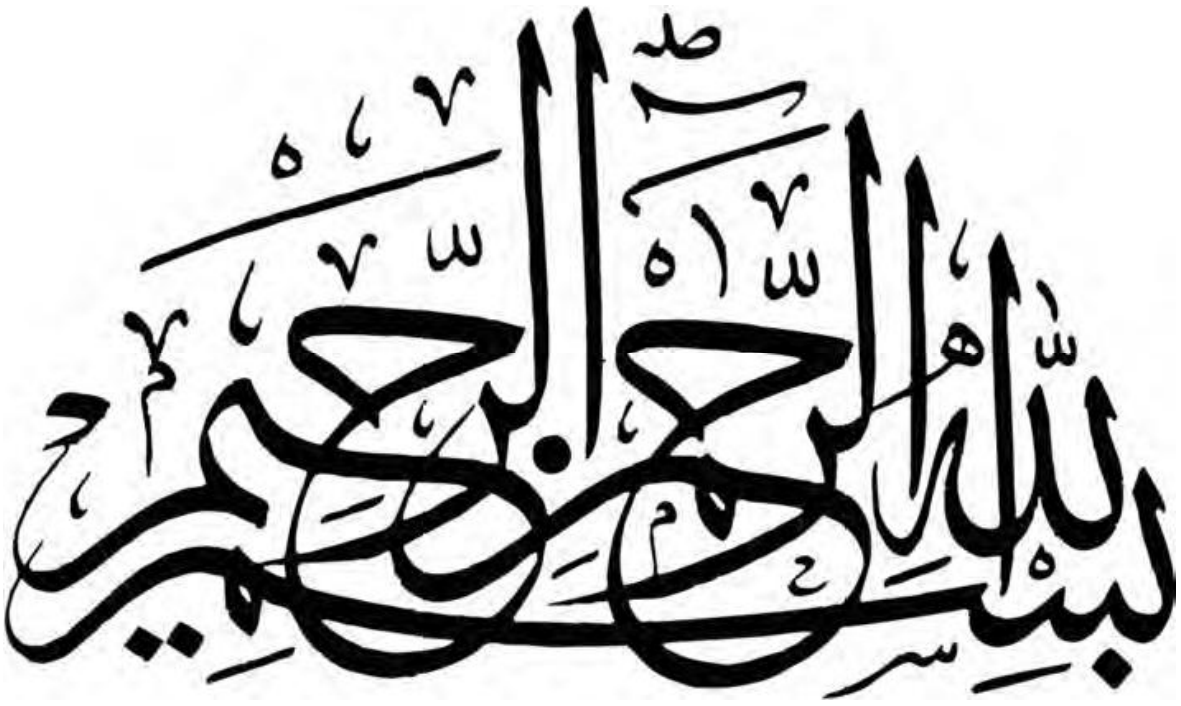

بسم الله الرحمن الرحيم

دانشگاه بیرجند

دانشکده مهندسی برق و کامپیوتر

پایان‌نامه دوره کارشناسی ارشد مهندسی برق- الکترونیک

ارائه یک روش ترکیبی جدید برای بازیابی تصویر مبتنی بر محتوا با استفاده از توصیف- گرهای بهینه

استاد راهنما:

دکتر ناصر مهرشاد

استاد مشاور:

دکتر سید محمد رضوی

نگارش:

محسن داودآبادی فراهانی

مهر ۱۳۹۲

دانشگاه هرمزگان
معاونت تحصیلات تکمیلی

صورتجلسه دفاع از پایان نامه کارشناسی ارشد

فرم شماره ۱۰

با تائیدات خداوند متعال جلسه دفاع از پایان نامه تحصیلی کارشناسی ارشد آقای محسن داودآبادی فراهانی به شماره دانشجویی : ۹۰۱۳۳۱۶۰۱۳ رشته : مهندسی برق گرایش : الکترونیک دانشکده: مهندسی برق و کامپیوتر

تحت عنوان:

ارائه یک روش ترکیبی جدید برای بازیابی تصویر مبتنی بر محتوا با استفاده از توصیفگرهای بهینه

به ارزش : ۶ واحد درساعت : ۱۹ روز : شنبه مورخ: ۱۳۹۲/۷/۲۰

با حضور اعضای محترم جلسه دفاع و نماینده تحصیلات تکمیلی به شرح ذیل تشکیل گردید:

سمت	نام و نام خانوادگی	رتبه علمی	اعضاء
استاد راهنمای اول	دکتر ناصر مهرشاد	دانشیار	
استاد راهنمای دوم			
استاد مشاور اول	دکتر سید محمد رضوی	دانشیار	
استاد مشاور دوم			
داور اول	دکتر سید حمید ظهیری	دانشیار	
داور دوم	دکتر ابوالفضل بنحاری	استادیار	
نماینده تحصیلات تکمیلی	دکتر رمضان هاونگی	استادیار	

نتیجه ارزیابی دفاع که منوط به ارائه اصلاحات پیشنهادی توسط هیئت داوران حداکثر ظرف مدت یکماه پس از تاریخ دفاع می باشد، به شرح زیر مورد تایید قرار گرفت:

☑ قبول (با درجه : و امتیاز : ۱۹) ☐ دفاع مجدد ☐ غیر قابل قبول

۱- عالی (۲۰- ۱۹) ۲- بسیار خوب (۱۸/۹۹- ۱۸) ۳- خوب (۱۷/۹۹- ۱۶) ۴- قابل قبول (۱۵/۹۹- ۱۴)

(بدیهی است عواقب آموزشی ناشی از عدم ارائه به موقع اصلاحات مزبور به عهده دانشجو می باشد)

تقدیم با بوسه بر دستان پدرم:

به او که نمی‌دانم از بزرگی‌اش بگویم یا مردانگی، سخاوت، سکوت، مهربانی و...
پدرم راه تمام زندگیست

تقدیم به مادر عزیزتر از جانم:

مادرم هستی من ز هستی توست

تقدیم به خواهرم و برادرم:

که وجودشان شادی‌بخش و صفایشان مایه آرامش من است

تشکر و قدردانی

بدون شک جایگاه و منزلت معلم، اجل از آن است که در مقام قدردانی از زحمات بی‌شائبه‌ی او، با زبان قاصر و دست ناتوان، چیزی بنگاریم. اما از آن‌جایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می‌کند و سلامت امانت‌هایی را که به دستش سپرده‌اند، تضمین،
برحسب وظیفه و از باب " من لم یشکر المتعم من المخلوقین لم یشکر الله عز و جل"
از پدر ومادر عزیزم، این دو معلم بزرگوارم که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت‌هایم گذشته اند و در تمام عرصه‌های زندگی یار و یاور بی‌چشم داشت برای من بوده‌اند.
از استاد باکمالات و شایسته جناب آقای دکتر ناصر مهرشاد که در کمال سعه‌صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان‌نامه را بر عهده گرفتند.
از استاد صبور و دلسوز جناب آقای دکتر سید محمد رضوی که زحمت مشاوره این پایان‌نامه را متقبل شدند، کمال تشکر و قدردانی را دارم، باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

محسن داودآبادی فراهانی

چکیده

با توجه به کاهش هزینه تهیه و نگهداری تصاویر دیجیتالی و گسترش روز افزون استفاده از این تصاویر، در سال‌های اخیر، حجم و تعداد مجموعه تصاویر فردی و عمومی بخصوص در شبکه جهانی اینترنت، رو به افزایش نهاده است. این امر، نیاز به جستجوی تصاویر دیجیتالی را مطرح ساخته است. در این میان، جستجوی تصاویر و توسعه سیستم‌های با قابلیت بازیابی تصاویر بر اساس محتوا و مشابه درک انسانی توجه محققین را به خود جلب نموده و به یکی از موضوعات مهم پردازش تصویر مبدل شده است. از آن-جا که معیار اولیه انسان برای تشخیص محتوا ویژگی دیداری است، تقریبا در تمامی این سیستم‌ها از این ویژگی‌ها استفاده شده است. در این پایان‌نامه، ساختاری برای پیاده‌سازی یک سیستم بازیابی تصاویر بر اساس ویژگی‌های اولیه انسانی (ویژگی‌های سطح پایین) ارائه می‌شود. در این سیستم به ویژگی‌های رنگ و بافت برای بازیابی تصاویر توجه می‌شود. به طوری که برای استخراج ویژگی رنگ از هیستوگرام رنگ و توصیف‌گر چیدمان رنگ (CLD) و برای استخراج ویژگی بافت از تبدیل موجک و بانک فیلتری استفاده شده است. پس از اعمال روش پیشنهادی بر روی دو پایگاه داده متفاوت، نتایج را با چندین روش بازیابی به کار رفته در مقالات و تحقیقات مقایسه کردیم و نشان دادیم که روش پیشنهادی بازدهی بهتری نسبت به سایر روش‌ها دارد.

کلید واژه‌ها: بازیابی تصویر مبتنی بر محتوا، ویژگی‌های سطح پایین، استخراج ویژگی، رنگ، بافت

فهرست مطالب

عنوان **صفحه**

فهرست جدول‌ها ..... د

فهرست شکل‌ها ..... ه

فصل ۱- مقدمه ۱

1-1- پیشگفتار ۱

۱-۲- تاریخچه و شیوه‌های اولیه‌ی بازیابی تصاویر ۱

۱-۳- انواع سیستم‌های بازیابی تصویر ۴

۱-۴- اهداف و کاربردهای تحقیق ۸

۱-۵- نوآوری تحقیق ۹

۱-۶- ساختار پایان‌نامه ۹

فصل ۲- سیستم بازیابی مبتنی بر محتوای تصاویر و مرور منابع ۱۰

۲-۱- سیستم بازیابی مبتنی بر محتوای تصویر ۱۰

۲-۲- مرور منابع ۱۱

۲-۳- ویژگی‌های مبتنی بر رنگ ۱۲

۲-۳-۱- توصیف فضاهای رنگ ۱۲

۲-۳-۲- ممان رنگ ۲۱

۲-۳-۳- هیستوگرام رنگ ۲۲

۲-۳-۴- هیستوگرام برچسب رنگ ۲۳

۲-۳-۵- بردار ارتباط رنگ ۲۴

۲-۳-۶- همبستگی‌نگار رنگ ۲۴

۲-۳-۷- توصیف‌گر چیدمان رنگ(CLD) ۲۴

۲-۳-۸- توصیف‌گر رنگ غالب (DCD) ۲۸

۲-۳-۹- توصیف‌گر رنگ مقیاس‌پذیر(SCD) ۳۱

۲-۴- ویژگی شکل ۳۲

۲-۵- ویژگی بافت ۳۳

۲-۵-۱- ویژگی‌های فیلتر گابور ۳۴

۲-۵-۲- ماتریس Co-occurrence ۳۴

۲-۵-۳- ویژگی‌های Tamura ۳۴
۲-۵-۴- تبدیل موجک ۳۵
۲-۶- جستجو ۳۶
۲-۷- معیارهای شباهت ۳۷

فصل ۳- روش پیشنهادی ۳۹

۳-۱- مقدمه ۳۹
۳-۲- سیستم CBIR پیشنهادی ۳۹
۳-۲-۱- هم اندازه کردن تصاویر ۴۱
۳-۲-۲- ایجاد صفحات ۴۱
۳-۲-۳- استخراج ویژگی هیستوگرام ۴۱
۳-۲-۴- آستانه گیری ۴۲
۳-۲-۵- تبدیل فوریه ۴۴
۳-۲-۶- عبور از فیلتر بانک ۴۵
۳-۲-۷- تبدیل فوریه معکوس ۴۶
۳-۲-۸- تبدیل موجک ۴۶
۳-۲-۹- حذف جزئیات تصویر ۴۹
۳-۲-۱۰- تبدیل موجک معکوس ۴۹
۳-۲-۱۱- توصیف‌گر چیدمان رنگ (CLD) ۴۹
۳-۲-۱۲- ترکیب طبقه بند ۵۳

فصل ۴- پایگاه‌های داده و نتایج ۵۵

۴-۱- پایگاه‌های داده ۵۵
۴-۱-۱- پایگاه داده ALOI ۵۵
۴-۱-۲- پایگاه داده Corel ۵۶
۴-۲- معیارهای ارزیابی و مقایسه ۵۷
۴-۲-۱- معیارهای دقت و فراخوانی ۵۷
۴-۳- نتایج روش هیستوگرام ۵۸
۴-۴- نتایج روش توصیف‌گر رنگ غالب (DCD) ۵۸
۴-۵- نتایج روش ممان رنگ ۵۸
۴-۶- نتایج روش CLD ۵۹
۴-۷- نتایج روش توصیف‌گر رنگ مقیاس‌پذیر (SCD) ۵۹

۴-۸- نتایج ماتریس Co-Occurrence ۵۹
۴-۹- نتایج روش پیشنهادی ۶۰
۴-۱۰- مقایسه روش‌ها با یکدیگر ۶۰
۴-۱۱- نمونه‌هایی از بازیابی تصویر در پایگاه‌های داده ۶۲
فصل ۵- نتیجه‌گیری و پیشنهادها ۶۴
۵-۱- نتیجه‌گیری ۶۴
۵-۲- پیشنهادها ۶۴

فهرست جدول‌ها

عنوان **صفحه**

جدول۲- ۱: رنگ‌های ادراکی برای نشان دادن تصویر بر اساس رنگ‌های غالب [۳۸] ۳۰

جدول۴- ۱: نتایج روش هیستوگرام ۵۸
جدول۴- ۲: نتایج روش DCD ۵۸
جدول۴- ۳: نتایج روش ممان رنگ ۵۸
جدول۴- ۴: نتایج روش CLD ۵۹
جدول۴- ۵: نتایج روش SCD ۵۹
جدول۴- ۶: نتایج روش ماتریس CO-Occurrence ۵۹
جدول۴- ۷: نتایج روش پیشنهادی ۶۰

فهرست شکل‌ها

عنوان **صفحه**

شکل۱- ۱: نحوه تقسیم‌بندی تصویر به مکان‌های هم‌ارزش و مقدار هر قسمت [۳]. ۲

شکل۲- ۱: شمای کلی یک سیستم بازیابی تصویر ۱۱
شکل۲- ۲: نمایش تصویر در فضای RGB ۱۴
شکل۲- ۳: (الف)شماتیک مکعب رنگ RGB که رنگ‌های اولیه و ثانویه را در رئوس نشان می‌دهد. (ب)مکعب رنگ ۱۴
شکل۲- ۴: نمایش تصویر در فضای YCbCr ۱۵
شکل۲- ۵: نمایش تصویر در فضای CMY ۱۶
شکل۲- ۶: (الف) شش ضلعی رنگ HSV (ب) مخروط شش ضلعی HSV ۱۷
شکل۲- ۷: نمایش تصویر در فضای HSV ۱۷
شکل۲- ۸: **فضای رنگی HMMD** ۱۸
شکل۲- ۹: نمایش تصویر در فضای رنگی HMMD ۱۹
شکل۲- ۱۰: نمایش تصویر در فضای NTSC ۲۰
شکل۲- ۱۱: نمایش تصویر در فضای HSI ۲۱
شکل۲- ۱۲: استخراج ویژگی CLD ۲۵
شکل۲- ۱۳: تقسیم تصویر به ۶۴ بلوک ۲۵
شکل۲- ۱۴: انتخاب رنگ نماینده ۲۶
شکل۲- ۱۵: اسکن کردن زیگزاگی ۲۷
شکل۲- ۱۶: استخراج ویژگی CLD ۲۷
شکل۲- ۱۷: کوانتیزه سازی با استفاده از الگوریتم FRCFE ۳۰
شکل۲- ۱۸: توابع موجک مادر ۳۶

شکل ۳- ۱: بلوک دیاگرام روش پیشنهادی ۴۰
شکل ۳- ۲: تصویر نمونه برای بدست آوردن هیستوگرام ۴۱
شکل ۳- ۳: هیستوگرام تصویر ۳-۲ ۴۱
شکل ۳- ۴: تصویر دسته‌بندی هیستوگرام شکل ۳-۳ ۴۲

شکل ۳- ۵: هیستوگرام مربوط به گروه اول (مقدار ماکزیمم حول صفر و دسته اول) ۴۳
شکل ۳- ۶: هیستوگرام مربوط به گروه اول (مقدار ماکزیمم حول ۲۵۵ و دسته آخر) ۴۳
شکل ۳- ۷: هیستوگرام مربوط به گروه دوم ۴۴
شکل ۳- ۸: تصویر نمونه برای به دست آوردن تبدیل فوریه ۴۵
شکل ۳- ۹: تصویر ضرایب حقیقی تبدیل فوریه دوبعدی شکل ۳-۸ ..... ۴۵
شکل ۳- ۱۰: بانک فیلتر استفاده شده در این پایان نامه ۴۶
شکل ۳- ۱۱: تجزیه موجک تصویر ۴۷
شکل ۳- ۱۲: نمایش چهار جزء ساخته شده توسط الگوریتم موجک ۴۸
شکل ۳- ۱۳: مراحل اجرای CLD ۴۹
شکل ۳- ۱۴: تقسیم‌بندی تصویر نمونه به ۶۴ بلوک ۵۰
شکل ۳- ۱۵: نمایش نحوه انتخاب رنگ نماینده ۵۰
شکل ۳- ۱۶: اجزای تصویر در فضای رنگ YCbCr ۵۱
شکل ۳- ۱۷: نشان دادن خروجی بدست آمده پس از استفاده از تبدیل DCT ۵۲
شکل ۳- ۱۸: نشان دادن خروجی رنگی بدست آمده پس از استفاده از تبدیل DCT ۵۲
شکل ۳- ۱۹: نمایش دو بردار به منظور توضیح نحوه عملکرد ترکیب طبقه بندها ۵۳
شکل ۳- ۲۰: نمایش دو بردار ساخته شده جدید به منظور توضیح نحوه عملکرد ترکیب طبقه بندها..... ۵۴

شکل۴- ۱: تصاویر نمونه از پایگاه داده ALOI ۵۶
شکل۴- ۲: تصاویر نمونه از پایگاه داده Corel ۵۶
شکل۴- ۳: مقایسه عملکرد روش‌های مختلف در پایگاه داده ALOI ۶۱
شکل۴- ۴: مقایسه عملکرد روش‌های مختلف در پایگاه داده Corel ۶۱
شکل۴- ۵: تصاویر بازیابی شده با استفاده از روش پیشنهادی در پایگاه داده ALOI ۶۲
شکل۴- ۶: تصاویر بازیابی شده با استفاده از روش پیشنهادی در پایگاه داده Corel ۶۳

فصل ۱– مقدمه

۱–۱– پیشگفتار

باپیشرفت تکنولوژی دردهه‌های اخیر، هزینه تولید و نگهداری تصاویر دیجیتالی به شدت کاهش یافته است به‌طوری که نسبت به تصاویرآنالوگ از هزینه کمتری برخوردار شده است. به همین سبب تمایل به نگهداری وکار با تصاویر دیجیتال به طور روز افزون و با سرعت بسیار در حال افزایش است. به طوری که مجموعه تصاویر خصوصی و عمومی بزرگی ایجاد شده است. از طرفی به سبب گویایی بالای تصویر نسبت به متن، حتی گاهی افراد از تصویر جهت بیان جملات خود استفاده می‌کنند. این تغییر رفتار نیز مانند هر تغییر رفتار دیگری باعث ایجاد تعدادی نیازمندی می‌گردد. یکی از نیازهای ایجاد شده پیدا نمودن تصویرهای مشابه و دسته‌بندی آن‌ها بر اساس محتوا بود. به همین جهت بازیابی تصاویر بر اساس محتوا مورد توجه محققین قرار گرفته و امروزه به یکی از مسائل اصلی پردازش تصویر تبدیل شده و هنوز راه حل جامع و کارایی نیز برای آن ارائه نگردیده است و اکثر راه حل‌های موجود نیز بر روی مجموعه تصاویر خاص و کوچک کارایی خوبی داشته‌اند و با بزرگتر شدن مجموعه تصاویر کارایی نیز افت محسوس و گاه زیادی از خود نشان می‌دهند.

۱–۲– تاریخچه و شیوه‌های اولیه‌ی بازیابی تصاویر

اولین سیستم بازیابی تصویر در اواخر دهه ۷۰ مطرح گردید. در آن سیستم دسته‌بندی تصاویر بدون توجه به ویژگی‌های دیداری آن‌ها و تنها بر اساس حاشیه‌نویسی متنی انجام می‌گرفت، به این ترتیب که هر تصویر توسط اپراتور بازدید شده و اپراتور تعدادی کلمه کلیدی برای آن عکس انتخاب می‌نمود و عمل بازیابی بر اساس آن کلمات انجام می‌گرفت .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 [1]گفته می‌شود [۱].

این سیستم‌ها با چند مشکل اساسی مواجه بودند:

۱- مضمون یک تصویر بیشتر از این است که بوسیله یک سری کلمات کلیدی بتوان آن را بیان کرد [۱].

۲- حاشیه نویسی به وسیله‌ی انسان کاری بسیار سخت، وقت‌گیر و مستلزم هزینه بسیار است.

۳- بی‌دقتی در حاشیه نویسی، که از ادراک و آگاهی انسان ناشی می‌شود [۲].

۴- نبودن عبارات و اصطلاح‌های دقیق در تفسیر مضمون تصاویر است.

[1] Text-based Image Retrieval

۵- یکسان نبودن مفاهیم موجود در یک تصویر از دید کاربران مختلف.

بنابراین حاشیه‌های تصاویر تمامی حیطه جستجو را نمی‌پوشاندند و این به این معنا است که جستجوهای مبتنی بر متن به میزان کافی کامل و گویا نیستند. پس از آن استفاده از مفاهیم دیداری مانند رنگ، شکل و بافت مورد استفاده قرار گرفت بدین صورت که کاربر یک یا چند ویژگی دیداری را انتخاب کرده و سپس مقدار یا محدوده ای برای آن‌ها تعیین می‌نمود. سپس سیستم با توجه به این مفاهیم عمل بازیابی را انجام می‌داد. بعدها در برخی از سیستم‌ها اولویت‌بندی این ویژگی‌ها نیز به سیستم اضافه گردید اما رضایت‌مندی کاربران و محققین حاصل نگردید. به همین جهت روش سیستم‌ها عوض گردید به‌طوری که امکاناتی در اختیار کاربر قرار گرفت تا بتواند شمایی از تصویر کلی مورد نظر را رسم نماید و بر اساس آن جستجو انجام گیرد. با وجود افزودن دریافت باز خورد کاربر و یادگیری به این سیستم‌ها به علت زمان‌بر بودن رسم تصویر و مشکلات آن بخش این سیستم‌ها نیز مورد توجه قرار نگرفتند و به سرعت به حیطه فراموشی سپرده شدند.

به ناچار محققین به سمت روش قبلی برگشتند و مفاهیم بیشتری را همراه با تحلیل به سیستم افزودند. به علت آنکه مفاهیم جدید همراه با تحلیل‌های انسانی بود مفاهیم دیداری، ویژگی‌های سطح پائین و مفاهیم جدید، ویژگی‌های سطح بالا نام‌گذاری گردید.

در سیستم‌های بازیابی تصاویر ابتدا تصویر به صورت یک‌جا و یک موجودیت در نظر گرفته شده و ویژگی‌های کل تصویر استخراج و مورد استفاده قرار می‌گرفت. پس از آن برای افزایش کارایی، تصویر به بخش‌های ثابت، یکسان و از پیش مشخص شده‌ای تقسیم گردید. با توجه به این‌که ارزش هر مکان در تصویر با مکان دیگر متفاوت است این روش به سرعت تبدیل به روش تقسیم تصویر به بخش‌های ثابت، از پیش مشخص شده و هم‌ارزش تبدیل شد و بسیاری از کارهای قبلی روش تقسیم‌بندی زیر که در [۳] بیان شده است را در کارهای خود استفاده نموده‌اند. اعداد درون هر قسمت نشان دهنده ضریب اهمیت آن قسمت هستند.

۱/۱۶	۱/۸	۱/۱۶
۱/۸	۱/۴	۱/۸
۱/۱۶	۱/۸	۱/۱۶

شکل۱- ۱: نحوه تقسیم‌بندی تصویر به مکان‌های هم‌ارزش و مقدار هر قسمت [۳].

این روش نیز اندکی کارایی را افزایش داد اما محققین را راضی ننمود و آن‌ها برای افزایش کارایی به سمت تقسیم تصویر به قسمت‌های از پیش نامعلوم سوق پیدا نمودند. به همین سبب قطعه‌بندی[1] و روش‌های آن به این موضوع راه پیدا نمودند. «قطعه‌بندی تصویر یک تکنیک و پردازشی است که تصویر را بر اساس ویژگی‌های مختلف به نواحی مختلفی تقسیم می‌کند» [۲]. ناحیه‌ها در واقع اشیاء متفاوت موجود در تصویر هستند که از نظر بافت یا رنگ تفکیک شده‌اند [۴]. «هدف از قطعه‌بندی، ساده‌سازی و یا تقسیم تصویر به قسمت‌های با معناتر و ساده‌تر است» [۳]. امروزه در بسیاری از سیستم‌های بازیابی تصویر بخشی جهت قطعه‌بندی درنظر گرفته شده است. به‌طوری که در ابتدا تصویر را به چند قطعه تقسیم می‌نمایند.

در اوایل سال ۱۹۹۰ افزایش سریع تصاویر با حجم بالا نظیر شبکه اینترنت و همچنین معایب مذکور، باعث معرفی سیستم‌های بازیابی مبتنی بر محتوای تصویر[2]شد. تکنیک‌های مبتنی بر محتوا در بازیابی تصویر بجای یادداشت‌گذاری بصورت متن، از خصوصیت‌های بصری برای تفسیر مفاد تصاویر استفاده می‌کند. اولین مشکل در هنگام ایجاد یک سیستم بازیابی تصویر فاصله معنایی است. این به عنوان « عدم انطباق بین اطلاعات مجازی دریافتی یک شخص با شخص دیگر تحت همان شرایط » تعریف می‌شود [۵]. همانطور که دو نفر توصیف‌های متفاوتی از یک منظره ارائه می‌دهند، یک توصیف محتوایی از یک تصویر می‌تواند با توصیف یک کاربر انسانی تفاوت زیادی داشته باشد، این عمل به CBIR مربوط است.

سیستم‌های CBIR توصیف‌های قابل تفسیر توسط ماشین را از ویژگی‌های فیزیکی یک تصویر تولید می‌کنند. این توصیف‌ها، به عنوان خصوصیات استخراج شده، می‌تواند به وسیله مقیاسی از تشابه مقایسه شود.

سیستم‌های بازیابی تصویر از سه بخش عمده تشکیل شده است. بخش اول، بخش استخراج ویژگی[3] است که بردار ویژگی[4] هر تصویر در پایگاه داده را تولید می‌کند و محتوای تصاویر را برای دسته‌بندی به نمایش می‌گذارد. بخش دوم اندیس‌گذاری[5] نامیده می‌شود که وظیفه دسته‌بندی تصاویر بر اساس ویژگی‌های استخراج شده را دارد. بخش سوم بازیابی[6] می‌باشد که تصویر مورد جستجو را پردازش می‌کند و با استفاده از معیار شباهت، بردارهای ویژگی تصویر مورد نظر و تصاویر پایگاه داده، فاصله بین تصویر مورد جستجو و تصاویر موجود در پایگاه داده را محاسبه می‌کند و تصاویر مشابه را نشان می‌دهد و واسط کاربری را فراهم می‌کند.

---

[1] Segmentation
[2] Image retrieval
[3] Feature extraction
[4] Feature vector
[5] Indexing
[6] Retrieval

۱-۳- انواع سیستم‌های بازیابی تصویر

سیستم‌های بازیابی تصویر به روش‌های مختلفی دسته‌بندی می‌گردند. جورز[1] سیستم‌های بازیابی تصویر را بر اساس کاربرد به سه دسته کلی تقسیم کرده است [۶]:

۱- **جستجو بر اساس تداعی**[2]: هدف از این روش بدست آوردن تصاویر مورد نظر کاربر از بین تصاویر موجود چند مجموعه تصویر نامشخص است.

۲- **جستجوی یک تصویر خاص**[3]: هدف از این روش بدست آوردن تصاویر مشابه با یک تصویر مشخص از مجموعه 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 وجود داشته باشد.

۳- **جستجوی تصویرهای یک دسته خاص**[4]: هدف از این روش بدست آوردن تصویری است که متعلق به دسته یا کلاس بخصوصی باشد. مانند تصاویر دکوراسیون منزل یا تصاویر رادیولوژی.

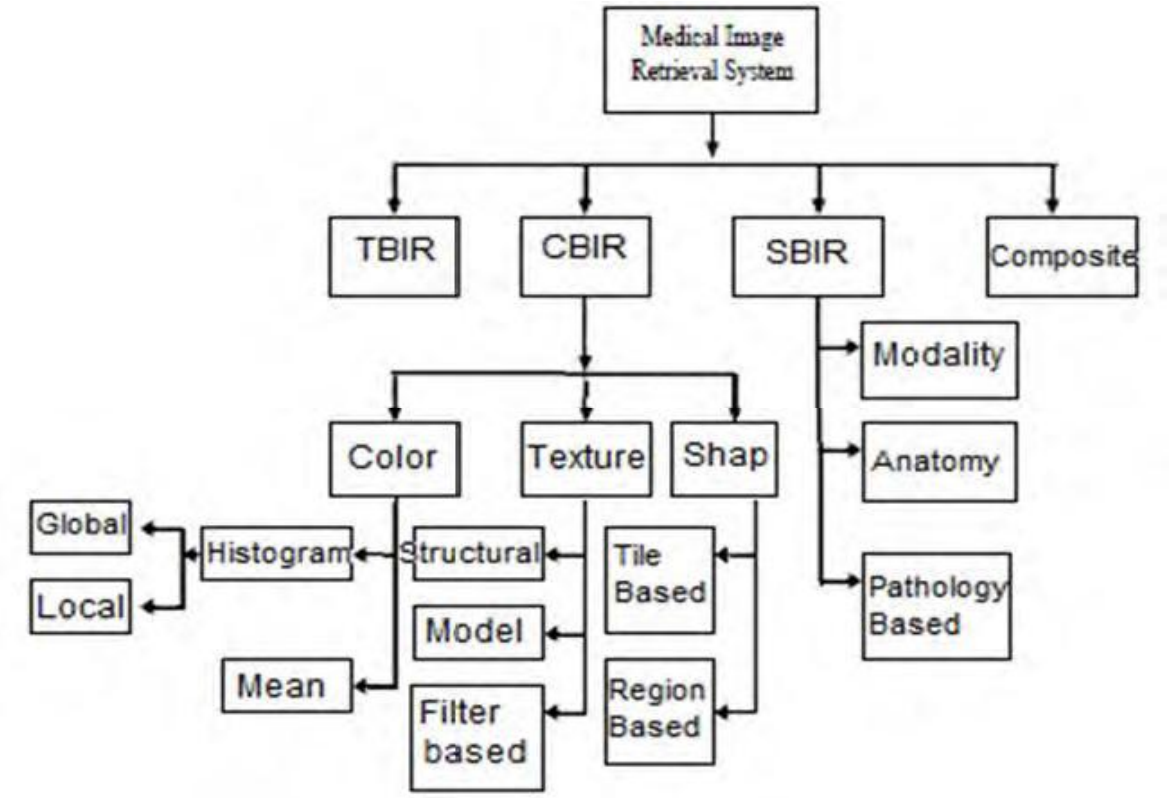
وی همچنین پایگاه داده تصاویر را بر اساس چگونگی میزان تغییر در ویژگی‌ها به دو دسته تقسیم کرده است:

۱- **دامنه محدود:** شامل تغییرات محدود در ویژگی‌های دیداری تصاویر (مانند تصاویر پزشکی)

۲- **دامنه وسیع:** شامل تغییرات وسیع و غیر قابل پیش‌بینی در تصاویر (مانند پایگاه تصاویر موجود در اینترنت)

میناکشی[5] انواع سیستم‌های بازیابی تصاویر را مطابق شکل ۱-۲ دسته‌بندی نموده است[۷]. همان‌طور که در شکل نیز مشاهده می‌شود، میناکشی سیستم‌های بازیابی تصویر را به چهار دسته تقسیم‌بندی نموده است. همان‌گونه که در شکل ۱-۲ قابل مشاهده است روش‌های مختلف بدست آوردن ویژگی‌های رنگ، بافت و شکل در دسته‌بندی سیستم‌های بازیابی تصویر قرار داده شده است.

[1] Gevers
[2] Search by association
[3] Target search
[4] Category search
[5] Meenakshi



شکل۱-۱: دسته‌بندی سیستم‌های بازیابی تصاویر به نقل از [۷]

پیش از شروع کار بازیابی، ویژگی‌های تصاویر پایگاه داده باید از آن‌ها استخراج شود. یک تصویر آرایه-ای دو بعدی از پیکسل‌ها است. برای سهولت دسترسی به ویژگی‌های آن تصویر، به جای استفاده از مقادیر پیکسل‌ها، با استفاده از روش‌هایی، ویژگی‌های تصویر را استخراج و به صورتی که قابل مقایسه باشد، نمایش می‌دهند. به این روش‌ها استخراج ویژگی می‌گویند. خروجی این روش‌ها عدد یا برداری است که به آن‌ها کلاس ویژگی یا توصیف‌گر می‌گویند(مانند هیستوگرام رنگ). به برداری که از کنار هم قرار دادن این کلاس‌های ویژگی بدست می‌آید، بردار ویژگی تصویر گفته می‌شود [۸].

از لحاظ توصیف تصویر ویژگی‌های تصویر به صورت زیر تقسیم می‌شود:

۱- ویژگی‌های عمومی یا سطح پایین: معرف کلیات تصویرند و توانایی توصیف اشیا یا مفاهیم موجود در تصویر را ندارند. رنگ[1]، بافت[2]، شکل[3] چند نوع از این ویژگی‌ها هستند. نخستین کار در این زمینه توسط چانگ در سال ۱۹۸۴ انجام شده است [۹].

۲- ویژگی‌های معنایی[1]: مفاهیم و اشیای موجد در تصویر را بیان می‌کند و اکثرا از روش‌های قسمت‌بندی تصویر برای این کار استفاده می‌کنند.

[1] color
[2] texture
[3] shape

۳- ویژگی‌های مختص کاربردهای خاص: نظیر ویژگی‌های توصیفی چهره در سیستم‌های بازیابی تصاویر مربوط به چهره [۷].

همان‌طور که گفته شد حضور بر چسب قابل اعتماد همراه با تصویر برای بازیابی تصویر بر پایه متن ضروری است. آینده سیستم‌های بازیابی در بستر بهره‌وری از دو نوع تکنولوژی متنی و محتوایی است. تا زمانی که نمونه‌های قبلی از نظر کاربر قابل اعتمادترند، پتانسیل زیادی سر ترکیب این دو وجود دارد که موجب ساخت موتور جستجویی قوی که بتواند به تصاویر پنهان در جاهایی از وب دسترسی پیدا کند می‌شود.

نمونه‌های زیادی برای طرح یک واسط به در خواست یک سیستم CBIR، پیشنهاد شده است. واسط دروازه‌ای به آرشیو تصویر است. یک واسط کاربری، توانایی مرور همه تصویرهای آرشیو را دارد و به کاربر امکان می‌دهد تصویر دلخواه را از آرشیو تصویر بازیابی کند. واسط‌های بعدی تصویرها را سازمان‌دهی می‌کنند یا قابلیت‌های بازیابی را فراهم می‌کنند. واسط‌ها را می‌توان به صورت زیر تقسیم کرد:

۱- مرورگری[۲]: این آسان‌ترین راه برای دسترسی به آرشیو تصویر است. در [۱۰] با پشت سر گذراندن یک ساختار مرورگری، تصویرها را در گروه‌هایی (گاهی با خوشه‌بندی) جهت بررسی کاربر سازمان دهی شده‌اند.

۲- دسته‌های سفارشی[۳]: تصویرها به صورت سلسله مراتبی، در دسته‌های مخصوص دامنه ساختاربندی می‌شوند. [۷] یک مثال از سیستم مبتنی بر هستی شناسی را توصیف می‌کند. سلسله مراتب می‌تواند شامل چندین دسته‌ی معنادار باشند. به طور مثال یک دسته سطح بالا می‌تواند وسایل نقلیه باشد و شامل دسته‌هایی مانند هواپیما و ماشین‌ها شود و هر یک نیز می‌تواند شامل زیردسته‌های دیگر باشد.

۳- درخواست با تصویر نمونه (QBE)[۴]: این نمونه‌ی جستجوی تصویر مبتنی بر محتوا به طور سنتی است. کاربر یک تصویر نمونه با هدف بازیابی تصویرهای مشابه از سیستم، فراهم می‌کند. نمونه‌ای از یک سیستم که امکان جستجو با این الگو را می‌دهد، درخواست با محتوای تصویر (QBIC)[۵] است [۱۱]. این شیوه نقاط ضعفی دارد. در واقع، بازیابی از طریق تصویر نمونه مستلزم این است که کاربر قبل از بازیابی، تصویری را به عنوان نماینده بازیابی کند. این تصویر شاید از مجموعه‌ی دیگری باشد که از طریق ابزارهای ماوراء سیستم CBIR خاص فراهم شده باشد. این می‌تواند یک دغدغه باشد زیرا کاربر باید از ابزارهای دیگری برای

---

[1] Semantic
[2] Browsing
[3] Customized categories
[4] Qurey By Example image
[5] Query By Image Content

جستجوی تصاویر استفاده کند. همینطور، شاید تصویر نمونه از همان آرشیو تصویر که بوسیله سیستم CBIR به کار رفته، باشد، در این حالت الگوی دیگری از بازیابی همانند مرورگری باید پیاده شود یا تصاویر درخواست شوند تا زمانی که کاربر انتخاب کند و با مثال تطابق دهد.

۴- درخواست با ناحیه تصویر[1]: درخواست‌ها می‌تواند براساس یک زیرمجموعه تعریف شده توسط کاربر یا سیستم از یک ناحیه سراسری شکل بگیرد. برای انجام این شیوه، به کاربر باید اجازه داده شود یک ناحیه از تصویر را به صورت دستی تعریف کند، یا یک شیوه‌ی قسمت‌بندی باید داخل سیستم قرار داده شود [۱۲و ۱۳]. هدف از این الگوی تصویر اصلاح نتایج بازیابی بوسیله درخواست بر مبنای مطابق‌ترین بخش داده شده است.

۵- درخواست به وسیله‌ی چندین تصویر نمونه(QBME)[2]: کاربر می‌تواند چندین تصویر نمونه برای سیستم فراهم کند [۱۴]. اشتراکات بین همه تصویرهای درخواست را می‌توان به عنوان مبنای جستجو به کار برد.

۶- درخواست با تعیین مستقیم ویژگی‌های بصری: این حرفه‌ای ترین روش است [۱۵]. کاربر باید مشخصات و مفهوم هر ویژگی بصری مشخص شده را درک کند. این شیوه برای کاربرانی که با طرح و عملکرد داخلی سیستم آشنا نیستند، ممکن است دشوار باشد.

۷- درخواست با طرح بصری: چندین پیاده‌سازی، ابزار ترسیم را برای کاربر جهت ایجاد یک تصویر دلخواه فراهم می‌کنند [۱۶و ۱۷]. این در غیاب تصویر نمونه مفید است. چالش این شیوه این است که اعتماد به توانایی‌های هنری کاربر، به یکی از واسطه‌های پر مطالبه منتهی می‌شود.

۸- درخواست با کلمه‌ی کلیدی: اگر تصویرها قبلا یادداشت‌گذاری شده‌اند یا متن موجود است می‌توان آن‌ها را با متن جستجو کرد. جستجوی تصویر گوگل نمونه موفق این روش است [۱۸]. جستجوی تصویر گوگل به‌طور خودکار تصویرها را با استفاده از متن پیرامون تصویر در صفحه وب نام‌گذاری می‌کند. همچنین سیستم‌های بازیابی با کلمه کلیدی می‌توانند بر نام-گذاری دستی تصویرهای شخصی به‌وسیله انسان تکیه کنند این شیوه‌ی موثری است که ممکن است در غیاب محتوای تصویر، سراسری عمل کند، اگر چه همیشه نام‌گذاری تصویرها با یک مدل مناسب میسر نیست. مثلا یک پایگاه داده بزرگ زمان زیادی برای نام‌گذاری دستی نیاز دارد.

[1] Query By image region
[2] Query By Multiple Example image

۹- باریابی چند کیفیتی[1]: روش‌هایی هستند که چند کیفیت (لامسه، صدا، حرکات بدن و ...) را ترکیب می‌کنند. ابداعات اخیر، همانند حوزه وسیع نمایشگرهای مالتی تاچ، بر افزایش کارکرد درخواست‌های مالتی مدال یا درخواست‌هایی که چندین شکل ورودی را پیوند می-دهند، دلالت می‌کند [۱۹].

در دهه‌های گذشته چندین نمونه محصول تجاری و سیستم تجربی توسعه داده شده‌اند که برخی از آن‌ها عبارتند از:QBIC بوسیله‌ی شرکت IB.M [۱۱]، Netra بوسیله‌ی دانشگاه کالیفرنیا [۱۳]، Photobook بوسیله‌ی آزمایشگاه رسانه‌ای MIT [۲۰]، [۲۱]، VisualSEEK بوسیله‌ی دانشگاه کلمبیا [۲۲]، SIMPLcity [۲۳].

۱-۴- اهداف و کاربردهای تحقیق

با وجود تلاش‌های انجام گرفته در سال‌های اخیر در زمینه بازیابی تصاویر، هنوز الگوریتمی قابل قبول و جامع برای مشخص کردن تصورات انسانی به خصوص در زمینه تفسیر تصاویر وجود ندارد. بخاطر طبیعت تکنولوژی CBIR، عملکرد آن موجب دو نوع مشکل می‌شود:

۱)چگونه یک تصویر را به صورت ریاضی بیان کنیم؟

۲) چگونه شباهت بین دو تصویر را که بر پایه توصیفات انتزاعی است بسنجیم؟

پرسش اول به این خاطر است که تصویر اصلی توسط آرایه‌ای از مقادیر پیکسل‌ها که با واکنش‌های دیداری هم‌خوانی کمی دارد به وجود می‌آید. حتی اگر فهم معنایی تصویر را به حال خود واگذاریم، ما برای اهداف بازیابی به توصیف ریاضی تصویر رجوع می‌کنیم که به عنوان ویژگی‌ها به‌کار می‌روند. چشم-انداز طرح از لحاظ استخراج ویژگی و محاسبات تصاویر مشابه به طور واضحی جدایی ناپذیر است. تدوین ویژگی‌هایی که به مقدار زیاد، حدودی را برای تعریف اندازه‌ی شباهت‌ها تعیین می‌کنند.

هدف این پایان‌نامه ارائه روشی سریع جهت بازیابی تصاویر طبیعی بر اساس ویژگی‌های اولیه وسطح پائین می‌باشد. همان‌طور که می‌دانید دستیابی به نتیجه زمانی مفید و کارا است که در بازه زمانی معقول و محدودی به جواب برسد. با توجه به کارایی محدود ویژگی‌های سطح پائین، برای افزایش بازدهی استفاده از ابزارها و ویژگی‌های دیگر مورد نیاز است. به‌طوری که اکنون علاوه بر ویژگی‌های سطح پائین از ویژگی‌های سطح بالا و یا برچسب‌گذاری اشیاء تصویر[2] استفاده شده ولی متاسفانه هنوز نتیجه ایده‌آل حاصل نگردیده است. به همین سبب یکی از زیر هدف‌های این پایان نامه دستیابی به کارایی قابل قبول در زمان بسیار اندک می‌باشد تا امکان افزایش ویژگی‌های سطح بالا و به طبع آن افزایش کارایی میسر گردد. دراین پایان‌نامه از دو ویژگی اصلی دیداری پر کاربرد یعنی رنگ و بافت استفاده شده است.

[1] Multimodal query
[2] Annotation