Archive of SID

مجله کامپیوتر و روباتیک ۱ (۱۳۸۷) ۱۵-۱

Computer & Robotics

بازیابی نیمه خودکار تصاویر با بهره گیری از برچسب های معنایی سطح بالا

شبنم اسبقی[۱]، محمدرضا کیوان پور[۲،*]

[۱] دانشگاه آزاد اسلامی واحدکرج
[۲] گروه کامپیوتر، دانشگاه تربیت مدرس

دریافت مهر۱۳۸۶، تجدید نظر فروردین۱۳۸۷، پذیرش اردیبهشت۱۳۸۷

چکیده

بازیابی مبتنی بر متن و بازیابی مبتنی بر محتوا دو رویکرد اساسی مطرح در حوزه بازیابی تصویر به شمار می‌روند. چالش های فراروی هر یک از این دو رویکرد محققین را به سمت استفاده از رویکردهای ترکیبی و بازیابی نیمه خودکار با مشارکت کاربر در چرخه بازیابی سوق داده است. براین اساس در این مقاله یک سیستم بازیابی تصویر معرفی می شود که امکان ارائه دو نوع پرس و جو براساس کلمه کلیدی و تصویر نمونه را برای کاربر فراهم می آورد. سیستم پیشنهادی، پس از بازیابی نتایج اولیه، به گونه ای نیمه خودکار با استفاده از بازخوردهای دریافتی از کاربر و برچسب های معنایی سطح بالای منتسب به تصاویر، عملیات بازیابی تعاملی تصویر را به گونه ای معنایی انجام می دهد. این سیستم با بهره گیری ازیک شبکه معنای سلسله مراتبی و انجام نوعی یادگیری از بازخوردهای کاربر قادر به پاسخگویی به پرس و جوهای متنوع در حوزه بازیابی تصویر می باشد. مطابق آزمون های انجام شده سیستم پیشنهادی از دقت قابل قبولی برخوردار است.

واژه های کلیدی: شبکه معنا، شباهت معنایی، بازیابی تصویر، یادگیری

۱. مقدمه

بازیابی تصویر، حوزه پژوهشی مهمی است که توجه بسیاری از محققین را در سالهای اخیر به خود جلب نموده است [۱]. دو رویکرد اساسی برای بازیابی تصویر از سوی پژوهشگران ارائه گردیده است: بازیابی مبتنی بر متن[۱] و بازیابی مبتنی بر محتوا[۲] [۳] [۲]. در بازیابی مبتنی بر متن فرآیند بازیابی با بهره گیری از کلمات کلیدی و انتساب آنها به تصاویر انجام می شود و مطابق با آن، قالب پرس وجوی[۳] دریافتی از کاربر به صورت کلمات می باشد.

بر این اساس معیار تشخیص شباهت میان تصاویر نیز، انطباق کلمات کلیدی منتسب به آنهاست [۴]. این رویکرد با دو چالش اساسی مواجه است [۶] [۳] :

- زمان بری و پرهزینه بودن فرایند انتساب دستی کلمات کلیدی به تصاویر
- گزینش سلیقه ای کلمات کلیدی معرف تصاویر توسط کاربر.

بازیابی مبتنی بر محتوا در پاسخ به نارسائی های موجود در رویکرد فوق ارائه گردید[۴] [۱]. در بازیابی مبتنی بر محتوا، فرایند بازیابی بر اساس ویژگی های بصری[۴] سطح پایین نظیر رنگ[۵]،

[1] Text-based Image Retrieval
[2] Content-based Image Retrieval
[3] Query

[4] Visual Features
[5] Color

www.SID.ir

Archive of SID

بافت[۶] و شکل[۷] انجام می شود. مهمترین برتری این روش نسبت به روش پیشین، قابلیت استخراج بردارهای ویژگی تصاویر بصورت خودکار[۸] می باشد [۴] [۵]. این امر، سرعت شاخص گذاری[۹] تصاویر در این روش را بصورت چشمگیری افزایش می دهد. با این وجود، بازیابی مبتنی بر محتوا با چالشی جدی مواجه است. این چالش عبارت است از وجود فاصله معنایی[۱۰] میان ویژگی های بصری سطح پائین و معانی سطح بالای موجود در تصاویر[۵] [۳]. وجود این چالش محققین را بسوی استفاده از رویکردهای ترکیبی و بازیابی نیمه خودکار[۱۱] با مشارکت کاربر در چرخه بازیابی سوق داده است[۸] [۷] [۶]. بر این اساس بازیابی نیمه خودکار با مشارکت کاربر در چرخه بازیابی و استفاده ترکیبی از کلمات کلیدی و ویژگی های بصری سطح پائین برای توصیف تصاویر، موضوع اصلی این پژوهش تلقی می شود. مشارکت کاربر در چرخه بازیابی، در قالب ارائه بازخوردهای ارتباطی[۱۲] ، به افزایش دقت و بهبود کارایی فرآیند بازیابی منتهی می‌شود [۱۲] [۱۱] [۱۰].در این مقاله، یک سیستم بازیابی تصویر معرفی می‌گردد که بصورت نیمه خودکار با کمک بازخورد های ارتباطی کاربر و برچسبهای معنایی سطح بالای منتسب به تصاویر عمل بازیابی تصویر را بگونه ای معنایی انجام می‌دهد.

از آنجائیکه امکان ارائه دو نوع پرس و جو براساس کلمه کلیدی و تصویر نمونه در این سیستم وجود دارد، عملکرد آن بنا به نوع پرس و جوی بکاررفته متفاوت خواهد بود. در واقع سیستم پیشنهادی با ارایه دو قالب پرس و جوی متفاوت، سعی در تکمیل فرآیند بازیابی معنایی به صورتی دارد که حتی در حالت پرس و جوی تصویری نیز، یک فرآیند جستجوی معنایی را جایگزین جستجوی محتوایی براساس ویژگی های بصری سطح پایین تصاویر می نماید.

سیستم پیشنهادی برای انجام نوعی یادگیری[۱۳] از کاربر، از تکنیک بازخورد ارتباطی و نیز یک شبکه معنای[۱۴] سلسله مراتبی شامل مجموعه های کلمات مترادف بعنوان گره های معنایی شبکه، برای افزایش کارایی خود در بازیابی تصاویر مرتبط[۱۵] استفاده می-نماید. در این راستا، فرآیند بازیابی تصویر با توجه به ساختار خاص شبکه معنا و نیز با در نظر گرفتن بازخوردهای ارتباطی کاربران که به صورت مرتبط و غیرمرتبط برای تصاویر بازیابی شده ارایه گردیده، نتایج را در بالاترین سطح نزدیکی به معنای مورد درخواست کاربران فراهم می نماید. قدرت روش پیشنهاد شده برای بازیابی معنایی، در حمایت از قابلیت بازیابی مترادفها[۱۶](کلمات کلیدی هم معنی) بیانگریک مفهوم معنایی و نیز رفع ابهام در مورد معانی کلمات با توجه به عضویت در مجموعه معنایی خاص به دلیل شبکه معنای استفاده شده می باشد و همچنین قدرت عملیاتی آن در خصوص بازیابی های معنایی، از طریق محاسبه شباهت معنایی میان گره های معنایی در شبکه معنا و استفاده از این شباهت ها برای تشکیل یک پرس و جوی جدید با توجه به معانی منتسب به تصاویر بازخورد شده نمود پیدا می کند. علاوه بر این، امکان بازیابی تصاویر مرتبط با پرس و جوی کاربر بر اساس انطباق دقیق و نیز انطباق غیر دقیق بصورت شباهت معنایی میان کلمات کلیدی منتسب به آنها، در صورت عدم وجود نتایج مستقیم به جای بازیابی تصادفی، در این سیستم وجود دارد. همچنین در سیستم پیشنهادی به دلیل وجود شناسه منحصر بفرد برای هر مجموعه مترادف معنایی در ساختار سلسله مراتبی شبکه، امکان استفاده از کلمه های چند معنایی نیز فراهم شده است.

در بخش دوم پیشینه تحقیق معرفی می شود. در بخش سوم روش پیشنهادی به تفصیل شرح داده می‌شود. بخش چهارم اختصاص به پیاده سازی و نتایج حاصل از آن دارد و در بخش پنجم نیز نتیجه گیری مقاله ارائه می‌گردد.

۲. پیشینه تحقیق

سیستم های بازیابی تصویر موجود یا از رویکرد بازیابی مبتنی بر متن (کلمات کلیدی) بهره می برند[۴] و یا بر اساس بازیابی مبتنی بر محتوا عمل می کنند[۱] [۵] . گروهی از سیستم های باز یابی تصویر نیز از ترکیب این دو رویکرد استفاده می کنند [۱۲] [۱۱]. استفاده از کلمات کلیدی برای بازیابی، به دلیل کاربر پسند بودن و ذات معنایی کلمات مورد توجه پژوهشگران بسیاری بوده است و جستجو از طریق آنها نتایج خوبی را به دنبال دارد، البته این تکنیک می‌تواند با تکنیکهای جستجوی محتوای بصری ترکیب

[6] Texture
[7] Shape
[8] Automatically
[9] Annotation
[10] Semantic Gap
[11] Semi-automatic
[12] Relevance Feedback
[13] Learning
[14] Semantic Net
[15] Relevant
[16] Synonyms

www.SID.ir

Archive of SID

گردد تا کاربر بتواند کنترل بیشتری روی جستجو داشته باشد. در همین راستا، بازخورد ارتباطی از تکنیکهای مکمل در زمینه بازیابی معنایی تصویر است که با امکان استفاده از نظرات کاربر، فرآیند بازیابی را به گونه‌ای معنایی به درخواست کاربر نزدیک می-کند[۱۳] [۲۱] [۲۲]. اساس عملکرد سیستم بازیابی ارایه شده در این پژوهش، بر استفاده از تکنیک بازخورد ارتباطی و انجام نوعی عملیات یادگیری در هر مرحله از بازیابی تعاملی استوار است. در [۱۷] به گونه‌ای مشابه با سیستم پیشنهادی، مسأله بازیابی در ادبیات یادگیری بر اساس تصاویر بازخورد شده به صورت مرتبط و غیر مرتبط مطرح شده است. در سیستم مذکور نوعی ارتباط یک به یک مابین کلاسهای معنایی و گروه‌های تصویری که در برچسب‌های معنایی مشترک می‌باشند انتشار می‌یابد. در واقع هدف سیستم مطرح شده، گروه گروه‌بندی تصاویر بر اساس بازخوردهای مرتبط و غیر مرتبط کاربر می‌باشد، که بدین ترتیب در حوزه یادگیری برای هر نشست بازیابی با روش پیشنهادی هم راستا و در نحوه استفاده از آن به دلیل مسایل خاص گروه‌بندی بطور غیرمشابه عمل می‌نماید. همچنین در زمینه یادگیری نظارت شده، مقاله [۲۰] نیز متد جدیدی را با استفاده از یک لغت‌نامه معنایی برای کاربرد سیستم‌های بازیابی ارایه نموده است که شامل یک راه حل یادگیری برای تخمین توزیع احتمالی مفاهیم در ارتباط با تصاویر پایگاه داده می‌باشد. فرآیند یادگیری در روش مذکور براساس کلاسهای معنایی منتسب به نواحی موجود در تصاویر عمل می نماید. علاوه بر این، پژوهش‌های [۶] و [۱۴] و [۱۵] و [۱۹] مشابه با سیستم پیشنهادی، از متدهای بازخورد ارتباطی برای تأمین نیازهای معنایی کاربران، استفاده می‌نمایند. لازم به ذکر است که روش‌های مذکور از بازخوردهای کاربران در هر دو زمینه متن و محتوای بصری تصاویر بهره برداری می‌نمایند و بر همین اساس عملکرد کاملا معنایی که مدنظر سیستم جاری است را ارایه نمی‌کنند. بطوریکه در [۶] از مجموع وزن دهی شده تشابه مبتنی بر کلمه کلیدی و تشابه مبتنی بر ویژگی بصری برای محاسبه درصد شباهت کلی یک تصویر به پرس و جو استفاده می-شود. نحوه عملکرد این سیستم برای پرس و جوی معنایی بر اساس انطباق دقیق کلمات می‌باشد که به این ترتیب امکان بازیابی تصاویر مرتبط با پرس و جو را، فقط بدلیل انتساب برچسب های مترادف دیگر آن تصاویر از دست می دهد. پژوهش‌های [۱۴] و [۱۵] نیز از یک معیار تشابه[17] برای اندازه گیری ارتباط کلی هر تصویر با منظور واقعی کاربر در خصوص کلمات کلیدی و محتوای سطح پایین استفاده می‌کنند. در این معیار، تشابه رابطه ای برای انطباق دقیق با کلمات پرس و جو همراه با میزان تشابه بصری و معنایی تصویر پایگاه داده با مجموعه های مرتبط و غیر مرتبط[18] نتیجه شده از تعامل کاربر بررسی می‌شود. مطابق با همین موضوع، در زمینه استفاده ترکیبی از معنا و ویژگی‌های بصری تصاویر، پژوهشی در[۱۸] ارایه شده است که چهارچوب استانداردی را با توجه به این توع استفاده فراهم می‌کند. عملکرد این سیستم براساس اختصاص ترکیباتی از معانی و محتوای بصری به قطعات سازمان یافته و سلسله مراتبی دلخواه از تصاویر می‌باشد.

در [۹] سیستمی معرفی شده است که با استفاده از تکنیک بازخورد ارتباطی، جستجوهای محتوایی براساس ویژگی‌های بصری را در رابطه با پرس و جوهای تصویری، برای مراحل بعدی تعاملی، به سمت جستجوهای معنایی هدایت می‌کند. در سیستم مذکور، تصاویر پایگاه داده به کلاسهای معنایی با سطوح مختلفی از درجه ارتباط تعلق دارند و جستجوی اولیه براساس ویژگی‌های محتوایی تصاویر انجام می‌گیرد. لازم به ذکر است که نحوه ارایه بازخوردهای ارتباطی در این سیستم به صورت نسبی و مقداری بین صفر و یک می‌باشد.

به همین ترتیب در [۲] نیز، روشی ارائه گردیده که از تکنیک بازخورد ارتباطی و پرس و جو، بصورت توام بر اساس کلمه کلیدی و محتوای بصری سطح پایین استفاده می‌کند. در این سیستم، مشابه با سیستم پیشنهادی، الگوریتمی برای یادگیری از تعامل با کاربر معرفی می‌گردد. الگوریتم مذکور با نام WARF[19] ارائه می-شود. در این تحقیق، تزاروس[20] (گونه ای از شبکه کلمات کلیدی) بر اساس آنالیز آماری رخداد کلمه و هم رخدادی کلمات مرتبط بصورت خودکار ساخته می‌شود. البته در این سیستم، قابلیت ساخت شبکه کلمات بر اساس فرمول WARF ، پس از ارائه بازخوردهای کاربر با بیش از یک تصویر مرتبط فراهم شده است. مشابه با سیستم پیشنهادی، مهم ترین منبع درک اطلاعات معنایی و درک منظور کاربر در این سیستم ، تعامل با اوست.

[17] Similarity metric
[18] Irrelevant
[19] Word Association via Relevance Feedback
[20] Thesaurus

www.SID.ir

Archive of SID

به گونه ای مشابه در زمینه شبکه معنای استفاده شده در سیستم پیشنهادی، تحقیق [۱۶]، یک سلسله مراتب معنایی پویا[21] و نیز یک معیار تشابه معنایی برای بهبود دقت انطباق معنایی در عملیات بازیابی را ارایه می‌دهد. شبکه معنای استفاده شده در این تحقیق شامل ارتباطات معنایی تصاویر با مجموعه ای از کلمات کلیدی مرتبط می‌باشد به گونه ای که تشابهات در این سیستم، به دو صورت تشابه کلمه با کلمه و تشابه برچسب ها با پرس و جوی کاربر مطرح می‌شوند. در ابتدا بازیابی توسط انطباق دقیق کلمات پرس و جو با کلمات شبکه معنا صورت می‌گیرد و سپس پس از تعامل با کاربر، کلمات کلیدی منتسب به تصاویر مرتبط و غیر مرتبط به عنوان شبه پرس و جوی دیگری برای یافتن نتایج بهتر عمل می‌کنند. بطور کلی تنوع قابل توجه تصاویر و معانی مندرج در آنها، کم توجهی به برچسب‌های متنی مترادف و چند معنایی ها، به عنوان نارسایی‌های مهم روش های ارائه شده در این حوزه، قابل ذکر می‌باشند[۲] [۱۱] [۱۶].

لازم به توضیح است که با مطالعه پیشینه تحقیق سیستم‌های بازیابی مبتنی بر معنا، می‌توان کاربرد بازخورد ارتباط را برای بهبود ارتباطات معنایی میان کاربر و سیستم از رایج‌ترین مفاهیم استفاده شده به شمار آورد. البته اکثر سیستم‌های مذکور، ویژگی‌های بصری سطح پایین را نیز در فرآیند بازیابی معنایی دخالت می‌دهند و همین امر یکی از نقاط اختلاف آنها با سیستم پیشنهاد شده در این پژوهش می‌باشد. علاوه بر این، ساختار خاص استفاده شده برای شبکه معنا با قابلیت بهبود نتایج بازیابی در ارتباط با کلمات هم‌معنی و چند معنایی‌ها و نیز امکان ارایه دو نوع پرس و جو هم براساس معنا و هم تصویر نمونه و به تبع آن انتقال از حوزه محتوا به معنا با کمک بازخوردهای ارتباطی بر کارایی سیستم پیشنهادی در حوزه یادگیری برای هر نشست بازیابی، می‌افزاید.

۳. روش پیشنهادی

با توجه به اهمیت استفاده ترکیبی از معنای سطح بالا و بازخورد ارتباطی کاربر در سیستم های بازیابی تصویر، در روش پیشنهادی سعی شده است تا با در نظر گرفتن ساختار مناسبی برای شبکه معنا و نیز استفاده بهینه از بازخورد های ارتباطی کاربر درک مناسب تری از درخواست او بنا بر نوع پرس و جوی ارائه شده، حاصل شود. مؤلفه های اصلی سیستم پیشنهادی عبارتند از شبکه معنا، موتور جستجو و واحد یادگیرنده. شکل (۱) مؤلفه های اصلی روش پیشنهادی و نحوه ارتباط میان این مؤلفه ها را نشان می دهد. مطابق شکل(۱)، در ابتدای کار، کاربر از طریق ارایه پرس و جو با واحد دریافت پرس و جو در واسط کاربری ارتباط برقرار می کند. در سیستم پیشنهادی، شبکه اطلاعاتی سیستم شامل پایگاه داده تصویری، شبکه معنای سلسله مراتبی و نیز یک شبکه ارتباطی برای اتصال معانی موجود در شبکه معنا با تصاویر پایگاه داده تصویری می باشد به گونه ای که موتور جستجو بنا بر نوع پرس و جو و شبکه اطلاعاتی، تصاویر را بازیابی می کند. در این حالت اگر پرس و جو براساس کلمه کلیدی باشد، واحد جستجوی معنایی و اگر پرس و جو براساس تصویر نمونه ارایه شده باشد، واحد جستجوی محتوایی توسط موتور جستجو راه اندازی می‌شود. پس از این مرحله تصاویر بازیابی شده به کاربر نمایش داده می شوند و کاربر از طریق واحد دریافت بازخورد ارتباطی، بازخوردهایی را برای تصاویر بازیابی شده بنا بر نزدیکی به معنای مورد درخواستش ارایه می دهد. واحد یادگیرنده، ابتدا براساس نوع پرس و جوی اولیه، عوامل مؤثر در ایجاد پرس و جوی جدید را تفکیک نموده و به واحد اصلاح پرس و جو ارایه می دهد.

برای پرس و جو براساس تصویر نمونه، مجموعه نمونه های مثبت و منفی بازخورد شده و برای پرس و جو براساس کلمه کلیدی، مجموعه نمونه های مثبت و منفی به علاوه پرس و جوی اولیه در ایجاد پرس و جوی جدید بکار می روند. واحد اصلاح پرس و جو، شباهت معنایی تصاویر پایگاه داده را با پرس و جوی اصلاح شده جدید برطبق عوامل مذکور و شبکه اطلاعاتی سیستم محاسبه نموده و بر این اساس تصاویر جدیدی را که به معانی مورد درخواست کاربر شبیه تر باشند را بازیابی می‌کند. در واقع عملیات واحد اصلاح پرس و جو، به انجام نوعی یادگیری از بازخوردهای کاربران برای بهبود نتایج نشست بازیابی جاری، بصورت ارزش دهی به تصاویر پایگاه داده براساس میزان شباهت معنایی به عوامل مؤثر در پرس و جوی جدید که از مرحله پیشین بازیابی تعاملی بدست آمده اند، اختصاص دارد. در ادامه مؤلفه های اصلی روش پیشنهادی و نحوه عملکرد آنها بطور مفصل تری تشریح می‌گردد.

۱،۳. شبکه معنا

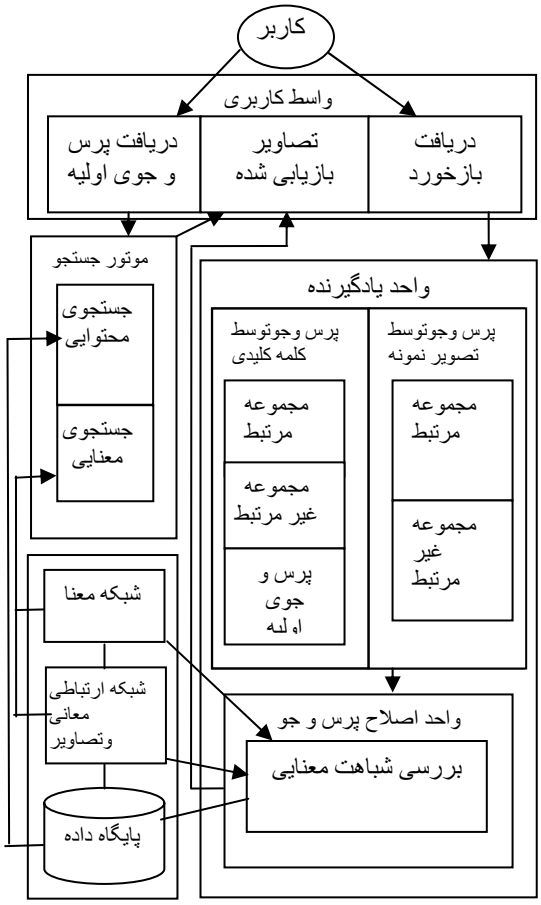
شبکه معنا یکی از مهمترین مؤلفه های روش پیشنهادی می-باشد. شبکه معنای استفاده شده در روش پیشنهادی را می‌توان همانند یک ساختار درختی دانست که گره های درخت بصورت

[21] Dynamic

www.SID.ir

Archive of SID

مجموعه ای از کلمات کلیدی مترادف (هم معنی) و ارتباطات میان گره ها، بر اساس سلسله مراتب صحیحی از روابط کلی– اختصاصی[۲۲] سازمان دهی شده است، بگونه ای که مفاهیم کلی در سطح بالای سلسله مراتب و مفاهیم خاص تر در سطوح پایین تر قرار می‌گیرند. در این شبکه معنا، هر گره مشتمل بر چندین کلمه مترادف می باشد و به هریک از کلمات موجود در یک گره، درجه عضویت یا وزن خاصی برای تعلق به آن گروه تعیین شده است.



شکل ۱. مؤلفه های اصلی سیستم پیشنهادی

- لازم به ذکر است که تمامی تلاش روش پیشنهادی، ارایه راهکاری کارآمد برای بازیابی تصاویر با معانی سطح بالا بوده است و از آنجاییکه مفاهیم مربوط به ابعاد ظاهری انسان همانند سن، قد، رنگ چهره و از این قبیل و همچنین ابعاد شخصیتی انسان همچون شادی، غم، خجالت، شرمندگی و ... در زمره مفاهیم سطح بالا قرار می گیرند، شبکه معنایی حاوی چنین مفاهیمی در سیستم بازیابی پیشنهادی مورد استفاده قرار گرفته است. شکل(۲) بخشی از شبکه معنای مورد استفاده در این پژوهش را نشان می دهد. در روش پیشنهادی هر تصویر به تک تک کلمات کلیدی موجود در شبکه معنا منتسب نمی‌شود بلکه هر تصویر با وزن خاصی با یک گره از شبکه مرتبط می‌گردد و چنانکه ذکر شد هر گره خود مشتمل بر مجموعه ای از کلمات مترادف است. بر این اساس در سیستم پیشنهادی در این بخش دو مجموعه وزنی تعریف و بکارگیری می شود:

۱- مقادیر وزنی بیان کننده درجه عضویت و اعتبار کلمات کلیدی هم معنی درون هر گره شبکه معنا.
۲- مقادیر وزنی بیان کننده ارتباط هر گره شبکه معنا با تصاویر درون پایگاه داده که نشان دهنده قدرت توصیف مجموعه مترادف در بیان محتوای معنایی تصاویر می‌باشد.

با استفاده از ساختار سلسله مراتبی علاوه بر مزیت بازیابی مترادفها و رفع ابهام در مورد معانی کلمات که با توجه به مجموعه قرار گرفته در آن به خوبی قابل تشخیص است، می‌توان شباهت کلمات کلیدی را نیز از روی رابطه کلی– اختصاصی و نیز تعلق به سلسله مراتبهای مختلف بدست آورد. به طوری که دو کلمه کلیدی که متعلق به یک سلسله مراتب نیستند شباهت معنایی نیز با یکدیگر ندارند. به بیانی دیگر، نوعی رابطه پدر- فرزندی میان سطوح شبکه معنا برقرار است و این امر بدیهی است که هر مجموعه معنایی در سطح پایین تر، که خود فرزند مجموعه معنایی دیگری در سطح بالاتر می باشد، خصوصیات پدر خود را به ارث خواهد برد. در مورد شباهت معنایی دو کلمه کلیدی در یک سلسله مراتب می توان گفت که کوچکترین جد مشترک آنها تعیین کننده شباهت میان آنان است و آن دو به دلیل رابطه سلسله مراتبی ارث بری، مشخصه ای از آن جد مشترک را بصورت مشترک دارا می-باشند. لازم به ذکر است شبکه معنای استفاده شده مشکل چند معنایی ها را نیز تا حدی حل نموده است و یک کلمه با معانی متفاوت می‌تواند در مجموعه های معنایی متفاوتی قرار بگیرد.

۱،۱،۳. دریافت و درج معانی جدید در شبکه معنا

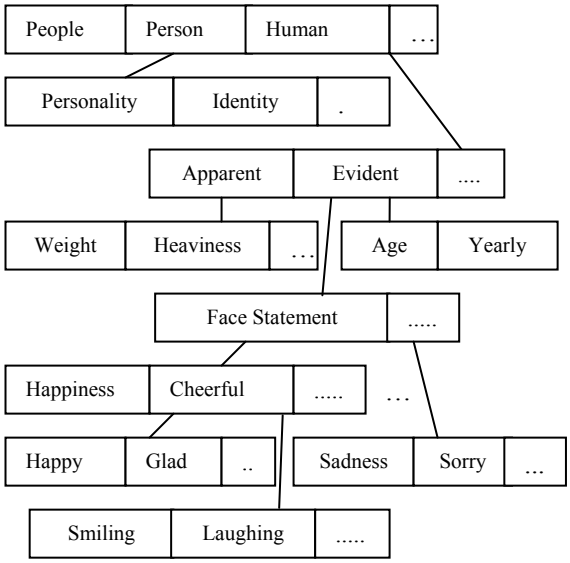
شبکه معنای مورد استفاده در روش پیشنهادی می تواند شامل چندین سلسله مراتب مختلف با حوزه های معنایی متفاوت باشد

[22] Generalization-Specialization

www.SID.ir

Archive of SID

که البته قابلیت گسترش در تمامی ابعاد، شامل تعداد کلمات هم معنی در یک گره شبکه و یا تعداد گره های معنایی در یک سلسله مراتب و همچنین تعداد سلسله مراتب‌های مختلف معنایی را داراست.



شکل۲. بخشی از شبکه معنای سلسله مراتبی

در ابتدای کار، شبکه معنا با کلمات هم معنی در گره‌های معنایی شیکه و نیز میزان اعتبار عضویت آنها در گره‌های مذکور به صورت دستی تنظیم شده است.

گسترش شبکه معنا به گونه‌ای نیمه خودکار با کمک کاربربدین ترتیب انجام می شود که در هنگام ورود کلمه جدید و یا در صورتی که کلمه کلیدی بکار رفته در پرس و جو در ساختار شبکه معنا وجود نداشته باشد، از کاربر درخواست دو کلمه کلیدی می شود که یکی از آنها بعنوان تعمیمی از کلمه جدید و دیگری مترادفی از آن خواهد بود. در صورت وجود مترادف کلمه مذکور در شبکه معنا، کلمه جدید بعنوان مترادف جدیدی در سطح گره مربوطه افزوده می‌گردد و در صورت وجود فقط تعمیمی از آن، کلمه مذکور بعنوان مشخصه جدیدی از گره معنایی تعمیم و به عنوان فرزند آن در شبکه معنا قرار خواهد گرفت و تمامی ارتباطات گره تعمیم به گره مشخصه نیز انتشار می یابد. در صورتی که مفاهیم معنایی مترادف و یا تعمیمی از کلمه جدید در ساختار شبکه نباشد، کلمه کلیدی جدید می‌تواند یک شاخه جدید را در ساختار سلسله مراتبی ایجاد نماید تا با ورود معانی دیگر در این حوزه معنایی، این سلسله مراتب تیز تکمیل شده و شبکه معنا را هرچه جامع تر نماید.

ویژگی های فوق باعث می شود که سیستم پیشنهادی از انعطاف مناسب برای توسعه شبکه معنا براساس نظرات کاربران برخوردار باشد. براین اساس اعمال تغییر در شبکه معنا نتیجه دریافت درخواست ها و بازخوردهای ارتباط کاربران است.

۲,۳. موتور جستجو

در روش پیشنهادی ارائه پرس و جو به دو صورت کلمه کلیدی و تصویرنمونه در شرایطی امکان پذیر می‌شود که هدف اصلی آن هدایت فرآیند بازیابی بر طبق معانی مورد درخواست کاربر می‌باشد. در واقع دو قالب پرس و جوی متفاوت، برای نمایاندن قدرت سیستم در ارایه بازیابی معنایی، حتی با پرس و جوی اولیه تصویری که هیچگونه معنایی را بطور مستقیم به سیستم معرفی نمی‌نماید فراهم شده است. بر این اساس در روش پیشنهادی موتور جستجو مشتمل بر دو بخش است: جستجوی معنایی و جستجوی محتوایی.

۱,۲,۳. جستجوی معنایی

در روش پیشنهادی کاربر امکان ارائه دو نوع پرس و جوی متنی ترکیبی را دارد، پرس و جوی متنی ترکیبی با استفاده از عملگر And و پرس و جوی متنی ترکیبی با استفاده از عملگر Or .در حالی که کاربر چندین کلمه کلیدی را با در نظر گرفتن رابطه Or میان آن ها جستجو می‌نماید بیان می‌کند که کلمات ذکر شده همسنگ می‌باشند و از نظر او دارای اعتبار یکسانی هستند. در این حالت برای بازیابی تصاویر مرتبط ابتدا کلمات کلیدی در شبکه معنا مورد جستجو قرار می‌گیرند و پس از بدست آمدن گره مشمول آنها، تصاویر مرتبط با این گره ها براساس رابطه (۱) ارزش گذاری می‌شوند.

$$Score(I) = W_p^k * W_I^k \qquad (۱)$$

Score (I): ارزش تصویر I.

W_p^k: وزن کلمه کلیدی p در گره معنایی k.

W_I^k: وزن ارتباطی تصویر I با گره معنایی k.

طبق رابطه(۱) برای بدست آوردن Score هر تصویر پایگاه داده در میزان شباهت با پرس و جو، دو مقدار وزنی W_p^k و W_I^k مورد استفاده قرار می گیرند، به بیانی دیگر هم درجه عضویت کلمه کلیدی p در گره مربوطه k و هم وزن ارتباطی تصویر پایگاه داده I با گره معنایی k، در تعیین شباهت میان تصویرI با کلمه کلیدی k

www.SID.ir

Archive of SID

تأثیر دارند. پس از محاسبه ارزش کلیه تصاویر پایگاه داده تمامی تصاویر به ترتیب ارزش های حاصل شده، از بالا به پایین مرتب می‌گردند تا بر اساس تعداد دلخواه بیان شده توسط کاربر ، با ارزش ترین تصاویر مجموعه بازیابی شوند. به دلیل استفاده از رابطه Or محدودیتی به جز کسب ارزش بالا در میان تصاویر برای بازیابی اعمال نمی‌گردد.

در مورد عملگر And به گونه ای متفاوت عمل می شود، بطوریکه با در نظر گرفتن رابطه And میان کلمات کلیدی پرس و جو، کاربر بطور واضح بیان می کند که خواستار تصاویری با معنای متشکل از کلیه این کلمات کلیدی می‌باشد. در این حالت ابتدا کلمات کلیدی بصورت جداگانه در شبکه معنا مورد جستجو قرار می گیرند و پس از بدست آمدن گره های مشمول آنها، تصاویر مرتبط با این گره ها بصورت جداگانه برای هریک از کلمات موجود در پرس و جو مطابق رابطه (۱) ارزش گذاری می‌شوند. اما از آنجا که کاربر بر ترکیب کلمات نظر دارد و خواستار تصاویری است که شامل کلیه کلمات کلیدی مورد نظرش باشند، ارزش نهایی هریک از تصاویر بر اساس رابطه (۲) محاسبه می‌شود.

$$Score(I) = \frac{\sum_{k=1}^{n} S_k(I)}{n} \qquad (۲)$$

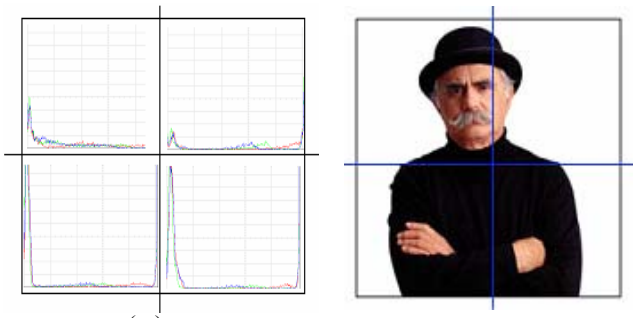
Score(I) :ارزش تصویر I (میزان پوشش کلیه کلمات کلیدی توسط تصویر I).
n: : تعداد کلمات کلیدی موجود در پرس و جوی کاربر.
$S_k(I)$: ارزش تصویر I در قبال گره معنایی K شامل یک کلمه پرس و جو.

پس از محاسبه ارزش نهایی هریک از تصاویر بر اساس رابطه (۲) تصاویری که ارزش بیشتری را کسب می‌کنند، بیان کننده میزان بیشتری از معنای درخواستی کاربر می‌باشند. بدین ترتیب تصاویر بر اساس ارزشهای بدست آمده، برای نمایش به کاربر رتبه بندی می گردند. بر این مبنا، تصاویری که در رتبه بندی، مرتبه بالاتری را کسب نمایند در مجموع حاوی درصد بالاتری از معانی مورد درخواست کاربر می‌باشند.

۲،۲،۳. جستجوی محتوایی

جستجوی محتوایی، قابلیت افزونه‌ای است که برای سیستم پیشنهادی در حالت پرس و جو براساس تصویر نمونه در نظر گرفته شده است. در این نوع جستجو، کاربر تصویری را بعنوان تصویر پرس و جو ارائه می نماید و سیستم بر اساس شباهت ویژگی های سطح پایین بصری تصاویر مشابه را بازیابی می‌نماید. برای یافتن شباهت بصری دو تصویردر سیستم پیشنهادی از عامل رنگ استفاده شده است. بر این اساس ابتدا هیستوگرام رنگ برای هر تصویر بصورت جداگانه محاسبه شده و در نهایت میزان شباهت دو تصویر از طریق محاسبه تفاوت میان هیستوگرامهای رنگ با استفاده از تابع فاصله اقلیدسی بدست می آید. هیستوگرام رنگ برای دو تصویر در این روش به صورت هیستوگرام رنگ سه بعدی محاسبه می گردد. در اینحالت به ازاء هر رنگ در فضای سه بعدی، بصورت مجزا اشتراک هیستوگرام در میزان حداقل پیکسلهای مشابه بدست می آید، سپس عدد حاصل در تعداد کل پیکسلهای تصویر تقسیم می شود تا بدین ترتیب اشتراک هیستوگرام برای هر رنگ بطور نسبی حاصل گردد. در مرحله آخر، میانگین برای سه فضای رنگی محاسبه شده و شباهت میان دو تصویر بدست می آید.

برای بالا بردن دقت در روش پیشنهادی، ابتدا تصویر به ۴ ناحیه با سایز برابر مطابق شکل (۳)- الف تقسیم شده و سپس هیستوگرام رنگ برای هر ناحیه بصورت جداگانه و مستقل طبق شکل (۳)- ب محاسبه می گردد. در نهایت درصد شباهت میان دو تصویر، از میانگین شباهت میان نواحی منطبق آنها براساس هیستوگرام های متناظرشان بدست می آید. بدین ترتیب تصاویر درون پایگاه داده بر اساس شباهت بصری با تصویر پرس و جو از بالا به پایین برای بازیابی رتبه بندی می‌شوند.



شکل۳. الف- تقسیم تصویر به ۴ ناحیه برابر. **ب**- محاسبه هیستوگرام رنگ سه بعدی برای هر ناحیه

۳،۳. واحد یادگیرنده

واحد یادگیرنده دیگر مؤلفه مهم سیستم پیشنهادی است. این واحد با دریافت بازخوردهای کاربر در طی یک نشست بازیابی، نوعی یادگیری از رفتار کاربر را در تشخیص معانی مندرج در تصاویر انجام می‌دهد. در سیستم پیشنهادی واحد یادگیرنده، قابلیت هدایت چرخه بازیابی را به‌گونه‌ای معنایی در هر دو حالت

www.SID.ir

Archive of SID

پرس و جو داراست و برهمین اساس، عملکردهای نسبتا متفاوت در مورد پرس و جو بر اساس کلمه کلیدی در دو حالت بکارگیری عملگرهای And و Or و نیز پرس و جو بر اساس تصویر نمونه دارد. در حوزه پرس و جو براساس کلمه کلیدی، واحد یادگیرنده با در نظر گرفتن پرس و جوی اولیه و نیز معانی منتسب به تصاویر بازخورد شده قادر است تا یک نشست بازیابی کاملا معنایی را ارایه نماید و همچنین در دیگر حالت پرس و جو براساس تصویر نمونه نیز، سیستم پیشنهادی، فرآیند بازیابی را پس از مرحله آغازین که به ناچار به صورت جستجوی محتوایی انجام پذیرفته است، توسط واحد یادگیرنده از حوزه محتوای بصری به حوزه معنا منتقل می نماید. در واقع واحد یادگیرنده در سیستم پیشنهادی، پرس و جوی کاملتری را که بصورت صریح نظرات کاربر را بیان کند، ایجاد می‌نماید و به این ترتیب این امکان را در اختیار کاربر قرار می‌دهد که نوعی اعمال هدایت بر فرآیند بازیابی داشته باشد.

۱،۳،۳. واحد یادگیرنده در حالت پرس و جو بر اساس کلمه کلیدی

روش کار در این بخش به این صورت است که پس از رتبه بندی تصاویر پایگاه داده بر اساس پرس و جوی کاربر به روش پیش گفته، تصاویری که بالاترین رتبه ها را دارا هستند به کاربر نمایش داده می‌شوند و کاربر برچسب های مرتبط یا غیر مرتبط را به هر یک از آنها منتسب می کند. واحد یادگیرنده با دریافت و پردازش بازخوردهای کاربر و یادگیری از او تغییراتی را در رابطه های (۱) و (۲) اعمال می‌کند. این فرایند طی جلسات متعدد تا حصول رضایت نسبی کاربر تکرار می‌شود.

در حالی که کاربر چندین کلمه را با در نظر گرفتن رابطه And میان آنها جستجو می کند، در واقع بیان می‌کند که خواستار تصاویری است که مشتمل بر کلیه کلمات کلیدی داده شده هستند. در این حالت سه عامل در پالایش پرس و جو و تعیین ارزش جدید هریک از تصاویر پایگاه داده مؤثر می باشند: الف) میزان شباهت هر تصویر به مجموعه تصاویر مرتبط ب) میزان شباهت هر تصویر به مجموعه تصاویر غیر مرتبط ج) میزان شباهت هر تصویر با پرس و جوی اولیه ارائه شده توسط کاربر. رابطه (۳) چگونگی پالایش پرس و جوی اولیه و محاسبه ارزش هریک از تصاویر پایگاه داده بر اساس مؤلفه های پیش گفته را در حالتی که از عملگر And برای ترکیب کلمات کلیدی استفاده شده است نشان می‌دهد. در این روش کوچکترین جد مشترک میان دو گره معنایی بیانگر شباهت معنایی میان آنهاست. درواقع در اینجا اساس کار بر بدست آوردن شباهت معنایی میان گره‌های معنا استوار است. مطابق این رابطه مساله مهم در مورد شباهت میان دو گره در شبکه معنا، عمق جد مشترک(DoLA[۲۳]) آنهاست که هر چند عمق بیشتری داشته باشد صورت کسر بیشتر شده و نشان دهنده شباهت بیشتر میان دو گره شبکه حتی در مفاهیم جزئی می‌باشد. لازم به توضیح است که در شرایط انطباق دو گره معنایی از دو تصویر، مقدار DoLA برابر با بیشینه عمق موجود در سلسله مراتب خواهد شد. بر این اساس، روابطی که از مفهوم DoLA استفاده می نمایند را می توان نمونه هایی با قابلیت محاسبه انطباق دقیق و همچنین شباهت معنایی در صورت عدم انطباق دقیق میان مجموعه های معنایی شبکه معنا در نظر گرفت. همانطور که ذکر شد اساس محاسبات برای بدست آوردن شبه پرس و جوی کاملتر بر طبق نظرات کاربر، بر شباهت معنایی میان کلمات کلیدی منتسب به تصاویر بازخورد با دیگر تصاویر پایگاه داده بنا بر کلمات کلیدی بیان کننده محتوای معنایی آنها تکیه دارد. روش کار بدین ترتیب است که در پایگاه داده تصویری، تصاویر به ترتیب طبق رابطه (۳) بر اساس پرس و جوی جاری و بازخوردهای ارتباطی در قالب برچسب های مرتبط و غیر مرتبط ارایه شده توسط کاربر ارزش گذاری می گردند. مطابق این رابطه، ابتدا بیشترین میزان شباهت معنایی میان هر گره تصویر مرتبط با گره های معنایی تصویر I، بدست آمده و سپس میانگین شباهت های مذکور برای تمامی گره های تصویر مرتبط محاسبه می شود. در نهایت، میانگین شباهت های معنایی میان کل تصاویر مجموعه مرتبط برای بخش اول رابطه حاصل می گردد و به همین ترتیب برای مجموعه غیرمرتبط نیز این مقدار میانگین محاسبه گشته و در آخر این دو مقدار بدست آمده با مقدار Similarity که بیانگرمیزان شباهت تصویر I با پرس و جوی کاربر است، جمع می شود. به بیانی دیگر رابطه (۳) به محاسبه شباهت معنایی تصویر I با کل مجموعه مرتبط و غیرمرتبط و نیز تمامی کلمات موجود در پرس و جوی اولیه در حالت استفاده از عملگر And می پردازد.

$$Score(I)= \qquad (۳)$$

$$\frac{1}{K}\sum_{K}\left\{\frac{1}{m}\sum_{j}^{1..m}\left\{\max_{i}^{1..n}\left(\frac{DoLA(I_i,K_j)}{Max.Depth}\right)\left(\frac{W_{Ii}}{\sum W_{I_i}}\right)\left(\frac{W_{Kj}}{\sum W_{K_j}}\right)\right\}\right\}-$$

[23] Depth of Lowest Ancestor

www.SID.ir

Archive of SID

$$\frac{1}{L}\sum_{L}\left\{\frac{1}{m}\sum_{j}^{1..m}\left\{\max_{i}^{1..n}\left(\frac{DoLA(I_i,L_j)}{Max.Depth}\right)\left(\frac{W_{Ii}}{\sum W_{I_i}}\right)\left(\frac{W_{Lj}}{\sum W_{L_j}}\right)\right\}\right\}$$
$$+Similarity(I,Q)$$

K: تعداد تصاویر مجموعه مرتبط.

L: تعداد تصاویر مجموعه غیر مرتبط.

I_i: معنای i منتسب به تصویر I پایگاه داده.

K_j: معنای j منتسب به تصویر K ام مجموعه مرتبط.

L_j: معنای j منتسب به تصویر L ام مجموعه غیر مرتبط.

Q_j: معنای j شامل یکی از کلمات کلیدی پرس و جو.

Max.Depth: ماکزیمم عمق موجود در سلسله مراتب معانی.

DoLA: عمق نزدیک ترین جد مشترک مجموعه‌های معنایی i و j در سلسله مراتب معانی.

همچنین رابطه (۴) نحوه محاسبه ارزش هریک از تصاویر پایگاه داده را از منظر میزان شباهت آنها با پرس و جوی اولیه کاربر نشان می‌دهد. برای این منظور از مقدار پیش فرض ۰/۰۵ استفاده شده است که برای بررسی اعتبار معنایی هر گره معنا در مورد تصویر مورد بررسی در حالت And ضروری است تا تصویر I در صورتی برای نمایش به کاربر بازیابی شود که بیان کننده تمامی معناهای مورد درخواست کاربر باشد. انتخاب مقدار فوق به عنوان پیش فرض براساس نتایج بدست آمده از آزمون های تجربی صورت گرفته است. در نهایت پس از بررسی تمامی تصاویر در حالت And برای بازیابی مفهومی تر، الزاما باید عملیات میانگین گیری انجام شود تا امتیاز تصویر از طریق رتبه بندی بر اساس Score تصاویر صورت گیرد و بازیابی دیگری بر اساس این رتبه بندی انجام گردد.

(۴)

$$\forall S_j > c(0.05), j = 1..\text{Qkey_num}|$$
$$S_j = \max_{i}^{1..n}\left(\frac{DoLA(I_i,Q_j)}{\text{maxDepth}}\right)*\left(\frac{W_{Ii}}{\sum W_I}\right)*\left(\frac{\text{W.of.key.in}\,Qj}{\sum \text{W.of.key.in}\,Qj}\right)\Rightarrow$$
$$\text{Score} = \text{Score} + S_j$$
$$\text{Similarity(I,Q)} = \frac{\text{Score}}{\text{Qkey_num}}$$

در حالتی که کاربر چندین کلمه کلیدی را با بهره گیری از عمگر Or به عنوان پرس و جوی اولیه ارائه نماید نیز سه عامل پیش گفته در تعیین ارزش جدید هریک از تصاویر پایگاه داده از منظر میزان انطباق آنها با نیاز کاربر مؤثر می‌باشند. البته در این بخش مقایسه ها کمی متفاوت انجام می‌شود و نیازی به انجام فرآیند میانگین گیری نمی‌باشد بطوریکه با استفاده از بدست آوردن مقدار بیشینه در تمامی بخش های مقایسه ای میان گره های معنایی بدلیل اینکه پرس و جو در حالت Or انجام می گیرد نتیجه بدست آمده، از لحاظ جستجوی مورد درخواست کاربر بهینه است زیرا که در حالت Or، شباهت معنایی با حتی یکی از کلمات مورد جستجو کفایت می‌کند. رابطه (۵) روش پالایش پرس و جوی اولیه و محاسبه ارزش جدید هر یک از تصاویر پایگاه داده، بر اساس مؤلفه های سه گانه پیش گفته را در حالتی که از عملگر Orبرای ترکیب کلمات کلیدی استفاده شده است نشان می‌دهد.

Score(I)= (۵)

$$\max_{K}\left\{\max_{j}^{1...m}\left\{\max_{i}^{1...n}\left(\frac{DoLA(I_i,K_j)}{\max Depth}\right)\left(\frac{W_{I_i}}{\sum W_{I_i}}\right)\left(\frac{W_{K_j}}{\sum W_{k_j}}\right)\right\}\right\}-$$

$$\max_{L}\left\{\max_{j}^{1...m}\left\{\max_{i}^{1...n}\left(\frac{DoLA(I_i,L_j)}{\max Depth}\right)\left(\frac{W_{I_i}}{\sum W_{I_i}}\right)\left(\frac{W_{L_j}}{\sum W_{L_j}}\right)\right\}\right\}$$
$$+Similarity(I,Q)$$

مطابق این رابطه، ابتدا بیشترین میزان شباهت معنایی میان هر گره j از تصویر k در مجموعه مرتبط و گره های معنایی منتسب به تصویر I پایگاه داده محاسبه شده و سپس بیشینه مقدار در میان شباهت های بدست آمده برای کل گره های معنایی تصویر k با معانی تصویر I محاسبه می گردد. در نهایت بیشترین میزان شباهت میان کل تصاویر مجموعه مرتبط و تصویر I برای بخش اول رابطه بدست می آید و به همین ترتیب در مورد بخش دوم رابطه برای مجموعه غیر مرتبط عمل می شود. به دیگر بیان اساس رابطه(۵) بر محاسبه ارزش تصویر I بر اساس شباهت معنایی به یکی از تصاویر مجموعه های مرتبط و غیر مرتبط از طریق شباهت

www.SID.ir

Archive of SID

به حتی یکی از مفاهیم منتسب به تصاویر و نیز شباهت معنایی به حتی یکی از کلمات کلیدی پرس و جو، استوار می باشد، که در حالت استفاده از عملگر Or می تواند جوابگو باشد. در این حالت برای بخش Similarity، ارزش هر یک از تصاویر پایگاه داده از منظر میزان شباهت آنها با پرس و جوی اولیه کاربر مطابق رابطه (۶) قابل محاسبه است

$$\text{Similarity}(I,Q) = \max_{j}^{1\ldots q_j}\left\{\max_{i}^{1\ldots n}\left(\frac{DoLA\ (I_i,Q_j)}{\max\ Depth}\right)\left(\frac{W_{I_i}}{\sum W_{I_i}}\right)\times\left(\frac{W\ .of\ .query\ \ .keyword\ \ .in\ .Q_j}{\sum W\ .of\ .keywords\ \ .in\ .Q_j}\right)\right\} \qquad (۶)$$

استفاده از مقدار بیشینه در روابط (۵) و (۶) به این دلیل است که شباهت بالای معنایی تصویر پایگاه داده حتی با یکی از تصاویر مجموعه مرتبط، مجموعه غیر مرتبط و پرس و جو به دلیل استفاده از عملگر Or برای کسب نتایج قابل قبول کفایت می کند.

۲،۳،۳. واحد یادگیرنده در حالت پرس و جو بر اساس تصویر نمونه

نحوه عملکرد واحد یادگیرنده در این بخش، مشابه با حالت پرس و جو بر اساس کلمه کلیدی، بر معانی بدست آمده از تصاویر بازخورد شده توسط کاربر استوار است با این تفاوت که در روش پیش گفته، ارزش معنایی تصویر پایگاه داده بر اساس تصاویر مجموعه های مرتبط و غیر مرتبط و نیز پرس و جوی اولیه مورد بررسی قرار می گرفت، در حالیکه در حالت پرس و جو توسط تصویر نمونه، ارزش نهایی تصویر پایگاه داده بر اساس معانی منتسب به تصاویر مجموعه مرتبط و مجموعه غیر مرتبط بدست می آید. در این مورد به دلیل اینکه پرس و جوی اولیه بصورت نمونه تصویری می باشد، امکان ارجاع مستقیم معنایی به پرس و جو در قالب کلمات کلیدی وجود ندارد و بدین ترتیب عملیات یادگیری برای این بخش، فقط از معانی مجموعه های مرتبط و غیر مرتبط استفاده می نماید. به دیگر بیان، در این حالت پرس و جو بر اساس تصویر نمونه، با پرس و جوی جدید بر اساس جستجوی معنایی جایگزین می گردد. در این بخش دو عامل در تعیین ارزش جدید هر تصویر پایگاه داده و مرتب سازی این تصاویر مؤثر می باشند:

الف) میزان شباهت معنایی هر تصویر به مجموعه مرتبط.

ب) میزان شباهت معنایی هر تصویر به مجموعه غیرمرتبط.

مطابق دو عامل مطرح شده و نیز عملیات ارزش دهی به تصاویر بر اساس شباهت معنایی بصورت کوچکترین جد مشترک میان مجموعه های معنایی دو تصویر، رابطه(۷) نحوه عملکرد واحد یادگیرنده را در محاسبه ارزش هر یک از تصاویر پایگاه داده در حالت پرس و جو توسط نمونه تصویری نشان می دهد. طبق این رابطه، فقط معانی منتسب به مجموعه های مرتبط و غیر مرتبط در تعیین ارزش هر تصویر پایگاه داده تأثیر می گذارند و بدین ترتیب کمتر به شباهت محتوایی میان تصویر نمونه و تصاویر پایگاه داده توجه می شود.

$$\text{Score}(I)= \max_{K}\left\{\frac{1}{m}\sum_{j}^{1..m}\left\{\max_{i}^{1..n}\left(\frac{DoLA(I_i,K_j)}{Max.Depth}\right)\left(\frac{W_{Ii}}{\sum W_{I_i}}\right)\left(\frac{W_{Kj}}{\sum W_{K_j}}\right)\right\}\right\} - \max_{L}\left\{\frac{1}{m}\sum_{j}^{1..m}\left\{\max_{i}^{1..n}\left(\frac{DoLA\ (I_i,L_j)}{Max\ .Depth}\right)\left(\frac{W_{Ii}}{\sum W_{I_i}}\right)\left(\frac{W_{Lj}}{\sum W_{L_j}}\right)\right\}\right\} \qquad (۷)$$

مطابق رابطه(۷)، ابتدا بیشینه مقدار برای شباهت میان هر گره معنایی تصویر مرتبط با گره های معنایی تصویر I پایگاه داده بدست آمده و سپس میانگین این مقادیر برای کلیه گره های معنایی تصویر مرتبط محاسبه می گردد. در نهایت بیشترین میزان شباهت میان تصاویر مرتبط و تصویر I به عنوان بخش اول رابطه محاسبه می شود. برای بخش دوم نیز، شباهت معنایی تصویر مذکور با مجموعه غیر مرتبط مورد بررسی قرار می گیرد. در واقع مطابق این رابطه، ارزش هر تصویر پایگاه داده بر اساس شباهت مفهومی به حداقل یکی از تصاویر مجموعه های مرتبط و غیر مرتبط، البته با بررسی تمامی معانی موجود در تصویر مربوطه محاسبه می گردد.

روابط بالا را می‌توان به صورت اشکال تعمیم یافته ای از فرمول Rocchio [۱۴][۱۶] در نظر گرفت. مقادیر وزنی بکار رفته در روابط بالا همگی نسبی می باشند و در واقع وزن نسبی هر گره را در بیان معنای تصویر بیان می کنند و در بخش Similarity نیز مقدار وزنی بکار رفته به اعتبار نسبی کلمه کلیدی پرس و جو در نود مشمولش اشاره دارد. بدین ترتیب واحد یادگیرنده براساس

www.SID.ir

Archive of SID

بازخوردهای ارتباطی ارایه شده در هر مرحله، بازیابی مفهومی که به معانی مورد درخواست کاربر نزدیکتر است را انجام می دهد.

۴. پیاده سازی و آزمون

یکی از مهمترین چالش های فراروی ارزیابی کارایی سیستم های بازیابی تصویر فقدان یک بانک اطلاعاتی جامع (محک[۲۴] استاندارد) در این زمینه است. اگرچه در سال های اخیر تلاش هایی جهت رفع این چالش انجام گرفته است لیکن به دلیل تنوع زیاد حوزه های کاربردی و نیازهای کاربران این سیستم ها، معضل مذکور همچنان باقی است. براین اساس در این تحقیق سعی شده است که از اصول و مراجع شناخته شده برای تدوین پایگاه داده مورد استفاده در آزمون روش استفاده شود. پایگاه داده تصویری استفاده شده در این بخش شامل ۱۰۰۰ تصویر از مجموعه شناخته شده Corel با موضوع انسان می باشد که در این رابطه سعی شده است تا تصاویری مورد گزینش قرار گیرند که از جامعیت لازم برای پوشش حالتهای مختلف چهره(ظاهری و شخصیتی) برخوردار باشند. همچنین در این بخش از یک شبکه معنای سلسله مراتبی مشتمل بر ۲۰۰ کلمه کلیدی در مجموعه های مترادف استفاده شده وپیاده سازی سیستم مذکور با نرم افزار MATLAB انجام گرفته است.

مراحل آزمون سیستم پیشنهادی در هر دو حالت پرس و جو براساس تصویر نمونه و کلمه کلیدی با انتخاب تصادفی کلمات و تصاویرصورت گرفته است. در این راستا به ازا هریک از گره های شبکه معنا در بیست مرحله مختلف یک کلمه کلیدی یا یک تصویر نمونه برای انجام آزمون انتخاب شده است. در نهایت میانگین نتایج بدست آمده برای هر مفهوم به عنوان نتیجه کلی عملکرد سیستم گزارش شده است.

آزمون های انجام شده در این بخش برای پرس و جو بر اساس کلمه کلیدی، درهر دو حالت And و Or و البته انتخاب تصادفی کلمات انجام پذیرفته است. در آغاز پرس و جو، کاربر کلمات کلیدی را وارد نموده و در هر مرحله بازیابی، نظرات خود را در قالب بازخوردهای ارتباطی، بصورت "مرتبط" و "غیرمرتبط" ارایه می نماید و به همین ترتیب برای جستجوی محتوایی، کاربر ابتدا تصویری را بعنوان تصویر نمونه برای پرس و جو انتخاب نموده و سپس مراحل تعامل خود را با سیستم، از طریق بازخورد تصاویر بصورت "مرتبط" و "غیرمرتبط" ادامه می دهد. در هر دو حالت کاربر تعداد مراحل تعامل را تا حصول رضایت از نتایج بازیابی تکرار می‌نماید. لازم به توضیح است که ساختار شبکه معنای استفاده شده، ساختاری پویا با قابلیت گسترش در قالب کلمات کلیدی مترادف و بطور کلی تر در حوزه های معنایی مختلف می باشد که بر همین اساس امکان کاربرد سیستم بازیابی پیشنهاد شده با پایگاه داده هایی با مفاهیم متفاوت معنایی فراهم می‌گردد.

شکل (۴) شمایی از سیستم بازیابی پیشنهادی در پاسخگویی به پرس و جو بر اساس کلمه کلیدی "Sad" را نشان میدهد.

در همین راستا، شکل (۵) شمایی از سیستم پیشنهادی را در حالت پرس و جو بر اساس تصویر نمونه نشان می دهد. در این حالت تصاویر بر اساس جستجوی محتوایی در قالب شباهت بصری با تصویر پرس و جو بازیابی شده اند.

در این سیستم کاربر با فشردن کلید " Continue"، بازیابی معنایی دیگری را آغاز می‌نماید. در آزمون های انجام شده ۲۰ پرس و جوی تصادفی در هر دو حالت پرس و جو بر اساس کلمه کلیدی با دو عملگر And و Or و پرس و جو بر اساس تصویر نمونه

شکل۴. سیستم بازیابی تصویر پیشنهادی در حالت پرس و جو بر اساس کلمه کلیدی

بررسی گردیده است. لازم به توضیح است که پرس و جو بر اساس تصویر نمونه، برای دو حالت جستجوی معنایی و جستجوی محتوایی برای مراحل تعاملی با کاربر مورد بررسی قرار گرفته است.

1- Benchmark

www.SID.ir

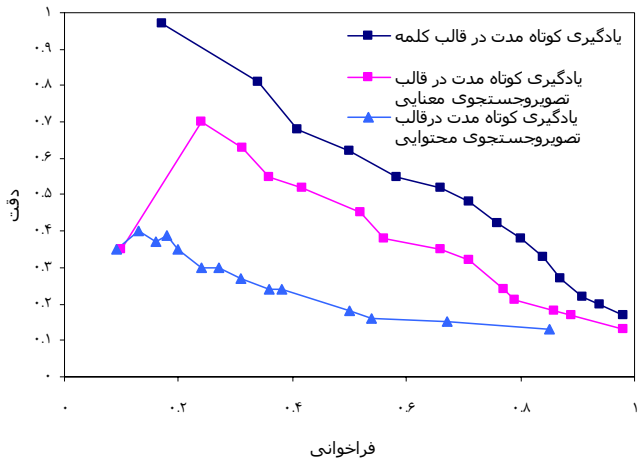
Archive of SID

در این بخش با توجه به هدف سیستم بازیابی پیشنهادی که ارایه نشست های بازیابی معنایی و جلب رضایت کاربران در معانی درخواست شده می باشد، آزمون های انجام شده در دو قالب معنا و محتوای بصری تصاویر برای نمایش کارایی سیستم مذکور در حوزه

شکل۵. سیستم بازیابی تصویر در حالت پرس و جو بر اساس تصویر

معنا انجام شده است. به دیگر بیان عملکرد سیستم پیشنهادی در یک حالت که با پرس و جو براساس کلمه کلیدی آغاز شده و بصورت کاملا معنایی با استناد به معانی مرتبط با تصاویر بازخورد شده از طریق واحد یادگیرنده ادامه می‌یابد، با دیگر حالت سیستم که با یک پرس و جوی تصویری راه اندازی شده و سپس توسط واحد یادگیرنده از حوزه تصویر و محتوای بصری سطح پایین به حوزه معنا از طریق بازخوردهای ارتباطی کاربر منتقل می‌گردد مقایسه شده است. برای تکمیل روند مقایسه میان بازیابی براساس معنا و محتوای بصری تصاویر، عملکرد سیستم با حالتی دیگر که به تمامی براساس جستجوی محتوایی انجام می‌پذیرد مورد ارزیابی و مقایسه قرار گرفته است. در این حالت پس از بازیابی اولیه تصاویر که براساس ویژگی سطح پایین رنگ صورت می پذیرد، مراحل بعدی تعاملی نیز بدون توجه به معانی مندرج در تصاویر بازخورد شده با استناد به محتوای بصری تصاویر همانند دیگر سیستم های مبتنی بر محتوا ادامه می‌یابد. مطابق نتایج بدست آمده، نمودار دقت - فراخوانی[25] سیستم پیشنهادی در طی چندین مرحله بازیابی معنایی برای سه حالت مذکور در شکل(۶) قابل نمایش است. چنانکه ملاحظه می شود در حالت پرس و جو بر اساس کلمه کلیدی، روش پیشنهادی از عملکرد قابل قبولی در بازیابی تصاویر برخوردار است و در حالت پرس و جو بر اساس تصویر نمونه، به دلیل دقت پایین جستجوی محتوایی برای عملیات بازیابی معنایی تصاویر، در اولین مرحله دقت بازیابی در نمودار دقت- فراخوانی ارائه شده در حد پایینی قرار گرفته و برای مراحل بعدی که بازیابی تصاویر بر اساس جستجوی معنایی ادامه می یابد دقت نتایج افزایش یافته است، در حالی که برای قسمت دیگر نمودار که به نوعی عملکرد سیستم های بازیابی مبتنی بر محتوا را به نمایش می گذارد و به پرس و جوی تصویری و عملیات پردازش محتوایی برای مراحل بعدی بازیابی تعاملی اختصاص دارد، نمودار دقت – فراخوانی نشان دهنده عدم قابلیت اعتماد به جستجوی محتوایی برای عملیات بازیابی معنایی تصاویر می باشد.

به دیگر بیان، در حالت پرس و جو بر اساس کلمه کلیدی، به دلیل ارجاع معنایی مستقیم به درخواست پرس و جو در قالب کلمات، عملکرد سیستم در تمامی مراحل بازیابی نسبت به پرس و جو بر اساس تصویر نمونه دقت مناسب تری را ارائه می دهد و همچنین در حالت پرس و جو توسط تصویر نمونه نیز، تکرار مراحل تعامل بر اساس معانی موجود در تصاویر بازخورد شده نتایج بهتری را نسبت به جستجوی محتوایی با استفاده از ویژگی های بصری تصاویر بازخورد شامل می شود.



شکل۶. نمودار دقت – فراخوانی سیستم پیشنهادی

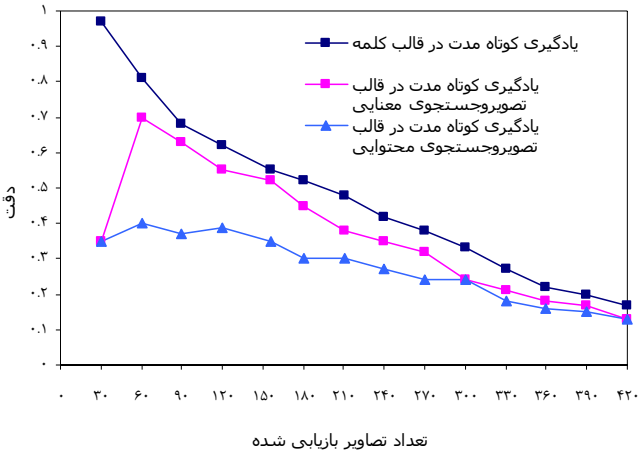
[25] Precision-Recall

www.SID.ir

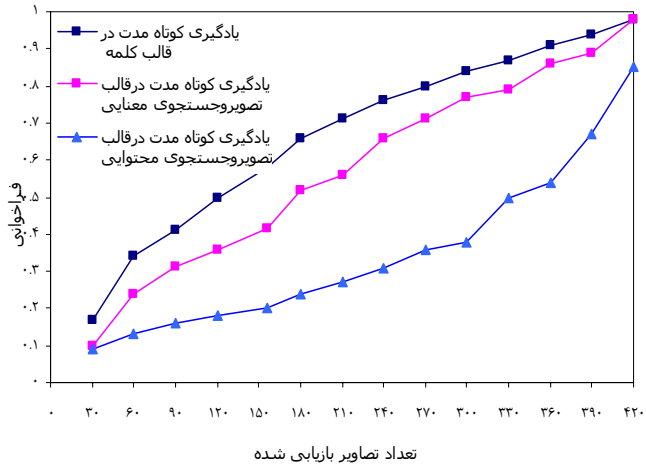
Archive of SID

شکل های(۷) و (۸) به ترتیب نمودار دقت و فراخوانی را در قبال تعداد تصاویر بازیابی شده برای دو حالت پرس و جو توسط کلمه کلیدی و تصویر نمونه نشان می‌دهند. همانطور که مشخص است، عملکرد سیستم در رابطه با پرس و جو توسط کلمه کلیدی نسبت به پرس و جو توسط تصویر نمونه از کارایی و دقت بالاتری برخوردار می باشد.

مطابق این نمودار ها میزان دقت و فراخوانی سیستم پیشنهادی در قبال تعداد تصاویر بازیابی شده طی یک نشست بازیابی تعاملی، برای پرس و جو بر اساس کلمه کلیدی در حد قابل قبولی می‌باشد. همچنین برای پرس و جو



شکل۷. نمودار دقت سیستم پیشنهادی



شکل۸. نمودار فراخوانی سیستم پیشنهادی

بر اساس تصویر نمونه که در ابتدا، کاربر تصویری را بعنوان درخواست پرس و جو ارائه می نماید و بدین ترتیب دقت نتایج بازیابی معنایی را به دلیل ماهیت ویژگی های بصری سطح پایین که حاوی هیچگونه معنایی در مورد تصویر نمی باشند کاهش می دهد، نمودار های دقت و فراخوانی سیستم پیشنهادی در حد پایین تری نسبت به نتایج حاصل شده برای حالت پرس و جو بر اساس کلمه کلیدی قرار می گیرند.

لازم به توضیح است که مطابق آزمون های انجام شده، عملکرد سیستم در ارایه راهکاری نیمه خودکار برای بازیابی معنایی تصاویر با استفاده از شبکه معنایی شامل مفاهیم سطح بالا هم در حوزه معنا و هم درحوزه انتقال تصویر به معنا از کارایی قابل قبولی برخوردار است.

۵. نتیجه گیری

از آنجاییکه درک نظرات کاربر در مورد نتایج سیستم های بازیابی از اهمیت بالایی برخوردار است، در این مقاله روشی ارایه می گردد که بصورت تعاملی بازخوردهای ارتباطی کاربر در هر مرحله تکرار بازیابی را پردازش نموده و پرس و جوی دیگری را بر اساس نظرات ارایه شده کاربر در هر مرحله ترتیب می دهد. عملکرد پردازشی این سیستم، تحت عنوان واحد یادگیرنده ارائه شده است. در واقع واحد یادگیرنده، با انجام نوعی عملیات یادگیری کوتاه مدت طی هر مرحله تعاملی، معانی مورد درخواست کاربر را از بازخوردهای او استخراج نموده و بازیابی معنایی دیگری را راه اندازی می‌کند.

سیستم پیشنهادی از یک شبکه معنای سلسله مراتبی در مجموعه های کلمات کلیدی مترادف شامل مفاهیم سطح بالایی مانند حالتهای چهره و ابعاد شخصیتی انسان استفاده می‌نماید و در طی هر مرحله یادگیری ، شباهت معنایی تصاویر را برای بازیابی های هرچه مفهومی تر در دو حالت And و Or برای پرس و جو براساس کلمه کلیدی، با توجه به پرس و جو و تصاویر ارایه شده بصورت "مرتبط" و "غیرمرتبط" محاسبه می‌نماید. علاوه بر این، سیستم پیشنهادی در حالت پرس و جو براساس تصویر نمونه، پس از بازیابی نتایج اولیه و برای مراحل بعدی بازیابی تعاملی، جستجوی محتوایی تصاویر را با جستجوی معنایی براساس شباهت معنایی با تصاویر بازخورد شده جایگزین می نماید. در واقع سیستم ارائه شده، در روشی سطح بالا و مفهومی، در هر دو حالت پرس و جو براساس کلمه کلیدی و تصویر نمونه، پاسخگوی نیازهای

www.SID.ir

Archive of SID

معنایی کاربران در حالتی است که یادگیری را بدون در نظر گرفتن ویژگی های بصری سطح پایین و فقط بر اساس محتوای معنایی تصاویر انجام می‌دهد و در همین راستا قابلیت انتقال جستجو را از قالب محتوایی به معنایی داراست. نتایج پیاده سازی، دقت قابل قبولی را برای این سیستم نشان می دهند به گونه‌ایکه تحت عملکرد واحد یادگیرنده، نتایج بدست آمده در حالت پرس و جو بر اساس کلمه کلیدی و جستجوی معنایی، بسیار معناگراتر از حالت پرس و جو براساس تصویر نمونه و عملیات جستجوی محتوایی می باشد و نتایج بهتری را حاصل می کند.

مراجع

[1] T. Gevers, A. W. M. Smeulders, Content-based image retrieval: An overview, G. Medioni, S. B. Kang (Eds.), Emerging Topics in Computer Vision. Prentice Hall, 2004.

[2] T.S.H. Xiang, S. Zhou, Unifying keywords and visual contents in image retrieval, IEEE Multimedia (2002) 23–33.

[3] A. Smeulders . Content-based image retrieval: the end of the early years, IEEE Transaction on Pattern Analysis and Machine Intelligence(PAMI) 22 (12) (2000) 1349–1380.

[4] H.Müller, W.Müller, et al. Performance Evaluation in Content-based Image Retrieval: Overview and Proposals, Pattern Recognition Letters 22 (5) (2001) 593–601.

[5] A. Vailaya, M.A. G.Figueiredo, A.K.Jain, H.J.Zhang, Image Classification for Content-based Indexing, IEEE Transaction on Image Processing 10 (1) (2001).

[6] L. Wenyin, S. Dumais, Semi-Automatic Image Annotation, In Proc. of Human-Computer Interaction (2001) 326–333.

[7] E. Hyvonen, M. Salminen, M. Junnila, Annotation of Heterogeneous Data Base Content for the Semantic Web, University of Helsinki. Department of Computer Science. In Proc. of the 4th International Workshop on Knowledge Markup & Semantic Annotation. Hiroshima, Japan. 2004.

[8] B. Shevade, H. Sundaram, Vidya: An Experiental Annotation System, Arts Media & Engineering Program, In ACM SIGMM Work. Experiental Telepresence,Berkeley, California, USA, 2003, pp. 91-98.

[9] Amin Shah-Hosseini, Semantic Image Retrieval Using Relevance Feedback And Transaction Logs, In proc. Of Graduate Faculty of the Louisiana State University and Agricultural and Mechanical Collage. The Interdepartmental Program in Engineering Science, Aug, 2007.

[10] J. Laaksonen, M. Koskela, S. Laakso, E. Oja, Self organizing maps as a relevance feedback technique in content-based image retrieval. In Proc. of Pattern Analysis & Applications, 2001.

[11] D. Brahmi, D. Ziou, Improving CBIR systems by integrating semantic features, In proc. of RIAO, Vaucluse, France, 2004, 291-305.

[12] Zhao Xu, Xiaowei Xu, Kai Yu, Volker Tresp. A Hybrid Relevance Feedback Approach to Text Retrieval. In Proc. of the 25th European Conference on Information Retrieval Research. Lecture Notes in Computer Science, 2633 (2003).

[13] P. Panagi, et al. A Learning Approach to Semantic Image Analysis. In Proc. of 2nd International Mobile Communications Conference (MobiMedia06). Alghero, Italy, (2006).

[14] X.Zhu, W. Liu, H. Zhang, Image retrieval and semi-automatic annotation scheme for large image databases on the Web. Microsoft Research. Fudan Univ, China (2000) 4311-21.

[15] Lu Y, et al. A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems. In Proc. of the 8th ACM International Conference on Multimedia (2000) 31-38.

[16] J. Yang, et al. Thesaurus-Aided Approach for Image Browsing and Retrieval. In proc. of IEEE International Conference on Multimedia and Exposition (ICME), Tokyo, Japan (2001) 313-316.

[17] G. Carneiro, N. Vasconcelos. A Database Centric View of Semantic Image Annotation and Retrieval. In Proc. ACM SIGIR Conf. Research and Development in Information Retrieval. Salvador, Brazil (2005).

[18] C. Hentschel, et al. SAFIRE: Towards Standardized Semantic Rich Image Annotation. In Proc. Of 4th International Workshop on Adaptive Multimedia Retrieval (AMR2006), Springer LNCS, Geneva, Switzerland (4398) (2006) 12-27.

[19] Euripides G. M. Petrakis, et al. Relevance feedback methods for logo and trademark image retrieval on the web. In Proc. of the 2006 ACM symposium on Applied computing.

www.SID.ir

Archive of SID

Dijon, France, 2006.

[20] G. Carneiro, N. Vasconcelos, Formulating Semantic Image Annotation as a Supervised Learning Problem, Appears in IEEE Conf. on Computer Vision and Pattern Recognition. San Diego , 2005.

[21] G. Carneiro, Antoni B. Chan, Pedro J. Moreno, Nuno Vasconcelos, Supervised Learning of Semantic Classes for Image Annotation and Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29 (3) (2007) 394-410.

[22] A. L. Ion, et al. Semantic Based Image Retrieval using Relevance Feedback, In proc. Of International Conference on Computer as a tool (EUROCON07), Warsaw, Poland (2007) 303-310.

www.SID.ir