بسم الله الرحمن الرحيم

دانشکده فنی
گروه مهندسی برق
گرایش الکترونیک

اصلاح روش‌های استخراج ویژگی مبتنی بر تبدیل موجک گسسته برای بازیابی تصویر مبتنی بر محتوا

از:
سیده لیلا نژادهاشمی

استاد راهنما:
دکتر بابک ناصرشریف
دکتر اسدا... شاه‌بهرامی

استاد مشاور:
دکتر رضا حسن‌زاده پاک رضایی

اسفند ماه ۱۳۹۰

با سپاس فراوان از خانواده مهربان و اساتید بزرگوارم که مرا در پیمودن این راه یاری نموده‌اند. سپاس و احترام ویژه‌ام را تقدیم می‌کنم به اساتید ارجمندم آقای **دکتر بابک ناصرشریف**، آقای **دکتر اسدا... شاه‌بهرامی** و آقای **دکتر رضا حسن‌زاده** که در تمام مراحل این پروژه همواره راهنما و مشوق من بودند.

فهرست مطالب

فهرست اشکال ..... **ح**

فهرست جداول ..... **د**

چکیده فارسی ..... **ذ**

چکیده انگلیسی ..... **ر**

فصل ۱: مقدمه **۱**

۱-۱- ساختار سیستم CBIR ۴

۱-۲- نمونه‌هایی از سیستم‌های CBIR تجاری و تحقیقاتی ۶

۱-۲-۱- QBIC ۷

۱-۲-۲- Virage ۷

۱-۲-۳- MARS ۷

۱-۲-۴- Photobook ۸

۱-۲-۵- Netra ۸

۱-۲-۶- VisualSEEk و WebSEEk ۸

۱-۲-۷- WBIIS ۹

۱-۲-۸- Simplicity ۹

۱-۳- جمع‌بندی ۱۰

۱-۴- معرفی ساختار پایان‌نامه ۱۰

فصل ۲: استخراج ویژگی **۱۲**

۲-۱- ویژگی‌های رنگ ۱۳

۲-۲- ویژگی‌های شکل ۱۵

۲-۳- ویژگی‌های بافت ۱۶

۲-۳-۱- ویژگی‌های حوزه مکان ۱۶

۲-۳-۲- ویژگی‌های حوزه تبدیل ۱۷

۲-۴- تبدیل موجک ۱۸
۲-۵- تبدیل موجک مختلط با درخت دوگانه (DT-CWT) ۲۲
۲-۵-۱-نواقص و مشکلات مربوط به DWT ۲۲
۲-۵-۲- شرط تأخیر نصف نمونه ۲۶
۲-۵-۳- طراحی فیلترهای DT-CWT ۲۸
۲-۵-۴- موضوعات مربوط به پیاده‌سازی ۲۸
۲-۵-۵- DT-CWT دو بعدی ۳۱
۲-۶- ساخت بردارهای ویژگی ۳۵
۲-۶-۱- ویژگی‌های مبتنی بر انرژی زیرباندها ۳۶
۲-۶-۲- ویژگی‌های مبتنی بر مدل‌سازی آماری ۳۷
۲-۵-۷- جمع‌بندی ۴۱
فصل ۳: روش‌های کاهش بعد ۴۲
۳-۱- روش‌های کاهش بعد خطی ۴۵
۳-۱-۱- تحلیل مؤلفه‌های اصلی (PCA) ۴۵
۳-۱-۲- تحلیل متمایزساز خطی (LDA) ۴۷
۳-۱-۳- LPP ۴۸
۳-۲- روش‌های کاهش بعد غیرخطی مبتنی بر هسته ۵۲
۳-۲-۱- Kernel PCA ۵۴
۳-۲-۲- Kernel LDA ۵۷
۳-۲-۳- Kernel LPP ۵۹
۳-۳- جمع‌بندی ۶۱
فصل ۴: اندازه‌گیری شباهت ۶۲
۴-۱- الگوریتم کلاس‌بندی k نزدیک‌ترین همسایه (KNN) ۶۵
۴-۲- الگوریتم کلاس‌بندی نایو بیز (NB) ۶۷
۴-۳- ماشین بردار پشتیبان (SVM) ۶۹
۴-۳-۱- نظریه آماری یادگیری ۶۹

۴-۳-۲- تشریح الگوریتم ۷۱
۴-۳-۳- ماشین بردار پشتیبان درحالت جدایی‌ناپذیر ۷۴
۴-۳-۴- ماشین بردار پشتیبان غیرخطی ۷۵
۴-۳-۵-ماشین بردار پشتیبان در حالت چندکلاسه ۷۶
۴-۴- جمع‌بندی ۷۷
فصل ۵: معرفی چارچوب پیشنهادی و ارائه نتایج آزمایش ها ۷۸
۵-۱- چارچوب پیشنهادی ۷۹
۵-۲- آزمایش‌ها و نتایج ۸۱
۵-۲-۱- محیط شبیه‌سازی و دادگان تصویری ۸۱
۵-۲-۲- نتایج بازیابی بر روی دو دادگان تصویری VisTex و Brodatz ۸۶
۵-۲-۲-۱- روش مبتنی بر انرژی زیرباندها ۸۶
۵-۲-۲-۲- روش مبتنی بر مدل‌سازی آماری ۹۰
۵-۲-۳- ارزیابی نتایج تجربی ۹۵
۵-۳- جمع‌بندی و پیشنهادها ۹۸
مراجع ۱۰۱

فهرست اشکال

شکل ۱-۱ دیاگرام بلوکی ساختار پایه‌ای یک سیستم CBIR..... ۴

شکل ۱-۲ دیاگرام بلوکی سیستم CBIR پیشنهادی. ۶

شکل ۲-۱ رنگ متوسط و رنگ اصلی: (الف) ناحیه اصلی، (ب) رنگ متوسط، (ج) رنگ اصلی. ۱۴

شکل ۲-۲ پیاده‌سازی بانک فیلتر دوکاناله با J طبقه. (الف)- بخش تحلیل. (ب) بخش سنتز. ۲۱

شکل ۲-۳ (الف)- پیاده‌سازی موجک گسسته دو بعدی با بانک فیلتر تفکیک‌پذیر در یک سطح. (ب)- تقسیم‌بندی طیف فرکانسی. ۲۱

شکل ۲-۴ بررسی اثر اعمال تبدیلات DWT و DT-CWT بر سیگنال آزمون x(n)=u(n=n0). (الف)- ضریب حقیقی d(0,8) به دست آمده از تبدیل DWT و (ب)- اندازه ضریب مختلط d(0,8) به دست آمده از تبدیل DT-CWT. ۲۳

شکل ۲-۵ جزءبندی حوزه فرکانس برای تبدیل موجک گسسته حقیقی به دست آمده از دو سطح تجزیه. ۲۳

شکل ۲-۶ بانک فیلتر آنالیز برای پیاده‌سازی مستقیم DT-CWT. ۲۴

شکل ۲-۷ بانک فیلتر سنتز برای پیاده‌سازی معکوس DT-CWT. ۲۵

شکل ۲-۸ بانک فیلتر آنالیز DT-CWT با مجموعه متفاوتی از فیلترها برای هر طبقه. ۲۹

شکل ۲-۹ جزءبندی حوزه فرکانس به دست آمده از دو سطح تجزیه برای DT-CWT. ۳۵

شکل ۲-۱۰ ۶ موجک جهتدار به دست آمده از تبدیل دو بعدی DT-CWT. (الف)- بخش حقیقی (ب)- بخش موهومی (ج)- اندازه. ۳۵

شکل ۲-۱۱ برازش GGD با توزیع ضرایب زیرباندهای DWT که به ترتیب از (الف)- سطح اول (ب)- سطح دوم (ج)- سطح سوم و (د) زیرباند تقریب در سطح سوم تجزیه به دست آمده‌اند. ۴۰

شکل ۲-۱۲ برازش GGD با توزیع ضرایب زیرباندهای DT-CWT که به ترتیب از (الف)- سطح اول (ب)- سطح دوم (ج)- سطح سوم تجزیه به دست آمده‌اند. بخش‌های حقیقی در سمت چپ و بخش‌های موهومی در سمت راست واقع شده‌اند. ۴۱

شکل ۳-۱ نمایشی از افزونگی بین داده‌های دو بعد در پایه‌های r1 و r2 ، خط r2=kr1 به صورت نقطه‌چین نمایش داده شده است. ۴۶

شکل ۳-۲ جهت‌های غیرعمود (x,y) و عمود یک سیستم نمونه. جهت با واریانس کم‌تر می‌تواند به عنوان نویز حذف شود..... ۴۶

شکل ۳-۳ ضرب داخلی در فضای ویژگی هسته $\mathcal{F}$ (ایده هسته). ۵۳

شکل ۳-۴ اثر kernel PCA. (الف) و (ج) دو مجموعه داده‌های مصنوعی را قبل از تبدیل نشان می‌دهند. (ب) و (د) به ترتیب توزیع به دست آمده بعد از kernel PCA خطی و غیرخطی را نشان می‌دهند. همچنین، توزیع نقاط داده در راستای محور x در زیر هر شکل نشان داده شده است. ۵۷

شکل ۳-۵ اثر kernel LDA. (الف) و (ج) دو مجموعه داده‌های مصنوعی را قبل از تبدیل نشان می‌دهند. (ب) و (د) به ترتیب نتایج به دست آمده بعد از اعمال kernel LDA بر داده‌های (الف) و (ج) را نشان می‌دهند. همچنین، توزیع کلاس‌ها در راستای محور x در زیر هر شکل نشان داده شده است. ۶۰

شکل ۴-۱ الگوریتم کلاس‌بندی k نزدکترین همسایه. ۶۶

شکل ۴-۲ ابر صفحهٔ بهینه در ماشین بردار پشتیبان SVM. ۷۲

شکل ۵-۱ بلوک استخراج ویژگی. ۸۰

شکل ۵-۲ بلوک کاهش بعد. ۸۰

شکل ۵-۳ بلوک اندازه‌گیری شباهت. ۸۰

شکل ۵-۴ نمونه‌هایی از تصاویر ۱۶ کلاس مختلف از مجموعه اول. ۸۲

شکل ۵-۵ نمونه‌هایی از تصاویر ۳۴ کلاس مختلف از مجموعه دوم. ۸۴

شکل ۵-۶ بهترین نتایج حاصل از الگوریتم‌های مختلف کلاس‌بندی بر بردارهای ویژگی F_{en} و F_{sm} در VisTex. ۹۵

شکل ۵-۷ بهترین نتایج حاصل از الگوریتم‌های مختلف کلاس‌بندی بر بردارهای ویژگی F_{en} و F_{sm} در Brodatz. ۹۵

فهرست جداول

جدول ۲-۱ میانگین آمار K-S برای هر یک از زیرباندهای حاصل از تجزیه تمامی تصاویر موجود در مجموعه آموزش با استفاده از DWT. ۳۹

جدول ۲-۲ میانگین آمار K-S برای هر یک از زیرباندهای حاصل از تجزیه تمامی تصاویر موجود در مجموعه آموزش با استفاده از DT-CWT. ۳۹

جدول ۵-۱ نرخ کلاس‌بندی متوسط به ازای تعداد سطوح تجزیه و پایه‌های موجک مختلف با استفاده از بردار ویژگی $F_{en,d}$ و کلاس‌بند NB. ۸۶

جدول ۵-۲ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند KNN در VisTex. ۸۷

جدول ۵-۳ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند KNN در Brodatz. ۸۷

جدول ۵-۴ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند NB در VisTex. ۸۸

جدول ۵-۵ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند NB در Brodatz. ۸۹

جدول ۵-۶ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند SVM در VisTex. ۸۹

جدول ۵-۷ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{en} و کلاس‌بند SVM در Brodatz. ۹۰

جدول ۵-۸ مقایسه عملکرد کلاس‌بند KNN با سه معیار فاصله مختلف بر روی بردار ویژگی F_{sm} در VisTex. ۹۱

جدول ۵-۹ مقایسه عملکرد کلاس‌بند KNN با سه معیار فاصله مختلف بر روی بردار ویژگی F_{sm} در Brodatz. ۹۱

جدول ۵-۱۰ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{sm} و کلاس‌بند KNN در VisTex. ۹۲

جدول ۵-۱۱ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{sm} و کلاس‌بند KNN در Brodatz. ۹۲

جدول ۵-۱۲ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{sm} و کلاس‌بند NB در VisTex. ۹۳

جدول ۵-۱۳ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{sm} و کلاس‌بند NB در Brodatz. ۹۳

جدول ۵-۱۴ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی F_{sm} و کلاس‌بند SVM در VisTex. ۹۴

جدول ۵-۱۵ نرخ کلاس‌بندی متوسط با استفاده از بردار ویژگی Fsm و کلاس‌بند SVM در Brodatz ۹۴

جدول ۵-۱۶ مقایسه نتایج و ویژگی‌های سیستم‌های CBIR معرفی شده در چندین مرجع مهم در زمینه استخراج ویژگی‌های چندمقیاسی از بافت. ۹۷

جدول ۵-۱۷- زمان محاسباتی متوسط. ۹

چکیده

اصلاح روش‌های استخراج ویژگی مبتنی بر تبدیل موجک گسسته برای بازیابی تصویر مبتنی بر محتوا

سیده لیلا نژادهاشمی

استخراج ویژگی و اندازه‌گیری شباهت در فضای ویژگی‌ها دو مرحله اصلی در یک سیستم بازیابی تصویر مبتنی بر محتوا هستند. از این رو در چنین سیستمی شناسایی روش‌های مناسب برای توصیف محتوای تصویر و همچنین تطبیق تصویر پرسش با تصاویر موجود در دادگان به روشی که به بهترین شکل منعکس‌کننده قضاوت انسان درباره میزان شباهت‌ها باشد، از اهمیت بسیاری برخوردار است. یکی از روش‌های استخراج ویژگی، استخراج ویژگی از بافت به عنوان یکی از عناصر اصلی تصویر است. در این راستا، تبدیل موجک گسسته با توجه به امتیازاتی از قبیل پیاده‌سازی کارآمد و امکان تحلیل تصویر در مقیاس‌های مختلف، پرکاربردترین روش برای استخراج ویژگی‌ها از بافت به شمار می‌آید. اما در سال‌های اخیر به دلیل کاستی‌هایی که مهمترین آن ضعف در تعیین جهات مختلف موجود در یک تصویر است، تبدیلات چندمقیاسی دیگری نیز مورد توجه قرار گرفته‌اند. تبدیل موجک مختلط با درخت دوگانه از جمله راهکارهایی است که برای غلبه بر کاستی‌های تبدیل موجک گسسته پیشنهاد شده است. هدف در این پایان-نامه ارائه روشی مبتنی بر تبدیلات موجک برای استخراج ویژگی‌هاست، به طوری که هم از لحاظ توانایی در توصیف محتوای بافت و هم از لحاظ زمان بازیابی کارآمد باشد. به این منظور ابتدا با استفاده از زیرباندهای به دست آمده از تبدیل موجک مختلط با درخت دوگانه، بردارهای ویژگی ساخته می شوند. این بردارها بر اساس انرژی زیرباندها و مدل‌سازی آماری توزیع ضرایب زیرباندها با استفاده از تابع گاوسی تعمیم‌یافته (GGD) به دست می‌آیند. سپس با اعمال تبدیلات کاهش بعد خطی و غیر خطی به بردارهای ویژگی چندمقیاسی مذکور، ضمن کمک به بهنگام نمودن فرایند بازیابی، توصیف مؤثرتری از تصاویر به دست می‌آید. این تبدیلات عبارتند از روش‌های خطی تحلیل متمایزساز خطی (LDA) و LPP و روش‌های غیرخطی kernel LDA و kernel LPP. در پایان به منظور کلاس‌بندی بردارهای ویژگی کاهش بعد یافته از سه کلاس‌بند K نزدیک‌ترین همسایه (KNN)، نایو بیز (NB) و ماشین بردار پشتیبان (SVM) استفاده شده است. نتایج به دست آمده بر روی دو دادگان تصویری استاندارد VisTex و Brodatz گویای کارآیی و کارآمدی چارچوب پیشنهادی برای سیستم بازیابی تصویر مبتنی بر محتوا در این پایان‌نامه است.

کلیدواژه: بازیابی تصویر مبتنی بر محتوا، استخراج ویژگی، تبدیل موجک گسسته، تبدیل موجک مختلط با درخت دوگانه، کاهش بعد، اندازه‌گیری شباهت.

Abstract

Improving Feature Extraction Methods based on Discrete Wavelet Transform for Content-Based Image Retrieval (CBIR)

Seyedeh Leila Nejadhashemi

Feature extraction and similarity measure in feature space are two basic steps in a content based image retrieval system. Therefore, identifying methods to describe image content properly and also approaches to model human similarity judgments are of great importance. Among different image descriptors, the important ones are those extracted from image texture. Discrete wavelet transform (DWT) has been widely used for texture feature extraction as a result of possibilities such as efficient implementation and image analysis in different scales. However, in recent years, its lack of directionality and some other drawbacks have drawn the researchers' attention to other directional and multiscale transforms. Dual tree complex wavelet transform (DT-CWT), a relatively recent enhancement to DWT, has been introduced to overcome DWT Shortcomings. This thesis aims to propose a wavelet based feature extraction approach which is efficient in two aspects of retrieval precision and computational time. To this end, first, we construct multiscale feature vectors from DT-CWT subbands in two different ways: traditional energy method using energy and standard deviation of wavelet subbands and statistical modeling based method using estimated generalized guassian density (GGD) parameters. Then, some linear and nonlinear dimensionality reduction transforms are applied to feature vectors to obtain more powerful image characteristics and decrease the computational costs during retrieval process. These transforms include linear discriminant analysis (LDA), locality preserving projections (LPP), Kernel LDA and Kernel LPP. Finally the classification algorithms of k nearest neighbors (KNN), Naïve Bayes (NB) and support vector machines (SVM) are utilized for measuring similarity between feature vectors. Experimental results on two standard image databases, VisTex and Brodatz, validate the efficiency of proposed framework for CBIR systems.

Keywords: Content based image retrieval, Feature extraction, Discrete wavelet transform, Dual tree complex wavelet transform, Dimensionality reduction, Similarity reduction.

فصل اول:

مقدمه

بازیابی متن و یا استفاده از آن به عنوان ابزاری برای بازیابی اطلاعات[1] در زبان‌های مختلف سال‌ها مورد توجه بوده است. اما زمانی که به ساماندهی تصاویر می‌رسیم، انسان در اکثر موارد بهتر از ماشین عمل می‌کند. یکی از دلایل این تمایز این است که متن ساخته دست بشر است. در حالیکه تصویر تنها نسخه‌ای دقیق از چیزی است که انسان از زمان تولد شاهد آن بوده است. لذا توصیف دقیق آن امری پیچیده می‌باشد. همچنین سیستم بینایی انسان در طول قرن‌های متمادی از نظر ژنتیکی رشد و توسعه یافته است. بنابراین تفسیر آنچه می‌بینیم دشوار و آموزش آن به سیستم دشوارتر و چالشی اساسی است. با وجود این، در طول دو دهه اخیر تلاش‌های بلندپروازانه‌ای برای قادر ساختن ماشین به یادگیری درک محتوای تصویر انجام شده است [1] .

در سال‌های اخیر شاهد رشد چشمگیری در کمیت، اهمیت و امکان دسترسی به تصاویر در همه جوانب زندگی بوده‌ایم. تصاویر نقش مهمی در زمینه‌های گوناگون از قبیل شناسایی جرم، تشخیص پزشکی، سنجش از راه دور[2] ، آموزش و بسیاری از موارد دیگر ایفاء می‌کنند. از این‌رو ابزارهای دسترسی و شاخص‌گذاری[3] مناسب به منظور استفاده بهینه از مقادیر فزاینده این داده‌های دیجیتال در پایگاه‌های داده از اهمیت بسزایی برخوردارند [2].

سیستم‌های اولیه بازیابی تصویر، تصاویر را با استفاده از توصیفات متنی از قبیل مؤلف، عنوان و کلمات کلیدی توصیف‌کننده محتوای تصویر یا به اختصار فراداده‌ها[4] نمایش می‌دادند. تصاویر به طور دستی توسط این توصیفات متنی حاشیه‌نویسی[5] می‌شدند و فرایند بازیابی بر اساس این دانش صریح و توسط تکنیک‌های رایج در بازیابی اطلاعات انجام می‌شد. اما این شیوه بازیابی مبتنی بر متن چند نقطه ضعف اساسی دارد. نخست اینکه حاشیه‌نویسی دستی کاری پرزحمت و غیرعملی است. پیدایش دادگان تصویری در مقیاس‌های عظیم این نقطه ضعف را چشمگیرتر نموده است. مشکل دوم چنین سیستم‌هایی از فاعلیت درک انسان ناشی می‌شود. افراد مختلف می‌توانند درک متفاوتی از محتوای یک تصویر یکسان داشته باشند و این امر فرایند بازیابی را غیرقابل اطمینان و پیچیده می‌نماید. در پایان، کلمات کلیدی در حاشیه‌نویسی دستی گاهی مبهم و ناتوان در توصیف کامل تصاویرند [2]. زیرا بر اساس یک ضرب‌المثل قدیمی " ارزش یک تصویر معادل هزار کلمه است. "

[1] Information Retrieval
[2] Remote Sensing
[3] Indexing
[4] Metadata
[5] Annotation

بازیابی تصویر مبتنی بر محتوا (CBIR)[1] یک روش جایگزین و یا مکمل برای بازیابی تصویر مبتنی بر متن است. CBIR روشی است برای بازیابی تصاویر مورد نظر از میان هزاران تصویر موجود در دادگان با استفاده از ویژگی‌های بصری از قبیل رنگ، بافت و شکل که به طور اتوماتیک از طریق تکنیک‌های پردازش تصویر و بینایی ماشین از تصاویر استخراج می‌شوند. بنابراین تصاویر بدون اتکاء به حاشیه‌نویسی دستی نمایه‌سازی و بازیابی می‌شوند. اولین استفاده از اصطلاح "بازیابی تصویر مبتنی بر محتوا" به مقاله‌ای منتشر شده در سال ۱۹۹۲ [3] بازمی‌گردد که در آن T. Kato و همکارانش تجربیاتشان در بازیابی اتوماتیک تصاویر از دادگان تصویری توسط ویژگی‌های رنگ و شکل را تحت این عنوان توصیف کردند. از آن پس این اصطلاح برای توصیف فرایند بازیابی اتوماتیک تصاویر مورد نظر از مجموعه‌های بزرگ بر اساس ویژگی‌های بصری به طور گسترده مورد استفاده قرار گرفته است.

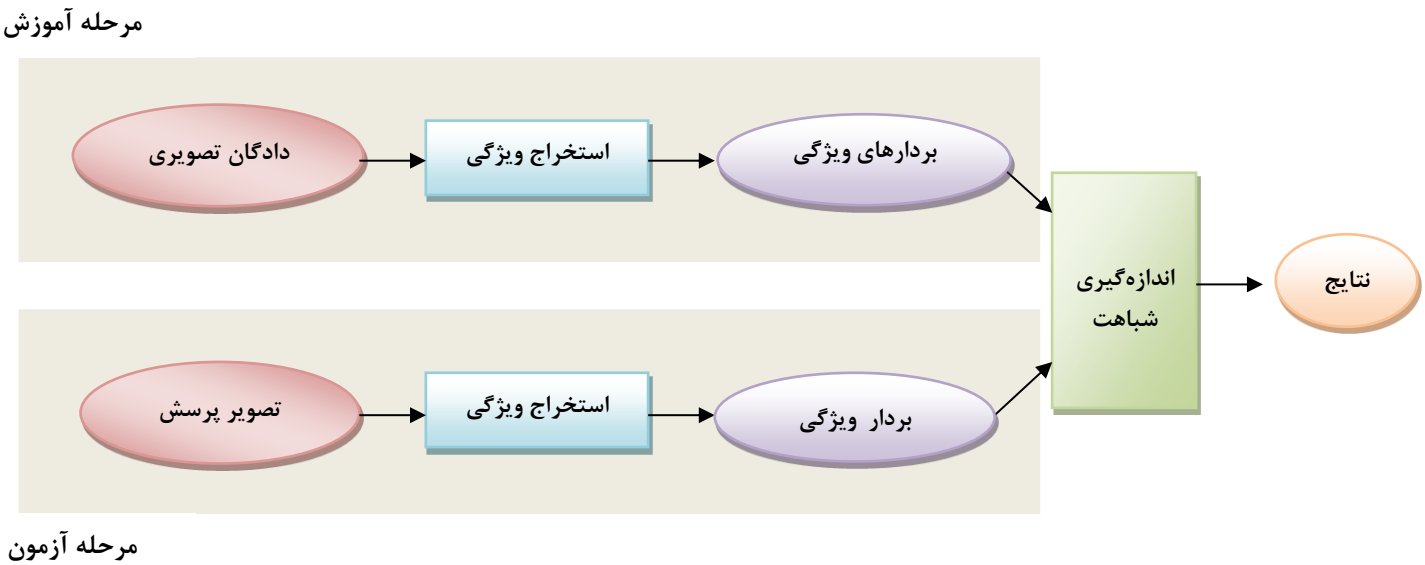
موضوعات تحقیقاتی در CBIR مجموعه‌ای از عناوین را که برخی از آنها با عناوین مطرح در بازیابی اطلاعات مشترک است، پوشش می‌دهند. تعدادی از مهمترین آنها عبارتند از [1]:

- درک نیازهای کاربر و خواسته او از سیستم
- شناسایی روش‌های مناسب در توصیف محتوای تصویر و استخراج ویژگی‌ها با استفاده از این روش‌ها از تصویر
- ذخیره فشرده‌سازی شده برای دادگان تصویری بزرگ
- تطبیق تصویر مورد پرسش و تصاویر موجود در دادگان به روشی که به بهترین شکل منعکس‌کننده قضاوت انسان درباره میزان شباهت‌ها باشد
- دسترسی مناسب به تصاویر ذخیره‌سازی شده با توجه به محتوای آنها
- فراهم کردن واسط‌های انسانی[2] مناسب برای سیستم‌های CBIR

[1] Content-based Image Retrieval
[2] Human Interfaces

۱-۱- ساختار سیستم CBIR

ساختار پایه‌ای یک سیستم بازیابی تصویر مبتنی بر محتوا از دو بخش اصلی تشکیل شده است. بخش اول، استخراج ویژگی است که در آن مجموعه‌ای از بردارهای ویژگی توصیف‌کننده محتوای تصاویر تولید می‌شود. این بخش هم در مرحله آموزش سیستم و در مورد تصاویر موجود در دادگان تصویری و هم در مرحله آزمون و تصاویر پرسش[1] یکسان است. دومین بخش اساسی در یک سیستم CBIR کلاس‌بندی تصاویر و یافتن شباهت میان تصویر پرسش و تصاویر موجود در دادگان با استفاده از بردارهای ویژگی استخراج شده و بازیابی N تصویر است که بیشترین شباهت را به تصویر پرسش داشته باشند. دیاگرام بلوکی این سیستم در شکل ۱-۱ نشان داده شده است.



شکل ۱-۱ دیاگرام بلوکی ساختار پایه‌ای یک سیستم CBIR.

شکل ۱-۱ سیستم CBIR ساده‌ای را نمایش می‌دهد که برای حوزه وسیعی از تصاویر کاربرد دارد. با توجه به اهداف سیستم طراحی شده ممکن است بخش‌های دیگری از قبیل ناحیه‌بندی تصویر[2] خصوصاً برای بازیابی ناحیه‌ای [4]، [5]، [6] و یا بازخورد

[1] Query Images
[2] Image Segmentation

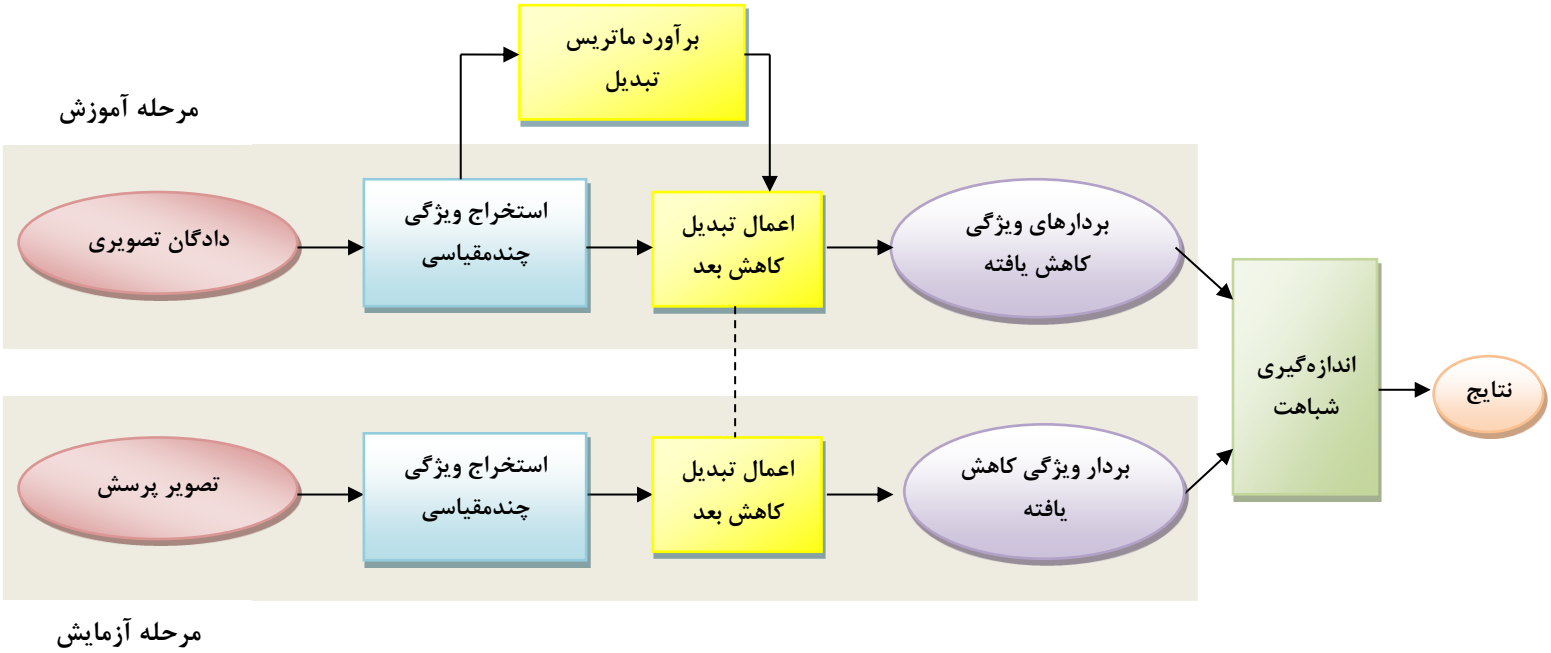
ارتباط[1] [7] به منظور افزایش تعامل میان کاربر و سیستم به آن اضافه شوند. با وجود این، استخراج ویژگی و اندازه‌گیری شباهت همچنان دو مرحله اصلی در هر سیستم CBIR طراحی شده با هر میزان پیچیدگی و یا هدفی می‌باشند. ویژگی‌های استخراج شده از تصویر اساس تصمیم‌گیری سیستم در رابطه با میزان شباهت‌هاست. اگرچه در طول سال‌ها تحقیقات بسیاری در این زمینه انجام شده است، انتخاب ویژگی‌هایی که به بهترین شکل محتوای موجود در یک تصویر را توصیف نمایند، هنوز هم در طراحی سیستم-های CBIR به عنوان یک چالش مطرح است. به طور سنتی ویژگی‌های به کار رفته در سیستم‌های CBIR در سه دسته جای می-گیرند: ویژگی‌های رنگ، شکل و بافت. از میان این ویژگی‌ها بازیابی مبتنی بر بافت از اهمیت بسیاری برخوردار است. زیرا بافت ویژگی ذاتی و طبیعی تمامی سطوح از قبیل ابر، گیاهان، درختان، آجرها و هر آنچه در در طبیعت موجود است، می‌باشد. یکی دیگر از دلایل این امر، طبیعت پیچیده بافت و نبود تعریف کمّی قابل قبولی از آن است. از میان رایج‌ترین و موفق‌ترین روش‌ها در استخراج ویژگی‌های بافت می‌توان به تبدیل موجک اشاره کرد. دلیل این امر امتیازاتی است که این تبدیل در مقایسه با دیگر روش-های رایج از قبیل روش‌های حوزه مکان دارد. از جمله این امتیازات می‌توان به پیاده‌سازی کارآمد با هزینه و زمان محاسباتی مناسب و امکان تحلیل تصویر در دقت‌ها و مقیاس‌های مختلف اشاره نمود.

هدف این پایان‌نامه بررسی کاربرد تبدیلات چندمقیاسی[2] و ارائه روشی مبتنی بر این تبدیلات برای استخراج ویژگی‌ها از تصاویر است، به طوری که هم از لحاظ توانایی در توصیف محتوای بافت و در نتیجه نتایج بازیابی و هم از لحاظ زمان مورد نیاز برای انجام فرایند بازیابی کارآمد باشد. استفاده از تبدیلات چندمقیاسی در کاربردهای پردازش تصویر تنها به تبدیل موجک گسسته محدود نمی‌شود. در سال‌های اخیر تبدیلات چندمقیاسی دیگری با قابلیت‌ها و توانایی‌های بیشتر در توصیف محتوای تصاویر معرفی شده و مسیر خود را در زمینه تحقیقاتی بازیابی تصویر گشوده‌اند. نتایج بازیابی گزارش شده از برخی از این تبدیلات بهتر از نتایج حاصل از تبدیل موجک گسسته است. اما در اکثر موارد این نتیجه بهتر در ازای افزایش طول بردار ویژگی استخراجی و احتمالاً افزایش هزینه و زمان بازیابی است. بنابراین در تحقیق ارائه شده، بخشی تحت عنوان کاهش بعد به ساختار پایه‌ای اضافه شده است تا با اعمال تعدادی از تبدیلات کاهش بعد به بردارهای ویژگی چندمقیاسی ضمن کمک به بهنگام نمودن فرایند، به توصیف مؤثرتری از تصاویر دست بیابیم. دیاگرام بلوکی این سیستم در شکل ۱-۲ نشان داده شده است.

در طول سال‌ها روش‌های کاهش بعد متنوعی با ویژگی‌ها و خصوصیات متمایز معرفی شده‌اند. برخی از این روش‌ها با کاهش همبستگی میان داده‌ها، برخی با تفکیک هرچه بیش‌تر داده‌های متعلق به کلاس‌های مختلف و برخی دیگر با سعی در یادگیری

[1] Relevance Feedback
[2] Multiscale Transforms

هندسه ذاتی داده‌ها و حفظ آن، قادرند ضمن کاهش بعد بردارهای ویژگی، کلاس‌بندی آنها را نیز بهبود بخشند. همان‌طور که در شکل (۱-۲) مشاهده می‌کنید، مرحله برآورد ماتریس تبدیل در این روش‌ها در مرحله آموزش انجام می‌گیرد و در نتیجه تأثیری در افزایش هزینه محاسباتی در مرحله آزمون سیستم نخواهد داشت.



شکل ۱-۲ دیاگرام بلوکی سیستم CBIR پیشنهادی.

۱-۲- نمونه‌هایی از سیستم‌های CBIR تجاری و تحقیقاتی

در این بخش تعدادی از سیستم‌های بازیابی تصویر موجود در بازار یا آزمایشگاه‌های تحقیقاتی را مرور می‌کنیم. آنچه در ادامه خواهیم دید، توضیح مختصری در رابطه با تعدادی از معروف‌ترین سیستم‌های موجود در این زمینه است.

۱-۲-۱- QBIC [1]

QBIC [8] اولین و احتمالاً معروف‌ترین سیستم تجاری بازیابی مبتنی بر محتوا برای تصویر و ویدئو است. این سیستم که توسط شرکت IBM در سال ۱۹۹۵ طراحی شده است، تأثیر قابل‌ملاحظه‌ای بر ساختار و چارچوب سیستم‌های CBIR بعد از خود نهاده است. پرسش در QBIC را می‌توان بر اساس رنگ، شکل، بافت، طرح ترسیمی و یا یک تصویر نمونه انجام داد.

در ابتدا، کاربرد رنگ در QBIC به هیستوگرام رنگ سراسری تصویر یا درصد رنگ در یک تصویر محدود می‌شد. اما نسخه‌های جدیدتر QBIC امکان پرسش بر اساس طرح مکانی رنگ را به کاربر می‌دهند. ویژگی بافت آن یک نسخه بهبودیافته از نمایش بافت Tamura یعنی ترکیبی از زبری[2]، تضاد[3] و جهت[4] است. آنچه به عنوان ویژگی‌های شکل در این سیستم استفاده شده است، سطح شکل[5]، مدور بودن[6]، دوری از مرکز[7]، جهت محور اصلی[8] و مجموعه‌ای از متغیرهای ممان جبری هستند.

۲-۲-۱- Virage

سیستم Virage [9] در سال ۱۹۹۷ توسط شرکت Virage طراحی شده است و همانند QBIC پرسش بر اساس رنگ، بافت و شکل را پشتیبانی می‌کند. همچنین Virage با فراهم کردن امکان استفاده از ترکیبی از عناصر فوق در یک تصویر یک گام فراتر از QBIC نهاده است. به طوری که کاربر می‌تواند وزن تخصیص‌یافته به هر عنصر را با توجه به نیاز خود تنظیم نماید.

۳-۲-۱- MARS [9]

MARS [10] (سیستم تحلیل و بازیابی چندرسانه‌ای) در دانشگاه ایلینویز واقع در آربانا چمپین طراحی شد. این سیستم با هدف بازیابی متن، تصویر و ویدئو، پرسش بر اساس رنگ، بافت، شکل و کلمات کلیدی را پشتیبانی می‌کند. MARS از روش‌های مبتنی

[1] Query By Image Content
[2] Coarseness
[3] Contrast
[4] Directionality
[5] Shape Area
[6] Circularity
[7] Eccentricity
[8] Major Axis Orientation
[9] Multimedia Analysis and Retrieval System

بر موجک برای تحلیل بافت استفاده می‌کند. این سیستم تلاش می‌کند تا به طور پویا با کاربرها و محیط‌های مختلف انطباق یابد. بدین منظور از یک ساختار بازخورد برای کمک به یادگیری و انطباق سیستم استفاده شده است.

Photobook -۴-۲-۱

Photobook [11] مجموعه‌ای است از ابزارهای تعاملی برای جستجوی تصاویر که در آزمایشگاه رسانه‌های MIT و در سال ۱۹۹۴ طراحی شده است. Photobook شامل سه زیرمجموعه است که در آنها به ترتیب ویژگی‌های شکل، بافت و چهره استخراج می‌شوند. با استفاده از یک عامل یادگیری تعاملی، ویژگی‌های استخراج شده بر اساس نمونه‌های مورد پرسش کاربر ترکیب شده و به کاربر در روشن کردن خواسته و نیازش برای سیستم کمک می‌شود.

Netra -۵-۲-۱

Netra [12] در دانشگاه کالیفرنیا واقع در سانتاباربارا برای پروژه کتابخانه دیجیتال الکساندریا[1] در سال ۱۹۹۷ طراحی شده است. در Netra ناحیه‌بندی و به دنبال آن استخراج رنگ، شکل، بافت و اطلاعات مکانی برای جستجو و بازیابی نواحی مشابه در دادگان تصویری انجام می‌شود. از موجک گابور[2] برای استخراج بردارهای ویژگی بافت و تبدیل فوریه سریع (FFT) برای محاسبه بردارهای ویژگی شکل استفاده شده است. پرسش از سیستم را می‌توان با استفاده از یک تصویر نمونه و یا با وارد کردن اطلاعات مکانی و رنگ به طور مستقیم انجام داد.

VisualSEEk و WebSEEk -۶-۲-۱

VisualSEEk [13] یک موتور جستجوی ویژگی بصری و WebSEEk [13] یک موتور جستجوی وب جهانگستر برای متن و تصویر هستند که هر دو در دانشگاه کلمبیا طراحی شده‌اند (سال ۱۹۹۷). ویژگی‌های تحقیقاتی این دو سیستم، ارتباط مکانی نواحی تصویر و استخراج ویژگی بصری در حوزه فشرده‌سازی شده است. ویژگی‌های بصری استفاده شده در این سیستم‌ها مجموعه

[1] Alexandria Digital Library (ADL)
[2] Gabor

رنگ[1] و ویژگی‌های بافت مبتنی بر تبدیل موجک هستند. همچنین به منظور سرعت بخشیدن به فرایند بازیابی، از الگوریتم‌های شاخص‌گذاری مبتنی بر درخت دودویی استفاده شده است. VisualSEEk از پرسش بر اساس ویژگی‌های بصری و ارتباط مکانی آنها پشتیبانی می‌کند. موتور جستجوگر وب WebSEEk از سه بخش تشکیل یافته است: ۱. بخش جمع‌آوری تصویر و ویدئو ۲. بخش کلاس‌بندی و شاخص‌گذاری موضوع ۳. بخش جستجو و بازیابی. این سیستم پرسش بر اساس کلمات کلیدی و محتوای بصری را پشتیبانی می‌نماید.

WBIIS -۷-۲-۱

WBIIS [14] یک سیستم CBIR است که در آن از موجک‌های دابشی[2] استفاده شده است (سال ۱۹۹۸). در WBIIS سیستمی برای شاخص‌گذاری و بازیابی تصاویر پیشنهاد شده است که از تبدیل موجک دابشی سه مؤلفه رنگ قرمز، سبز و آبی برای استخراج ویژگی‌ها استفاده می‌کند. شاخص‌گذاری تصاویر با ذخیره انحراف معیارها و ضرایب موجک سطح چهارم و پنجم تبدیل موجک دابشی انجام می‌شود. مزیت WBIIS تغییرناپذیری آن نسبت به اندازه تصاویر است. زیرا همه تصاویر در ابتدا و پیش از استخراج بردارهای ویژگی به اندازه یکسانی نرمالیزه می‌شوند.

Simplicity -۸-۲-۱

در سیستم Simplicity [15] (سال ۲۰۰۱)، یک مرحله پیش‌پردازش قبل از استخراج بردارهای ویژگی در نظر گرفته شده است که در آن هر تصویر در یکی از دو گروه بافت در برابر غیربافت یا گراف در برابر عکس کلاس‌بندی می‌شود. همچنین، تصاویر به چند ناحیه بخش‌بندی می‌شوند. تعدادی از ویژگی‌ها مستقیماً از مؤلفه‌های رنگ LUV و تعدادی دیگر از تبدیل موجک مؤلفه L استخراج می‌شوند. مزیت Simplicity توانایی آن در محدود کردن فضای جستجو از طریق گروه‌بندی تصاویری است که به یک طبقه یکسان تعلق دارند. اما اشکال آن زمان لازم برای انجام کلاس‌بندی است. این کار با محاسبه یک بردار ویژگی برای هر بلوک پیکسل ۴×۴ و آنگاه خوشه‌بندی آنها برای یافتن گروهی که تصویر به آن تعلق دارد، انجام می‌شود.

[1] Color Set
[2] Daubechies