بسم الله الرحمن الرحيم

دانشگاه بیرجند

دانشکده مهندسی

پایان‌نامه دوره کارشناسی ارشد مهندسی برق - مخابرات

بازیابی تصاویر با استفاده از تبدیل ویولت وفقی

سجاد محمدزاده

استاد راهنما:

دکتر حسن فرسی

استاد مشاور:

دکتر ناصر مهرشاد

تابستان ۱۳۹۱

تقدیم به ساحت مقدس حضرت فاطمه (س) و امام زمان (عج)

اللّهُمَّ عَجِّل لِوَلِیِّکَ الفَرَج

پدر و مادر عزیزم تشکر بی مقدار است مقابل گذر بهارهای پیاپی عمرتان، گاهی معانی در قالب واژه ها نمی گنجد. هرچند واژه ها ناتوانند برای بیان سال ها مهر، سال ها تلاش و سال ها تحمل؛ برای بیان لحظاتی که با گذر هر ثانیه اش تندیس مقدس وجودتان ترک می خورد و به آینه های چشمانتان غبار می نشست اما امید دارم اکنون که فرصتی فراهم است که هرچند عاجزانه ولی از صمیم قلب سپاسگزار زحمات بی شائبه و بی دریغتان باشم تنها برای لحظه ای لبخند به لبانتان بیاید.

همسر عزیزم، که سایه مهربانیت سایه سار زندگیم است، تقدیم به تو که به صمیمیت بارانی، تو که پناه خستگیم و امید بودنم هستی.

خواهر مهربانم، تقدیم به تو که وجودت شادی بخش و صفایت مایه آرامش من است.

تشکر و قدردانی

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت های او ندانند و کوشندگان، حق او را گزاردن نتوانند. و سلام و دورد بر محمّد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است.

با سپاس و قدردانی ازپدر و مادر عزیزم، که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاوری بی چشم داشت برای من بوده اند و با تشکر فراوان از استاد با کمالات و شایسته، جناب آقای دکتر حسن فرسی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان نامه را بر عهده گرفتند و نیز با سپاس گذاری از استاد صبور و با تقوا، جناب آقای دکتر ناصر مهرشاد، که با صبر و محبت مشاوره این تحقیق را تقبل فرمودند. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

سجاد محمدزاده

چکیده

افزایش روزافزون تصاویر دیجیتال از اماکن و محیط ها و اشیای مختلف و هزینه پائین ذخیره سازی و سهولت کار در اینترنت و وب سایت ها، نیاز به طراحی یک ماشین جستجوگر تصویر، با توانایی پرداخت به تمام نیازهای کاربر، در سال های اخیر به شدت افزایش داده است. بازیابی تصویر رشته ای است که با جستجوی دیجیتال تصاویر از پایگاه های داده تصاویر سروکار دارد. امروزه تلاش های بسیاری برای افزایش سرعت و بازدهی بازیابی تصاویر انجام می شود. هدف این پایان نامه، ارائه سیستمی مناسب برای بازیابی تصاویر دیجیتال مبتنی بر محتوای تصاویر است که در پایگاه های داده متفاوت بازدهی بیشتری داشته باشد و تصاویر مشابه با تصویر مورد سوال را به بهترین نحو افزایش دهد. در این پایان نامه چند سیستم جدید برای بازیابی تصاویر مبتنی بر محتوا با ترکیب ویژگی های رنگ و بافت و استفاده از تکنیک های مختلف ارائه شده است. از تبدیل موجک گسسته و تبدیل کسینوسی گسسته برای استخراج ویژگی های بافت تصویر و از روش های کلاسه بندی و روش های کاهش بعد در سیستم های بازیابی مبتنی بر محتوای پیشنهادی استفاده کرده ایم. پس از اعمال روش های پیشنهادی بر روی سه پایگاه داده متفاوت، نتایج را با چند روش بازیابی بکار رفته در استاندارد MPEG-7 و تحقیقات و مقالات مقایسه کرده ایم. در بخش نتایج نشان داده ایم که در هر پایگاه داده، یکی از روش های بازیابی بازدهی بهتری نسبت به دیگر روش ها دارد به این معنا که یک روش وجود ندارد که در هر سه پایگاه داده بهترین بازدهی را از خود نشان دهد. در ادامه در هر پایگاه داده سه روشی را که بازدهی بهتری نسبت به بقیه روش ها دارد لیست کرده ایم که نشان می دهد روش پیشنهادی ترکیب HDWT2 و DCT جزء سه روش برتر در هر پایگاه داده است.

کلید واژه‌ها: استخراج ویژگی، بازیابی مبتنی بر محتوای تصاویر، تبدیل موجک گسسته، ماتریس هادامارد

فهرست مطالب

عنوان **صفحه**

فهرست جدول‌ها ..... ه

فهرست شکل‌ها ..... و

فصل ۱- مقدمه ۱

۱-۱- پیشگفتار ۱

۱-۲- تاریخچه و شیوه های اولیه و نوین بازیابی تصاویر ۱

۱-۳- اهداف از انجام تحقیق ۷

۱-۴- نوآوری تحقیق ۸

۱-۵- ساختار پایان نامه ۸

فصل ۲- بازیابی مبتنی بر محتوای تصویر و روش های مطالعه شده ۱۰

۲-۱- سیستم بازیابی مبتنی بر محتوای تصویر ۱۰

۲-۲- روش های مطالعه شده ۱۱

۲-۳- ویژگی های مبتنی بر رنگ ۱۲

۲-۳-۱- توصیف فضاهای رنگی ۱۲

۲-۳-۱-۱- فضای رنگ RGB (قرمز، سبز، آبی) ۱۳

۲-۳-۱-۲- فضای رنگ HSV ۱۴

۲-۳-۱-۳- فضای رنگ HMMD ۱۶

۲-۳-۱-۴- فضای رنگ YCbCr ۱۷

۲-۳-۱-۵- فضای رنگ CMY و CMYK (CMY – سیاه) ۱۷

۲-۳-۱-۶- فضای رنگ L*a*b* ۱۸

۲-۳-۲- ممان رنگ ۱۸

۲-۳-۳- هیستوگرام رنگ ۱۹

۲-۳-۴- توصیف گر چیدمان رنگ (CLD) ۲۱

۲-۳-۴-۱- تقسیم تصویر به بلوک ها ۲۲

۲-۳-۴-۲- انتخاب رنگ نماینده ۲۲

۲-۳-۴-۳- اجرای DCT ۲۳

۲-۳-۴-۴- انتخاب ویژگی ۲۳

۲-۳-۴-۵- تطبیق شباهت ۲۴
۲-۳-۵- توصیف گر رنگ غالب (DCD) ۲۵
۲-۳-۵-۱- روش GLA ۲۵
۲-۳-۵-۲- روش FRCFE ۲۶
۲-۳-۵-۳- تطبیق شباهت ۲۸
۲-۳-۶- توصیف گر رنگ مقیاس پذیر (SCD) ۲۹
۲-۴- ویژگی های بافت ۲۹
۲-۴-۱- ماتریس Co – Occurrence ۳۰
۲-۴-۱-۱- لبه یابی ۳۱
۲-۴-۱-۲- استخراج بافت از لبه ۳۲
۲-۴-۲- ویژگی های Tamura ۳۲
۲-۴-۳- تبدیل موجک ۳۳
۲-۴-۳-۱- تبدیل موجک گسسته ۳۳
۲-۵- ویژگی شکل ۳۶
۲-۶- معیارهای شباهت ۳۶
فصل ۳- سیستم پیشنهادی برای بازیابی تصویر مبتنی بر محتوای ۳۹
۳-۱- مقدمه ۳۹
۳-۲- سیستم CBIR پیشنهادی با ویژگی HDWT1 ..... ۳۹
۳-۲-۱- کوچک کردن تصاویر ۳۹
۳-۲-۲- تولید صفحات ۳۹
۳-۲-۳- استخراج ویژگی HDWT1 ..... ۴۰
۳-۲-۳-۱- اجرای تبدیل موجک گسسته ۴۱
۳-۲-۳-۲- ساخت مولفه تغییر یافته با ماتریس هادامارد ۴۳
۳-۲-۳-۳- اجرای معکوس تبدیل موجک ۴۴
۳-۲-۳-۴- کاهش نمونه ها ۴۴
۳-۲-۳-۵- اجرای دوباره مراحل ۴۴
۳-۲-۴- مقیاس شباهت ۴۵
۳-۳- استفاده از شبکه عصبی در روش HDWT2 ..... ۴۵
۳-۳-۱- کوچک کردن تصاویر ۴۵
۳-۳-۲- تولید صفحات ۴۶
۳-۳-۳- استخراج ویژگی HDWT2 ..... ۴۶

۳-۳-۴- طبقه بندی شبکه عصبی ۴۷
۳-۳-۴-۱- مروری بر شبکه های عصبی ۴۷
۳-۳-۴-۲- شبکه عصبی به کار رفته ۵۰
۳-۴- ترکیب روش HDWT2 و تبدیل کسینوسی گسسته ۵۱
۳-۴-۱- استخراج ویژگی روش HDWT2 ۵۲
۳-۴-۲- استخراج ویژگی DCT پیشنهادی ۵۳
۳-۴-۲-۱- کوچک کردن تصویر ۵۳
۳-۴-۲-۲- اجرای DCT ۵۴
۳-۴-۲-۳- تقسیم تصویر به بلاک ها ۵۵
۳-۴-۲-۴- بردار ویژگی DCT ۵۶
۳-۴-۳- کاهش بعد بوسیله آنالیز مولفه اصلی ۵۶
۳-۴-۳-۱- کاهش ابعاد داده مبتنی بر انتخاب ویژگی ۵۷
۳-۴-۳-۲- کاهش ابعاد داده مبتنی بر استخراج ویژگی ۵۷
۳-۴-۴- طبقه بند k- همسایه نزدیکتر ۶۰
فصل ۴- نتایج بدست آمده و پایگاه های داده ۶۲
۴-۱- پایگاه های داده ۶۲
۴-۱-۱- پایگاه داده Corel ۶۳
۴-۱-۲- پایگاه داده ALOI ۶۳
۴-۱-۳- پایگاه داده MPEG-7 ۶۴
۴-۲- معیارهای ارزیابی ۶۵
۴-۳- نتایج روش ممان رنگ ۶۶
۴-۳-۱- نتایج روش Intensity- ممان رنگ ۶۶
۴-۳-۲- نتایج روش RGB- ممان رنگ ۶۶
۴-۳-۳- نتایج روش ممان رنگ در پایگاه داده MPEG-7 ۶۶
۴-۴- نتایج روش هیستوگرام رنگ ۶۷
۴-۴-۱- نتایج روش Intensity- هیستوگرام ۶۷
۴-۴-۲- نتایج روش RGB- هیستوگرام ۶۷
۴-۴-۳- نتایج روش هیستوگرام رنگ در پایگاه داده MPEG-7 ۶۷
۴-۵- نتایج روش CLD ۶۸
۴-۶- نتایج روش DCD ۶۸

۴-۷- نتایج روش SCD ۶۸
۴-۸- نتایج ماتریس Co – Occurrence ۶۹
۴-۹- نتایج روش HDWT1 ..... ۶۹
۴-۹-۱- نتایج روش Intensity-HDWT1 ..... ۶۹
۴-۹-۲- نتایج روش RGB-HDWT1 ..... ۷۱
۴-۹-۳- نتایج روش HMMD-HDWT1 ..... ۷۲
۴-۹-۴- نتایج روش HDWT1 در پایگاه داده MPEG-7 ..... ۷۳
۴-۱۰- نتایج روش HDWT2 ..... ۷۴
۴-۱۰-۱- نتایج روش Intensity-HDWT2 ..... ۷۴
۴-۱۰-۲- نتایج روش RGB-HDWT2 ..... ۷۴
۴-۱۰-۳ نتایج روش HDWT2 در پایگاه داده MPEG-7 ..... ۷۵
۴-۱۱- نتایج روش HDWT2 با شبکه عصبی ۷۵
۴-۱۱-۱ نتایج روش Intensity-HDWT2 با شبکه عصبی ۷۵
۴-۱۱-۲ نتایج روش RGB-HDWT2 با شبکه عصبی ۷۶
۴-۱۱-۳ نتایج روش HDWT2 با شبکه عصبی در پایگاه داده MPEG-7 ..... ۷۷
۴-۱۲- نتایج روش DCT ۷۷
۴-۱۲-۱- نتایج روش Intensity-DCT ۷۷
۴-۱۲-۲- نتایج روش RGB-DCT ۷۷
۴-۱۲-۳- نتایج روش DCT در پایگاه داده MPEG-7 ..... ۷۸
۴-۱۳- نتایج ترکیب روش های HDWT2 و DCT ۷۸
۴-۱۴- مقایسه تمام روش ها ۷۸
۴-۱۵- مثال هایی از بازیابی تصویر در پایگاه های داده ۸۰

فصل ۵- نتیجه‌گیری و پیشنهادها ۸۳

۵-۱- نتیجه‌گیری و جمع بندی ۸۳
۵-۲- پیشنهادات ۸۴

فهرست مراجع ۸۵

واژه نامه فارسی به انگلیسی ۹۰

واژه نامه انگلیسی به فارسی ۹۳

فهرست جدول‌ها

عنوان **صفحه**

جدول ۴- ۱: نتایج Intensity- ممان رنگ در پایگاه های داده. ۶۶
جدول ۴- ۲: نتایج RGB- ممان رنگ در پایگاه های داده. ۶۶
جدول ۴- ۳: نتایج ممان رنگ در پایگاه داده MPEG-7 . ۶۷
جدول ۴- ۴: نتایج Intensity- هیستوگرام در پایگاه های داده. ۶۷
جدول ۴- ۵: نتایج RGB- هیستوگرام در پایگاه های داده. ۶۷
جدول ۴- ۶: نتایج هیستوگرام رنگ در پایگاه داده MPEG-7 . ۶۸
جدول ۴- ۷: نتایج روش CLD در پایگاه های داده ۶۸
جدول ۴- ۸: نتایج روش DCD در پایگاه های داده ۶۸
جدول ۴- ۹: نتایج روش SCD در پایگاه های داده ۶۹
جدول ۴- ۱۰: نتایج روش ماتریس Co – Occurrence در پایگاه های داده ۶۹
جدول ۴- ۱۱: نتایج روش Intensity-HDWT1 در پایگاه های داده. ۷۰
جدول ۴- ۱۲: نتایج روش RGB-HDWT1 در پایگاه های داده. ۷۱
جدول ۴- ۱۳: نتایج روش HMMD-HDWT1 در پایگاه های داده. ۷۳
جدول ۴- ۱۴: نتایج روش HDWT1 در پایگاه های داده MPEG-7. ۷۴
جدول ۴- ۱۵: نتایج روش Intensity-HDWT2 در پایگاه های داده. ۷۴
جدول ۴- ۱۶: نتایج روش RGB-HDWT2 در پایگاه های داده. ۷۵
جدول ۴- ۱۷: نتایج روش HDWT2 در پایگاه داده MPEG-7. ۷۵
جدول ۴- ۱۸: نتایج روش Intensity-HDWT2 با شبکه عصبی در پایگاه های داده. ۷۶
جدول ۴- ۱۹: نتایج روش RGB-HDWT2 با شبکه عصبی در پایگاه های داده. ۷۶
جدول ۴- ۲۰: نتایج روش HDWT2 با شبکه عصبی در پایگاه داده MPEG-7. ۷۷
جدول ۴- ۲۱: نتایج روش Intensity-DCT در پایگاه های داده. ۷۷
جدول ۴- ۲۲: نتایج روش RGB-DCT در پایگاه های داده. ۷۸
جدول ۴- ۲۳: نتایج روش DCT در پایگاه داده MPEG-7. ۷۸
جدول ۴- ۲۴: نتایج ترکیب روش های HDWT2 و DCT در پایگاه های داده. ۷۸

فهرست شکل‌ها

عنوان **صفحه**

شکل ۱- ۱: خلاصه توصیف تصویری ساختار پایان نامه ۸

شکل ۲- ۱: فرآیند سیستم بازیابی تصویر. ۱۱

شکل ۲- ۲: نمایش تصویر در فضای رنگی RGB. ۱۳

شکل ۲- ۳: (الف) شماتیک مکعب رنگ RGB که رنگ های اولیه و ثانویه را در رئوس نشان می دهد. نقاط در امتداد قطر اصلی، دارای مقادیر خاکستری، از سیاه در مبدا تا سفید در نقطه (۱،۱،۱) است. (ب) مکعب رنگ RGB. ۱۴

شکل ۲- ۴: (الف) شش ضلعی رنگ HSV (ب) مخروط شش ضلعی HSV. ۱۵

شکل ۲- ۵: نمایش تصویر در فضای رنگی HSV. ۱۵

شکل ۲- ۶: فضای رنگی HMMD. ۱۶

شکل ۲- ۷: نمایش تصویر در فضای رنگی HMMD. ۱۷

شکل ۲- ۸: نمایش تصویر در فضای رنگی YCbCr. ۱۷

شکل ۲- ۹: استخراج ویژگی CLD. ۲۱

شکل ۲- ۱۰: تقسیم بندی تصویر. ۲۲

شکل ۲- ۱۱: انتخاب رنگ نماینده. ۲۲

شکل ۲- ۱۲: اسکن کردن زیگزاگی. ۲۴

شکل ۲- ۱۳: استخراج ویژگی CLD. ۲۴

شکل ۲- ۱۴: کوانتیزه سازی با استفاده از الگوریتم FRCFE. ۲۸

شکل ۲- ۱۵: توابع موجک مادر. ۳۵

شکل ۳- ۱: بلوک دیاگرام اولین سیستم CBIR پیشنهادی. ۴۰

شکل ۳- ۲: فلوچارت استخراج ویژگی HDWT سطح- p. ۴۱

شکل ۳- ۳: تجزیه موجک تصویر. ۴۲

شکل ۳- ۴: سطح های اول تبدیل موجک. ۴۳

شکل ۳- ۵ : بلوک دیاگرام دومین سیستم CBIR پیشنهادی. ۴۶

شکل ۳- ۶: بلوک دیاگرام سومین سیستم CBIR پیشنهادی. ۵۲

شکل ۳- ۷: بلاک دیاگرام DCT پیشنهادی. ۵۳

شکل ۳- ۸: تصویر در فضای RGB. ۵۳
شکل ۳- ۹: صفحه Red. ۵۴
شکل ۳- ۱۰: تصویر با اندازه ۳۲×۳۲. ..... ۵۴
شکل ۳- ۱۱: تصویر ضرایب DCT. ۵۵
شکل ۳- ۱۲: تصویر ضرایب DCT بصورت رنگی. ۵۵
شکل ۳- ۱۳: تقسیم به بلوک ها. ۵۵
شکل ۴- ۱: تصویرهای نمونه از پایگاه داده Corel. ۶۳
شکل ۴- ۲: تصویرهای نمونه از پایگاه داده ALOI. ۶۴
شکل ۴- ۳: تصویرهای نمونه از پایگاه داده MPEG-7. ..... ۶۴
شکل ۴- ۴: نمودارهای دقت و فرا خوانی روش Intensity-HDWT1 در پایگاه داده Corel. ۷۰
شکل ۴- ۵: نمودار های دقت و فرا خوانی روش RGB-HDWT1 در پایگاه داده Corel. ۷۲
شکل ۴- ۶: نمودار های دقت و فرا خوانی روش HMMD-HDWT1 در پایگاه داده Corel. ۷۳
شکل ۴- ۷: مقایسه عملکرد روش های مختلف برای پایگاه داده Corel. ۷۹
شکل ۴- ۸: مقایسه عملکرد روش های مختلف برای پایگاه داده ALOI. ۷۹
شکل ۴- ۹: مقایسه عملکرد روش های مختلف برای پایگاه داده MPEG-7. ..... ۸۰
شکل ۴- ۱۰: تصاویر بازیابی شده با استفاده از روش RGB-HDWT1 در سطح-۵ در پایگاه داده Corel. ۸۱
شکل ۴- ۱۱: تصاویر بازیابی شده با استفاده از روش HMMD-HDWT1 در سطح-۵ در پایگاه داده ALOI. ۸۲
شکل ۴- ۱۲: تصاویر بازیابی شده با استفاده از روش HDWT2 در سطح-۵ در پایگاه داده MPEG-7. ۸۲

فصل ۱– مقدمه

۱–۱– پیشگفتار

اختراع دوربین دیجیتال به ما این امکان را می دهد که از جهان اطرافمان عکس بگیریم و به راحتی آن را در اختیار دیگران قرار دهیم. امروزه پیشرفت تکنولوژی ایجاد مجموعه تصاویری از خانواده ها در کنار هم، گردش در پارک ها و تفریحگاه ها و اماکن زیارتی و سیاحتی را بسیار راحت کرده است. هزینه پائین ذخیره سازی و سهولت کار در وب سایت ها از تغییرات نسبت به گذشته تا به امروز این تکنولوژی است. امروزه، تنوع منابع موجود و گستردگی مفاهیم از گرداگرد جهان و رشد سریع و چشم گیر تکنولوژی، قابلیت جستجو در تمام زمینه ها را به ما داده است. در عین حال کاربرد تکنولوژی در جهان واقعی با محدودیت هایی روبرو است. طراحی یک ماشین جستجوگر تصویر، با توانایی پرداخت به تمام نیازهای کاربر و یا تعامل بین دو کاربر ملزوم وجود قابلیت های منحصر به فرد است. بازیابی تصویر رشته ای است که با جستجوی تصاویر دیجیتال از پایگاه های داده تصاویر سروکار دارد. امروزه تلاش های بسیاری برای افزایش سرعت و بازدهی بازیابی تصاویر انجام می شود.

۱–۲– تاریخچه و شیوه های اولیه و نوین بازیابی تصاویر

تحقیق بر روی بازیابی تصاویر از دهه ۱۹۷۰ میلادی آغاز شد که نظر محققان را در رشته های مختلف پردازش تصویر، چند رسانه ای، کتابخانه های دیجیتال، دریافت از راه دور، ستاره شناسی، مد، پیشگیری از جرم و جنایت، چاپ و نشر، پزشکی، معماری و سایر حوزه های مربوط جلب کرده است. یک سیستم بازیابی تصویر موثر، این قابلیت را دارد تا در یک مجموعه از تصاویر، تصویر مناسب را با توجه به تصویر خواسته شده با دقتی نزدیک به ادراک انسان بازیابی کند.

دو زمینه عمده ی تحقیقاتی، مدیریت پایگاه داده (بر پایه متن) و بینایی ماشین (بر پایه ویژگی بصری)، بازیابی تصویر را از جهات مختلف مورد مطالعه قرار می دهند. سابقه ی بازیابی متن به سال ۱۹۷۰ باز می گردد [۱]. در این قبیل سیستم ها، تصاویر به صورت دستی و با توصیف گرهای متنی حاشیه نویسی می شوند، سپس به وسیله ی سیستم مدیریت پایگاه داده برای اجرای بازیابی تصویر استفاده می گردند. اما عیب های این روش عبارتند از:

۱- مضمون یک تصویر خیلی بیشتر از این است که بوسیله یک سری کلمات کلیدی بتوان آن را بیان کرد [۲].

۲- حاشیه نویسی به وسیله ی انسان کاری بسیار سخت، وقت گیر و مستلزم هزینه بسیار است.

۳- بی دقتی در حاشیه نویسی، که از ادراک و آگاهی انسانی ناشی می شود [۳].

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

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

در اوایل سال ۱۹۹۰ افزایش سریع تصاویر با حجم بالا نظیر شبکه اینترنت و همچنین معایب مذکور، سیستم بازیابی مبتنی بر متن را ناکارآمد کرد و باعث معرفی سیستم های بازیابی مبتنی بر محتوای تصویر (CBIR)[1] شد. تکنیک های مبتنی بر محتوا در بازیابی تصویر بجای یادداشت گذاری دستی بصورت متن، از خصوصیت های بصری برای تفسیر مفاد تصاویر بهره می گیرند. در سال های اخیر، بازیابی تصویر مبتنی بر محتوا (CBIR) به عنوان گزینه مناسب تر نسبت به بازیابی تصویر مبتنی بر متن مورد استفاده قرار گرفته است. اولین چالش در هنگام ساخت یک سیستم بازیابی تصویر فاصله معنایی است. این به عنوان "عدم انطباق بین اطلاعاتی که یکی می تواند از اطلاعات مجازی خارج کند و تفسیری که دیگری از همان شرایط داده شده دارد" تعریف می شود [۴]. یک توصیف محتوایی از یک تصویر می تواند با توصیف یک کاربر انسانی تفاوت شایانی داشته باشد، همان طور که دو نفر توصیف های متفاوتی از یک منظره ارائه می دهند، این عمل به CBIR مربوط است. سیستم های CBIR توصیف های قابل تفسیر توسط ماشین را از ویژگی های فیزیکی یک تصویر تولید می کنند. این توصیف ها، به عنوان خصوصیات استخراج شده، می تواند به وسیله مقیاسی از تشابه مقایسه شود. سپس شباهت میان تصویر داده شده و هر تصویری در بایگانی تصاویر به وسیله سیستم CBIR محاسبه می شود. سپس نتایج به صورت نزولی بر اساس شباهت نمایش داده می شوند. در ادامه به توضیح بیشتر سیستم های CBIR می پردازیم.

سیستم های بازیابی تصویر از سه بخش عمده تشکیل شده است. بخش اول و مهم ترین بخش سیستم های بازیابی، بخش استخراج ویژگی[2] است که وظیفه ی تولید بردار ویژگی[3] هر تصویر از پایگاه داده و نمایش محتوای تصویر برای دسته بندی تصاویر را دارد. اندازه بردار ویژگی باید بسیار کوچکتر از تصویر اولیه باشد. در نتیجه استخراج ویژگی تصاویر پایگاه داده باعث حداقل کردن زمان جستجو، ساده کردن جستجو و حتی الامکان تصاویر مشابه به سریعترین وجه قابل بازیابی باشند. بخش دوم اندیس گذاری[4] نامیده می شود که وظیفه ی دسته بندی تصاویر براساس ویژگی های استخراج شده را دارد.

[1] Content Based Image Retrieval
[2] Feature extraction
[3] Featur vector
[4] Indexing

بخش بازیابی، آخرین بخش سیستم های بازیابی است که تصویر مورد سوال کاربر را پردازش و با استفاده از معیار شباهت، بردارهای ویژگی تصویر مورد نظر و تصاویر پایگاه داده، فاصله بین تصویر مورد سوال و تصاویر موجود در پایگاه داده را محاسبه می کند و تصاویر مشابه را نشان می دهد و واسط کاربری را فراهم می کند.

Gevers در [۵] روش های بازیابی تصاویر را براساس کاربردهای مختلف در سه دسته طبقه بندی کرده است:

۱- جستجو براساس شباهت تصاویر[1]: هدف این روش به دست آوردن تصاویر مورد نظر کاربر از بین تصاویر موجود در یک گالری است که در مراحل پی درپی و با استفاده از بازخورد کاربر صورت می گیرد.

۲- جستجو به هدف یافتن تصویر خاص[2]: 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[3]: هدف از این روش به دست آوردن تصویری است که متعلق به دسته یا کلاس بخصوصی باشد، مانند تصاویر دکوراسیون منزل و یا تصاویر رادیولوژی.

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

۱- دامنه محدود: شامل تغییرات محدود در ویژگی های دیداری تصاویر (مانند پایگاه داده تصاویر مربوط به پزشکی)

۲- دامنه وسیع: شامل تغییرات وسیع و غیرقابل پیش بینی در تصاویر (مانند پایگاه داده تصاویر موجود در اینترنت).

پیش از شروع کار بازیابی، ویژگی های تصاویر پایگاه داده باید از آنها استخراج شوند. یک تصویر آرایه ای دو بعدی از پیکسل ها است. برای سهولت دسترسی به ویژگی های آن تصویر، به جای استفاده از مقادیر پیکسل ها، با استفاده از روش هایی خاص، ویژگی های تصویر را استخراج و در قالبی که قابل مقایسه باشد، نمایش می دهند. به این روش ها استخراج ویژگی می گویند. خروجی این روشها عدد یا

[1] Search by association
[2] Target search
[3] Category search

برداری است که به آن ها کلاس ویژگی یا توصیف گر می گویند (مانند هیستوگرام رنگ). به برداری که از کنار هم قرار دادن این کلاس های ویژگی بدست می آید، بردار ویژگی تصویر گفته می شود [٦ و ٧].

ویژگی های تصویر را می توان از لحاظ نحوه توصیف تصویر به سه دسته تقسیم کرد:

١- ویژگی های 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 ویژگی ها است. نخستین کار در این زمینه توسط چانگ در سال ١٩٨۴ انجام شده است [٨].

٢- ویژگی های معنایی[4]: به بیان مفاهیم و اشیای موجود در تصویر می پردازند و غالبا از روش های قسمت بندی[5] تصویر برای اینکار استفاده می کنند.

٣- ویژگی های مختص کاربرد های خاص: نظیر ویژگی های توصیفی چهره در سیستم های بازیابی تصاویر مربوط به چهره [٩].

همانطور که گفته شد حضور بر چسب قابل اعتماد همراه با تصویر برای بازیابی تصویر بر پایه متن ضروری است. ما معتقدیم که آینده دنیای واقعی سیستم های بازیابی در بستر بهره وری از دو نوع تکنولوژی متنی و محتوایی است. تا زمانی که نمونه های قبلی از نظر کاربر قابل اعتمادترند، پتانسیل زیادی سر ترکیب این دو وجود دارد که موجب ساخت موتور جستجوی قوی می شود که می تواند به تصاویر پنهان شده در جاهایی از وب دسترسی پیدا کنند که این تلاش در طی سال های آتی به امید پیروزی عملی خواهد بود.

نمونه های زیادی برای طرح یک واسط به درخواست یک سیستم CBIR، پیشنهاد شده است. واسط دروازه ای به آرشیو تصویر است. یک واسط کاربری حداقلی، به طور ساده توانایی مرور همه تصویر های آرشیو به صورت حرفه ای، لزوم وجود یک سیستم بازیابی تصویر را می رساند که به کاربر امکان می دهد تصویر دلخواه را از آرشیو تصویر بازیابی کند. اما از لحاظ زمانی که برای مرور دستی آرشیو نیاز است، بسیار بهینه است. واسط های بعدی تصویر ها را سازمان دهی می کنند یا قابلیت های بازیابی را فراهم می کنند. واسط ها را می توان اینگونه تقسیم کرد:

---

[1] Color
[2] Texture
[3] Shape
[4] Semantic
[5] Segmentation

۱- مرورگری[1]: این آسان ترین راه برای دسترسی به آرشیو تصویر است. با پشت سر گذراندن یک ساختار مرورگری، همواره تصویر ها می توانند در گروه هایی (گاهی با خوشه بندی) جهت بررسی کاربر سازمان دهی شوند. مانند CLUE [۱۰].

۲- دسته های سفارشی[2]: تصویر ها به صورت سلسله مراتبی[3]، در دسته های مخصوص دامنه ساختاربندی می شوند. یک مثال از سیستم مبتنی بر هستی شناسی در [۱۱] توصیف شده است. سلسله مراتب می توانند شامل چندین دسته ی معنادار باشند. (به طور مثال یک دسته سطح بالا می تواند وسایل نقلیه باشد و شامل دسته هایی مانند هواپیما و ماشین ها شود و هر یک نیز می تواند شامل زیر دسته های دیگر باشد).

۳- درخواست با تصویر نمونه (QBE)[4]: این نمونه ی جستجوی تصویر مبتنی بر محتوا به طور سنتی است. کاربر یک تصویر نمونه با هدف بازیابی تصویر های مشابه از سیستم، فراهم می کند. نمونه ای از یک سیستم که امکان جستجو با این الگو را می دهد، درخواست با محتوای تصویر (QBIC)[5] است [۱۲]. 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 مجموعه ی دیگری باشد که از طریق ابزارهای ماوراء سیستم CBIR خاص فراهم شده باشد. این می تواند یک دغدغه باشد زیرا کاربر باید از ابزارهای دیگری برای جستجوی تصاویر استفاده کند. متناوباً، 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 تصادفی درخواست شوند تا زمانی که کاربر انتخاب کند و با مثال تطابق دهد.

۴- درخواست با ناحیه تصویر[6]: درخواست ها می تواند براساس یک زیر مجموعه 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 طور دستی تعریف کند، یا یک شیوه ی قسمت بندی باید داخل سیستم قرار داده شود. Blobworld [۱۳] و NeTra [۱۴] دو سیستمی هستند که امکان بازیابی بوسیله ناحیه تصویر را می دهند. هدف از این الگوی تصویر اصلاح نتایج بازیابی بوسیله درخواست بر مبنای مطابق ترین بخش داده شده است.

---

[1] Browsing
[2] Customized categories
[3] Hierarchical
[4] Query By Example image
[5] Query By Image Content
[6] Query by image region

۵- درخواست بوسیله ی چندین تصویر نمونه (QBME)[1]: کاربر می تواند چندین تصویر نمونه برای سیستم فراهم کند، همانند [۱۵]. اشتراکات بین همه تصویر های درخواست را می توان به عنوان مبنای درخواست بکار برد.

۶- درخواست با طرح بصری: چندین پیاده سازی، ابزار ترسیم را برای کاربر جهت ایجاد یک تصویر دلخواه فراهم می کنند همانند Santini و Jain [۱۶] و Retrievr [۱۷]. این در غیاب تصویر نمونه مفید است. چالش این شیوه این است که اعتماد به توانایی های هنری کاربر، به یکی از واسطه های پر مطالبه منتهی می شود.

۷- درخواست با تعیین مستقیم ویژگی های بصری: این حرفه ای ترین شیوه است، همانند Webseek [۱۸]. کاربر باید مشخصات و مفهوم هر ویژگی بصری مشخص شده را درک کند . این شیوه برای کاربرانی که ناآشنا با طرح و عملکرد داخلی سیستم هستند، ممکن است دشوار باشد.

۸- درخواست با کلمه ی کلیدی: اگر تصویر ها قبلاً یادداشت گذاری شده اند یا متن موجود است می توان آن ها را با متن جستجو کرد. جستجوی تصویر گوگل نمونه موفق این روش است [۱۹]. جستجوی تصویر گوگل به طور خودکار تصویر ها را با استفاده از متن پیرامون تصویر در صفحه وب نامگذاری می کند. همچنین سیستم های بازیابی با کلمه کلیدی می توانند بر نامگذاری دستی تصویر های شخصی بوسیله انسان تکیه کنند این شیوه ی موثری است که ممکن است در غیاب محتوای تصویر، سراسری عمل کند، اگر چه همیشه نامگذاری تصویر ها با یک مدل مناسب میسر نیست (مثلاً یک پایگاه داده بزرگ زمان زیادی برای نامگذاری دستی نیاز دارد).

۹- بازیابی چند کیفیتی[2]: روش هایی هستند که چند کیفیت (مثلاً لامسه، صدا، حرکات بدن و غیره) را ترکیب می کنند. ابداعات اخیر، همانند حوزه وسیع نمایشگرهای مالتی تاچ، بر افزایش کارکرد درخواست های مالتی مدال یا درخواست هایی که چندین شکل ورودی را پیوند می دهند، دلالت می کند. در [۲۰] چندین کاربر می توانند از همه ی انگشتانشان به طور طبیعی برای دستکاری و سازماندهی تصویر ها استفاده کنند.

در دهه های گذشته چندین نمونه محصول تجاری و سیستم تجربی توسعه داده شده اند که برخی از آن ها عبارتند از : QBIC [۱۲]، Photobook [۲۱]، Virage [۲۲]، VisualSEEK [۲۳]، Netra [۱۴] و

[1] Query By Multiple Example images

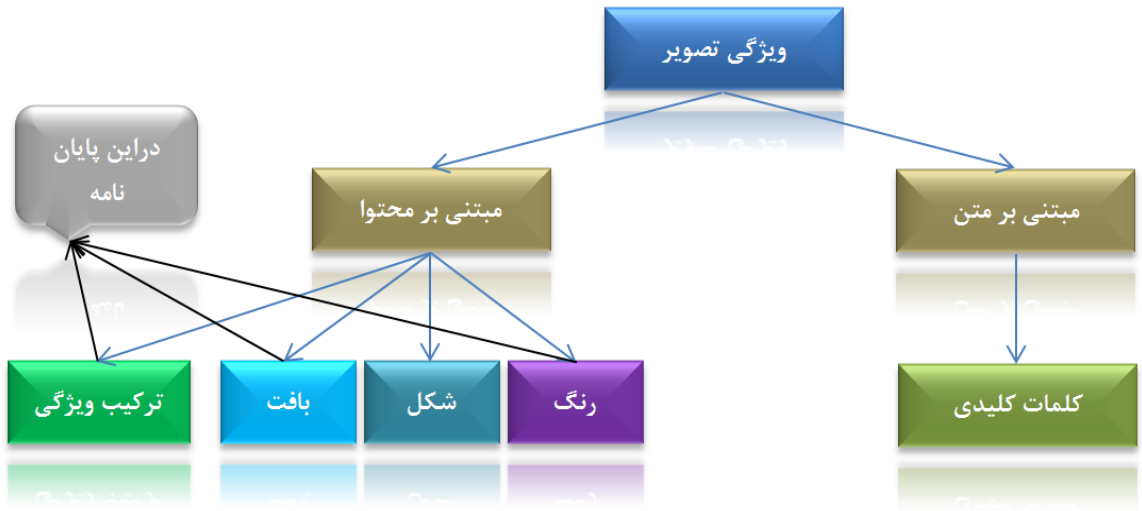
[2] Multimodal query

SIMPLcity [۲۴]. به طور مثال QIBC که در مرکز تحقیقات IBM در Almadeu بدست آمد. برخلاف سیستم مبتنی بر متن در این سیستم خصوصیت های بصری از خود تصویر استخراج می شود.

۱-۳- هدف از انجام تحقیق

با وجود تلاش های انجام شده در سال های اخیر در جستجوی بازیابی تصویر، هنوز الگوریتمی قابل قبول و جامع برای مشخص کردن تصورات انسانی به خصوص در زمینه ی تفسیر تصاویر وجود ندارد. بخاطر طبیعت تکنولوژی CBIR، عملکرد آن موجب دو نوع مشکل ذاتی می شود: ۱) چگونه یک تصویر را به صورت ریاضی بیان کنیم؟ ۲) چگونه شباهت بین دو تصویر را که بر پایه توصیفات انتزاعی است بسنجیم؟ اولین پرسش به این خاطر به وجود می آید که تصویر اصلی توسط آرایه ای از مقادیر پیکسل ها که با واکنش های دیداری همخوانی کمی دارد به وجود می آید. حتی اگر فهم معنایی تصویر را به حال خود واگذاریم، ما برای اهداف بازیابی به توصیف ریاضی تصویر رجوع می کنیم که به عنوان «ویژگی ها» به کار می روند. چشم انداز طرح از لحاظ استخراج ویژگی و محاسبات تصاویر مشابه به طور واضحی جدایی پذیر نیست. تدوین ویژگی هایی که به مقدار زیاد، حدودی را برای تعریف اندازه ی شباهت ها تعیین می کنند.

هدف از این پژوهش یافتن الگوریتم هایی مناسب جهت جستجو در پایگاه های داده ی متفاوت است که علاوه بر دقت بالای بازیابی، دارای محاسبات کمتری از نظر زمان بازیابی باشند. مطالعات ما روی ویژگی های سطح پایین رنگ و بافت، روش های گوناگون استخراج ویژگی که در CBIR وجود دارند و توسط محققین و استاندارد MPEG-7 [۲۵] استفاده شده اند، انجام شده است. این روش ها براساس دقت بازیابی، زمان اجرا و مناسب بودن برای یک سیستم بازیابی انتخاب و مقایسه شده اند. در این خصوص، چند سیستم CBIR پیشنهاد شده است و ۳ پایگاه داده مختلف و کاملا متنوع با اندازه های متفاوت در نظر شده است. هر پایگاه داده خاصیت های گوناگونی نسبت به پایگاه داده دیگر دارد و در سیستم های بازیابی تصاویر مورد استفاده قرار می گیرد که در فصل ۴ بطور مفصل شرح داده خواهند شد. شکل ۱- ۱ ساختاری کلی از پایان نامه و ویژگی های مبتنی بر متن و مبتنی بر محتوای تصویر را نشان می دهد.



شکل ۱- ۱: خلاصه توصیف تصویری ساختار پایان نامه

۱-۴- نوآوری تحقیق

ما در این تحقیق بصورت کاملا متفاوت ویژگی های رنگ و بافت را ترکیب کرده ایم و با تولید صفحات رنگی و استخراج ویژگی بافت از این صفحات، بردار ویژگی تصاویر را بدست آورده ایم. از تبدیل موجک گسسته (DWT)[1] و تبدیل کسینوسی گسسته (DCT)[2] به عنوان ویژگی بافت، برای استخراج ویژگی تصویر استفاده کرده ایم. از شبکه های عصبی، آنالیز مولفه اصلی و طبقه بند k-همسایه نزدیکتر (kNN)[3] در این پایان نامه برای بدست آوردن نتایج مطلوب بازیابی و ارائه روش های جدید بازیابی تصاویر استفاده کرده ایم و نتایج هر روش را در ۳ پایگاه داده نشان داده ایم.

۱-۵- ساختار پایان نامه

این پایان نامه شامل پنج فصل است که فصل اول، تاریخچه، اهداف و نوآوری تحقیق در بازیابی تصویر را شرح می دهد. در فصل دوم خلاصه ای از سیستم های مختلف و روش های به کار رفته در بازیابی، آورده شده است. این فصل، شامل معرفی ابزارهای اساسی و لازم جهت استخراج ویژگی های مهم یک تصویر و شرح بیشتر روش های مهم استفاده شده در مقالات علمی با دقت خوب و مناسب در پایگاه های داده مختلف بازیابی تصاویر آورده شده است. فصل سوم سیستم های CBIR پیشنهادی را شرح می دهد که ترکیبی از چند روش پیشنهادی، شبکه های عصبی، طبقه بند k-همسایه نزدیکتر و همچنین آنالیز

[1] Discret Wavelet Transform
[2] Discret Cosine Transform
[3] k-Nearest Neighbor