بسم الله الرحمن الرحيم

دانشگاه شهید باهنر کرمان
دانشکده فنی و مهندسی
بخش مهندسی برق

رساله برای دریافت درجه دکتری رشته برق گرایش مخابرات

بازیابی تصویر مبتنی بر محتوا با استفاده از تلفیق یادگیری کوتاه مدت و یادگیری بلند مدت در فضای معناها

مولف:
عصمت راشدی

استاد راهنما:
دکترحسین نظام آبادی پور

استاد مشاور:
دکتر سعید سریزدی

مرداد ماه ۱۳۹۲

این رساله به عنوان یکی از شرایط احراز درجه دکتری به

بخش مهندسی برق
دانشکده فنی و مهندسی
دانشگاه شهید باهنر کرمان

تسلیم شده است و هیچگونه مدرکی به عنوان فراغت از دوره مزبور شناخته نمی‌شود.

دانشجو: عصمت راشدی

استاد راهنما: دکتر حسین نظام آبادی پور

استاد مشاور: دکتر سعید سریزدی

داور ۱: دکتر محمد باقر شمس اللهی (دانشگاه صنعتی شریف)

داور ۲: دکتر ناصر مزینی (دانشگاه علم و صنعت)

داور ۳: دکتر مریم پور محی آبادی

داور ۴: دکتر مهدی افتخاری

نماینده تحصیلات تکمیلی دانشکده در جلسه دفاع: دکتر مجید رحیم پور

معاونت آموزشی و پژوهشی دانشکده: دکتر مریم احتشام زاده

حق چاپ محفوظ و مخصوص به دانشگاه شهید باهنر است.

با تقدیم به:

پدر و مادر مهربان

و خواهران عزیزم:

الهام و الهه

تشکر و قدردانی:

از زحمات و راهنمایی‌های ارزنده جناب آقای پرفسور نظام آبادی پور متشکرم. نظرات و راهنمایی های ایشان نقش مهمی در پیشبرد این رساله داشته است. همچنین از راهنمایی های جناب آقای دکتر سریزدی ممنونم. از خداوند متعال سلامتی و بهروزی اساتید محترم دانشگاه شهید باهنر کرمان و خانواده ایشان را تمنا دارم. خصوصا از اساتید محترم: خانم‌ها مهندس ارجمند، دکتر مغفوری و آقایان دکتر صیدنژاد، دکتر طالبی و دکتر حکیمی، که از محضر ایشان استفاده کرده‌ام، نهایت تشکر را دارم. همچنین برای روح بزرگ جناب آقای مهندس افضلی پور بنیانگذار دانشگاه شهید باهنر کرمان طلب آمرزش و مغفرت می‌کنم.

از اساتید بزرگواری که زحمت داوری رساله فوق را بر عهده داشته اند، کمال تشکر را دارم.

والدین و خواهران عزیزم الهام و الهه همواره در امر تحصیل مشوق و پشتیبان من بوده‌اند. خداوند به ایشان سلامتی و عمر با عزت عطا کند.

این پایان نامه از پشتیبانی مالی مرکز تحقیقات مخابرات ایران در چهارچوب قرارداد شماره ۵۰۰/۱۹۲۴۵/ت مورخ ۹۰/۱۲/۲۸ بین دانشگاه شهید باهنر کرمان و آن مرکز محترم برخوردار بوده است.

چکیده

در این رساله، یک سامانه بازیابی تصویر طراحی شده است که از هر دو نوع یادگیری کوتاه مدت و بلند مدت برای بازیابی تصاویر بهره برده است. در سامانه فوق، سه رویکرد جهت ترکیب اطلاعات کوتاه مدت و بلند مدت شامل ترکیب در سطح تصاویر بازیابی شده، در سطح رتبه تصاویر و در سطح تابع شباهت ارائه شده است. از دستاوردهای دیگر این رساله ارائه‌ی دو روش جدید در یادگیری بلند مدت با نام‌های مبتنی بر مورد (CB) و مبتنی بر خوشه‌بندی معنایی آماری (SSC) است. همچنین، روش یادگیری کوتاه مدت مبتنی بر بهبود تابع شباهت با استفاده از گرادیان نزولی، بهبود داده شده است.

سامانه پیشنهادی این رساله، در یک پایگاه تصویر با ۱۰۰۰۰ تصویر از ۸۲ گروه معنایی مختلف آزموده شده است. نتایج با دو سامانه مبتنی بر ویژگی‌های مجازی (VFB) و مبتنی بر خوشه‌بندی معنایی (DSC) پویا مقایسه شده‌اند. دقت روش CB، ۱۳ درصد بیشتر از دقت روش VFB و در حد دقت روش DSC است. دقت روش SSC با ترکیب در سطح تابع شباهت، ۱۵ درصد بیشتر از روش VFB و دو درصد بیشتر از روش DSC است. روش SSC با ترکیب در سطح تصاویر بازیابی شده یا در سطح رتبه تصاویر، دقت کمتری نسبت به سایر روش‌ها دارد اما زمان بازیابی و اثر اغتشاش را بهبود داده است. در مقایسه روش‌ها، علاوه بر مقایسه گراف‌های دقت در حالت بدون اغتشاش و با اغتشاش، گراف‌های پیشرفت روش‌ها در طولانی مدت نیز استخراج شده اند.

دراین رساله، روش یادگیری کوتاه مدت بهبود یافته، با کاهش زمان بازیابی به نصف، دقت بازیابی را نسبت به روش‌های مشابه ۴ درصد ارتقاء داده است. همچنین، روش انجام همزمان انتخاب و بهبود ویژگی‌ها ضمن کاهش تعداد ویژگی‌ها به نصف، دقت بازیابی یک سامانه متداول را ۱۰ درصد بهبود داده است.

کلید واژه: بازیابی تصویر مبتنی بر محتوا، یادگیری کوتاه مدت، یادگیری بلند مدت، ترکیب اطلاعات، بهبود ویژگی، انتخاب ویژگی.

فهرست

فصل اول: مقدمه ۱

۱-۱- بازیابی تصویر ۱

۱-۲- ساختار کلی یک سامانه بازیابی تصویر بر اساس محتوا ۱

۱-۳- شکاف معنایی ۳

۱-٤- بازخورد ربط و یادگیری های کوتاه مدت و بلند مدت ٤

۱-٥- چالش ها ۸

۱-٦- اهداف رساله ۱۰

۱-۷ سازمان رساله ۱۲

فصل دوم: مروری بر اجزای سامانه بازیابی تصاویر بر اساس محتوا و کارهای انجام شده ۱٤

۲-۱- مقدمه ۱٤

۲-۲- استخراج ویژگی ۱٤

۲-۳- معیارهای عدم شباهت ۱٥

۲-٤- معیارهای ارزیابی ۱۷

۲-٥- پایگاه تصویر ۱۸

۲-٦- بهبود در سطح ویژگی های سطح پایین و ترکیب و انتخاب ویژگی ها ۱۸

۲-۷- مروری بر روش های یادگیری کوتاه مدت ۲۰

۲-۷-۱- روش های مبتنی بر طبقه بندی ۲۱

۲-۷-۲- روش‌های مبتنی بر بهبود تابع شباهت ۲۱

۲-۷-۳- روش های مبتنی بر بهبود بردار پرس و جو ۲٤

۲-۷-٤- روش های چند پرسشی ۲٦

۲-۷-٥- روش های ترکیبی ۲۹

۲-۸- مروری بر روش های یادگیری بلند مدت ۲۹

۲-۸- ۱- روش های مبتنی بر همبستگی ۳۲

۲-۸-۲- روش های مبتنی بر ویژگی های سطح پایین ۳۷

۲-۸-۳- روش های مبتنی بر ویژگی های سطح بالا ٤۱

۲-۹- جمع بندی ٤۳
فصل سوم: بهبود سامانه در سطح ویژگی های دیداری ٤٥
۳-۱- مقدمه ٤٥
۳-۲- انتخاب ویژگی ٤٥
۳-۳- بهبود ویژگی ٤٦
۳-٤- بهبود و انتخاب ویژگی به صورت همزمان ٤۸
۳-٥- الگوریتم جستجوی گرانشی ٤۹
۳-٥-۱- الگوریتم جستجوی گرانشی حقیقی ٥۰
۳-٥-۲- الگوریتم جستجوی گرانشی باینری ٥۲
۳-٥-۳- الگوریتم جستجوی گرانشی تلفیقی ٥۳
۳-٦- توصیف ویژگی های دیداری ٥٤
۳-٦-۱- هیستوگرام رنگ ٥٥
۳-٦-۲- تبدیل چند دقتی ٥٦
۳-۷- ارزیابی روش پیشنهادی ٥۸
۳-۷-۱- بهبود ویژگی ها ٦۰
۳-۷-۲- انتخاب ویژگی ها ٦۱
۳-۷-۳- بهبود و انتخاب ویژگی ها به صورت همزمان ٦٤
۳-۸- جمع بندی ٦٥
فصل چهارم: سامانه بازیابی تصویر با تلفیق یادگیری های کوتاه مدت و بلند مدت ٦٦
٤-۱- مقدمه ٦٦
٤-۲- ساختار کلی سامانه بازیابی تصویر مشارکتی پیشنهادی ٦٦
٤-۳- یادگیری کوتاه مدت مبتنی بر بهبود تابع شباهت به کمک گرادیان نزولی ٦۹
٤-٤- یادگیری بلند مدت مبتنی بر خوشه بندی آماری ۷٥
٤-٤-۱- روش کار ۷٦
٤-٥- یادگیری بلند مدت مبتنی بر مورد ۸۱
٤-٥-۱- روش کار ۸۲
٤-٥-۲- یادگیری بلند مدت مبتنی بر مورد در یک سامانه با روش یادگیری کوتاه ۸٥

مدت مبتنی بر بهبود تابع شباهت
٤-٦- ترکیب اطلاعات یادگیری های کوتاه مدت و بلند مدت ۹۳
٤-٦-١- ترکیب اطلاعات در سطح تصاویر برگردانده شده ۹۳
٤-٦-٢- ترکیب اطلاعات در سطح رتبه تصاویر ۹٤
٤-٦-٣- ترکیب اطلاعات در سطح تابع شباهت ۹٥
٤-٦-٤- استخراج اطلاعات از یادگیری های کوتاه مدت و بلند مدت ۹٦
٤-٥- جمع بندی ۹۹
فصل پنجم: پیاده سازی سامانه پیشنهادی و نتایج آزمایش‌ها ۱۰۰
٥-١- مقدمه ۱۰۰
٥-٢- سامانه پایه ۱۰۰
٥-٣- ارزیابی روش یادگیری کوتاه مدت پیشنهادی ۱۰۳
٥-٤- ارزیابی سامانه بازیابی با ترکیب اطلاعات یادگیری های کوتاه مدت و بلند مدت ۱۰۹
٥-٤-١- کارایی روش ها ۱۱۰
٥-٤-٢- تعقیب نتایج بازیابی در طولانی مدت ۱۲۷
٥-٤-٣- تحلیل پارامترها و توابع ۱۳۳
٥-٥- جمع بندی ۱۳٥
فصل ششم: جمع بندی و پیشنهاد ها ۱۳۷
٦-١- جمع بندی تحقیق ۱۳۷
٦-٢- پیشنهادها ۱۳۹
مراجع فارسی ۱٤۳
مراجع انگلیسی ۱٤٤
واژه نامه فارسی- انگلیسی ۱٥٤
واژه نامه انگلیسی - فارسی ۱٥۸
مقالات مستخرج از رساله ۱٦۲

فهرست شکل‌ها

فصل اول:

شکل ۱-۱- ساختار یک سامانه تعاملی بازیابی تصویر بر اساس محتوا و ارتباط زیر سامانه‌ها در آن [Nez04]. ۳

شکل ۱-۲- تصاویری که ویژگی‌های دیداری متفاوتی دارند اما دارای مفهوم مشترک "فیل" می‌باشند. ۳

شکل ۱-۳- مجموعه تصاویری که نشان می‌دهد مجموعه تصاویر مرتبط با هم معنا را می‌رسانند. ۵

شکل ۱-٤- بازیابی به کمک بازخورد ربط. ٦

فصل دوم:

شکل ۲-۱- دسته‌بندی انواع روش‌های یادگیری کوتاه مدت. ۲۰

شکل ۲-۲- دسته‌بندی روش‌های یادگیری بلند مدت. ۳۲

شکل۲-۳- شبکه اتصالات وزن‌دار بین تصاویر برای نمایش ارتباطات معنایی [Han05]. ۳۳

شکل ۲-٤- شبکه معنایی و یادگیری بلند مدت در [Lu03]. ۳٤

شکل ۲-۵- سامانه بازیابی با استفاده از قوانین فازی و بازخورد ربط [Lak08]. ٤۰

شکل ۲-٦- بازیابی تصویر به کمک الگوهای معنایی و روش WordNet [Zhu00]. ٤۲

فصل سوم:

شکل ۳-۱- انتخاب ویژگی به کمک الگوریتم ابتکاری باینری جهت بهبود دقت نتایج بازیابی. ٤٦

شکل ۳-۲- بهبود ویژگی به کمک الگوریتم ابتکاری حقیقی جهت بهبود دقت نتایج بازیابی. ٤۷

شکل ۳-۳- انتخاب ویژگی و بهبود ویژگی به صورت همزمان و به کمک الگوریتم ابتکاری جهت بهبود دقت بازیابی. ٤۸

شکل ۳-٤- بلوک دیاگرام انتخاب ویژگی و بهبود ویژگی به صورت همزمان و به کمک الگوریتم‌های ابتکاری جهت بهبود دقت بازیابی. ٤۹

شکل ۳-۵- تبدیل موجک دو بعدی. ۵۷

شکل ۳-٦- نمایش اجسام در روش بهبود ویژگی به کمک الگوریتم جستجوی گرانشی. ٦١
شکل ۳-۷- گراف دقت پس از بهینه‌سازی ویژگی رنگ، بهینه‌سازی ویژگی بافت، و بهینه‌سازی ویژگی رنگ و بافت در مقایسه با گراف دقت سامانه متداول بازیابی تصویر. ٦٢
شکل ۳-۸- گراف دقت بهینه‌سازی ویژگی رنگ و بافت در مقایسه با گراف دقت سامانه متداول بازیابی تصویر پس از انتخاب ویژگی. ٦٣
شکل ۳-۹- گراف دقت سامانه متداول بازیابی تصویر پس از انتخاب ویژگی به کمک روش‌های BGSA و IBGSA. ٦٣
شکل ۳-۱۰- ساختار اجسام در الگوریتم MGSA برای تنظیم پارامترها و انتخاب ویژگی به صورت همزمان. ٦٤
شکل ۳-۱۱- گراف دقت روش پیشنهادی بهبود و انتخاب ویژگی به صورت همزمان در مقایسه با گراف دقت سامانه متداول بازیابی تصویر و سامانه متداول بازیابی تصویر پس از انتخاب ویژگی. ٦٥

فصل چهارم:

شکل ٤-۱- ساختار کلی سامانه مشارکتی پیشنهادی جهت بازیابی تصویر با ترکیب اطلاعات حاصل از یادگیری‌های کوتاه مدت و بلند مدت. ٦٧
شکل ٤-۲- شبه کد مرحله آموزش در روش یادگیری بلند مدت SSC. ٧٩
شکل ٤-۳- شبه کد مرحله ادغام در روش یادگیری بلند مدت SSC. ٧٩
شکل ٤-٤- شبه کد مرحله بهره‌برداری در روش یادگیری بلند مدت SSC. ٨٠
شکل ٤-٥- ساختار الگوهای معنایی پیشنهادی. ٨٤
شکل ٤-٦- شبه کد فرایند آموزش در روش یادگیری بلند مدت CB. ٨٥
شکل ٤-۷- سامانه پیشنهادی بازیابی تصویر مبتنی بر محتوا با روش یادگیری کوتاه مدت مبتنی بر بهبود تابع شباهت و یادگیری بلندمدت CB. ٨٧
شکل ٤-۸- ساختار الگوهای معنایی پیشنهادی در فضای ویژگی. ٨٨
شکل ٤-۹- به روز رسانی الگوهای معنایی در انتهای پرس و جوی کاربر. ٩٠
شکل ٤-۱۰- شبه کد فرایند آموزش در پایان هر نشست برای یک سامانه با روش یادگیری کوتاه مدت مبتنی بر بهبود تابع شباهت و یادگیری بلندمدت CB. ٩١

شکل ٤-١١- بهبود تابع شباهت در ابتدای هر نشست به کمک اطلاعات پایگاه دانش بلند مدت. ۹۲
شکل ٤-١٢- شبه کد فرایند بهره‌برداری در هر نشست برای یک سامانه با روش یادگیری کوتاه مدت مبتنی بر بهبود تابع شباهت و یادگیری بلندمدت CB. ۹۲
شکل ٤-١٣- ترکیب اطلاعات در سطح تصاویر برگردانده شده. ۹٤
شکل ٤-١٤- ترکیب اطلاعات در سطح رتبه تصاویر. ۹٥
شکل ٤-١٥- ترکیب اطلاعات در سطح تابع شباهت. ۹٥

فصل پنجم:

شکل ٥-١- نمونه‌هایی از تصاویر گروه‌های معنایی در پایگاه تصویر. ۱۰۱
شکل ٥-٢- گراف دقت در روش‌های مختلف یادگیری کوتاه مدت. ۱۰٤
شکل ٥-٣- تعداد ۲۵ تصویر برگردانده شده برای یک تصویر نمونه در تکرار صفر و تکرار آخر هر روش یادگیری کوتاه مدت. ۱۰۷
شکل ٥-٤- گراف دقت در روش‌های یادگیری بلند مدت با استفاده از ۵۰۰، ۱۰۰۰ و ۱۵۰۰ پرس و جوی ثبت شده. ۱۱۲
شکل ٥-٥- تعداد ۲۵ تصویر برگردانده شده برای یک تصویر نمونه در تکرار آخر هر روش یادگیری بلند مدت. ۱۱۷
شکل ٥-٦- تعداد ۲۵ تصویر برگردانده شده برای یک تصویر نمونه در تکرار آخر هر روش یادگیری بلند مدت. ۱۱۹
شکل ٥-٧- گراف دقت در روش‌های یادگیری بلند مدت در حالت با و بدون اغتشاس. ۱۲۲
شکل ٥-٨- گراف‌های پیشرفت. ۱۲۸
شکل ٥-٩- مقایسه گراف‌های دقت برای مقادیر مختلف *th* در روش‌های یادگیری بلند مدت. ۱۳٤
شکل ٥-١٠- گراف دقت در روش SSC با رویکرد ترکیب در سطح رتبه تصاویر با توابع مختلف ترکیب رتبه‌ها. ۱۳٥

فهرست جدول‌ها

جدول ۱-۱- مقایسه روش‌های یادگیری بلند مدت. ۱۰

جدول ۵-۱- زمان بازیابی یک نشست و میانگین دقت درتکرار آخر بازخورد ربط برای روش‌های یادگیری کوتاه مدت مورد مقایسه. ۱۰۶

جدول ۵-۲- سامانه پیشنهادی با روش‌های یادگیری بلند مدت CB و SSC با رویکردهای ترکیب مختلف در مقایسه با دو روش VFB و DSC. ۱۲۵

فصل اول

مقدمه

۱-۱- بازیابی تصویر

تعداد تصاویر در پایگاه‌های داده روز به روز در حال رشد بوده و برچسب زنی و دسته‌بندی آنها برای انسان وقت گیر است. از سوی دیگر، سامانه‌های جستجو مانند گوگل، جستجوی تصویر را با برچسب‌هایی که توسط انسان به تصویرها داده شده انجام می‌دهد. در حالی که بسیاری از تصاویر دارای برچسب نیستند. سامانه‌های بازیابی تصویر بر اساس محتوا به منظور خودکار کردن چنین جستجوهایی به کار گرفته می‌شوند. در این سامانه‌ها، با یادگیری خودکار معناها و ارتباط بین تصویرها، یک موتور جستجو قدرت‌مندتر می‌شود. در این رساله، یک سامانه بازیابی تصویر بر اساس محتوا پیاده سازی و رویکردهای جدیدی برای یادگیری معناها در سامانه ارائه و ارزیابی خواهد شد.

یک سامانه بازیابی تصویر شامل قسمت‌های متنوعی چون استخراج ویژگی، طبقه‌بندی معنایی، یادگیری کوتاه مدت، یادگیری بلند مدت، برچسب زنی تصویر و مباحث متنوع دیگر می‌باشد. در این رساله آنچه بیشتر مورد توجه است، ارائه‌ی راه‌کارهایی جهت کاهش فاصله معنایی می‌باشد. در ادامه، ساختار کلی یک سامانه بازیابی تصویر و مباحث کلیدی آورده شده است.

۱-۲- ساختار کلی یک سامانه بازیابی تصویر بر اساس محتوا

در همه سامانه‌های بازیابی تصویر، ویژگی‌های سطح پایین تصاویر بطور خودکار استخراج و جهت نمایه‌سازی تصاویر استفاده می‌شوند. از اوایل دهه ۹۰، بازیابی تصویر بر اساس محتوا، یک زمینه فعال

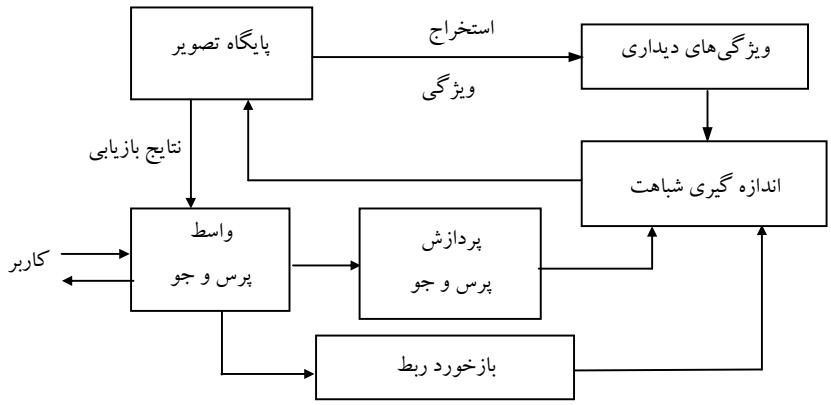
برای تحقیقات قلمداد می‌شود. تا بحال سامانه‌های بازیابی تصویر مبتنی بر محتوای (CBIR[1]) متعددی ایجاد و معرفی شده است. در همه این سامانه‌ها، ویژگی‌های سطح پایین تصاویر بطور خودکار استخراج شده و تصاویر با آنها نمایه‌سازی می‌شوند. بطور کلی، سامانه‌های بازیابی تصویر در دو مرحله عمل می‌کنند. در مرحله اول، ویژگی‌های دیداری تصاویر پایگاه بصورت خودکار استخراج شده و تصاویر با آنها نمایه‌سازی می‌شوند. در مرحله دوم، پس از دریافت تصویر پرس و جوی[2] کاربر، ویژگی‌های سطح پایین[3] یا ویژگی‌های دیداری آن استخراج شده و پایگاه ویژگی‌های دیداری برای یافتن نزدیکترین تصاویر به تصویر پرس و جو، جستجو می‌شود. در یک سامانه بازیابی تصویر، برای اجرای مرحله دوم، به یک زیر سامانه برای پردازش تصویر پرس و جو و استخراج ویژگی‌های سطح پایین آن نیاز است. ساختار یک سامانه CBIR در شکل ۱-۱ به تصویر کشیده شده است [Nez04].

در شکل ۱-۱، پایگاه تصویر شامل تصاویری است که بازیابی در بین آنها انجام می‌شود. ویژگی‌های دیداری (یا سطح پایین) مناسب از این تصاویر استخراج شده و در پایگاه ویژگی‌های دیداری نگهداری می‌شوند. این سامانه شامل یک رابط گرافیکی یا واسط پرس و جو است که با استفاده از آن با کاربر ارتباط برقرار کرده و با او در تعامل است. رابط گرافیکی تصاویر بازیابی شده را به کاربر نشان داده و اطلاعات لازم برای ادامه بازیابی را از کاربر دریافت و به سامانه ارائه می‌کند. زیر سامانه پردازش پرس و جو، ویژگی‌های مناسب را از تصویر پرس و جو استخراج می‌کند. زیر سامانه اندازه-گیری شباهت، بردار ویژگی پرس و جو را با بردارهای ویژگی تصاویر پایگاه مقایسه کرده و نزدیکترین تصاویر به تصویر پرس و جو را پیدا می‌کند. این تصاویر از طریق رابط گرافیکی به کاربر ارائه می‌شوند و با دریافت نظر کاربر راجع به آنها، تعامل با کاربر تا رسیدن به تصاویر دلخواه او ادامه می‌یابد. این فرآیند، به بازخورد ربط (RF[4]) معروف است.

سامانه‌های مختلف بازیابی تصویر، ممکن است از روش‌های مختلفی با کاربر ارتباط برقرار کنند که عبارتند از: پرس و جوی اتفاقی، ارائه یک مثال، ارائه یک طرح ساده تصویری، و پرس و جوی متنی. در پرس وجوی اتفاقی، ابتدا چند تصویر از پایگاه داده انتخاب شده و به عنوان نمونه به کاربر نمایش داده می‌شوند، سپس کاربر تصویر یا تصاویر دلخواه خود را انتخاب کرده و به سامانه اعلام می‌کند. در پرس و جو با ارائه مثال، کاربر یک تصویر از نوع آنچه به دنبال آن است به سامانه ارائه می‌کند. برای

[1] Content Based Image Retrieval (CBIR)
[2] Query image
[3] Low level
[4] Relevance Feedback (RF)

پرس و جو از نوع ارائه طرح ساده، کافی است که کاربر طرحی از آنچه که می‌خواهد بکشد و آنرا در پایگاه داده جستجو کند. پرس و جوی متنی نیز بر پایه توصیف متنی تصویر توسط کاربر استوار است.



شکل ۱-۱- ساختار یک سامانه تعاملی بازیابی تصویر بر اساس محتوا و ارتباط زیر سامانه‌ها در آن [Nez04].

۱-۳- شکاف معنایی

عموماً کاربران هنگام پرس و جوی تصویر بدنبال ویژگی‌های سطح بالا یا همان ویژگی‌های معنایی[1] هستند. در این موارد اغلب سامانه‌های بازیابی تصویر، عملکرد ضعیفی دارند چرا که بین ویژگی‌های دیداری سطح پایین و ویژگی‌های معنایی سطح بالا فاصله زیادی وجود دارد که به فاصله معنایی[2] مشهور است. یک مثال در شکل ۱-۲ آورده شده است. در این شکل، کلیه تصاویر دارای مفهوم مشترک فیل هستند اما ویژگی‌های بصری متفاوتی دارند. برای رفع این مشکل، می‌توان از بازخورد ربط و طبقه‌بندی معنایی تصاویر استفاده کرد.

شکل ۱-۲- تصاویری که ویژگی‌های دیداری متفاوتی دارند اما دارای مفهوم مشترک "فیل" می‌باشند.

[1] Semantic feature
[2] Semantic gap

از سوی دیگر، سلیقه افراد مختلف با یکدیگر متفاوت است. وقتی که کاربر احساس می‌کند که دو تصویر مشابه‌اند، در اغلب موارد بدین معنی است که این دو تصویر دارای معانی یکسان هستند، نه اینکه ویژگی‌های دیداری یکسان دارند. برای رفع این محدودیت‌ها لازم است عمل بازیابی با تعامل کاربر همراه باشد.

طبق [Liu07]، روش‌های متداول برای کاهش فاصله معنایی شامل هستان شناسی شیء [1]، استفاده از روش‌های یادگیری ماشین و بازخورد ربط می‌باشند. در روش‌های مبتنی بر هستان شناسی شیء، ویژگی‌های مفاهیمی چون آسمان، سبزه و غیره تعیین و در هنگام بازیابی معانی مانند آسمان آبی، در تصاویر جستجو می‌شوند. این ویژگی‌ها شامل رنگ، شکل، بافت، اندازه و غیره هستند. روش‌های مبتنی بر یادگیری ماشین به دو دسته با نظارت و بدون نظارت تقسیم می‌شوند. هدف از این روش‌ها برچسب زدن به تصاویر با مجموعه‌ای از مقادیر اندازه گیری شده می‌باشد. در روش‌های با نظارت، از طبقه‌بندهایی چون ماشین بردار پشتیبان یا طبقه بند بیز و در روش‌های بدون نظارت می‌توان از خوشه‌بند k-میانگین استفاده نمود. در روش‌های مبتنی بر بازخورد ربط، کاربر تصاویر مرتبط و غیر مرتبط را تعیین کرده و سامانه با یادگیری سعی در بهبود نتایج بازیابی دارد.

۱-٤- بازخورد ربط و یادگیری های کوتاه مدت و بلند مدت

در بازخورد ربط، کاربر و کامپیوتر تا رسیدن به تصاویر دلخواه کاربر با یکدیگر تعامل دارند. این تعامل به سامانه کمک می‌کند تا ارتباطی معنی دار بین ویژگی‌های سطح پایین و ویژگی‌های سطح بالا برقرار کرده و فاصله معنایی را کاهش دهد. در فرآیند بازخورد ربط، سامانه باید پرس و جوی کاربر را بیاموزد تا بتواند کاربر را در رسیدن سریع و دقیق به اهدافش کمک کند. یادگیری پرس و جوی مورد نظر کاربر نیز یکی از اهداف مهم استفاده از بازخورد ربط است که کمک شایانی به سرعت و دقت سامانه می‌کند. بازخورد ربط اولین بار توسط Rui در سال ۱۹۹۸ [Rui98] به کار گرفته شد و پس از آن محبوبت زیادی بین محققان پیدا کرد به طوری که امروزه از بازخورد ربط جهت یادگیری و طبقه‌بندی استفاده زیادی می‌شود.

در بازخورد ربط، کاربر تقاضای خود را به سامانه بازیابی تصویر اعلام می‌کند. سامانه تصاویر مرتبط اولیه را بازیابی کرده و به کاربر نمایش می‌دهد. کاربر تصاویر مرتبط و غیر مرتبط را مشخص می‌کند

[1] Object ontology

و سامانه با استفاده از اطلاعات بازخورد ربط سعی در یادگیری معناها کرده و نتایج را بهبود می‌بخشد. این کار می‌تواند چند مرحله تکرار شود.

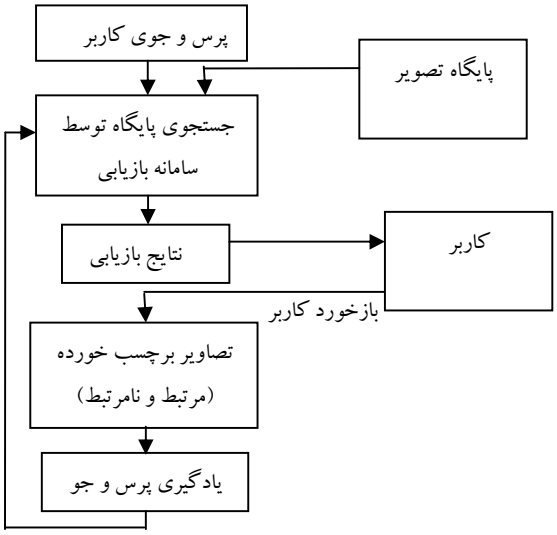
در زمینه استخراج ویژگی‌های سطح پایین تاکنون کارهای فراوانی انجام شده که نتایج آن رضایت بخش است. اما رسیدن به تصاویر دلخواه در یک مرحله کار با سامانه بازیابی، کاری دشوار و دور از ذهن است. علت این است که کاربر به دنبال جستجوی معنایی تصویر است در حالیکه سامانه، پایگاه داده را بر اساس ویژگی‌های سطح پایین جستجو می کند. همچنین، تصاویر بسیاری وجود دارد که دارای ویژگی‌های سطح پایین یکسانی هستند ولی ویژگی‌های معنایی متفاوتی دارند. در چنین شرایطی، بازخورد ربط برای کاهش شکاف معنایی بسیار کمک کننده است.

در این زمینه مشکلاتی نیز وجود دارد مانند: عدم توانایی در استخراج ویژگی‌های معنایی، کم بودن بازخوردهای داده شده توسط کاربران در پرس و جوها و نیاز به ذخیره اطلاعات بازخورد ربط برای بهبود جستجوهای آینده (چرا که دریافت معناها از ویژگی‌های سطح پایین به تنهایی میسر نیست). همچنین عموماً تعداد تصاویر مرتبط از تصاویر غیر مرتبط کمتر است.

در تعامل با کاربر، معنای تصویر پرس و جو بر این اساس تعیین می‌شود که چه تصاویری با هم به عنوان مرتبط شناخته شوند. یک تصویر با یک تعداد از تصاویر مرتبط، در یک معنای خاص اشتراک دارد. همین تصویر که با مجموعه تصاویر دیگری مرتبط شده است، در معنای دیگری مشترک است [Jia05]. به عنوان مثال در شکل ۲-۳، چنانچه کلیه تصاویر مرتبط شناخته شوند معنای "خانه" مورد پرس و جو بوده است اما اگر دو تصویر سمت راست غیرمرتبط و دو تصویر سمت چپ مرتبط شناخته شوند معنای "کوه" مورد نظر بوده است.

شکل ۱-۳- مجموعه تصاویری که نشان می‌دهد مجموعه تصاویر مرتبط با هم معنا را می‌رسانند. چنانچه همه تصاویر مرتبط شناخته شوند، مفهوم "خانه" مورد نظر کاربر بوده، اما اگر تنها دو تصاویر سمت چپ مرتبط شناخته شوند، مفهوم "کوه" مورد نظر کاربر بوده است.

طرح کلی استفاده از بازخورد ربط در سامانه بازیابی تصویر در شکل ۱-٤ آورده شده است. همانطور که در شکل دیده می‌شود، پس از هر مرحله از جستجوی پایگاه توسط سامانه بازیابی تصویر، نتایج بازیابی به کاربر ارائه شده و تصاویر مرتبط و غیرمرتبط طی بازخورد ربط کاربر برچسب می-خورند. با توجه به اعلام نظر کاربر، سامانه کار یادگیری پرس و جوی کاربر را انجام می‌دهد و مجدداً کار جستجو را انجام می‌دهد. حلقه بازخورد ربط می‌تواند تا اعلام رضایت کاربر ادامه داشته باشد. به فرایند کامل شدن یک پرس و جوی کاربر، یک نشست[1] گفته می‌شود.



شکل ۱-۴- بازیابی به کمک بازخورد ربط.

به منظور بهبود نتایج بازیابی معنایی تصاویر و کاهش شکاف معنایی، تاکنون رویکردهای مختلفی ارائه و مورد آزمایش قرار گرفته‌اند. این رویکردها در دسته‌های زیر قابل طبقه‌بندی هستند: بهبود و انتخاب ویژگی‌ها، یادگیری کوتاه مدت، یادگیری بلند مدت و برچسب‌زنی خودکار تصاویر. یکی از قسمت‌های اصلی در سامانه‌های بازیابی تصویر، قسمت استخراج ویژگی است. بنابراین در تحقیقات مختلف سعی شده با پیشنهاد ویژگی‌های جدید یا بهبود ویژگی‌های موجود، نتایج بهبود یابد. از راهکارهای دیگر برای افزایش کارایی سامانه، انتخاب ویژگی‌های مناسب است.

یادگیری کوتاه مدت در ارتباط با یک جستجوی کاربر و با هدف هدایت سامانه به سمت خواسته مطلوب کاربر انجام می‌شود. در یادگیری کوتاه مدت (STL[2]) یادگیری در خلال یک پرس و جوی

[1] Session
[2] Short Term Learning (STL)

خاص[1] صورت می‌گیرد. در تعامل هر کاربر با سامانه، سامانه به کمک یادگیری کوتاه مدت سعی در کشف معنای مورد نظر کاربر می‌کند.

در یادگیری بلند مدت، با جمع‌آوری دانش به دست آمده در نشست‌های متفاوت، از این دانش برای بهبود دقت و سرعت پاسخگویی سامانه در نشست های بعد استفاده می‌شود. به این نوع یادگیری، یادگیری بلند مدت گفته می‌شود. در یادگیری بلند مدت (LTL)[2]، یادگیری روی مجموعه زیادی از پرس و جوها [3] انجام می‌شود. در یادگیری بلند مدت، هدف جمع آوری دانش تولید شده در نشست-های قبل برای استفاده در نشست‌های آینده است. یادگیری بلند مدت، توسعه یافته یادگیری کوتاه مدت است. با استفاده از دانش به دست آمده از جستجوهای قبلی، کارایی بازیابی در جستجوهای فعلی و آینده بهبود می‌یابد.

در گذشته، تحقیقات زیادی در مورد یادگیری کوتاه مدت انجام شده است اما یادگیری بلند مدت از موضوعات مورد علاقه محققین در سالهای اخیر است. ایده استفاده از یادگیری بلند مدت از آنجا است که اگر کاربران مختلف، موارد پرس و جو شبیه به هم داشته باشند، ممکن است به دنبال مفاهیم مشترکی باشند. در نتیجه، می‌توان با ذخیره و یادگیری اطلاعات بازخورد کاربران، برای بهبود دقت بازیابی در مواجهه با کاربران در نشست های آینده استفاده کرد. یادگیری کوتاه مدت به یادگیری در یک پرس وجوی کاربر می‌پردازد و یادگیری بلند مدت به یادگیری در طول پرس و جو ها می‌پردازد. این روند به سامانه اجازه می‌دهد که به کمک یادگیری کوتاه مدت، تصاویر مرتبط‌تری را به کاربر برگرداند و به کمک یادگیری بلند مدت، معناها را در طول پرس و جوها بیاموزد.

در روش‌های برچسب زنی تصویر[4]، سامانه با برچسب زنی خودکار تصاویر، به کاربر اجازه جستجوی تصویر با کلمه کلیدی را می‌دهد. در روش‌های پیشنهاد شده در این زمینه، برچسب تعدادی از تصاویر برچسب‌دار به سایر تصاویر توسعه داده می‌شود که می‌تواند جهت بهبود نتایج بازیابی مورد استفاده قرار بگیرد. همچنین جستجوی کاربر با کلمه کلیدی قابل انجام خواهد بود.

---

[1] Intra-query
[2] Long Term Learning (LTL)
[3] Inter-query
[4] Image annotation