

## چالش‌های بازیابی معنایی تصاویر و کاربرست نوین اصطلاحنامه‌ها



- فرحناز صدوقی<sup>1</sup> | دانشیار گروه مدیریت اطلاعات بهداشتی درمانی، دانشگاه علوم پزشکی تهران
- علی ولی‌نژادی\*<sup>2</sup> | دانشجوی دکتری مدیریت اطلاعات سلامت، دانشگاه علوم پزشکی تهران
- حافظ محمدحسن‌زاده<sup>2</sup> | کارشناس ارشد کتابداری و اطلاع‌رسانی پزشکی دانشگاه علوم پزشکی تبریز
- حمید بورقی<sup>3</sup> | مربی گروه کتابداری و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی همدان
- پرینا پاسیار<sup>4</sup> | کارشناس ارشد کتابداری و اطلاع‌رسانی پزشکی، کتابخانه ملی ایران

دریافت: 1388/10/02 | پذیرش: 1389/03/18

فصلنامه علمی پژوهشی  
پژوهشگاه علوم و فناوری اطلاعات ایران  
شاپا(چاپی) 2251-8223  
شاپا(الکترونیکی) 2251-8231  
نمایه در SCOPUS و LISA و ISC  
<http://ijpm.irandoc.ac.ir>  
دوره 27 | شماره 3 | صص 641-666  
بهار 1391

نوع مقاله: مروری

**چکیده:** امروزه، گسترش فناوری‌های رقومی، دسترسی گسترده‌تری را به منابع تصویری موجب شده است. با این حال، گسترش پرشتاب فناوری‌های تصویری، رشد شگفت‌آور حجم تصاویر رقومی و رشد روزافزون علائق کاربران برای اطلاعات دیداری، با پیشرفت‌های قابل‌قیاس در حوزه بازیابی تصاویر، نظام‌های نمایه‌سازی و تعامل با کاربران، پشتیبانی نشده است. در این مطالعه، ابتدا ارزش ذاتی تصاویر و تفاوت‌های عمده واژه و توصیف واژگانی با تصویر و توصیف تصویری مورد بحث قرار گرفته است، سپس ناکارآمدی فنون بازیابی مبتنی بر واژگانی تصاویر به‌علت وابستگی عمیق آنها به ویژگی‌های نحوی (نحوه ترکیب عناصر دیداری) و نیز کاربردهای اصطلاحنامه‌های مبتنی بر واژگان در بازیابی تصاویر و خلأ عمده آنها در محیط‌های اطلاعاتی جدید با عنوان "شکاف معنایی" بررسی شده است. در ادامه، محتوای نحوی و محتوای معنایی تصاویر تبیین شده و توجه به خصوصیات ذاتی تصاویر و محتوای آنها در به‌کارگیری اصطلاحنامه‌های مبتنی بر خصوصیات ذاتی تصاویر مورد تأکید قرار گرفته است. در نهایت، کاربرست رویکرد اصطلاحنامه ریختاری، یعنی آزمودن چارچوبی اصطلاحنامه‌ای برای شناسایی روابط بین اشیای<sup>1</sup> درون تصاویر جهت ارتقاء فرایند بازیابی تصاویر به‌منظور رفع شکاف معنایی در فرایند بازیابی معنایی تصاویر، پیشنهاد شده و به تفصیل ارائه شده است.

**کلیدواژه‌ها:** بازیابی تصاویر، نمایه‌سازی تصاویر، بازیابی معنایی تصاویر، شکاف معنایی، اصطلاحنامه ریختاری

۱. [sadoughi.f@gmail.com](mailto:sadoughi.f@gmail.com)

\* [a\\_valinejad@razi.tums.ac.ir](mailto:a_valinejad@razi.tums.ac.ir)

۲. [hafezhassanzadeh70@gmail.com](mailto:hafezhassanzadeh70@gmail.com)

۳. [hamid.bouraghi@gmail.com](mailto:hamid.bouraghi@gmail.com)

۴. [p.pasayar@gmail.com](mailto:p.pasayar@gmail.com)

## 1. مقدمه

انسان امروزی در همه جا و هر حالتی در محاصره تصاویر قرار دارد و آن را در همه جنبه‌های زندگی خود به کار می‌گیرد (Smith ۲۰۰۱). تصاویر ابزار قدرتمندی برای انتقال اطلاعات محسوب می‌شوند و بیشتر در ترکیب با متون می‌توانند اطلاعاتی را که با کلمات قابل توصیف نیستند، منتقل کنند (Hove ۲۰۰۴). یک تصویر می‌تواند هزاران بار با ارزش‌تر از کلمات باشد، اما گاه هزار توصیفگر نیز برای توصیف محتوای یک پایگاه اطلاعاتی تصویری کافی نیست. ماهیت چندبعدی بودن تصاویر باعث می‌شود که هر کاربر با توجه به نیاز خود، نگاه ویژه‌ای به یک تصویر داشته باشد و حتی با هدفی جدا از هدف خالق تصویر، آن را به کار گیرد. برای مثال، تصویری از یک خیابان شلوغ در قرن 19 ممکن است مورد استفاده چندین نفر قرار گیرد و هر کدام از کاربران از یک بعد به آن توجه کنند، یعنی معمار از نظر ساختمان‌سازی و تاریخدان از دیدگاه تغییراتی که در این دوره انجام شده است به آن توجه کند. این مسأله به خصوص زمانی که تصاویر رمزی و نمادین هستند، بسیار جدی‌تر می‌نماید (کرمی 1385). درحقیقت، مفاهیم موجود در هر تصویر با توجه به جنبه‌های فرهنگی، آموزشی، پژوهشی، علمی، و ... متفاوت است. وجود یک تصویر در ترکیب با متن می‌تواند به درک و انتقال مفاهیم موجود در آن کمک کند، حال آنکه همان تصویر در متنی دیگر مفهوم دیگری را می‌تواند تداعی کند (Hassan and Zhang ۲۰۰۱).

امروزه، گسترش فناوری‌های رقومی امکان دسترسی سریع به منابع دیداری را فراهم کرده است. در دهه‌های اخیر، مؤسسات فرهنگی عهده‌دار طرح‌های رقومی‌سازی کلان‌سی شده‌اند تا تصاویر موجود در مجموعه‌های خود را به صورت رقومی ارائه کنند. رشد چشمگیر و روزافزون تصاویر رقومی نیاز به روش‌های مؤثر برای ذخیره و جستجو و بازیابی چنین اطلاعاتی را موجب شده است. کاربران برای به دست آوردن تصاویر مورد نیاز باید از ابزارهای جستجوی کارآمد و قابل اعتمادی استفاده کنند تا بتوانند در میان حجم عظیم مجموعه‌های تصاویر آنچه را که می‌خواهند، بازیابی کنند (Smith ۲۰۰۱)، چرا که به دلیل حجم بالای اطلاعات، این امر بیشتر شبیه به جستجوی یک سوزن در انبار کاه است (Hassan and Zhang ۲۰۰۱). محیط وب امکان دسترسی به تصاویر رقومی و دستکاری و جستجوهای چندگانه آنها را فراهم می‌کند (Smith ۲۰۰۱).

در سویی دیگر، بدون داشتن یک نمایه مناسب برای همه تصاویر در اینترنت، یافتن یک تصویر خاص فرایندی بسیار دشوار و زمان‌بر است، زیرا اطلاعات استخراج شده از تصاویر به طور معمول، پیچیده و چندبعدی هستند و بیشتر عدم ارتباط بین آنچه از یک تصویر مجسم شده و آنچه که موضوع واقعی یک تصویر بوده است، مشکل آفرین می‌شود (Hassan and Zhang ۲۰۰۱). با اذعان به ارزش ذاتی تصاویر، باید در نظر داشت که در واقع زیرساخت‌های آرشیو و سازماندهی

یعنی فهرست‌نویسی و به‌خصوص نمایه‌سازی تصاویر است که طریقه بهره‌گیری و نتایج حاصل از آنها را مشخص می‌کند. واژگان نمایه‌سازی تصاویر نه تنها باید عنوان و خالق و ناشر آنها را پوشش دهد، لازم است که مفاهیم مورد نظر خالق از خلق آن اثر را نیز مورد نظر قرار دهد. به‌عبارت دیگر، وقتی در یک پایگاه اطلاعاتی یک مقاله کوتاه، با حدود 20 واژه نمایه‌سازی می‌شود، تصاویر آن را نیز باید دست‌کم با همان عمق مورد تجزیه و تحلیل قرار داد (Roberts ۲۰۰۱).

کار نمایه‌سازی و سازماندهی تصاویر زمانی مطلوب و مناسب شمرده می‌شود که رضایت کاربران را در بازیابی تصاویر به‌دنبال داشته باشد. به این ترتیب، شناخت رفتار اطلاع‌یابی کاربران، دانستن انتظارات آنها، انواع پرسش‌های آنها از پایگاه‌های اطلاعاتی، و نحوه قضاوت‌شان درباره موارد بازیابی‌شده، ضروری به‌نظر می‌رسد. یک نظام مطلوب بازیابی تصاویر باید از تمامی انواع پرسش‌های کاربران در تمامی سطوح آن پشتیبانی کند و تا حد امکان توانایی پاسخگویی به نیازهای بالقوه کاربران را داشته باشد. برای تحقق این مطلوب، چنین نظامی نیازمند روش‌های بهینه توصیف، رده‌بندی تمامی سطوح و ویژگی‌ها، و فنون مورد نیاز است (Chu ۲۰۰۱).

پل ارتباطی میان فرایند ذخیره‌سازی اطلاعات تصویری و فرایند بازیابی آنها، نمایه‌سازی و ابزار آن، اصطلاحنامه است. وجود اصطلاحنامه‌ها در فرایند نمایه‌سازی باعث می‌شود که با توجه به انواع روابط معنایی با مفهوم مورد نظر، ذخیره‌سازی اطلاعات تصویری تا حد امکان به بهترین وجه صورت گیرد. از طرف دیگر، حضور مستقیم و یا غیرمستقیم اصطلاحنامه‌ها در فرایند بازیابی، موجب برقراری ارتباط میان واژگان آزاد مورد استفاده توسط کاربران و واژگان کنترل‌شده نظام می‌گردد. رویکردهای متفاوت فرایند ذخیره و بازیابی تصاویر باعث به‌وجود آمدن اصطلاحنامه‌های متفاوتی شده است. این مقاله سعی دارد با تبیین رویکردهای مختلف ذخیره و بازیابی تصاویر، شیوه‌های متفاوت کاربست اصطلاحنامه‌ها را در بازیابی تصاویر مورد بررسی قرار دهد و درنهایت، طرح پیشنهادی اصطلاحنامه ریختاری<sup>۱</sup> را ارائه نماید.

## 2. نمایه‌سازی و ذخیره‌سازی تصاویر

در گذشته، دسترسی به مجموعه‌های تصویری آرشیوها، موزه‌ها، و کتابخانه‌ها تا حدود زیادی به نحوه ترتیب و چیدمان آنها و نیز نحوه نمایه‌سازی بستگی داشت و کاربر به‌منظور دسترسی به منابع، لازم بود که با ساختار این نهادها آشنا باشد تا استفاده بهینه‌ای داشته باشد (Roberts ۲۰۰۱). در حال حاضر نیز بخش زیادی از آرشیوها هنوز به‌صورت کاغذی هستند و برای بازیابی هر موضوع زمان زیادی نیاز است. از سوی دیگر، فرایند نمایه‌سازی نیز بسیار پرهزینه است. پس از

۱. Shape thesaurus

آنکه تصاویر از یک بسته بزرگ انتخاب شدند، بر روی آنها برچسب<sup>۱</sup>های کلیدواژه‌ای چسبانده می‌شود و برحسب موضوعات موجود و موضوعات جدید طبقه‌بندی و سپس، سازماندهی می‌شوند. به این ترتیب، نیاز به رویکردهای نوینی در نمایه‌سازی تصاویر است تا فرایندهای سنتی که فقط به تورق تصاویر منجر می‌شد را ارتقا بخشند (Encinas, Liorens, and Adoracion ۱۹۹۹).

این اواخر پیشرفت‌هایی در زمینه انواع رده‌بندی‌ها، تحلیل محتوی، و استخراج ویژگی‌ها به صورت خودکار به دست آمده که موجب ارتقاء قابلیت‌های جستجو، پالایش و دسترسی هرچه مؤثرتر به ابعاد مختلف تصاویر از جمله ساختار و ویژگی‌های معنی‌شناختی و ادراکی شده است (Smith ۲۰۰۱). نظام‌های جدید بر روی تصاویر رقومی و توالی تصاویر یک پایگاه داده متمرکزند. هر چند که بخشی از آنها درحقیقت گسترشی از رویکردهای سنتی اطلاعات است که برای نیازهای امروز طراحی شده‌اند، به نوآوری‌ها نیز باید توجه داشت (Hassan and Zhang ۲۰۰۱). رویکردهای نمایه‌سازی را می‌توان در دو بخش رده‌بندی مردمی و رده‌بندی سامانه‌ای مورد بررسی قرار داد.

## 1-2. رده‌بندی مردمی

در این روش، کاربران کلیدواژه‌هایی که به نظرشان برای هر تصویر مناسب است را در قالب برچسب در درون نظام قرار می‌دهند. آنها همچنین می‌توانند از نحوه به کارگیری برچسب‌ها از سوی کاربران دیگر برای مفاهیم گوناگون تصاویر دیگر آگاه شوند. اشتراک و بازخوردهای سریع باعث می‌شود که رده‌بندی مردمی از سایر نمایه‌سازی‌ها که به طور معمول، فقط توسط یک گروه خاص و بدون دخالت کاربران راهبری می‌شود، متمایز باشد.

رده‌بندی مردمی توانایی توصیف و بازیابی تصاویر را در اختیار قرار می‌دهد و نظام درون‌پیوندی<sup>۲</sup> برچسب‌ها، عملکردهای تورق و یافتن یک تصویر در محیط رقومی را پشتیبانی می‌کند. مهمترین ویژگی این نوع رده‌بندی، ارتباط نزدیک آن با کاربران و زبان آنهاست. رده‌بندی مردمی به طور مستقیم از انتخاب عبارات، اصطلاحات، و صراحت کاربران تأثیر می‌پذیرد. سایت‌های اشتراک تصاویر مثل "فلیکر"<sup>۳</sup> به کاربران اجازه می‌دهند تا تصاویر را بارگذاری کنند و آنها را با استفاده از کلمات خود مقوله‌بندی نمایند. نمایه ساخته شده توسط

۱. Tag

۲. Interlinked system

۳. Flickr

(از وبسایت‌های نظیر آن می‌توان del.icio.us, librarything, Atpic, Fotki, ipernity, Panoramio, Picasa, Zoomr, Snepi, Smug Mug به اشاره کرد.)

کاربران که بیشتر با نام "فولکسونومی"<sup>۱</sup> و یا "رده‌بندی مردمی" شناخته می‌شود، مورد توجه بسیاری قرار گرفته است و علاقه‌مندان از آن به‌عنوان انقلابی در طبقه‌بندی علم و هنر تعبیر نموده‌اند (سبزی‌پور ۱۳۸۷).

اما، برخی از پژوهشگران به ماهیت آشفته و نامرتب برجسب‌گذاری‌های مردمی به‌خصوص در مقایسه با نظام‌های رسمی رده‌بندی اشاره کرده‌اند. حلقه‌های بازخورد از طریق ابردادها شکلی از ارتباطات نامتقارن را به‌وجود می‌آورند. تحلیل برجسب‌ها ویژگی‌های سازماندهی محتوی را در این روش نشان می‌دهد. برخی از این ویژگی‌ها مثل عدم کنترل مترادف‌ها، وجود غلط‌های املائی، استفاده از اشکال جمع و مفرد، کدگذاری‌های نادرست و برجسب‌های خصوصی، نشان‌دهنده محدودیت‌های رده‌بندی مردمی برای اهداف بازیابی است. با این حال، جنبش رده‌بندی مردمی بحثی را در جوامع کتابخانه‌های رقومی در باب استفاده از کاربردهای شبکه‌های مردمی، درگیر کردن کاربران، و ساختن جوامع مجازی آغاز کرده است (Matusiak ۲۰۰۶).

## 2-2. رده‌بندی سامانه‌ای

هر چند رده‌بندی مردمی امکان دستیابی به دیدگاه‌های مختلف درباره مفهوم یک تصویر را فراهم می‌آورد، محدودیت‌های موجود در این نوع رده‌بندی نشان می‌دهد که نمایه‌سازی تصاویر همچنان نیاز به یک رده‌بندی سامانه‌ای و کنترل‌شده دارد. برای یافتن تصاویر در محیط‌های رقومی باید تصاویر رقومی را دقیق‌تر توصیف کرد. فهرست‌نویسان حرفه‌ای در موزه‌ها و کتابخانه‌ها، ابردادهای توصیفی را با توجه به استانداردها به‌وجود می‌آورند و از ابزارهای زبان کنترل‌شده استفاده می‌کنند. این روش، جایگزین روش سنتی نمایه‌سازی مدرک‌گرا<sup>۲</sup> شده است.

- جستجو در متون مرتبط، دو روش جدید را برای نمایه‌سازی سامانه‌ای تصاویر نشان می‌دهد:
- (۱) روش "مبتنی بر توصیف"<sup>۳</sup> که در آن صفات تصاویر و محتوای معنایی آنها مشخص می‌شود و به‌صورت تحت‌اللفظی توسط نمایه‌سازان انسانی توصیف می‌شوند.
  - (۲) روش "مبتنی بر محتوی"<sup>۴</sup> که در آن ویژگی‌های تصاویر مثل رنگ و شکل و بافت به‌صورت خودکار توسط نرم‌افزارهای رایانه‌ای مشخص و استخراج می‌شوند.

## 2-2-1. نمایه‌سازی مبتنی بر توصیف

نمایه‌سازی مبتنی بر توصیف، راهیابی ذهنی به محتوای دیداری یک تصویر را فراهم می‌آورد و شامل ترجمه اطلاعات دیداری به توصیفات متنی است تا مشخص کند که تصویر

۱. Folksonomy  
۲. Document-oriented indexing  
۳. Concept based approach  
۴. Content based approach

درباره چیست و چه چیزی را ارائه می‌کند. علاوه بر توصیف موضوعی، ابرداده‌هایی که در ارتباط با تصویر هستند، می‌توانند اطلاعاتی درباره اعتبار و منشأ و تولیدکننده تصویر را دربرگیرند. ابرداده‌های توصیفی بر پایه یک طرح ابرداده‌ای استاندارد مثل "دوبلین کور" ساخته شده‌اند که از ابزارهای زبان کنترل‌شده و یا زبان طبیعی برای محتوای ابرداده‌ها استفاده می‌کنند. نمایه‌سازی مبتنی بر توصیف، نیازمند عوامل انسانی است تا معنای تصاویر را تفسیر کند، سرعنوان‌های موضوعی را مشخص کند و از عنوان تصاویر و یادداشت‌ها یا گزارمان‌های متنی آنها رونوشت بردارد (Matusiak ۲۰۰۶). فرایند ترجمه محتوای یک تصویر به صورت تحت‌اللفظی با چالش‌های مهمی در مقابل نمایه‌سازی مبتنی بر محتوی روبرو است. برخی پژوهشگران به این توافق رسیده‌اند که حتی نمایه‌سازی گسترده مبتنی بر متن نیز برای مواجهه با نیازهای کاربران و فراهم‌آوری یک بازیابی مؤثر، بیشتر اوقات ناکافی است.

## 2-2-2. نمایه‌سازی مبتنی بر محتوی

در بیانی ساده، رویکرد "مبتنی بر محتوی" به فنون نمایه‌سازی و بازیابی تصاویر بر پایه پردازش خودکار اطلاعات متنی اطلاق می‌شود. ویژگی‌های تحلیلی تصویر در رویکرد مبتنی بر محتوی به سه سطح تقسیم می‌شوند:

- 1) ویژگی‌های ابتدایی، مانند: رنگ و ریخت و بافت،
  - 2) ویژگی‌های استدلالی، مانند: اصالت اشیای نمایش داده‌شده، و
1. ویژگی انتزاعی<sup>۱</sup> مانند مفاهیم معنی‌دار صحنه‌های به تصویر کشیده شده (Encinas, Liorens, and Adoracion ۱۹۹۹).

نظام‌های مبتنی بر محتوی این امکان را برای کاربران فراهم می‌آورند که از توصیفگرهایی بهره‌گیرند که از خود تصاویر گرفته شده‌اند و در ارتباط با همان تصاویر هستند.

برای شروع جستجو، کاربر به‌طور معمول، در اندیشه و ذهن خود یک تصویر ایده‌آل و یا مطلوب ذهنی دارد. نحوه بیان تصویر ذهنی و ایده‌آل یکی از چالش‌های طراحی پرسش کاربر است. نظام‌های مبتنی بر محتوی به کاربران اجازه می‌دهند که پرسش‌های خود را از طریق یک یا چند روش زیر طراحی نمایند:

- 1) انتخاب یا ارائه یک تصویر نمونه یا تمثیلی،
- 2) پیش‌طراحی ساده و گرافیکی یک تصویر ذهنی،
- 3) بیان پرسش در یک زبان ساختاریافته پرسش مانند "اس کیوال"<sup>۲</sup>، و

۱. Abstract

۲. SQL

(4) قراردادن پرسش به‌همراه مثال‌هایی در آن زمینه درون یک فایل (Encinas, Liorens, and Adoracion ۱۹۹۹).

باید در نظر داشت که در هر مورد، پرسش مطرح‌شده فقط تا حدودی با تصویر ذهنی و ایده‌آل کاربر تطابق دارد. پس از ارائه پرسش از جانب کاربر، سامانه مبتنی بر محتوی به تصاویر دسترسی پیدا می‌کند و عناصر پرسش را با اطلاعات توصیفی پایگاه‌های اطلاعاتی تطابق می‌دهد و در نهایت، تصاویر حاصل از جستجو را براساس میزان شباهت به تصویر ذهنی کاربر، رتبه‌بندی می‌کند. فرایند بازیابی مبتنی بر محتوای تصاویر بر نحوه شکل‌گیری تصویر ذهنی کاربر از تصویر مورد درخواست، نحوه شکل‌دهی کاربر به پرسش خود و نیز معیارهای ارزیابی ربط تصاویر حاصل از جستجو به تصویر ذهنی کاربر تأکید دارد. این مسائل از نحوه توصیف اطلاعات تصویری تأثیر پذیرفته و با ابعاد متنوعی نظیر معنی‌شناسی و ساختار و ویژگی‌ها مورد تطابق قرار می‌گیرند.

۱. **معنی‌شناسی:** از جمله نمونه‌های معنایی می‌توان اشیا و رویدادها و اعمال را برشمرد. در معنی‌شناسی، بازنمون مناسب اطلاعات معنایی برای توصیف تصاویر و نیز پرسش کاربر اهمیت بسیاری دارد. باید در نظر داشت که معنی‌شناسی به‌میزان زیادی وابسته به دامنه موضوعی مربوط به آن است.

۲. **ساختار:** ترکیب فضایی محتوای یک تصویر از عناصر سازنده آن را ساختار تصویر می‌نامند. از جمله نمونه‌هایی از مفاهیم ساختاری می‌توان به مناطق، مقاطع، و ... اشاره کرد. مناطق مختلف یک تصویر از طریق روابط ساختاری از جمله تراکم فضایی، ترکیب، و پیوستگی با هم در ارتباط هستند.

۳. **ویژگی:** عبارت است از جنبه‌های سطح پایین تصاویر که برای شناسایی، تشخیص، تمیز، و رده‌بندی آنها مورد استفاده قرار می‌گیرد. رنگ و بافت و ریخت، ویژگی‌های مهم تصاویر محسوب می‌شوند. رنگ یک بُعد مهم در درک انسان از تصویر است که فرایند تشخیص و تمیز را برای وی مقدور می‌سازد. بافت تصویر نیز یک عنصر مهم در درک انسانی و مربوط به الگوهای تشابه و هم‌جنسی یا ترکیب فضایی پیکسل<sup>۱</sup>هایی است که عامل رنگ به اندازه کافی از عهده توصیف آن برنیامده است. بافت تصاویر ممکن است خصائص آماری، خصائص ریختاری یا هر دوی آنها را دارا باشد. بافت درحقیقت، راهنمایی برای درک عمیق صحنه و سمت و سوی سطح فراهم می‌آورد. ریخت تصویر ممکن است از طریق بازنمون خطوط مرزی بین اشیای قابل توصیف باشد (Smith ۲۰۰۱).

هر چند شیوه‌های دسترسی و بازیابی از طریق رنگ و ریخت و بافت تصاویر، جالب توجه است و دانشمندان علوم رایانه توجه ویژه‌ای به ادغام این فنون با برخی فنون نمایه‌سازی سنتی داشته‌اند که نمونه‌ای از آنها استفاده از هستی‌شناسی‌های گسترده است (Jørgensen et al. ۲۰۰۱)، مطالعات نشان می‌دهد که در حوزه علوم رایانه، افراد بر رویکرد مبتنی بر محتوی تمرکز دارند در

۱. Pixel

حالی که جامعه علوم اطلاع‌رسانی از جمله علوم کتابداری تمرکز خود را بر روش مبتنی بر توصیف قرار داده‌اند. به هر حال، بدیهی است که هر کدام از این دو رویکرد جایگاه ویژه خود را دارد و در فرایند نمایه‌سازی و بازیابی تصاویر، نقاط ضعف و قوت خاص خود را داراست. پژوهش در نمایه‌سازی و بازیابی تصاویر به‌خصوص در حوزه مبتنی بر محتوی در حال رشد و توسعه است. اولین نظام‌های تجاری بازیابی تصاویر بر پایه نمایه‌سازی مبتنی بر محتوی در سال 1995 به کار گرفته شدند (دستغیب 1384). به نظر می‌رسد رویکرد مبتنی بر محتوی در حال حاضر نسبت به رویکرد مبتنی بر توصیف، استفاده گسترده‌تری دارد. با این حال، رویکرد مبتنی بر توصیف همواره در نظام‌های عملیاتی گزینه انتخابی اول است (Chu 2001).

### 3. بازیابی تصاویر

حوزه‌های مختلف پژوهشی دیدگاه‌های گوناگونی نسبت به تصویر و بازیابی آن دارند و به تبع آن رویکردها و مبانی نظری متفاوتی برای این موضوع قائل‌اند. در دهه‌های اخیر، فنون بازیابی تصاویر به‌طور معمول بر پایه فنون بازیابی اطلاعات متنی قرار داشتند، اما به‌تازگی از فونونی چون هوش مصنوعی و پردازش علائم نیز در بازیابی تصاویر استفاده شده است. فنون بازیابی تصاویر را می‌توان به دو بخش تقسیم کرد:

- 1) بازیابی مبتنی بر گزارمان‌های متنی تصویر، و
- 2) بازیابی مبتنی بر ویژگی نحوی (نحوه ترکیب عناصر دیداری) و الگوی داده‌ای تصویر (Hove 2004).

#### 3-1. فنون بازیابی مبتنی بر گزارمان‌های متنی

این رویکرد به بازیابی مبتنی بر توصیف واژگانی محتوای مجموعه‌های تصویری اطلاق می‌شود. این روش بر پایه تشابه میان مجموعه‌ای از جستجوهای متنی و کلیدواژه‌ها یا یادداشت‌های متن آزاد تصاویر درون پایگاه‌های اطلاعاتی قرار دارد. پرسش‌های متنی امکان استفاده از تمامی فنون سنتی بازیابی اطلاعات از جمله فنون مبتنی بر متن را برای بازیابی تصاویر میسر می‌سازند. با استفاده از پرسش‌های متنی می‌توان به‌راحتی نیاز اطلاعاتی کاربر را ترجمه کرد و به تصاویر دارای یادداشت گزارمانی دست یافت و بیشتر تصاویر مرتبط را بازیابی نمود.

اما، برخی مواقع پرسش‌هایی مطرح می‌شوند که مبتنی بر ویژگی‌های نحوی تصویر است. این گونه پرسش‌ها مشکل آفرین‌ترین نوع پرسش‌ها در استفاده انحصاری از نظام‌های بازیابی مبتنی بر متن هستند. باید توجه داشت که تصاویر، همواره دارای ماهیت دیداری هستند و بسیاری از نیازهای اطلاعاتی به‌هیچ وجه از طریق واژه‌ها قابل بیان نیستند. مشکل اصلی در



استفاده از فنون بازیابی مبتنی بر متن، ریشه در چالش‌های "حجم"<sup>1</sup> و "موضوع"<sup>2</sup> دارد (در صفحات بعد توضیح داده خواهند شد)، چرا که تصاویر درون مجموعه‌ها باید به‌طور کامل و صحیح دارای یادداشت‌های گزارمانی باشند. حتی اگر پرسش کاربر به‌خوبی فرمول‌بندی شده باشد، یادداشت‌های گزارمانی نادرست و ناکافی تصویر نتایج جستجوی ضعیفی در پی خواهد داشت. به‌علاوه، از آنجایی که کاربر به‌ضرورت نمی‌تواند واژگان مورد استفاده گزارمان‌نویس تصویر را به‌کار بندد، ممکن است نتیجه، عدم تطابق بین معیارهای پرسش کاربر و واژگان مورد استفاده در گزارمان‌نویسی تصویر باشد.

### 2-3. فنون بازیابی مبتنی بر ویژگی نحوی و الگوی داده‌ای

این روش بازیابی تصویر برپایه استفاده از توصیفگرهای ویژگی نحوی تصویر برای جستجوی تشابه‌های عناصر درون پرسش و تصاویر درون مجموعه‌ها قرار دارد. در کل، در این روش سه رویکرد مختلف برای اختصاصی کردن تصویر جهت بازیابی آن وجود دارد:

- (1) پرسش مبتنی بر ویژگی‌های نحوی ارائه‌شده در تصویر،
- (2) پرسش مبتنی بر مثال، و
- (3) پرسش مبتنی بر طرح اولیه.

پرسش‌های مبتنی بر ویژگی‌ها یا الگوی داده‌ای تصویر، برپایه تشابه بین پرسش و الگوهای موجود در مجموعه تصاویر قرار دارد. موارد بازیابی‌شده، براساس میزان تشابه میان توصیفگرها و الگوهای تصویر از طرفی و الگوهای پرسش از طرف دیگر حاصل می‌شود. رویکرد دوم برپایه روش "پرسش مبتنی بر مثال"<sup>3</sup> قرار دارد. استفاده از این روش در بازیابی تصویر جهت تخصصی کردن پرسش، یکی از روش‌های مورد استفاده گسترده در تخصصی کردن پرسش دیداری است. این رویکرد در مواقعی که کاربر ذهنیت کامل روشنی از تصویر مورد درخواست خود دارد یا اینکه در گذشته تصویری مرتبط با آن داشته است، می‌تواند به‌کار گرفته شود. رویکرد سوم، استفاده از طرح اولیه اساس جستجوی تصویر مشابه است. این رویکرد، "پرسش از طریق طرح اولیه"<sup>4</sup> نامیده می‌شود (Hove ۲۰۰۴).

پرسش‌های مبتنی بر سه رویکرد اشاره‌شده، برپایه میزان تشابه بین یک پرسش و تصاویر درون مجموعه قرار دارند. این میزان تشابه از طریق اندازه‌گیری عملکردهای تشابه‌سازی میان

۱. Volume  
 ۲. Subjectivity  
 ۳. Query Example Based (QEB)  
 ۴. Query Based Sketch (QBS)

توصیفگرهای ویژگی نحوی تصویر محاسبه می‌شود. حال آنکه تعداد زیاد و متنوعی از عملکردهای تشابه‌سازی جهت محاسبه انواع مشابهت‌ها وجود دارند که اساس بیشتر آنها بر طراحی بردارهای دوگانه‌ای برای ویژگی‌ها و تبدیل آنها به عددی خاص است که این عدد نمایانگر میزان تشابه میان دو تصویر است.

مشکل عمده فنون بازیابی مبتنی بر الگوی داده‌ای و ویژگی نحوی تصویر در این مسأله است که تصاویر، بازنمونی دوبعدی از موجودیت‌هایی سه‌بعدی هستند. حتی با در نظر گرفتن مسائلی چون مقیاس، جهت، چرخش تصویر، و... ضروری است که به منظور مقایسه هر چه مؤثرتر شباهت‌های مبتنی بر الگوی داده‌ای، اشیای متصور شده دارای درجه بالایی از "تغییرناپذیری"<sup>1</sup> باشند (Sclaroff and Liu ۲۰۰۱).

در نهایت، بایستی اشاره شود که آنچه کاربر از یک مجموعه تصویر می‌خواهد، به تصاویری که از لحاظ دیداری شباهت‌هایی با پرسش کاربر دارند محدود نمی‌شود، بلکه نتیجه مرتبط، تصویری است که به لحاظ معنی‌شناختی نیز برای آن پرسش خاص، معنی‌دار باشد. بنابراین، ضروری می‌نماید پژوهش‌های این حوزه به سمت ایجاد و گسترش ابزار و الگوریتم‌هایی برای شناسایی و تشخیص اشیای بنیادی و مفاهیم معنایی درون تصاویر حرکت کنند.

#### 4. چالش‌های نظام‌های جاری بازیابی تصاویر

مشکلات بازیابی در پایگاه‌هایی که داده‌های خود را با استفاده از عبارات زبان طبیعی برچسب‌گذاری نموده‌اند، ناشی از کنترل اندکی است که بر واژه‌ها و ساختار عبارات صورت می‌گیرد و این نظام‌ها در تحلیل دقیق زبان طبیعی به‌ویژه در رفع مشکلات "تبارشناسی"<sup>2</sup> واژه‌ها و همچنین، تعیین مرز دقیق بین آنها به‌هنگام تغییر و اصلاح آنها دچار چالش‌های زیادی هستند (Audrey, Clement, and Leung ۲۰۰۱).

هرچند بازیابی تصاویری که با زبان طبیعی نمایه شده‌اند بسیار پیچیده می‌نماید، فنون بازیابی مبتنی بر متن و فنون بازیابی مبتنی بر ویژگی تصاویر نیز با مشکلات خاص خود روبرو است.

#### 4-1. مشکلات فنون بازیابی مبتنی بر متن تصاویر

شاید بتوان گفت تصاویر به‌مثابه ابتدای دوران کودکی و زبان و واژگان به‌منزله خرد بالغ هستند. در بیشتر موارد، متون در مرکز توجه و برعکس، تصاویر در حاشیه و اهمیت پایین قرار داشته‌اند. در میان دانشمندان تاریخ هنر، افرادی که به مطالعه تصاویر می‌پردازند نیز به‌طور عمده بر

۱. Invariance

۲. Antecedent

روی مباحث لغوی تمرکز داشته‌اند. علاوه بر آن، قسمت بیشتر تاریخ هنر به‌میزان قابل توجهی تکیه بر واژگان و بهره‌گیری از آنها در جهت مباحث خود داشته است، حال آنکه در صورت عدم پشتیبانی واژگانی از تصویر، آن تصویر از تجلی‌گاه خود به‌دور خواهد بود. به‌نقل از والتر بنجامین تجلی‌گاه یک تصویر عبارت از است از "جوهره تمامی آنچه که از ابتدای پیدایش تصویر قابل انتقال است" و نیز اشاره به زمان، مکان، عملکرد اصلی، و موقعیت متغیر آن دارد (Roberts ۲۰۰۱). بدون مستندسازی و پشتیبانی واژگانی، تصاویر همواره به‌مثابه جسم شناور بر روی آب و جدا از زمان و مکان و اصالت خود خواهند بود. جهت درک بیشتر ماهیت تجلی‌گاه تصویر و ارتباط آن با توصیف واژگانی می‌توان آن را به‌مانند عددی دانست که به‌صورت مجرد فاقد بار معنایی است، اما در زمینه‌ای از آمار می‌تواند معنایی خاص خود داشته باشد. در عمل هم مشاهده می‌شود که بسیاری از پایگاه‌های اطلاعاتی تصاویر، از عناصر واژگانی کمتری بهره می‌گیرند. اما، این روش یک خطر جدی می‌تواند داشته باشد و آن اینکه فقط از طریق مشاهده یک تصویر نمی‌توان در مورد آن به درک و برداشتی همه‌جانبه دست یافت.

از طرف دیگر، بدیهی است که واژه‌ها به‌تنهایی میزان دقت پایینی دارند، این امر به‌ویژه هنگام بازیابی مواردی از میان مجموعه‌های بزرگ نمود بیشتری می‌یابد که خود به‌دلیل نقصان در نحو آنها به‌دست می‌آید. جستجوی کلیدواژه‌ای در ارتباط دادن بین صفت‌ها و اسامی ناتوان است (Audrey, Clement, and Leung ۲۰۰۱) و در نتیجه همین ناتوانی است که در جستجوی تصویری یک واژه ساده در نظام بازیابی اطلاعات تصویری (مانند AltaVista و Photo finder) هزاران تصویر که بیشتر نامرتب هستند بازیابی می‌شود و همین جریان سیل‌گونه نتایج بی‌ربط، نارسایی رویکرد جستجوی مبتنی بر کلیدواژه‌ای محض را هر چه بیشتر نمایان می‌سازد.

یکی از واقعیت‌های مسلم پژوهش در بازیابی تصاویر، این است که تفاوتی ساختاری میان داده‌های متنی و داده‌های دیداری وجود دارد و به‌همین دلیل بدیهی است که استفاده از ابزارها و فوننی که برای بازیابی داده‌های متنی ایجاد شده‌اند، جهت بازیابی داده‌های دیداری و نیز برعکس آن، مشکلاتی را ایجاد خواهد نمود.

در متون مختلفی چنین ادعا شده است: "در حالی که متن ساختار یافته است، تصاویر در کل بدون ساختار هستند". چنین ادعایی بر این اساس مطرح شده است که متون به‌طور منطقی توسط واژه‌ها، جملات، پاراگراف‌ها، و در شکل الکترونیکی آن از طریق کدهای نمادین نظیر "اسکی"<sup>۱</sup>، ساختار یافته‌اند. اما، نکته حائز اهمیت این است که چنین کدها و ساختارهایی امکان درک معنایی یا تفسیر یک متن توسط رایانه را مقدور نمی‌سازند. از طرف دیگر، تصاویر دارای حجم زیادی از

<sup>۱</sup> ASCII

داده‌هایی هستند که آنها را درون خود به صورت فضایی سازماندهی کرده‌اند. بنابراین، آیا ممکن نیست که اطلاق ساختگی صفت "ساختاریافته" به داده‌های متنی و صفت "بدون ساختار" به داده‌های ارائه‌شده درون یک تصویر، بازتابی از این واقعیت باشد که چارچوب‌های موجود در بازیابی متون به پردازش و بازیابی تصویر تسری داده می‌شوند؟ (Audrey, Clement, and Leung 2001) و اینکه در صورت وجود تفاوتی واقعی میان این دو مفهوم، آیا توسعه نابرابر فنون تشخیص و شناسایی تصاویر در برابر توسعه چنین فونی برای بازیابی متون دلیل اصلی این تفاوت نبوده است؟

مطالعات انتقادی متعددی در مورد عدم توانایی واژه‌ها در بیان واقعی آنچه که کاربر در واقع در جستجوی آن است، منتشر شده‌اند. اما، باید توجه داشت که چنین چالش‌هایی در مورد ویژگی‌های تصاویر نیز مطرح است. پرسش‌های سطح بالا همیشه نیازمند قوه ابتکار و نیز تفکر جانبی است تا در سوی جستجوگر انسانی، گستره بین آنچه مطلوب کاربر است و آنچه قابل جستجو است را تعیین کند.

به این ترتیب، این واقعیت که نویسندگان از تصاویر و متن در کنار هم برای انتقال ایده خود استفاده می‌کنند، به دلیل اعتقاد به این حقیقت است که همواره تصاویر می‌توانند مطالبی را بیان کنند که متون فاقد این توانایی هستند. علاوه بر آن، ایشان معتقدند که نه تنها تصاویر می‌توانند اطلاعات را به طریقی متفاوت از متن انتقال دهند، می‌توانند حتی اطلاعات متفاوتی را انتقال دهند (Layne 1994).

#### 2-4. مشکلات بازیابی معنایی تصاویر

مشکلات بازیابی معنایی اطلاعات تصویری را در نقل قولی از پیرس<sup>1</sup> می‌توان بهتر دریافت:

- (1) مسأله مستقل بودن مفهوم از هر عامل دیگر،
- (2) مسأله ارتباط و میانکنش با عوامل دیگر، و
- (3) مسأله میانجیگری، در زمان برقراری ارتباط میان موارد "یک" و "دو".

ویژگی‌های سطح پایین (مثل رنگ و ریخت و بافت) موارد مستقل از عوامل دیگر هستند و در "بند یک" قرار می‌گیرند. روابط معنایی به دلیل وابسته بودن به پیش‌زمینه کاربر در "بند دو" قرار می‌گیرند. اما، آنچه در مورد آن گسست احساس می‌شود، مورد سوم یعنی فقدان قواعد یا نظام‌های استدلالی جهت برقراری ارتباط میان موارد یک و دو یا به‌طور ایده‌آل ایجاد مورد دو از مورد یک است. بنابراین، برای رفع این مشکلات، نیازمند تفسیری معنایی از داده‌های دیداری برای

1. Peirse

یادداشت‌نویسی گزارمانی آنها هستیم (Hove ۲۰۰۴).

عمده چالش‌های بازیابی معنایی تصاویر را می‌توان در سه مقوله "حجم" و "موضوع" و "توصیف"<sup>۱</sup> بیان کرد.

#### 3-4. چالش حجم و موضوع

بازیابی تصاویر درون نظام‌ها براساس یادداشت‌هایی گزارمانی است که درباره هر تصویر در قالب کلیدواژه‌های توصیفگر نوشته شده است. افراد مختلف، محتوای یک تصویر را به گونه‌های متفاوتی درک می‌کنند. دانشجوی هنر یک سبک نقاشی را به گونه‌ای متفاوت نسبت به دانش‌آموز دبیرستان توصیف می‌کند. به علاوه، درک افراد مختلف از معنای یک کلیدواژه می‌تواند متفاوت باشد. این مشکلات مربوط به حجم و موضوع هستند. تاکنون راه‌حل مناسبی برای تولید خودکار ابر داده معنایی ارائه نشده است و این نوع ابر داده‌ها همچنان به صورت دستی اعمال می‌شود. این مقوله تا زمانی که اندازه مجموعه تصاویر کوچک باشد، مشکل چندانی را ایجاد نمی‌کند. اما، هنگامی که اندازه مجموعه تصاویر بزرگتر می‌شوند، یادداشت‌های گزارمانی دستی، مستعد مشکلات و چالش‌های حجم و موضوع می‌گردند (Brathen ۲۰۰۹; Rui, Huang, and Mehrotra ۱۹۹۸). این دو نوع مشکل با این واقعیت در ارتباط هستند که یادداشت‌نویسی گزارمانی یک تصویر، فرایندی زمان‌بر است و نیز قرار گرفتن محتوای پیچیده تصویر در کنار تفاوت‌های نوع برداشت انسانی از آن، موجب واگرایی تفاسیر یک تصویر واحد می‌شود.

در دهه‌های اخیر از بُعد رایانه‌ای، بازیابی مبتنی بر محتوای تصاویر (سی بی آی آر)<sup>۲</sup> به عنوان یک راه‌حل ممکن برای این مشکلات مطرح شده است. سازوکارهای نظام‌های مبتنی بر محتوای، با یک درجه مشخص می‌توانند تصاویر را برحسب ویژگی‌های نحوی آنها مثل رنگ، بافت، ریخت، و محل قرارگیری شیء در تصویر به‌طور موفقیت‌آمیزی مقایسه و بازیابی کنند. اما، بازیابی خودکار تصاویر، برپایه درجه بالاتری از محتوای نظیر محتوای معنایی آنها هنوز دچار چالش است (Hove ۲۰۰۴; Hartvedt ۲۰۰۷)، در حالی که بازیابی برپایه ویژگی‌های نحوی برای برخی از حوزه‌های موضوعی کافی به نظر می‌رسد. اما، اکنون نیازمند نوعی از نظام‌های بازیابی هستیم که قابلیت نمایه‌سازی و بازیابی تصاویر را برپایه درجه بالاتری از محتوای دارا باشند.

#### 4-4. چالش توصیف

باوجود اینکه رده‌بندی مبتنی بر متن دارای قدرت گویایی بالایی است، برخی محدودیت‌ها

۱. Explicability

۲. Content-Based Image Retrieval (CBIR)

در نحوه برخورد با اشیای دیداری ظاهر می‌شوند. توصیف بعضی ویژگی‌های نحوی تصویر از طریق واژه بسیار دشوار است. برای مثال، اگرچه مجموعه‌ای از واژه‌ها برای توصیف رنگ‌های مختلف وجود دارد، هیچ کدام از آنها آن‌طور که باید و شاید نمی‌توانند توصیف‌کننده آن رنگ باشند، چرا که هر رنگی دارای طیف گسترده‌ای با شدت و حدت‌های خاص خود است. هرچند بیشتر افراد توانایی تشخیص دو رنگ با اختلاف جزئی را درون یک طیف رنگ خاص دارا هستند، بیان این تفاوت‌ها از طریق کلمات بدون استفاده از واژگان فازی نظیر "بیشتر" و "کمتر" بسیار دشوار است. علاوه بر آن، ارائه توصیف‌های متنی واقعی و عینی از بافت و ریخت تصویر نیز بسیار مشکل‌آفرین است. این‌گونه مشکلات را چالش توصیف می‌نامند (Hove ۲۰۰۴; Eakins and Graham ۱۹۹۹).

#### ۵. شکاف معنایی

هرچند نمایه‌سازی و رده‌بندی مبتنی بر واژگانی تصاویر به‌عنوان ابزار قدرتمندی قادر به توصیف محتوای تصاویر است، این روش مستعد مشکلاتی چون حجم و موضوع و توصیف است. بنابراین، برای استفاده در مجموعه‌های بسیار بزرگ تصویری ایده‌آل نیست. فنون نمایه‌سازی و بازیابی مبتنی بر ویژگی نحوی و الگوی داده‌ای به‌عنوان رویکردی جایگزین برای بازیابی تصاویر پیشنهاد شده و به‌عنوان فنی کارآمد و مفید مورد اثبات قرار گرفته است.

با این وجود، بازیابی مبتنی بر ویژگی‌های نحوی تصویر از پاسخگویی به پرسش‌ها در تمامی سطوح پشتیبانی نمی‌کند. در واقع، بین آنچه کاربران از یک نظام بازیابی تصاویر می‌خواهند و آنچه که نظام‌های جاری بازیابی اطلاعات قادر به ارائه آن هستند، یک اختلاف وجود دارد و این خلاء - بین آنچه ممکن است و آنچه ما می‌خواهیم - از سوی بسیاری از پژوهشگران با عنوان "شکاف معنایی" نام برده شده است (Eakins and Graham ۱۹۹۹). خلاصه کلام اینکه به‌دلیل عدم درک معنایی محتوای یک تصویر، مشاهده می‌شود که موتور بازیابی تصاویری را بدون هیچ‌گونه رابطه معنایی با تصویر ذهنی کاربر بازیابی می‌کند و به این ترتیب، همواره مشکلات زیر پیش می‌آید:

(۱) عدم بازیابی اشیای مورد نظر به‌دلیل نقصان در تشخیص و شناخت اشیاء،

(۲) عدم بازیابی اشیای مورد نظر به‌دلیل نقصان در تشخیص و شناخت روابط، و

(۳) بازیابی نادرست اشیاء به‌دلیل تشخیص و شناخت نادرست اشیاء (Hove ۲۰۰۴).

وجود یک نظام بازیابی تصاویر که بتواند تصاویری را که هم از نظر محتوای معنایی و هم از نظر محتوای ساختاری به هم شبیه هستند بازیابی نماید، راه‌حلی است که برای کاهش این

۱. Semantic gap

مشکلات و رفع خلأ معنی‌شناختی پیشنهاد می‌شود. چنین پیشنهادی می‌تواند کاربست ابزاری قدیمی تحت عنوان اصطلاحنامه برای کاربردی نوین یعنی بازیابی معنایی تصاویر به‌ویژه در محیط‌های رقومی جدید باشد. ادامه مقاله در ارتباط با این طرح پیشنهادی است.

## 6. اصطلاحنامه در نظام‌های بازیابی تصاویر

یونی‌سیست<sup>۱</sup> اصطلاحنامه را چنین تعریف می‌کند:

«یک اصطلاحنامه ممکن است به‌لحاظ عملکرد یا ساختار آن تعریف شود. به‌لحاظ عملکرد، اصطلاحنامه یک ابزار کنترل واژه‌شناختی برای ترجمه از زبان طبیعی مدارک، نمایه‌سازان و کاربران به یک زبان تحمیل‌شده سامانه‌ای است، و به‌لحاظ ساختار، اصطلاحنامه مجموعه‌ای از واژگان کنترل‌شده و پویاست که به‌نوعی و به‌لحاظ معنایی به واژگانی که حوزه خاصی از دانش را پوشش می‌دهند، ارتباط دارد» (Foskett ۱۹۹۷).

بسته به دیدگاه یک بیننده، محتوای یک تصویر می‌تواند در سطوح مختلف انتزاعی قابل توصیف باشد. برای داشتن درکی روشن از محتوای یک تصویر، نیاز به بررسی عناصر مفهومی و ساختاری آن است. مفهوم یک تصویر با بازنمون یا ایده‌ای انتزاعی و عمومی از نمونه‌های ادراک‌شده آن تصویر در ارتباط است. در واقع، مفهوم تصویر، بر دانش پیش‌زمینه‌ای و تفسیری ذاتی از آنچه ادراک شده است، دلالت دارد.

همچنین، در بررسی محتوای تصویر، ترکیب یا نحو آن در مقابل معنی‌شناختی تصویر قرار دارد. در حالی که ادراک تصویر، مربوط به باوری است که بیننده از طریق حواس خود بدان نایل شده است، نحو یا ترکیب تصویر، مربوط به خود عناصر دیداری و روش نظم و چیدمان آنهاست. حال آنکه معنی‌شناسی تصویر، فقط با معنی و مفهوم عناصر نحوی و روش نظم و چیدمان آنها در ارتباط است. در کل، محتوای نحوی تصویر شامل عناصر ساختاری آن مانند رنگ، بافت، ریخت، و ترکیب فضایی است و در مقابل، محتوای معنایی شامل مفهوم یک تصویر از جمله دلالت‌هایی ضمنی نمادین، شماتیک، ذهنی، و احساسی و رای موضوع صریح آن است (Jaimes and Chang ۲۰۰۲).

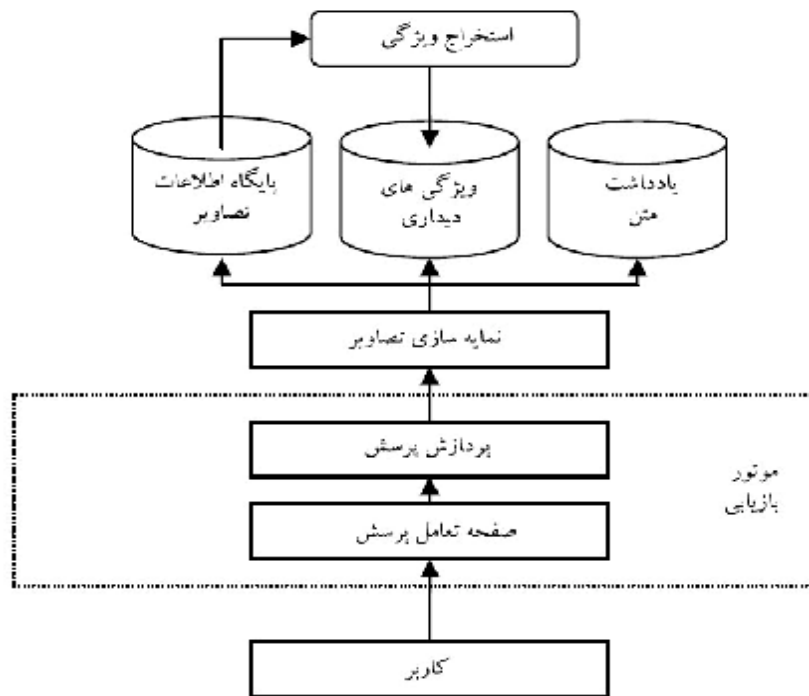
برداشت ما از محتوای معنایی یک تصویر، بر پایه استفاده از دانش ما در تفسیر ادراکات و اشارات دیداری تصویر قرار دارد. این اشارات دیداری از محتوای ساختاری تصویر تشکیل یافته است. رنگ‌ها، ریخت‌ها، و سایر ویژگی‌های نحوی در کنار هم، اشیای دیداری قابل شناسایی تصویر را ایجاد می‌کنند. این اشیای دیداری، واحد اساسی معنی‌شناختی در یک تصویر است و

۱. UNISIST

توانایی شناسایی اشیای مجسم شده در متن را مقدر می سازد. اما، زمانی که بحث جستجوی معنایی تصاویر مطرح می شود، در ارتباط با فنون بازیابی آنها با مسائل و چالش های زیادی روبرو می شویم. رویکردهای جدید باید شکاف معنایی موجود بین نظام های بازیابی جاری و کاربردهای مورد نیاز کاربر را کاهش دهند. به سبب همین انگیزه، رویکردی نوین جهت بازیابی تصاویر پیشنهاد شده است. در این روش، فنون بازیابی مبتنی بر ویژگی و مبتنی بر داده تصاویر با ساختار ارائه شده توسط یک اصطلاحنامه، ترکیب و در قالب اصطلاحنامه ریختاری مطرح می شوند.

### 7. اصطلاحنامه ریختاری

هوانگ و رویی در سال 1999 معماری زیر را برای نظام های بازیابی تصاویر، پیشنهاد کرده اند:



شکل 1. معماری نظام بازیابی تصاویر (منبع: Huang and Rui ۱۹۹۹)

این معماری دارای سه عنصر اصلی موتور بازیابی، یک ساختار نمایه سازی، مجموعه ای از مجموعه های داده ای و نیز سازوکارهایی برای جستجو و شناسایی اشیای مبتنی بر ویژگی های نحوی است. این معماری همچنین، شامل سه نوع مجموعه داده ای نیز است که عبارتند از: مجموعه



داده‌های خام تصویری برای اهداف نمایش و بازنمایی، مخزنی از ویژگی‌های دیداری استخراج‌شده از تصاویر برای پشتیبانی از جستجوهای مبتنی بر ویژگی تشابه، و مخزنی از یادداشت‌های گزارمانی متنی شامل توصیفات کلیدواژه‌ای و متن آزاد از تصاویر. همچنین، سازوکارهایی به منظور استخراج ویژگی‌های نحوی تصاویر و ذخیره آنها در مخزن ویژگی‌ها وجود دارد. نمایه‌سازی چندبُعدی، از بازیابی سریع تصاویر و همچنین از قابلیت سنجش و اندازه‌گیری مجموعه‌های بزرگ تصاویر پشتیبانی می‌کند. موتور بازیابی شامل یک رابط پرسش و یک واحد پردازش پرسش است. رابط پرسش، فنون جلوه‌های گرافیکی و دست‌کاری مستقیم را به کار می‌گیرد، اطلاعات را از کاربران جمع‌آوری و درنهایت، نتایج بازیابی شده را ارائه می‌کند (Huang and Rui ۱۹۹۹). بنابراین، به نظر می‌رسد که این معماری توصیف‌شده، یک چارچوب کلی مطلوب را از عناصر اصلی یک نظام بازیابی اطلاعات تصویری ارائه می‌کند. درنهایت، اینکه بر مبنای معماری این نظام، پیشنهاد چارچوب نهایی اصطلاحنامه‌ای ریختاری به‌عنوان ابزاری برای کاهش برخی چالش‌های مرتبط با شکاف معنایی ارائه می‌گردد.

یک اصطلاحنامه مبتنی بر ریختار تصاویر، این گونه می‌تواند تعریف شود:

- (1) سیاه‌های از پیش تدوین‌شده از ریختارهای مهم که اشیای دیداری را در حوزه‌ای از دانش بازنمون می‌کنند؛
  - (2) توصیفگرهای ویژگی‌هایی که ریختارها را توصیف می‌کنند؛
  - (3) توصیف متنی / معنایی این ریختارها؛ و
  - (4) مجموعه‌ای از ریختارهای مرتبط با هر کدام از ریختارهای موجود (Hove ۲۰۰۴).
- اما، همچنان که پیشتر نیز اشاره شد، دو مشکل عمده در ارتباط با فنون جاری بازیابی مبتنی بر ویژگی‌های نحوی و نیز مبتنی بر الگوی داده‌ای که به عدم بازیابی تصاویر مطلوب منتهی می‌شود، عبارتند از:
- (5) عدم شناسایی و تشخیص اشیا؛ و
  - (1) عدم درک روابط بین اشیا.

با نگاهی به روش سنتی بازیابی مبتنی بر متن، به نظر می‌رسد "اصطلاحنامه" فقط ابزاری است که می‌تواند روابط معنایی را بین واژگان مختلف برقرار سازد. این روش در بازیابی داده‌های متنی، بازیابی متونی را که به لحاظ معنایی با پرسش کاربر مرتبط هستند، ممکن می‌سازد.

اما آنچه مورد نظر است، ارزیابی قابلیت بهره‌گیری از یک ساختار مبتنی بر یک اصطلاحنامه متنی و یا حتی یک ساختار مشابه اصطلاحنامه متنی در ارتقای فرایند بازیابی تصاویر، به‌ویژه بررسی امکان استفاده از چنین ساختاری برای برقراری و حفظ پیوندها بین ریختارهای به‌ظاهر

نامشابه، ولی به لحاظ محتوای معنایی مشابه است. همچنین هدف دیگر آن، بررسی امکان توانایی چنین ساختاری برای شناسایی و تشخیص اشیای دیداری درون تصاویر است. در صورت نایل شدن به چنین هدفی، به نظر می‌رسد که بتوانیم در رفع شکاف معنایی گامی برداشته باشیم، بنابراین ارائه طرح پیشنهادی اصطلاحنامه ریختاری دو انگیزه عمده دارد:

الف) کمک به نظام‌های بازیابی تصاویر در تشخیص و شناخت ریختارهایی که در عین دارا بودن کمترین تشابه ساختارهای نحوی، دارای تشابه محتوای معنایی - در بعضی موارد، زیاد- هستند؛ و

ب) کمک به نظام‌های بازیابی تصاویر در بازیابی تصاویری که شامل ریختارهایی هستند که به لحاظ معنی‌شناختی به محتوای معنایی ریختاری دیگر مرتبط هستند.

#### 1-1. ویژگی‌های ریختاری تصویر

تاکنون مشاهده شد که "ریخت" یکی از مهمترین ویژگی‌های نحوی توصیف‌کننده محتوای تصویر و اشیای دیداری است. اما، ریخت چیست؟ چگونه نمایش داده می‌شود؟ و چگونه در نظام‌های بازیابی اطلاعات مورد استفاده قرار می‌گیرد؟ عینیت خارجی یا حالت مرئی شی، مرز میان اشیای درونی و اشیای بیرونی، طرح کلی و صفت ممیزه پیکره‌بندی سطح شی یا به عبارتی کلی‌تر، آنچه پس از پالایش تصویر براساس ویژگی‌هایی چون مقیاس، موقعیت، و وضعیت گردش آن باقی می‌ماند، "ریخت" تصویر نامیده می‌شود (Kimia ۲۰۰۲).

استفاده از ریخت به عنوان یک راهنمای دیداری در بازیابی تصویر نسبت به استفاده از رنگ و بافت تصویر در چنین فرایندی، کمتر گسترش یافته است. علت اصلی این موضوع را می‌توان در پیچیدگی ذاتی نمایش ریخت جستجو کرد. با این حال، به نظر می‌رسد که بازیابی تصاویر با استفاده از ویژگی ریختاری آنها، برای تبدیل شدن به یک فن جستجوی بسیار کارآمد تصاویر در بیشتر حوزه‌های کاربردی دارای پتانسیل زیادی است، چرا که "ریخت" تصویر، یکی از مهمترین راهنماهای برداشت انسانی جهت تفسیر ادراکات دیداری است.

هرچند که جزء اول اصطلاحنامه ریختاری یعنی سیاهه ریختارها، فهرستی از ریختارهای مهم را فراهم می‌نماید، هرکدام از این ریختارها باید توسط مجموعه‌ای از توصیفگرها توصیف شوند. به نظر می‌رسد این توصیفگرها در واقع مهمترین جزء اصطلاحنامه ریختاری باشند، چرا که تشخیص و شناسایی مناسب واژگان اصطلاحنامه‌ای، وابستگی کامل به قابلیت یافتن تشابه‌های میان محتوای تصویر و ریختارهای اصطلاحنامه دارد.

انتخاب توصیفگرهای مناسب برای نمایش ریختارها، در کنار الگوریتم‌هایی برای مقایسه

ریختارها، مهمترین تصمیم‌گیری در فرایند ایجاد اصطلاحنامه ریختاری است. به‌نسبت تخصصی‌تر شدن توصیفگرها، الگوریتم‌های دامنه موضوعی نیز جزئی‌تر می‌شوند. این مسأله نشانگر آن است که ضرورت دارد انتخاب هر توصیفگری درون اصطلاحنامه ریختاری، بر مبنای ویژگی‌های حوزه کاربردی آن توصیفگر ویژه صورت پذیرد (Sclaroff and Liu ۲۰۰۱).

### 2-7. برچسب معنایی<sup>۱</sup>

در صورتی که قرار باشد اصطلاحنامه ریختاری به‌طور مستقیم در دسترس کاربر قرار گیرد، باید یک توصیف متنی/معنایی از اشیای درون تصاویر وجود داشته باشد. گرچه ممکن است که چنین توصیفی از ریختارها از طریق نسبت دادن یک نام معنی‌دار به هر کدام از آنها، در سیاهه ریختارها وجود داشته باشد، این احتمال وجود دارد که درک برخی از واژه‌ها دشوار باشد و بعید به‌نظر می‌رسد که توصیفگرها بتوانند اطلاعات مورد نیاز را به‌نحو مقتضی انتقال دهند. بنابراین، ضروری به‌نظر می‌رسد که یک برچسب معنایی نیز به هر یک از واژه‌ها الحاق شود (Hove ۲۰۰۴).

### 3-7. روابط اصطلاحنامه‌ای

در اصطلاحنامه ریختاری مجموعه‌ای از روابط بین اشیای تعریف‌شده وجود دارد. درحالی‌که اصطلاحنامه مبتنی بر متن بیشتر دارای سه نوع رابطه اصطلاح اعم، اصطلاح اخص، و اصطلاح وابسته است، اصطلاحنامه ریختاری این روابط را با افزودن دو نوع رابطه دیگر گسترش داده و انواع روابط اصطلاحنامه‌ای را در دو بخش اصلی روابط بین اشیا و روابط بین ریختارها، سازماندهی کرده است:

- ۱) روابط بین اشیا: این نوع روابط بین اشیای دیداری که ریختارها را نمایش می‌دهند، برقرار است. روابط بین اشیا از طریق برقراری ارتباط معنایی بین محتوای تصاویر و فقره‌های مورد درخواست درون پرسش کاربر - هرچند به‌طور کامل هم منطبق نباشند - فرایند بازیابی تصاویر را پشتیبانی می‌کنند. درواقع، این روابط، مشابه روابط موجود در اصطلاحنامه‌های واژگانی است و همانند آنها امکان گسترش پرسش و همچنین پالایش آن را فراهم می‌آورد.
- ۲) روابط بین ریختارها: این نوع روابط بین ریختارهایی که اشیای دیداری را نمایش می‌دهند، برقرار است. روابط بین ریختارها از طریق بازنمایی بخش‌های زیرمجموعه‌ای یا مغایر اشیای دیداری مشابه، فرایند بازیابی تصاویر را پشتیبانی می‌کنند. درواقع، این نوع روابط، مشکل عدم

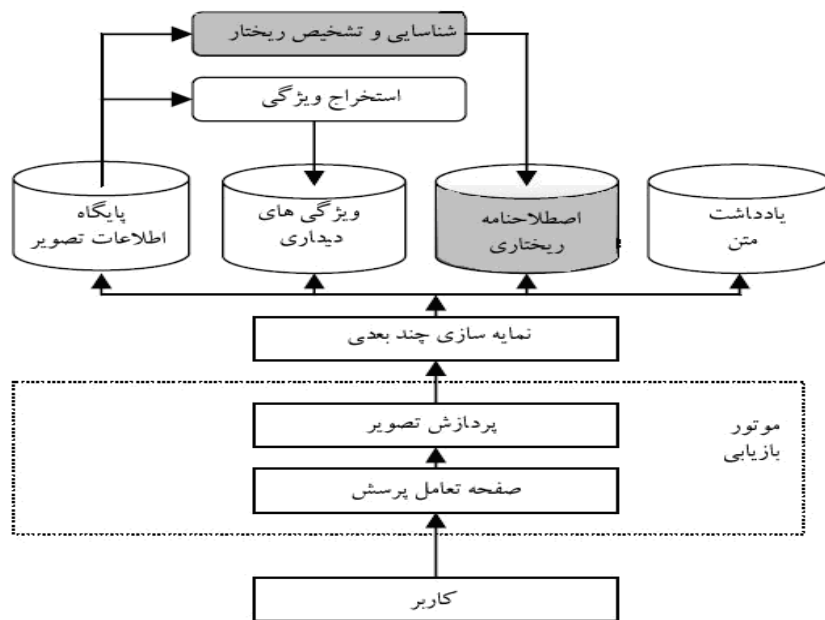
۱. Semantic label

تغییرناپذیری را مورد توجه قرار می‌دهند. رابطه "بخشی - از"<sup>1</sup> بازیابی تصاویری که در آنها، فقره موجود را در پرسش به صورت کامل آشکار زیرمجموعه‌ای از یک فقره دیگر است، مقدور می‌سازد. همچنین، رابطه "مغایر - با"<sup>2</sup>، بازیابی تصاویری را که در پرسش کاربر برای اشیای یکسان، ریختارهای مختلف و مغایری وجود دارند، ممکن می‌سازد. علاوه بر آن، روابط "بخشی - از" و "مغایر - با" می‌توانند برای شناسایی اشیای دیداری نمایش داده شده در ریختارها مورد استفاده قرار گیرند. از طریق تطابق یک ریختار درون تصویر با ریختارهای موجود در اصطلاحنامه، نظام بازیابی تصاویر می‌تواند موفق به بازیابی تصاویر مرتبط شود (Hove ۲۰۰۴).

باید توجه داشت که اصطلاحنامه ریختاری تا زمانی قادر به پشتیبانی از این فعالیت‌ها خواهد بود که خود توسط عملکردهای "تشابه‌سازی"<sup>3</sup> کارآمدی پشتیبانی شود. در صورتی که عملکردهای تشابه‌سازی قادر به پردازش تصویر، شناسایی ریختارها و مقایسه آنها با توصیفگرهای ریختارهای اصطلاحنامه نباشد، به تبع آن، اصطلاحنامه ریختاری نیز نخواهد توانست از نظام بازیابی تصویر پشتیبانی کافی داشته است. این موضوع در واقع، تأییدی دوباره بر اهمیت توصیفگر و عملکردهای تشابه‌سازی مناسب است.

شکل 2 مراحل مختلف فرایند بازیابی تصویر را نشان می‌دهد و چگونگی پشتیبانی اصطلاحنامه ریختاری از مراحل مختلف این فرایند را به تصویر می‌کشد.

۱. Part-of  
۲. Invariant-with  
۳. Similarity functions



شکل 2. معماری نظام بازیابی تصاویر با استفاده از اصطلاحنامه ریختاری (منبع: Hove ۲۰۰۴)

#### 4-7. تفسیر پرسش

ضروری است که اصطلاحنامه ریختاری بتواند پرسش‌ها را در دو شکل دیداری و متنی دریافت کند. فرایند تفسیر پرسش در مورد پرسش‌های متنی، از طریق فنون بازیابی مبتنی بر متن اطلاعات صورت می‌گیرد، به این ترتیب که متن ورودی با برجسب‌های معنایی اشیای اصطلاحنامه‌ای مقایسه می‌شود و در صورت وجود تطابقی بین آنها، نظام می‌تواند گام‌های بعدی را برپایه این داده‌ها پایه‌گذاری کند (Lu ۱۹۹۹).

اما، در مورد پرسش‌های دیداری (پرسش از طریق مثال و پرسش از طریق طرح اولیه)، تفسیر پرسش وارد فرایند پیچیده‌ای می‌شود. برای اثبات مفید بودن اصطلاحنامه، وجود سازوکاری جهت مقایسه تصویر مورد درخواست پرسش با توصیفگرهای ریختارها ضروری است. موفقیت یا شکست این سازوکار بستگی به قابلیت تجزیه و تحلیل تصاویر و همچنین، شناسایی و تشخیص اشیای موجود در اصطلاحنامه دارد. در صورتی که نظام در شناسایی و تشخیص صحیح عناصر ارائه‌شده در یک تصویر با شکست روبرو شود، نخواهد توانست از اصطلاحنامه بهره‌ای ببرد.

در مورد پرسش‌های دیداری، تصاویر اصلی یا هسته ممکن است دارای یک یا چند شی در درون خود باشند. علاوه بر آن، هر تصویر در درون خود، به‌احتمال، دارای برخی عناصر فرعی نیز

باشد. هر چند می‌توان کاربران را ترغیب کرد که اشیای درون پرسش تصویری را از طریق ارائه مثال‌هایی مشخص‌تر سازند، بعید به نظر می‌رسد که این عمل در مورد تمامی تصاویر اصلی قابل اجرا باشد. الگوریتم‌های جداساز مناسبی به منظور جداسازی و تشخیص اشیای درون تصاویر مورد نیاز است. بنابراین، ریکتارهای جدا از هم حاصل از کاربرست این الگوریتم‌ها باید به‌عنوان اساس کار برای مقایسه ریکتارهای شناسایی و تشخیص داده‌شده، مدنظر قرار گیرند (Hove ۲۰۰۴).

#### 5-7. بسط پرسش

پس از آنکه اشیای موجود در پرسش شناسایی شدند، روابط اصطلاحنامه‌ای می‌توانند برای بسط پرسش مورد استفاده قرار گیرند. در واقع، این روابط برای شناسایی و تشخیص اشیایی از اصطلاحنامه که دربرگیرنده اشیای درون پرسش هستند به کار گرفته می‌شود. البته در این بین، تعیین اینکه تا چه سطحی از این روابط باید مورد بررسی قرار گیرد، نیازمند مجموعه‌ای از معیارهاست. این موضوع می‌تواند از طریق پارامترهایی پیش‌گزیده، یا تعیین شده توسط کاربر قبل از طرح پرسش یا در طی آن صورت پذیرد (Hove ۲۰۰۴).

در مرحله بعد، روابط بین ریکتارها می‌توانند برای شناسایی و تشخیص اشیایی که به‌لحاظ معنایی به واژگان شناسایی شده مرتبط هستند، به کار گرفته شوند. در اینجا ضروری است که بیشترین دقت اعمال شود، چرا که باید توجه داشت به تبع افزایش تعداد اشیاء، نتایج زیادی نیز حاصل خواهد شد. بعید به نظر می‌رسد که از روابط بین اشیاء بتوان در بسط پرسش برای گسترده‌تر یا خاص کردن نتایج بهره جست.

#### 6-7. بازیابی تصویر از طریق اصطلاحنامه

در طی مرحله بازیابی تصویر، از اصطلاحنامه می‌توان برای بازیابی تصاویر شامل اشیایی که در طی تفسیر پرسش مورد شناسایی و تشخیص قرار گرفته‌اند، استفاده کرد. در صورتی که تصاویر درون یک مجموعه توسط متن یا واژگانی یادداشت‌نویسی گزارمانی شده‌اند، استفاده از برچسب‌زنی معنایی برای اشیای شناسایی و تشخیص داده‌شده به‌عنوان معنای اساسی برای جستجوهای مبتنی بر شباهت‌های واژگانی، مقدور است. تصاویری که توسط واژه‌هایی مشابه برچسب معنایی تصویر، یادداشت‌نویسی گزارمانی شده‌اند، به احتمال زیادی به پرسش اصلی کاربر مرتبط هستند. برای استفاده کامل از اصطلاحنامه، نیاز به برقراری برخی "همبستگی‌ها"<sup>۱</sup> میان واژگان اصطلاحنامه و مجموعه‌های تصاویر است. بدون وجود همبستگی، کارآمدی یک

۱. Association

اصطلاحنامه ریختاری محدود به جستجوهای مبتنی بر متن خواهد بود. دو نوع از وابستگی‌های متفاوت میان اصطلاحنامه ریختاری و مجموعه تصاویر در اینجا مورد پیشنهاد و بحث قرار می‌گیرد.

#### 6-7-1. جستجوی تشابه‌های ریختاری

در توضیح جستجوی تشابه‌های ریختاری باید یادآور شد که از عملکرد تشابه‌سازی استفاده شده در شناسایی ریختارها، در اینجا نیز می‌توان برای بازیابی تصاویر شامل ریختارهای مشابه استفاده کرد.

این رویکرد بر استفاده از عملکردهای تشابه‌سازی برای مقایسه توصیفگرهای توصیف‌کننده اشیای اصطلاحنامه با توصیفگرهای توصیف‌کننده تصاویر درون مجموعه‌ها استوار است. نقطه قوت اصلی این رویکرد نسبت به توصیفگرهای نمایه‌سازی، توانایی آن در یافتن تصاویر بدون هیچ‌گونه پیش‌پردازش تصویر است (Kimia ۲۰۰۲).

#### 6-7-2. پیوندهای مستقیم

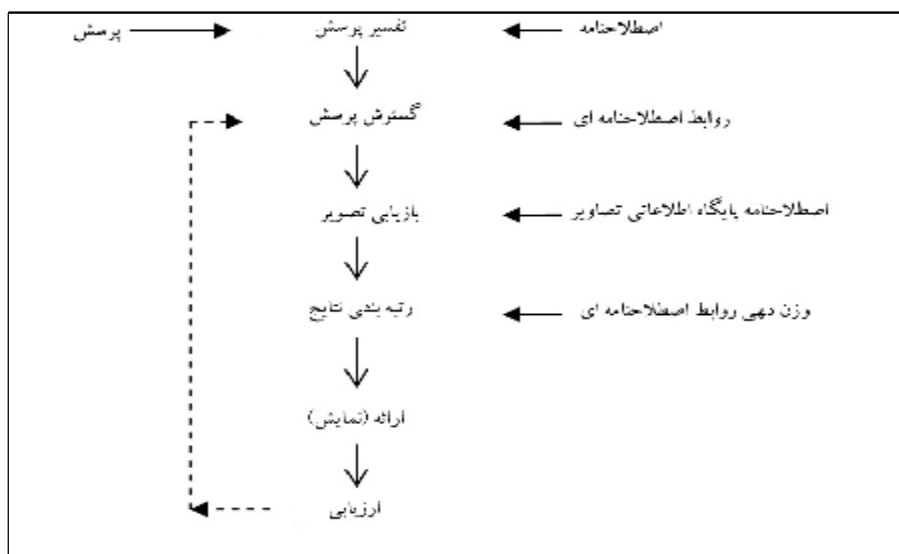
اشیای درون اصطلاحنامه می‌توانند به صورت مستقیم به تصاویر حاوی توصیفات آنها پیوند داشته باشند. بنابراین، چنین پیوندهایی برای بازیابی سریع تصاویر حاوی این اشیای مورد استفاده قرار می‌گیرند. در اینجا سه فنّ مختلف برای برقراری این پیوندها مورد بحث قرار می‌گیرد:

نخست پیوندهایی که می‌توانند به طور دستی برای تصاویر موجود در پایگاه اطلاعاتی یا الحاقی به آن، ایجاد گردند. چرا که عامل انسانی در قضاوت محتوای تصویر می‌تواند بسیار بهتر از فرایندی خودکار عمل کند. این رویکرد در شناسایی و تشخیص صحیح و کامل اشیای دارای درجه اطمینان بالاتری است. فنّ دوم، برقراری پیوند به صورت خودکار از طریق عملکردهای تشابه‌سازی ریختارهاست. پیونددهی خودکار نسبت به مشکلات حجم و موضوع حساس است. همچنین، یک فرایند خودکار توانایی اداره تعداد زیادی از تصاویر را در زمان کوتاهی داراست. علاوه بر آن، تشخیص و شناسایی خودکار ریختارها نسبت به تشخیص و شناسایی اشتباه حساس است و در نهایت، حاصل کار برقراری پیوندهای صحیح خواهد بود. فنّ سوم، استفاده از روش نیمه خودکار برای برقراری پیوندهاست. این روش برپایه تعامل میان نظام بازیابی تصاویر و کاربران آن برای برقراری پیوندهای مبتنی بر ساختار ربط/بازخورد قرار دارد. هرچند این روش در معرض چالش پیونددهی دستی است، با استفاده از میانگین وزن‌دهی شده‌ای از بازخوردهای تمامی کاربران می‌توان به راه‌حلی برای غلبه بر مشکلات پیونددهی دستی و نیز چالش موضوع و خطاهای فردی دست یافت (Li and Kuo ۲۰۰۲).

ترکیبی از این رویکردهای ارائه شده به احتمال زیاد می تواند رویکرد مناسبی را برای تحقق یک چارچوب اصطلاحنامه ریختاری فراهم آورد. از طرف دیگر، بعید به نظر می رسد که تمامی تصاویر به طور کامل مورد تجزیه و تحلیل و محتوای تمامی آنها با موفقیت مورد تشخیص و شناسایی قرار گیرند، مگر در مواقعی که مجموعه تصاویر کوچک باشد. این مسأله بار دیگر ضرورت جستجوهای مبتنی بر تشابه را یادآور می شود.

### 7-7. رتبه بندی نتایج

پس از اینکه نظام تصاویری را مرتبط یافت، بازیابی می کند. نتایج باید براساس میزان ربط به تصویر مورد درخواست پرسش کاربر رتبه بندی شوند. رتبه بندی این نتایج که به کمک استفاده از اصطلاحنامه ریختاری به دست آمده است (شکل 3)، می تواند براساس قرابت معنایی درک شده میان پرسش و تصاویر بازیابی شده باشد (Hove 2004).



شکل 3. بازیابی تصاویر با استفاده از اصطلاحنامه ریختاری

### 8. نتیجه گیری

محیط های الکترونیک جدید و از جمله آنها وب، قابلیت های زیادی را فراروی تدوین کنندگان اصطلاحنامه برای گنجاندن اطلاعات و مفاهیم بیشتر با ایجاد پیوندهایی به



پایگاه‌های داده‌های دیگر، به‌منظور غنی‌سازی شبکه مفهومی اصطلاحنامه، قرار داده است. این مسأله به‌خصوص در زمینه بازیابی رسانه‌های غیرمتنی از جمله تصاویر بسیار بااهمیت است. زیرا با آنکه تصاویر بیشتر اوقات اطلاعات مفیدی را برای پژوهشگران رشته‌های مختلف فراهم می‌آورند، برای افراد متفاوت می‌توانند معانی مختلفی را تداعی کنند و حتی لایه‌های معنایی متفاوتی داشته باشند، به‌طوری که بتوان از آنها برای اهدافی استفاده کرد که حتی خالق آنها نیز پیش‌بینی نکرده است. به این ترتیب، مسائل مربوط به زبان مبهم و مشکلات نمایه‌سازی تصاویر مطرح می‌شود، زیرا درباره ویژگی‌ها و صفاتی که درباره یک تصویر باید نمایه شود، توافقی وجود نداشته و سطوح نمایه‌سازی نیز تعریف نشده است. از سوی دیگر، در هنگام بازیابی، ترکیب واژه‌هایی که کاربران وارد نظام می‌کنند با واژگانی که در نمایه‌ها وجود دارد، هماهنگی ندارد و بنابراین، مشکلات مربوط به تفاوت مدل‌های ذهنی کاربران و مدل‌های ذهنی نمایه‌سازان درباره محتوای تصاویر، باقی است. اصطلاحنامه ریختاری سعی دارد تا این شکاف‌های معنی‌شناختی را با ترکیب فنون مبتنی بر متن و فنون مبتنی بر محتوی و نیز نزدیک کردن دیدگاه‌های نمایه‌سازان و متخصصان رایانه به یکدیگر برطرف سازد، تا بازیابی‌های کاربران به‌طور مؤثرتری صورت گیرد و بازیابی تصاویر نامرتبط، به‌میزان زیادی کاهش یابد. در صورت تحقق چنین اهدافی است که کاربران خواهند توانست در میان انبوهی از تصاویر موجود در وب، پاسخ واقعی نیاز اطلاعاتی خود را در یک جستجوی سریع و موفق به‌دست آورند.

## 9. منابع

- دستغیب، محمدباقر. 1382. مروری بر نمایه‌سازی متنی، محتوایی و فازی تصاویر رقمی: نظام بازیابی تصاویر رقمی. فصلنامه کتاب 16 (4): 111-116.
- سبزی‌پور، مجید. 1387. کاربرد "فولکسونومی" در بازنمون تصاویر دیجیتالی: رویکردی نوین در نمایه‌سازی کاربرمدار. کتابداری و اطلاع‌رسانی 11 (2): 143-160.
- کرمی، افسانه. 1385. نمایه‌سازی تصاویر در پایگاه‌های اطلاعاتی. مجله الکترونیکی نما (ارتباط علمی) 3 (6). <http://ejournal.irandoc.ac.ir> (دسترسی در 1388/6/2).
- Audrey, M., T. Clement, and H. C. Leung. ۲۰۰۱. Structured natural-language descriptions for semantic content retrieval of visual materials. *Journal of the American Society for Information Science and Technology* ۵۲ (۱۱): ۹۳۰-۹۳۷.
- Brathen, J.-E. ۲۰۰۶. An analysis of image folksonomy generation. Master's thesis. Faculty of Social Sciences, Department Of Information Science and Media Studies, Bergen University, Norway.
- Chu, H. ۲۰۰۱. Research in image indexing and retrieval as reflected in the literature. *Journal of the American Society for Information Science and Technology* ۵۲ (۱۲): ۱۰۱۱-۱۰۱۸.
- Eakins, J. P., and M. E. Graham. ۱۹۹۹. Content based image retrieval: a report to the JISC technology applications program. Newcastle, Institute for Image Data Research, University of Northumbria.
- Encinas, J., J. Liorens, and A. D. Miguel. ۱۹۹۹. A classification-matching combination for image retrieval. *Online & CD-ROM Review* ۲۳ (۱): ۱۱-۱۸.

- Foskett, D. J. ۱۹۹۷. Thesaurus. In *Readings in Information Retrieval*. Sparck Jones, K., Willet, P. (eds.), ۱۱۱-۱۳۴. San Francisco, California: Morgan Kaufmann Publishers.
- Hartvedt, C. ۲۰۰۷. Utilizing context in ranking results from distributed image retrieval – the CAIRANK Prototype. Master's thesis. Faculty of Social Sciences, Department Of Information Science and Media Studies, Bergen University, Norway.
- Hassan, I., and J. Zhang. ۲۰۰۱. Image search engine feature analysis. *Online Information Review* ۲۵ (۲): ۱۰۳ – ۱۱۴.
- Hove, L-J. ۲۰۰۴. Improving image retrieval with a thesaurus for shapes, the VORTEX prototype. Master's thesis. Faculty of Social Sciences, Department Of Information Science and Media Studies, Bergen University, Norway.
- Huang, T., and S. Y. Rui. ۱۹۹۹. Image retrieval: Current techniques, promising directions and open issues. *Journal of Visual Communication and Image Representation* ۱۰ (۴): ۳۹-۶۲.
- Jaimes, A., and S. F. Chang. ۲۰۰۲. Concepts and Techniques for Indexing Visual Semantics. In *Image Databases: Search and Retrieval of Digital Imagery*. V. Castelli and L. D. Bergman, ۴۹۷-۵۶۵ New York, John Wiley & Sons, Inc.
- Jørgensen, C., A. Jaimes, A. B. Benitez, and S-F. Chang. ۲۰۰۱. A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science* ۵۲ (۱۱): ۹۳۸-۴۷.
- Kimia, B. B. ۲۰۰۲. Shape representation for image retrieval. In *Image Databases: Search and Retrieval of Digital Imagery*. V. Castelli and L. D. Bergman, ۳۴۰-۳۷۰ New York, John Wiley & Sons, Inc.
- Layne, S. S. ۱۹۹۴. Some Issue in the indexing of images. *Journal of the American Society for Information Science* ۴۵ (۸): ۵۸۳-۵۸۸.
- Li, Y., and C.-C. J. Kuo ۲۰۰۲. Introduction to content-based image retrieval-overview of key techniques. In *Image Databases: Search and Retrieval of Digital Imagery*. V. Castelli and L. D. Bergman, ۲۶۱-۲۸۴. New York, John Wiley & Sons, Inc.
- Lu, G. ۱۹۹۹. *Multimedia database management systems*. Norwood, Artech House INC.
- Matusiak, K. ۲۰۰۶. Towards user- centered indexing in digital image collections. *International Digital Library Perspectives* ۲۲ (۴): ۲۸۳-۲۹۸.
- Roberts, H. E. ۲۰۰۱. A picture is worth a thousand words: art indexing in electronic databases. *Journal of the American Society for Information Science and Technology* ۵۲ (۱۱): ۹۱۱- ۹۱۶.
- Rui, Y., Huang, T. S., and S. Mehrotra. ۱۹۹۸. Relevance feedback techniques in interactive content-based Imagere trieval. In *Storage and Retrieval for Image and Video Databases VI*, Ishwar K. Sethi and Ramesh C. Jain eds., ۲۵-۲۶. SPIE Proceedings.
- Sciaroff, S., and L. Liu. ۲۰۰۱. Deformable shape detection and description via model-based region grouping. *IEEE Transactions on Pattern Analysis and Machi Intelligence* ۲۳ (۵): ۴۷۵-۷۸۹.
- Smith, J. R. ۲۰۰۱. Quantitative assessment of image retrieval effectiveness. *Journal of the American Society for Information Science and Technology* ۵۲ (۱۱): ۹۶۹-۹۷۹.

# Image Semantic Retrieval Challenges and Thesauri Modern Applications

**Farahnaz Sadoughi<sup>۱</sup>**

Associate Professor of Health Information Management  
Department, School of Health Management and Information  
Sciences of Tehran University of Medical Sciences

**Ali Valinejadi\***

PhD student of Health Information Management Department,  
School of Health Management and Information Sciences, and  
Exceptional Talent Center of Tehran University of Medical  
Sciences

**Hafez Mohammadhassanzadeh<sup>۲</sup>**

MS in Medical Library and Information Science, Tabriz  
University of Medical Sciences

**Hamid Bouraghi<sup>۳</sup>**

Lecturer of LIS Department, Hamadan University of Medical  
Sciences

**Parisa Pasyar<sup>۴</sup>**

MLIS in National Library of Iran

Iranian Journal of  
**Information  
Processing &  
Management**

**Abstract:** Today, images are used in almost every career, and a different range of professional groups use images in their work. They have always been used as a means to convey information. Images are very often used either alone or in compound with text, as they are able to conveying information difficult to depict in words. However, image retrieval based on syntactic features alone does not provide support for various types of queries. There is a mismatch between what the current solutions for this are able to providing, and what we demand from an image retrieval system. This gap between what is possible and what we require, by many authors has been named the "Semantic gap". The focus of this paper was how to obtain successful semantic image retrieval from large image collections. The research investigated some of the difficulties of getting correlate results when searching in image collections. A new tool, the shape thesaurus, was suggested as a step into improving image retrieval. The principal motivation behind this suggestion has been to assess the possibilities of narrowing the Semantic gap.

**Keywords** Image retrieval, image indexing, image semantic retrieval, semantic gap, shape thesaurus

Iranian Research Institute  
For Science and Technology

ISSN ۲۲۵۱-۸۲۲۳

eISSN ۲۲۵۱-۸۲۳۱

Indexed in LISA, SCOPUS & ISC

Vol. ۲۲ | No. ۳ | pp: ۶۴۱-۶۶۶

spring ۲۰۱۲

<sup>۱</sup>. [sadoughi.f@gmail.com](mailto:sadoughi.f@gmail.com)

\* **Corresponding Author:** [a\\_valinejad@razi.tums.ac.ir](mailto:a_valinejad@razi.tums.ac.ir)

<sup>۲</sup>. [hafezhassanzadeh۲@gmail.com](mailto:hafezhassanzadeh۲@gmail.com)

<sup>۳</sup>. [hamid.bouraghi@gmail.com](mailto:hamid.bouraghi@gmail.com)

<sup>۴</sup>. [p.pasyar@gmail.com](mailto:p.pasyar@gmail.com)