

# مروری بر سیستم‌های بازیابی تصویر مبتنی بر محتوا در کاربردهای عام

زهرا منصوری

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

[z\\_mansoori@ce.sharif.edu](mailto:z_mansoori@ce.sharif.edu)

منصور جم‌زاد

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

[jamzad@sharif.edu](mailto:jamzad@sharif.edu)

**چکیده:** امروزه مجموعه‌های حجیم تصاویر دیجیتال در زمینه‌های متنوع نظیر تجاری، دولتی، علمی و پزشکی موجود است. این مجموعه‌ها به منظور جستجو عموماً بر اساس متنی الصاق شده به تصاویر اندیس گذاری می‌شوند. برای بهبود جستجو روش‌هایی ابداع شده‌اند که قادرند تصاویر را بر اساس محتوای دیداری آن‌ها اندیس‌گذاری کنند. در این مقاله سعی بر آن است تا دیدی کلی نسبت به این سیستم‌ها و ساختار شکل دهنده آن‌ها ارائه دهد.

**واژه‌های کلیدی:** بازیابی تصویر مبتنی بر محتوا، رنگ، بافت، اندیس‌گذاری

## ۱- مقدمه

در سال‌های اخیر، به دلیل گسترش نیاز به بازیابی بهینه تصاویر در پایگاه‌های داده حجیم، تحقیقات گسترده‌ای در زمینه بازیابی تصاویر صورت گرفته است. اولین ایده‌ها در اوایل ده ۷۰ میلادی مطرح شد که در آنها دسته‌بندی تصاویر بدون توجه به ویژگی‌های دیداری آنها و تنها بر اساس حاشیه‌نویسی<sup>۱</sup> متنی انجام می‌گرفت؛ به صورتی که مفاهیم موجود در تصویر توسط اپراتور تشخیص داده شده و در پایگاه داده‌ای به عنوان کلمات کلیدی آن تصویر ذخیره می‌شد. 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 می‌شود.

سیستم‌های فوق‌الذکر با مشکلاتی مواجه بودند: نخست اینکه حاشیه‌نویسی 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 حاشیه‌های الصاق شده به تصاویر تمامی حیطه پرس‌وجو را نمی‌پوشاندند؛ و این به این معنا است که پرس‌وجوهای مبتنی بر متن به میزان کافی کامل و گویا نیستند. در ابتدای دهه ۹۰، با افزایش حجم تصاویر موجود در پایگاه‌های داده نظیر شبکه جهانی اینترنت و برای غلبه بر مشکلات

سیستم‌های مبتنی بر متن، توسعه سیستم‌های بازیابی تصویر مبتنی بر محتوا<sup>۲</sup> یا به اختصار CBIR پیشنهاد شد که وظیفه استخراج تصاویر به صورت اتوماتیک و با استفاده از مفاهیم دیداری نظیر رنگ، بافت و طرح‌بندی<sup>۳</sup> تصویر را بر عهده داشت [۱، ۲]. سیستم‌های اولیه از این رده، برای بازیابی تصویر از کاربر درخواست می‌کرد که یک یا چند ویژگی دیداری را انتخاب و محدوده‌ای برای مقدار آن‌ها تعیین کند، سپس عمل بازیابی بر اساس این اطلاعات صورت می‌گرفت. در سیستم‌های پیشرفته‌تر، کاربر قادر بود تا درجه اهمیت ویژگی‌های انتخاب شده را تعیین کند. به دلیل ناتوانی در جلب رضایت کاربر، سیستم‌های آتی دریافت تصویر به عنوان الگوی بازیابی را در سیستم خود امکان‌پذیر ساختند. به این شیوه که کاربر در کنار قابلیت‌های ذکر شده، می‌توانست 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 سیستم‌های پیشرفته کاربر قادر بود تا میزان اهمیت هر یک از ویژگی‌ها را خود تعیین نماید، همچنین قابلیت بهبود نتایج جستجو بوسیله تعامل با کاربر و اعمال نظر وی در جستجوهای بعدی نیز به این سیستم‌ها اضافه شد [۳، ۴].

در این مقاله سعی بر آن است تا راه‌کارهایی که تاکنون برای توسعه سیستم‌های بازیابی تصاویر انجام شده‌اند معرفی گردد. ساختار مقاله بدین صورت است: ابتدا سیستم‌های بازیابی تصویر به صورت کلی بررسی می‌شود و سپس اجزای آن‌ها ذکر شده و توضیحاتی در هر زمینه ارائه می‌گردد، سپس برخی پایگاه‌های داده موجود برای بررسی سیستم‌های بازیابی تصویر ذکر شده و در انتها جمع‌بندی کلی صورت می‌گیرد.

## ۲- خصوصیات سیستم‌های بازیابی تصویر

در شکل (۱) شمای کلی یک سیستم بازیابی تصویر آمده است. در این سیستم‌ها پایگاه داده‌ای وجود دارد که شامل کلیه تصاویر قابل بازیابی

<sup>۲</sup> Content-based Image Retrieval

<sup>۳</sup> Layout

<sup>۱</sup> Annotation

### ۳- استخراج ویژگی‌های تصویر

پیش از شروع کار بازیابی، ویژگی‌های تصاویر بانک داده باید از آن‌ها استخراج شوند. برای سهولت دسترسی به ویژگی‌های آن تصویر، به جای استفاده از مقادیر پیکسل‌ها، با استفاده از روش‌هایی خاص ویژگی‌های تصاویر 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، «بردار ویژگی» تصویر گفته می‌شود [۳، ۷، ۱۱].

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 تصویر هستند و توانایی توصیف اشیاء یا مفاهیم موجود در تصویر را ندارند. رنگ و بافت دو نوع از این ویژگی‌ها است. ویژگی‌های معنایی به بیان مفاهیم و اشیاء موجود در تصویر می‌پردازند و غالباً از روش‌های قسمت بندی<sup>۵</sup> تصویر برای اینکار استفاده می‌کنند. دسته سوم مربوط به ویژگی‌های مربوط به کاربردهای خاص است، نظیر ویژگی‌های توصیفی چهره مورد استفاده در سیستم‌های بازیابی تصاویر چهره [۱۲].

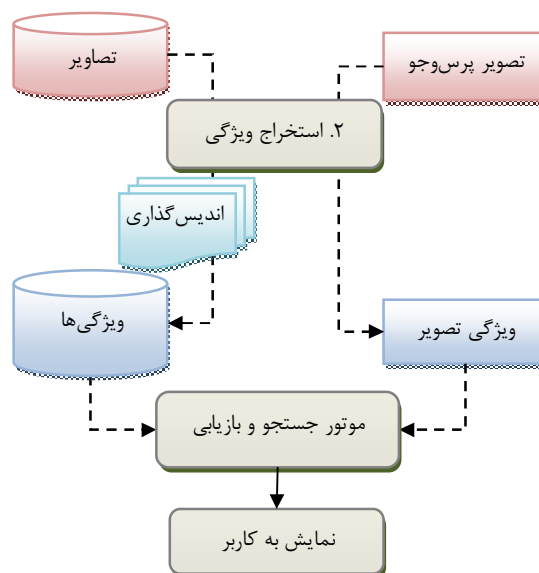
### ۳-۱ استخراج رنگ

رنگ یکی از متداول‌ترین و تعیین‌کننده‌ترین ویژگی دیداری در زمینه بازیابی تصویر است؛ چرا که نسبت به تغییرات مربوط به اندازه، جهت و دورنمایی<sup>۶</sup> و اغتشاش تصویر پایدار است [۷].

از دید شخص، رنگ از سه مولفه تشکیل شده است: قرمز، سبز و آبی؛ که این سه یک فضای رنگ<sup>۷</sup> 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<sup>۸</sup> 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].

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 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 کلاس‌بندی شوند.

سیستم‌های بازیابی نوعاً از دو بخش تشکیل می‌شوند: ۱. بخش استخراج ویژگی که اطلاعات دیداری تصویر را برای دسته‌بندی آنها استخراج می‌کند؛ ۲. بخش بازیابی که پرس‌وجوی کاربر را پردازش می‌کند.



شکل (۱) شمای کلی یک سیستم بازیابی تصویر (اقتباس شده از [۴، ۱۰])

از سیستم‌های بازیابی تصویر موجود به عنوان نمونه می‌توان از QBIC [۵]، VisualSEEK [۶]، EFF۲ [۷] و PicToSeek [۸] و Viper [۹] نام برد. سیستم QBIC یکی از مهمترین سیستم‌های بازیابی به شمار می‌رود، چرا که سایر سیستم‌های بازیابی از آن به عنوان مدل اولیه توسعه استفاده کرده‌اند.

در فصول آتی، دو بخش عمده یک سیستم‌های بازیابی محتوا یعنی بخش استخراج ویژگی و بخش اندیس‌گذاری به تفصیل توضیح داده خواهند شد.

<sup>4</sup> Semantic

<sup>5</sup> Segmentation

<sup>6</sup> Perspective

<sup>7</sup> Color Space

<sup>8</sup> Perceptual Uniformity

فضاهای رنگی متنوعی برای استفاده وجود دارد، که مهمترین آنها به منظور استفاده در سیستم‌های مذکور ذکر می‌شود. فضای RGB به دلیل سادگی استفاده و درک و نیز پشتیبانی سخت‌افزاری به‌طور گسترده مورد استفاده قرار می‌گیرد؛ اما دارای یکنواختی ادراکی نیست. فضای HSV یک فضای ادراکی است، به همین دلیل برای استفاده در سیستم‌های بازیابی مناسب است [۷]. 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 تغییرات نور و جهت تصویربرداری است و بنابراین برای سیستم‌های بازیابی بسیار مناسب است. با این حال H به تغییرات روشنایی به شدت حساس است. 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. [۱، ۴، ۷].

از آنجا که هر یک از این فضاهای رنگ برای استخراج برخی ویژگی‌ها کاراتر از سایر فضاها است، 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 [۷] 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. در ادامه برخی روش‌های استخراج ویژگی رنگ ذکر می‌شود.

## 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 اربیبی رنگ<sup>۹</sup> [۱۳]، توزیع رنگ تصویر را به صورت کارآمد و بهینه ارائه می‌دهند. 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. این ویژگی اصولاً قدرت جداسازی کمی دارد و بنابراین از آن برای فیلتر کردن تصاویر در مراحل

اولیه جستجو استفاده می‌شود [۱]. در [۷] ویژگی‌های دیگری نظیر انرژی و آنتروپی آمده است. QBIC از این سه ممان اول در سه فضای RGB، YIQ و CIE Lab استفاده کرده است [۵].

## هیستوگرام رنگ

هیستوگرام معرف توزیع تجمعی<sup>۱۰</sup> 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 بازیابی تصویر به صورت گسترده استفاده می‌شود، همچنین نسبت به تغییر اندازه کم، تغییر زاویه، دورنمایی و اغتشاش تصویر پایدارتر است [۲، ۷].

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 بازیابی را افزایش نمی‌دهد. مشکل دیگری که در پایگاه‌های داده تصویر بسیار بزرگ رخ می‌دهد، مساله اشباع شدن هیستوگرام<sup>۱۱</sup> 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 [۱۴] 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 [۵].

## هیستوگرام برچسب رنگ<sup>۱۲</sup>

در برخی از 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 سه بعد (مربوط به سه مولفه رنگ)، از یک بعد (برچسب رنگ) تشکیل شده است.

Li در [۱۱] 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»<sup>۱۳</sup> حائز اهمیت است.

<sup>10</sup> Joint Distribution

<sup>11</sup> Histogram Saturation

<sup>12</sup> Color Label Histogram

<sup>13</sup> Dimensionality Curse

<sup>9</sup> Skewness

## بردار ارتباط رنگ<sup>۱۴</sup>

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، ۱۳].

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 [۱۱] روشی از نوع آماری 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<sup>۱۸</sup> یا AR، مدل گوسی مارکوف یا RMF و مدل Gibbs RMF می‌شوند.

روش‌های تبدیلی شامل ویژگی‌های Tamura، تجزیه Wold<sup>۱۹</sup>، تبدیل موجک<sup>۲۰</sup> 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 [۱۳] یا GLCM<sup>۲۱</sup> [۱۵] نیز معروف است، احتمال تجمعی توزیع جفت پیکسل‌ها را محاسبه می‌کند 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، تباین و همگنی از آن استخراج می‌شود که می‌توانند معرفی برای بافت موجود در تصویر باشند [۱۵]. Arvis به همراه سایرین در [۱۶] روش دیگری برای استخراج بافت از این ماتریس ارائه کرده است.

### ویژگی‌های Tamura

این ویژگی‌های ویژگی‌هایی از بافت را نظیر درشتی<sup>۲۲</sup>، تباین<sup>۲۳</sup>، جهت‌دار بودن<sup>۲۴</sup>، خطی بودن<sup>۲۵</sup>، نظم<sup>۲۶</sup> و زبری<sup>۲۷</sup> 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

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  $\alpha_i$  معرف پیکسل‌های مرتبط در مولفه  $i$  ام هیستوگرام (مربوط به رنگ  $i$  ام) و  $\beta_i$  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  $\langle \alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_N + \beta_N \rangle$  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<sup>۱۵</sup>

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  $i$  ام در فاصله  $k$  از یک پیکسل از رنگ  $j$  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].

### ۲-۳ 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<sup>۱۶</sup> و گراف همجاری<sup>۱۷</sup> است که بافت را بر اساس ویژگی‌های ساختاری و قواعد مربوط به نحوه قرار گرفتن پیکسل‌ها و اصطلاحاً بر اساس ریزبافت‌ها

<sup>18</sup> Auto-Regressive

<sup>19</sup> Wold Decomposition

<sup>20</sup> Wavelet

<sup>21</sup> Gray-Level Co-occurrence Matrices

<sup>22</sup> Coarseness,

<sup>23</sup> Contrast

<sup>24</sup> Directionality

<sup>25</sup> Line-likeness

<sup>26</sup> Regularity

<sup>27</sup> Roughness

<sup>14</sup> Color Coherence Vector

<sup>15</sup> Color Correlogram

<sup>16</sup> Morphological Operator

<sup>17</sup> Adjacency Graph

اول در سیستم‌های بازیابی نتایج بسیار موفق‌تری داشته و در برخی از سیستم‌های اولیه مانند QBIC استفاده شده‌اند [۱، ۵، ۱۵].

### 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 [۱۷]. Andrysiak در [۱۸] روشی برای استخراج بردار ویژگی بافت از فیلتر گابور آورده است. به گونه‌ای که بردار ویژگی حاصل می‌نیم طول ممکن را دارا باشد.

### ویژگی‌های تبدیل موجک

تبدیل موجک گسسته یک بعدی یا  $DWT^28$ ، یک سیگنال گسسته  $f(x)$  را بر حسب مقادیر تابع مقیاس  $\varphi(x)$  و تابع شیفیت یافته و گسترش یافته  $\psi(x)$  تجزیه می‌کند. رابطه (۱) معرف این تبدیل است.

$$f(x) = \sum_{l \in \mathbb{Z}} S_{j_0, l} \varphi_{j_0, l}(x) + \sum_{j \geq j_0, l} C_{j, l} \psi_{j, l}(x) \quad (1)$$

که در آن  $\varphi_{j, l}$  و  $\psi_{j, l}$  و به ترتیب توسط روابط (۲) و (۳) به دست می‌آیند.

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

$$\psi_{j, l}(x) = 2^{\frac{j-l}{2}} \psi(2^j x - l) \quad (3)$$

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<sup>۲۹</sup> و کوتاه، تعامد، تقارن، گشتاورهای صفر و غیره است؛ اما تمامی خصوصیات فوق کاملاً در موجک‌های کلاسیک برآورده نمی‌شوند. به همین منظور موجک‌های چنددقتی<sup>۳۰</sup> 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 موجک-

های 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*، *Daubechies* که از این میان، فیلتر *Haar* به دلیل سهولت در سیستم‌های بازیابی بیشتر استفاده می‌شود [۱۹].

### 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 خورا<sup>۳۱</sup> [۱۹] مشخصات تصاویر مطلوب خود را اعلام کند. از روش دوم در اکثر سیستم‌های بازیابی تصویر نظیر *PicToSeek* [۸]، *EFF2* [۷]، *QBIC* [۵]، *VisualSEEK* [۶] و *Viper* [۹] استفاده شده است.

در غالب سیستم‌ها برای جستجوی تصاویر و پیدا کردن تصاویر مشابه از جستجوی نزدیک‌ترین همسایه<sup>۳۲</sup> استفاده می‌شود، که بنابه تعریف *Chiueh* در [۱۹]، 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<sup>۳۳</sup> به تصویر مطلوب کاربر استخراج می‌شود، که  $k$  می‌تواند عددی از پیش تعیین شده باشد و یا کاربر آن را مقداردهی کند. در *VisualSEEK* [۱۹] 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<sup>۳۴</sup>.

*Einarsson* در سیستم بازیابی خود به نام *Eff2* [۷] برای پیدا کردن این تصاویر از جستجوی پشت‌سرهم<sup>۳۵</sup> 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 بر محاسباتی بالا عنوان کرده است.

<sup>31</sup> Seed

<sup>32</sup> Nearest-neighbor Search

<sup>33</sup> K-nearest Neighbor

<sup>34</sup> Give me more Query

<sup>35</sup> Sequential

<sup>28</sup> Discrete Wavelet Transform

<sup>29</sup> Compact Support

<sup>30</sup> Multiresolution

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<sup>۳۶</sup> 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 بعضا دارای ابعاد زیاد هستند، باعث بروز مشکل ابعاد بالا می‌شود [۲۰].

هدف از اندیس‌گذاری تصاویر، مرتب کردن بردارهای ویژگی و در نتیجه تصاویر متناظر 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 [۵] یا Gevers SS-Tree در سیستم PicToSeek [۸] ذکر کرده که به دلیل بار محاسباتی مربوط به جستجوی این درخت‌ها، از ساختار SR-tree استفاده کرده است. همچنین نشان داده است که کارایی SR-tree از R\*-tree بیشتر است.

Chiueh و سایرین در [۲۱] 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<sup>۳۷</sup> به عنوان مرکز ثقل فضای ویژگی استفاده کرده است و محاسبات فاصله بر اساس آن انجام گرفته است. 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 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.

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 [۲۲] 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 (۴) تبعیت می‌کنند که به فرم فاصله مینکوفسکی<sup>۳۸</sup> 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.

$$D_M^\rho(\bar{k}, \bar{l}) = \left( \sum_{i=1}^n |k_i - l_i|^\rho \right)^{1/\rho} \quad (۴)$$

فاصله مرتبه اول<sup>۳۹</sup> که برای محاسبه اختلاف هیستوگرام دو تصویر از آن استفاده می‌شود [۲۲، ۱۱] با قرار دادن  $\rho = 1$  در رابطه (۴) به دست می‌آید. استفاده از این معیار در هیستوگرام تصویر باعث بروز منفی نابجا<sup>۴۰</sup> 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 [۲۲] آمده است استفاده می‌شود.

فاصله مرتبه دوم<sup>۴۱</sup> یا فاصله اقلیدسی با قرار دادن  $\rho = 2$  در رابطه (۴) 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 [۱۱].

فاصله کسینوسی<sup>۴۲</sup> که در رابطه (۵) آمده است، بردارهای ویژگی دو تصویر را مقایسه می‌کند و کسینوس زاویه بین دو بردار را به دست می‌دهد.

$$D_C(\bar{k}, \bar{l}) = 1 - \cos\theta \quad (۵)$$

$$\cos\theta = \frac{\bar{k} \cdot \bar{l}}{\|\bar{k}\| \|\bar{l}\|} \quad (۶)$$

<sup>38</sup> Minkowski-form

<sup>39</sup> L<sub>1</sub>-Distance

<sup>40</sup> False Negative

<sup>41</sup> L<sub>2</sub>-Distance

<sup>42</sup> Cosine Distance

<sup>36</sup> Offline

<sup>37</sup> Vantage Point

که در اینجا  $\theta$  زاویه بین دو بردار  $\bar{k}$  و  $\bar{l}$  است. اگر این دو هم جهت باشند، فاصله (۵) برابر یک خواهد شد.  $\Theta$  می تواند تابعی از  $\bar{k}$  و  $\bar{l}$  باشد (رابطه (۶)). 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 (۷)). پیش از به دست آوردن این فاصله، دو هیستوگرام باید نرمالسازی شوند. [۴].

$$D_H(\bar{k}, \bar{l}) = 1 - \frac{\sum_{i=1}^n \min(k_i, l_i)}{\sum_{i=1}^n k_i} \quad (7)$$

### ۳-۴ امتیازدهی به تصاویر

Li در [۲، ۱۱] 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 (۸) آمده است.

$$D_{sim}(\bar{k}, \bar{l}) = \frac{D_{color}(\bar{k}, \bar{l})}{MAX_{d \in \{1, k\}}(D_{color}(\bar{k}, \bar{d}))} + \frac{D_{texture}(\bar{k}, \bar{l})}{MAX_{d \in \{1, k\}}(D_{texture}(\bar{k}, \bar{d}))} \quad (8)$$

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 (۸)، 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 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 ZuBuD شامل ۱۰۰۵ تصویر از ۲۰۱ خانه تشکیل شده که هر خانه از ۵ زاویه مختلف عکس‌برداری شده است.

### ۶- جمع‌بندی

در این مقاله سعی شده است تا روش‌های موجود در زمینه بازیابی تصویر مبتنی بر محتوا با ذکر نمونه بررسی شود. ساختار یک سیستم بازیابی تصویر به دو دسته کلی استخراج ویژگی و جستجو تقسیم می‌شوند. بخش استخراج ویژگی، مربوط به استخراج ویژگی‌های رنگ و بافت می‌باشد. در بخش جستجو، روش‌های الگوریتمی موجود برای ارتقای نتایج جستجو در سیستم‌های بازیابی تصویر بررسی شده است؛ این روش‌ها مربوط به بهینه‌سازی نتایج جستجو با استفاده از روش‌های متنوع اندیس‌گذاری تصاویر و نیز نحوه جستجو در پایگاه داده است. در انتها چند نمونه پایگاه داده که در بررسی نهایی سیستم‌های بازیابی تصویر مورد استفاده قرار گرفته‌اند ذکر شده است.

هرچند روش‌های موجود در توسعه بازیابی تصاویر مبتنی بر محتوای کنونی کارایی مطلوبی نشان داده‌اند، اما این زمینه از جمله زمینه‌های تحقیقاتی بازی است که در مسائلی نظیر افزایش سرعت بازیابی در پایگاه‌های داده اطلاعاتی با حجم تصاویر بسیار زیاد، و نیز افزایش دقت بازیابی به تلاش قابل ملاحظه‌ای نیازمند است.

- [17] Deselaers Th., "Features for Image Retrieval," 2003.
- [18] Bhagavathy S. ; Tesic J. and Manjunath B. S., "On the Rayleigh Nature of Gabor Filter Outputs," in Intl. Conf. on Image Processing (ICIP), 2003.
- [19] Smith J. R. and Chang S., "Tools and Techniques for Color Image Retrieval," in SPIE, 1996, pp. 1630-1639.
- [20] 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.
- [21] Chiueh T., "Content-based image indexing," in Proceedings of VLDB '94, Santiago, Chile, 1994, pp. 582-593.
- [22] Stricker M. A. and Orengo M., "Similarity of Color Images," in SPIE, 1995, pp. 381--392.
- [23] L. W. Kheng, "Color Spaces and Color-Difference Equations," 2002.
- [1] 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.
- [2] 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.
- [3] Einarsson S. H., "Data structures for intermediate search results in the Eff2 image retrieval system," Reykjavik University, technical report 2004.
- [4] Gevers Th. and Smeulders A.W.M., "Image Search Engines, An Overview," The International Society for Optical Engineering (SPIE), vol. VIII, pp. 327--337, 2003.
- [5] Veltkamp and Tanase, "Content-Based Image Retrieval Systems: A Survey," Dept. of Computing Science, Utrecht University, Technical Report 2000.
- [6] Smith J. R. and Chang S. F., "VisualSEEK: A fully automated content-based image query system," in ACM Multimedia Conference. Boston, MA, USA, 1996.
- [7] Einarsson S. H. ; Grétarsdóttir R. Ý. ; Jónsson B. Þ. and Amsaleg L., "The EFF 2 Image retrieval System Prototype," in ASTED Intl. Conf. on Databases and Applications (DBA), Innsbruck, Austria, 2005.
- [8] Gevers Th. and Smeulders A. W. M. , "The PicToSeek WWW Image Search System " in IEEE ICMCS, 1999.
- [9] Squire D. ; Muller W. and Muller H., "Relevance feedback and term weighting schemes for content-based image retrieval," Huijsmans and Smeulders vol. 5, pp. 549-556, 1998.
- [10] Schettini R. ; Ciocca G. and Zuffi S., "A Survey of Methods for Color Image Indexing and Retrieval in Image Databases."
- [11] 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.
- [12] 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. 39-62, 1999.
- [13] Materka A. and Strzelecki M. , "Texture Analysis Methods – A Review," Technical University of Lodz, Institute of Electronics, Brussels, COST B11 1998.
- [14] T. Gevers, "Robust Histogram Construction from Color Invariants," IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI), vol. 26, pp. 113-118, 2004.
- [15] Howarth P. and Ruger S., "Evaluation of Texture Features for Content-Based Image Retrieval," in Third International Conference, CIVR 2004, Dublin, Ireland, 2004.
- [16] 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, 2004.