- 27 Coarseness,
- 28 Contrast
- 29 Directionality
- 30 Line-likeness
- 31 Regularity
- 32 Roughness
- 33 Discrete Wavelet Transform
- 34 Compact Support
- 35 Multiresolution
- 36 Resolution
- 37 Query
- 38 Seed
- 39 Nearest-neighbor Search
- 40 K-nearest Neighbor
- 41 Give me more Query
- 42 Sequential
- 43 Offline
- 44 Vantage Point
- 45 Minkowski-form
- 46 L_1 -Distance
- 47 False Negative
- 48 L_r -Distance
- 49 Cosine Distance
- 50 Cut-point