

مدلی به منظور بازیابی تصاویر مبتنی بر چند درخواست

مریم تقی‌زاده^۱، دانشجوی دکتری؛ عبدالله چاله‌چاله^۲، استادیار

۱- دانشکده مهندسی - گروه کامپیوتر - دانشگاه رازی - کرمانشاه - ایران - taghizadehmail@gmail.com

۲- دانشکده مهندسی - گروه کامپیوتر - دانشگاه رازی - کرمانشاه - ایران - chalechale@razi.ac.ir

چکیده: جستجوی مبتنی بر تصویر در پایگاه‌داده‌های تصویری، یکی از موضوعات کاربردی در بینایی ماشین می‌باشد. جستجو بر اساس یک یا چند درخواست تصویری^۱ با محتویات یکسان در چند دهه گذشته بسیار مرسوم بوده اما بازیابی با تکیه بر چند درخواست تصویری با محتویات متفاوت چندان مورد توجه نبوده و تحقیقات چندانی در این زمینه انجام نگرفته است. در این مقاله برای بازیابی تصویر مبتنی بر چند درخواست متفاوت یک چارچوب جدید مبتنی بر ترکیب محتوای تصاویر پیشنهاد شده و بر پایه این روش، چندین عملگر جدید جستجو معرفی شده است. روش پیشنهادی شامل چهار مرحله است: در مرحله اول برای هر یک از درخواست‌های تصویری صادر شده یک بازیابی بر مبنای ویژگی‌های سطح پایین انجام می‌شود و از نتایج این بازیابی یک بردار باینری متناظر با مؤلفه‌های تصویر به صورت خودکار به دست می‌آید. این بردار متناظر با نواحی و اشیاء برجسته درون تصویر است. سپس بردارهای هر درخواست تصویری با یک عملگر منطقی ترکیب و یک بردار جدید حاصل می‌شود. در مرحله سوم و چهارم به ترتیب عملیات بازیابی و نمایش نتایج بر اساس همین بردار جدید انجام می‌گردد. از ویژگی‌های برجسته این شیوه ارائه یک روش کارا برای ترکیب تصاویر به منظور انجام بازیابی بر اساس چند درخواست تصویری با محتویات متفاوت بوده که کاهش محاسبات پیچیده و زمان‌گیر و پیاده‌سازی آسان عملگرهای جستجو را فراهم می‌نماید. این عملگرها مبتنی بر تئوری مجموعه‌ها و در قالب عملگرهای منطقی پیشنهاد و پیاده‌سازی شده‌اند. نتایج به دست آمده از آزمایش‌های تجربی کارایی روش پیشنهادی را اثبات می‌کند. در مقایسه با روش‌های پیشین مقدار کارایی ۲۷٪ افزایش داشته است.

واژه‌های کلیدی: بازیابی تصویر، چند درخواست تصویری، عملگر جستجو، عملگر منطقی، ویژگی سطح پایین.

A Model to Image Retrieval Based on Multiple-Query

M. Taghizadeh¹, PhD Student; A. Chalechale², Assistant Professor

1- Faculty of Engineering, Computer Engineering Department, University of Razi, Kermanshah, Iran, Email: taghizadehmail@gmail.com

2- Faculty of Engineering, Computer Engineering Department, University of Razi, Kermanshah, Iran, Email: Chalechale@razi.ac.ir

Abstract: Image-based search is one of the most applied issues in computer vision. In the past decades, content based image retrieval (CBIR) systems using one query or multiple queries with the same semantic information were well investigated. In contrast, the CBIR systems based on multiple queries with different semantics have rarely considered in the literature. This article presents a novel framework in order to combine the content of multiple queries for image retrieval. The proposed framework introduces several new search operators. The proposed approach comprises four stages: in the first stage, for each issued query, a retrieval process based on low-level features is applied and according to the results, a binary component vector is automatically computed for each issued query. This vector corresponds to regions and salient objects inside the query image. The vectors are then combined by logical operators to generate one binary component vector. In the third and fourth stages, retrieving and showing results are performed based on the binary component vector, respectively. The remarkable advantages of the proposed method are to efficiently combine different images for retrieval purposes; also it decreases complicated and time-consuming computations as well as simplifying the implementation of several search operators. The operators correspond to the set theory and known logical operators. The experimental results distinctly exhibit the effectiveness of the proposed method. In comparison to the previous methods, it improves 27% of the performance achievement.

Keywords: Image retrieval, multiple-query, search operator, logical operator, low level feature.

تاریخ ارسال مقاله: ۱۳۹۵/۰۲/۱۰

تاریخ اصلاح مقاله: ۱۳۹۵/۰۴/۱۴

تاریخ پذیرش مقاله: ۱۳۹۵/۰۶/۲۱

نام نویسنده مسئول: عبدالله چاله‌چاله

نشانی نویسنده مسئول: ایران - کرمانشاه - باغ ابریشم - دانشگاه رازی - دانشکده مهندسی - گروه کامپیوتر.

۱- مقدمه

جدی در این نوع جستجو معرفی کرد. در این روش مقدار فاصله هر درخواست با تمام تصاویر پایگاه داده به طور جداگانه محاسبه شده و در نهایت مجموعه زیادی داده معادل با هر درخواست به دست می آید. حال با به کارگیری شیوه های بهینه سازی، مجموعه زیادی تصویر در خروجی نتیجه می شود و تصاویری که در میانه این مجموعه قرار دارند نتیجه نهایی مسئله می باشند و محتویات تصاویر خروجی به همه درخواستها مرتبط است. با توجه به اینکه این روش، حجم محاسبات را تا حدودی کاهش داده است دو ایراد را می توان به آن وارد کرد. اول اینکه، تصاویری که در ابتدا در خروجی نمایش داده می شوند فقط به یک درخواست مرتبط است و با درخواستهای دیگر تا حدودی غیر مرتبط می باشد. باید تعداد زیادی تصویر به دست آید تا در میانه آنها خروجی های مورد نظر کاربر مشاهده شود. دوم اینکه با افزایش تعداد درخواستها، محاسبات به فضاهای با ابعاد بزرگ تر تغییر یافته و حجم محاسبات به طور چشم گیری افزایش می یابد. از این رو، به دنبال روش نوینی هستیم که مسئله بازیابی را با کم کردن پیچیدگی محاسبات و همچنین داشتن کارایی مناسب پاسخ دهد.

نکته دیگری که این پژوهش به دنبال آن هست طراحی یک سامانه هوشمند با قابلیت زیر است: یک سامانه هوشمندی را فرض نمایید که قادر به پیدا کردن خودکار تصاویر در یک آلبومی از تصاویر باشد به گونه ای که تصاویری را فراهم نماید که شامل شیء «گره» باشد و شیء «سگ» را شامل نشود یا به دنبال تصاویری باشد که دارای میز و صندلی باشد [۱۰]. به طور کلی این مقاله به دنبال طراحی روشی هست که قادر به بازیابی تصاویر بر پایه بیش از یک درخواست با محتوای متفاوت باشد و هدف جستجوی ترکیبی است. روشی که در این مقاله پیشنهاد داده شده علاوه بر بازیابی مورد نظر، قادر به انجام عملگرهای جستجوی متفاوتی است که بر پایه اطلاعات کنونی می توان گفت اولین بار است که چنین تحقیقی انجام شده است. با توجه به نوع عملگر جستجوی پیشنهادی؛ تصویر نهایی بر اساس نوع عملگر، مرتبط با تمام درخواستها خواهد بود. در مجموع ۳ نوع عملگر مجزا پیشنهاد شده است. روش پیشنهادی برای بازیابی تصویر شامل چهار مرحله مجزا است. اولین مرحله شامل مجموعه درخواستهای تصویری است که بازیابی بر اساس آنها انجام می شود. در این مرحله نواحی و اشیاء تصویر محاسبه شده، سپس در مرحله دوم همه درخواستها ترکیب شده تا یک درخواست به دست آید. روش ترکیب کردن این درخواستها به صورت عملگرهای منطقی و مبتنی بر نوع جستجو پیاده سازی شده است. در مرحله سوم، بازیابی بر اساس فقط یک درخواست نتیجه شده انجام می گیرد و تعدادی تصویر نهایی به عنوان نتیجه بازیابی نمایش داده می شوند. در این تحقیق برای پیاده سازی عملگرهای مختلف، هر تصویر را بر اساس یک بردار باینری مبتنی بر مؤلفه های تصویر نمایش داده و توصیف می کنیم. این بردار «بردار مؤلفه باینری»^۶ نامگذاری شده است. برای محاسبه این بردار باید مؤلفه های تصویر را تعیین نموده که بدین منظور از شیوه های استخراج

با توجه به رشد قابل توجه فن آوری های تصویر برداری، فراگیر بودن ابزارهای مرتبط با تصویر برداری در زندگی انسان و همچنین فراهم بودن حافظه های با ظرفیت و سرعت بالا، جستجوی تصاویر را به یک مسئله اولیه و حیاتی تبدیل کرده است. یکی از راه حل های اولیه برای جستجوی تصویر، بازیابی بر پایه متن (TBIR)^۲ بوده است. در این روش هر تصویر دارای کلمات کلیدی متنی است که آن تصویر را توصیف می کند. سپس برای جستجو، کاربران با به کارگیری این کلمات، عملیات بازیابی را انجام می دهند. یکی از معایب این روش، زمان بردن حاشیه نویسی تصاویر است و ممکن است توصیف صحیحی توسط کاربر برای تصویر وجود نداشته باشد و یا اینکه کاربران متعدد توصیف های متفاوتی را برای یک تصویر ارائه دهند. اما از امتیازات این شیوه می توان به سادگی عملگر جستجو در آن اشاره کرد. برای رفع معایب این شیوه، بازیابی مبتنی بر محتویات تصویر (CBIR)^۳ پیشنهاد داده شد که برای توصیف تصویر از ویژگی های سطح پایین مانند شکل، بافت و رنگ استفاده می کند. روش های متعددی برای بازیابی تصویر مبتنی بر محتویات و بهبود کارایی آن پیشنهاد شده است. برای نمونه یک روش بازیابی بر پایه ترکیب ویژگی رنگ و بافت تصویر پیشنهاد شده است [۱]. در تحقیق دیگری یک روش بازیابی مطابق با ویژگی شکل و هیستوگرام رنگ در فضای YCbCr معرفی گردیده است [۲]. همچنین در [۳] یک مطالعه مروری نسبتاً فراگیر در زمینه سامانه های CBIR انجام شده است. البته هنوز فاصله زیادی میان توصیف معنایی تصویر و ویژگی های سطح پایین آن وجود دارد. محققان برای حداقل کردن این فاصله و افزایش کارایی نیز روش هایی را پیشنهاد کرده اند. برای مثال، بعد از انجام بازیابی، کاربر نتایج به دست آمده را امتیاز می دهد و بر اساس نظر کاربر تعدادی از تصاویر با امتیاز بالا انتخاب شده و برای انجام دوباره بازیابی به سامانه جستجو اعمال می شوند. بنابراین نتایج بازیابی مجدد دقت بیشتری را به دنبال خواهد داشت [۴-۵]. این روش با نام «بازخورد مرتبط»^۴ شناخته می شود. یکی دیگر از پیشنهادها در این زمینه به کارگیری چند تصویر برای جستجو است که به طور یقین افزایش کارایی را در نتایج بازیابی خواهد داشت [۶-۸]. در این شرایط هر کدام از تصاویر جنبه های متفاوتی از شیء یا تصویر مورد جستجو را در نظر گرفته و با اعمال تمام درخواستها نتایج با دقت بیشتری به دست خواهد آمد. در این روش معمولاً درخواستها دارای مفاهیم معنایی مشابه و یکسانی هستند.

مسئله متفاوتی که در این مقاله به آن پرداخته شده است بازیابی بر اساس ترکیبی از «چند درخواست تصویری»^۵ با محتویات متفاوت است به گونه ای که نتایج برگشتی ترکیبی از تمام درخواستهای مورد جستجو بوده و به همه درخواستها نزدیک و مرتبط باشند. این مسئله به ندرت مورد مطالعه قرار گرفته شده است. در [۹] مسئله جستجو بر پایه چند درخواست با محتویات متفاوت برای اولین بار در سال ۲۰۱۵ پیشنهاد شده است و می توان آن را به عنوان اولین کار

(QE)^۲ [۱۴-۱۲] تصویر درخواست اصلی با تعداد بیش‌تری از تصاویر که دارای رتبه بالایی هستند جایگزین شده و در حقیقت مسئله بازیابی با یک تصویر ساده به یک بازیابی با چند درخواست تبدیل خواهد شد. نتیجه آن افزایش کارایی بازیابی است اما به‌هرحال، این روش نیز مناسب با هدف این پژوهش نیست؛ زیرا در این روش همه درخواست‌ها به هم شبیه هستند.

در [۱۵] به‌منظور بازیابی یک شیء خاص در یک مجموعه تصاویر از قابلیت چند درخواست استفاده می‌کند. در واقع به‌دنبال تمام تصاویری هست که شامل شیء خاصی باشد و برای این کار از مجموعه‌ای درخواست استفاده می‌کند. در این روش نقطه شروع جستجو بر اساس متن قرار داده شده است. در ابتدا برای جستجوی شیء اولیه، از واژه‌های متنی استفاده می‌کند. برای مثال در موتور جستجوی گوگل تصویرهای مرتبط با شیء موردنظر را بازیابی کرده و کاربر به‌آسانی تعداد دلخواهی از تصویرهای مرتبط با شیء موردنظر خود را انتخاب می‌کند. سپس مجموعه انتخاب‌شده از نتایج این مرحله برای ادامه عمل بازیابی به‌کار گرفته می‌شوند. روش‌های مختلفی مانند Joint-Avg، MQ-Max و MQ-Avg معرفی شده است که به دو نوع متفاوت پیاده‌سازی شده‌اند. در پیاده‌سازی نوع اول، به‌صورت مستقل برای هر درخواست تعدادی تصویر فرض شده، سپس برپایه عملگر خطی مناسبی نتایج با یکدیگر ترکیب و یک درخواست جدید و ساده ایجاد می‌شود. سپس بازیابی نهایی بر اساس این درخواست به‌دست‌آمده انجام می‌گردد. در نوع دوم، بازیابی به‌صورت مستقل برای هر درخواست انجام شده و سپس نتایج با یک عملگر خطی مناسب ترکیب شده و تعداد دلخواهی از نتایج به‌عنوان تصاویر نهایی بازگردانده می‌شوند. به‌طورکلی این دو رویکرد متفاوت برای پیاده‌سازی سامانه‌های بازیابی مبتنی بر چند درخواست می‌باشد که مراحل آن‌ها در شکل ۱ ترسیم شده است.



شکل ۱: دو رویکرد مختلف بازیابی تصویر بر اساس چند درخواست

روش Joint-Avg، از نمایش «کیفی از کلمات»^۸ (BoW) برای همه تصاویر در مجموعه درخواست استفاده می‌کند و از همه آن‌ها میانگین می‌گیرد. نمایش BoW [۱۶] یکی از روش‌های فراگیر در بازیابی تصویر است و مبتنی بر ایده کوانتیزه کردن ویژگی‌های تصویر به خوشه‌هایی است که آن‌ها را کلمات دیداری می‌نامند. سپس از روش‌های استاندارد بازیابی اطلاعات که برای بازیابی متن پیشنهاد شده است، استفاده می‌کند. در این روش یک فرهنگ لغات از لغات

ویژگی سطح پایین استفاده می‌شود و به‌طور خودکار این کار انجام می‌گردد. این عمل برای تمام درخواست‌های موردنظر اجرا شده، سپس عملگرهای جستجو را در این مرحله اعمال کرده و یک بردار باینری متناظر با یک درخواست ساده با عملگرهای منطقی تولید می‌شود. حال این بردار مؤلفه‌ها به سامانه بازیابی تصویر اعمال شده و نتایج موردنظر به‌تعداد دلخواه نمایش داده می‌شود. با روش ارائه‌شده عملگرهای جستجوی ترکیبی به‌آسانی قابل پیاده‌سازی هستند و نتایج آزمایش‌ها کارایی روش را نشان می‌دهد.

در ادامه، مقاله بدین‌صورت سازمان‌دهی شده است: در بخش دوم، شیوه‌های مرتبط با CBIR مبتنی بر بیش از یک درخواست مرور شده‌اند. بخش بعدی رویکرد پیشنهادشده به همراه عملگرهای جستجو تشریح شده است. ارزیابی روش و نتایج به‌دست‌آمده به‌ترتیب در بخش چهارم و پنجم آورده شده‌اند. در پایان نتیجه‌گیری و کارهای آینده معرفی شده‌اند.

۲- مروری بر روش‌های بازیابی مبتنی بر بیش از یک درخواست

در سامانه‌های CBIR معمولاً روش‌های پیشنهادشده با چند درخواست به‌منظور افزایش کارایی و بهبود دقت بازیابی تصویر می‌باشد. در حقیقت همه درخواست‌ها دارای مفاهیم برابر بوده و شیء یکسانی در آن‌ها به‌کار گرفته شده است یا اشیاء متفاوتی از یک کلاس مشابه هستند. در [۷] یک سامانه بازیابی مبتنی بر محتویات تصویر بر اساس استخراج هیستوگرام رنگ برپایه چند درخواست معرفی شده است. هدف اصلی بر این پایه استوار است که هر تصویر به‌تنهایی قادر به نمایش تمام ویژگی‌های موردنظر نیست پس بهتر است که از مجموعه‌ای از تصاویر استفاده شود که ویژگی‌های اصلی و مهم را نمایان نموده و ویژگی‌های غیرمرتبط را نیز برجسته نماید. این تصاویر به‌عنوان درخواست «مثبت» و «منفی» در نظر گرفته می‌شوند. نتایج خروجی باید تاجایی که امکان دارد به مجموعه تصاویر مثبت نزدیک بوده و از تصاویر منفی دور باشد. برای این کار هیستوگرام هر دو تصویر محاسبه شده و بر اساس یک ماتریس فاصله، فاصله میان دو هیستوگرام محاسبه خواهد شد و برای انجام محاسبات از عملگر تفریق استفاده می‌شود. در حقیقت مؤلفه‌های رنگ غیرمرتبط را از تصاویر منفی مشخص کرده و به مجموعه تصاویر مثبت اعمال نموده و با عملگر تفریق یک درخواست ساده محاسبه و بازیابی بر اساس آن انجام می‌شود. در [۱۱] یک سامانه بازیابی بر اساس چند درخواست پیشنهاد شده است که در واقع، همان یک درخواست است و چندتایی بودن را به این دلیل به‌کار برده است که از دو هیستوگرام جداگانه بافت و رنگ برای هر درخواست استفاده نموده و اشتراک آن‌ها را محاسبه می‌کند. سپس میزان تشابه میان درخواست به‌دست‌آمده با تصاویر پایگاه‌داده اندازه‌گیری شده و نزدیک‌ترین تصاویر به درخواست به‌دست می‌آید. این کار نیز مرتبط با هدف این مقاله نیست. در روش بسط درخواست

در [۹] است که یک الگوریتم جدیدی برای بازیابی بر اساس محتویات تصویر پیشنهاد داده و هدف آن نیز پیدا کردن تصاویری است که مرتبط با تمام درخواست‌های صادر شده باشند. همچنین درخواست‌ها دارای معنای متفاوتی هستند. در این روش نشان داده شده است که ترکیب تصاویر بازیابی شده بر اساس بهینه‌سازی پاریتو بسیار بهتر از ترکیبات خطی جبری [۱۵] هستند. در این روش الگوریتم بهینه‌سازی «پاریتو اول»^۹ همراه با رتبه‌دهی چندتایی کارا (EMR)^{۱۱} [۱۸] به کار گرفته شده است. در حقیقت هر درخواست را صادر نموده و تمام تصاویر پایگاه داده را بر اساس غیرتشابهات آن‌ها با درخواست صادر شده رتبه‌بندی می‌نماید و این عمل به‌طور مستقل برای هر درخواست انجام خواهد شد. سپس الگوریتم EMR بر روی تمام درخواست‌های رتبه‌بندی شده اجرا می‌گردد. از EMR نقاط پاریتو ایجاد خواهد شد و در هر تکرار اجرا نقاط غیرغالب مرتب شده تا آن جایی که تعداد نمونه‌های کافی به کاربر برگشت داده شوند. حال نقاط وسط پاریتو برای بازیابی مبتنی بر چند درخواست از اهمیت زیادی برخوردار است و نشان‌دهنده نتایج خروجی است. نتایج پایانی این ایده در مقایسه با روش‌های ترکیب خطی جبری [۱۵] کارایی بهتری را پیشنهاد می‌دهد. اما محاسبات بالایی را به سامانه تحمیل می‌نماید. به‌ویژه اگر تعداد درخواست‌ها به بیش‌تر از ۲ افزایش داده شود، حجم محاسبات تحمیل شده به سامانه افزایش می‌یابد. این مقاله به دنبال روشی هست که در مقایسه با راه‌حل‌های پیشین کارایی خوبی را ارائه نموده و همچنین محاسبات را کاهش دهد. در روش پیشنهادی این مقاله، بعد از انجام یک مرحله بازیابی مبتنی بر استخراج ویژگی‌های سطح پایین، درخواست تصویری با یک بردار باینری مبتنی بر مؤلفه‌های تصویر که از اشیاء و نواحی تصویر حاصل شده، نمایش داده می‌شود. سپس برای دیگر درخواست‌ها نیز بردار باینری مؤلفه محاسبه شده و بازیابی نهایی بر اساس ترکیب منطقی این بردارهای باینری انجام می‌شود. از مزایای این روش سادگی آن و محاسبات ساده منطقی را می‌توان ذکر کرد. همچنین طراحی و پیاده‌سازی عملگرهای کاربردی جدید جستجو در چارچوب ارائه شده از دیگر قابلیت‌های جذاب این تحقیق است. به‌طور یقین پیاده‌سازی این‌گونه عملگرهای جستجو در چارچوب‌های دیگر می‌تواند پیچیده و هزینه محاسباتی بالایی داشته باشد.

۳- روش پیشنهادی

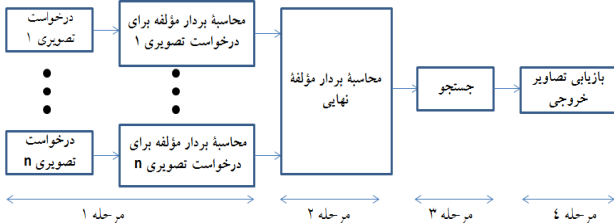
مفهوم مرتبط با مسئله بازیابی که هدف اصلی این مقاله می‌باشد در شکل ۲ به تصویر کشیده شده است. بر اساس این شکل هر درخواست دارای تعدادی شیء و ناحیه است. در عملیات جستجو به دنبال بازیابی تصاویری هستیم که شامل همه مؤلفه‌ها از همه درخواست‌ها باشد. در حقیقت تصویر خروجی باید به همه درخواست‌ها مرتبط باشد. تفاوتی که مسئله تعریف شده در این مقاله با [۹] دارد این است که همه اشیاء

دیداری به دست آمده و هر توصیف‌گر به نزدیک‌ترین کلمه دیداری‌اش تخصیص داده می‌شود. برای تولید فرهنگ لغات می‌توان از روش‌های k-means و روش‌های مشابه با آن بهره برد. در Joint-Avg بعد از محاسبه میانگین از تمام درخواست‌ها یک درخواست ساده به دست می‌آید و بر اساس آن بازیابی انجام می‌شود. در این الگوریتم نتایج به دست آمده معمولاً به یک درخواست بسیار نزدیک بوده و به درخواست دیگر ممکن است مرتبط نباشد. از این رو برای مسئله مطرح شده در این مقاله این راه‌حل نمی‌تواند پاسخ‌گو باشد. دو روش MQ-Avg و MQ-Max مشابه هم هستند فقط نوع عملگر به کار گرفته برای ترکیب نتایج متفاوت است. در روش MQ-Max برای هر درخواست تعدادی نتیجه رتبه‌بندی شده به صورت مستقل محاسبه و نمایش داده می‌شود. سپس نتایج از طریق امتیازدهی هر تصویر بر اساس ماکزیمم امتیازات به دست آمده از هر درخواست ترکیب می‌شوند. در حقیقت تصاویر بر اساس ماکزیمم رتبه آن‌ها در کنار هم به عنوان خروجی قرار خواهند گرفت. در روش بعدی، میانگین تمام امتیازات به دست آمده از هر درخواست را محاسبه نموده و از این طریق آن‌ها را باهم ترکیب می‌کند. در این دو راه‌حل نیز نتایج به یکی از درخواست‌ها بیش‌تر شبیه و نزدیک خواهد شد. از این رو برای مسئله تعریف شده در این مقاله نمی‌توانند راه‌حل کامل و مؤثری باشند. البته با توجه به برجسته بودن این سه روش، در کارهای پیشنهاد شده اخیر در بازیابی تصویر، این روش‌ها [۱۵] مبنای ارزیابی و مقایسه پژوهش‌های جدید قرار گرفته است. بنابراین در این مقاله روش پیشنهادی با این سه روش مقایسه خواهد شد.

در پژوهش ارائه شده در [۱۷] از چند درخواست برای بازیابی تصویر استفاده نموده و مشابه [۱۵] است. در آن از واکاوی الگوی بدون ناظر برای شناسایی شیء درون تصویر استفاده می‌کند. در ابتدا کیفی از کلمات محلی از هر همسایگی کوچک در اطراف هر نقطه کلیدی ایجاد می‌نماید تا اطلاعات ساختاری و مکانی محلی را استخراج کند. با استفاده از این هیستوگرام‌های محلی، تصاویر درخواست را واکاوی می‌نماید تا یک مجموعه از الگوهای سطح میانی را کشف نماید که تصاویر درخواست را بهتر توضیح می‌دهند. این الگوریتم روش قدرتمندی برای بازیابی تصویر است و کارایی خوبی دارد. این روش نیز با هدف این مقاله در یک راستا نیست زیرا بر اساس یک شیء خاص است.

به‌طور کلی تمام راه‌حل‌هایی که در اینجا مرور شدند، مناسب مسئله تعریف شده در این مقاله نیستند. در حقیقت تصاویری که در نهایت در خروجی مشاهده می‌شوند، معمولاً بخشی از آن‌ها فقط به یکی از درخواست‌ها مرتبط است و دیگر تصاویر خروجی نیز به همین صورت فقط به درخواست‌های دیگر مرتبط خواهند بود. اولین تحقیق در ارتباط با مطالعه مورد نظر در این مقاله، پژوهش ارائه شده

و نواحی برجسته یک تصویر برای نمایش و توصیف تصویر به‌کار گرفته می‌شود. با توجه به اینکه هر تصویر معمولاً از اشیاء و نواحی متعددی تشکیل شده است، برای هر تصویر تعدادی مؤلفه مجزا در نظر گرفته می‌شود که مؤلفه‌ها متناظر با نواحی و اشیاء درون تصویر هستند. برای مثال یک تصویر می‌تواند دارای مؤلفه‌های «انسان»، «رود» و «کوه» باشد یا تصویر دیگر دارای مؤلفه‌های «اسب»، «چمن‌زار»، «گل» و «درخت» باشد.



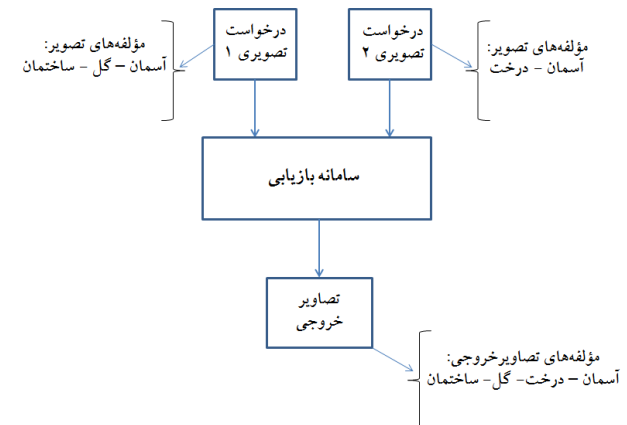
شکل ۳: مراحل الگوریتم پیشنهادی بازیابی تصاویر مبتنی بر چند درخواست

در این مقاله بر اساس ویژگی‌های سطح پایین تصویر، سعی می‌شود بردار مؤلفه مناسبی برای هر درخواست محاسبه شود و این عمل به‌صورت خودکار انجام می‌شود. برای ارزیابی بهتر کار از چهار روش بسیار ساده برای استخراج ویژگی‌های سطح پایین تصویر استفاده نموده و بردار باینری مؤلفه‌های تصویر محاسبه می‌شود. این روش‌ها شامل محاسبه هیستوگرام شدت، لبه‌های تصویر، بافت تصویر بر پایه عملگر LBP [۱۹] و هیستوگرام رنگ هستند. در این مرحله برای پیاده‌سازی، از اطلاعات محلی تصویر استفاده شده است. در هر چهار روش، تصویر به ۱۶ بلوک ثابت تقسیم شده و از هر بلوک ویژگی تصویر استخراج شده، سپس از کنار هم قرار دادن این مقادیر در کنار هم بردار ویژگی نهایی تصویر به‌دست می‌آید. این ویژگی‌ها بر اساس چهار روش اشاره شده در بالا استخراج می‌شوند. در الگوریتم اول هیستوگرام مقادیر برای هر بلوک محاسبه می‌شود. در الگوریتم دوم برای بردار ویژگی از اطلاعات لبه در هر بلوک محلی تصویر استفاده می‌شود. برای محاسبه لبه روش سوبل اعمال شده است. در الگوریتم سوم، در هر بلوک از تصویر عملگر بافت LBP به‌کار گرفته شده است. سپس هیستوگرام مقادیر نتیجه‌شده از اعمال LBP محاسبه و در نهایت بردار ویژگی تصویر تعیین شده است. در الگوریتم چهارم، هیستوگرام رنگ در هر سه مؤلفه RGB برای هر بلوک محاسبه و از کنار هم قرار دادن آن‌ها یک بردار ویژگی استخراج شده است. البته می‌توان از هر نوع الگوریتم استخراج ویژگی استفاده نمود و به‌طور یقین هرچه میزان کارایی این الگوریتم بالاتر باشد بردار مؤلفه باینری محاسبه‌شده دقیق‌تر خواهد بود. البته این تحقیق بر روی نوع الگوریتم استخراج ویژگی و انتخاب ویژگی تاکید نکرده و جزء تمرکز اصلی این پژوهش نبوده است.

همان‌گونه که اشاره شد، در این پیاده‌سازی هدف اولیه محاسبه خودکار بردار باینری مؤلفه تصاویر است و در این روش از رویکرد اول طبق شکل ۱ استفاده شده است. برای محاسبه این بردار برای هر درخواست تصویری به‌صورت مستقل یک مرحله بازیابی مبتنی بر بردار ویژگی‌های استخراج شده انجام می‌گردد و تعدادی تصویر

و نواحی برجسته یک تصویر برای نمایش و توصیف تصویر به‌کار گرفته می‌شود.

با توجه به اینکه هر تصویر معمولاً از اشیاء و نواحی متعددی تشکیل شده است، برای هر تصویر تعدادی مؤلفه مجزا در نظر گرفته می‌شود که مؤلفه‌ها متناظر با نواحی و اشیاء درون تصویر هستند. برای مثال یک تصویر می‌تواند دارای مؤلفه‌های «انسان»، «رود» و «کوه» باشد یا تصویر دیگر دارای مؤلفه‌های «اسب»، «چمن‌زار»، «گل» و «درخت» باشد.



شکل ۲: مفهوم پیشنهاد بازیابی تصویر بر پایه چند درخواست

در ابتدا برای کل مجموعه تصاویر پایگاه‌داده یک بردار باینری که هر درایه آن متناظر با یک مؤلفه است در نظر گرفته می‌شود. در صورتی که در تصویر مؤلفه خاصی وجود داشته باشد متناظر با آن مؤلفه در بردار باینری مؤلفه مقدار «یک» و در غیر این صورت مقدار «صفر» ذخیره می‌شود. برای مثال اگر تصاویر مورد بررسی در مجموع حداکثر دارای n مؤلفه زیر باشند: $\{x_1, x_2, \dots, x_n\}$ برای هر تصویر یک بردار باینری n تایی با مقادیر صفر و یک به‌صورت $L = [0, 1]^n$ تعریف می‌شود. مقدار هر درایه L_i در صورتی یک است که مؤلفه x_i در تصویر وجود داشته باشد و هر تصویر می‌تواند همزمان بیش از یک مؤلفه داشته باشد.

الگوریتم پیشنهادی شامل چهار مرحله است: در مرحله اول بردار باینری مؤلفه‌های تصویر برای هر درخواست محاسبه می‌شود. سپس در مرحله دوم، بردارهای باینری همه درخواست‌ها بر اساس نوع عملگر جستجو ترکیب می‌شوند البته در اجرای این روش از دو تا درخواست استفاده کرده‌ایم اما می‌تواند برای تعداد بیش‌تری درخواست نیز به‌کار گرفته شود. بردار نهایی به‌دست‌آمده را به سامانه جستجو در مرحله بعدی داده و بر پایه یکی از روش‌های اندازه‌گیری تشابه، تعدادی تصویر را به‌عنوان نتیجه بازیابی ارائه می‌کند. در روش‌های بازیابی مبتنی بر تصویر معمولاً از فاصله اقلیدسی، کانبرا یا مانه‌امتان برای اندازه‌گیری تشابهات استفاده می‌شود. یکی از مشخصات این روش‌های اندازه‌گیری، محاسبات سنگین و حجیم می‌باشد. حال الگوریتم پیشنهادی در این پژوهش، دارای این قابلیت است که از عملگرهای منطقی بسیار ساده با

۴- ارزیابی

برای ارزیابی روش‌های بازیابی تصویر، معیارهای مختلفی وجود دارد که در این مقاله از «رشد تجمعی تخفیف داده‌شده» (DCG) استفاده شده است، که از معیارهای استاندارد در جامعه بازیابی هست [۲۰]. پارامتر DCG میزان ارتباط تصویر بازگردانده‌شده را با درخواست اندازه‌گیری می‌کند. در حقیقت این پارامتر یک اندازه‌گیری از کیفیت رتبه‌بندی است و در حوزه‌هایی که دارای ارتباطی چندسطحی است، کاربرد دارد. بنابراین مفاهیمی مانند «دقت»^{۱۲} و «فراخوانی»^{۱۳} که در بازیابی در سطح باینری کاربرد دارند در این نوع بازیابی به‌طور صریح قابل استفاده نیستند.

در سامانه‌های بازیابی مبتنی بر یک درخواست، امتیاز محاسبه‌شده برای هر تصویر خروجی یک عدد باینری است که مقدار «۱» یا «۰» را دارد. مقدار یک مرتبط بودن تصویر خروجی با درخواست مورد جستجو را نشان می‌دهد در غیراین‌صورت تصویر غیرمرتبط خواهد بود. اما برای سامانه چند درخواست که هر تصویر می‌تواند مؤلفه‌های متفاوتی داشته باشد، امتیازدهی باینری نمی‌تواند مناسب باشد، از این‌رو باید از یک معیار جدید برای امتیازدهی به تصاویر خروجی استفاده شود. این معیار مشابه با پارامتر MQUR^{۱۴} که در [۹] معرفی شده است، می‌باشد که بر اساس نیاز ما در این مسئله تغییراتی به آن اعمال شده است. این پارامتر کسری از مؤلفه‌های تصویر درخواست را که به‌وسیله مؤلفه‌های تصویر بازیابی‌شده پوشانده می‌شوند را به‌طور منحصر به‌فرد اندازه‌گیری می‌کند. روابط (۱) و (۲) روش محاسبه آن را بیان می‌کند.

$$\beta = y^1 \vee y^2 \vee \dots \vee y^n \quad (1)$$

$$MQUR_i(X) = \frac{|l_i \wedge \beta|}{|\beta|}, \quad \text{for } i = 1, 2, \dots, k \quad (2)$$

مقدار y^1, y^2, \dots و y^n به‌ترتیب بردارهای باینری مؤلفه برای درخواست‌های مورد جستجو $\{q_1, q_2, \dots, q_n\}$ هستند و مقادیر بردار l_i معادل بردار باینری مؤلفه‌های تصویر بازیابی شده است. عملگرهای « \vee » و « \wedge » به‌ترتیب معادل عملگرهای OR و AND هستند. اندازه $|l_i|$ تعداد درایه‌های غیرصفر بردار l_i را نشان می‌دهد. مقدار k تعداد بازیابی‌های مرحله پایانی الگوریتم پیشنهادی است.

برای هر جفت تصاویری که باید تصاویر مرتبط با آن‌ها بازیابی شوند، بردارهای باینری مؤلفه‌های آن‌ها و سپس بعد از پایان بازیابی تصویر در مرحله نهایی مقدار MQUR محاسبه می‌شوند. همان‌گونه که در بالا ذکر شد؛ برای ارزیابی کارایی الگوریتم بازیابی چند درخواست از «رشد تجمعی تخفیف داده‌شده» (DCG) استفاده می‌شود که یکی از استانداردها در بازیابی است. پارامتر DCG ارتباط تصویر بازگردانده‌شده را با درخواست موردنظر اندازه‌گیری می‌کند. این پارامتر بر اساس رابطه (۳) محاسبه می‌شود. عبارت rel_i مقدار نمره‌بندی‌شده نتیجه در مکان i ام است، که در این الگوریتم همان مقدار MQUR است. مقدار

بازیابی می‌شود. برای اندازه‌گیری شباهت تصاویر، از فاصله اقلیدسی استفاده نموده و تعداد دلخواهی که کم‌ترین فاصله را داشته باشند به‌عنوان نتیجه بخشی از مرحله اول در نظر گرفته می‌شود. این تصاویر بازیابی‌شده هر کدام دارای بردار باینری هستند که از قبل به‌صورت غیربرخط مشخص شده است. حال از تصاویر بازیابی‌شده، بردارهای مؤلفه‌های باینری آن‌ها را بررسی کرده و از ترکیب آن‌ها یک بردار نهایی برای هر درخواست تصویری صادرشده به‌دست می‌آید. بنابراین برای هر درخواست یک بردار باینری متناظر وجود دارد. این بردارهای به‌دست‌آمده را با یک عملگر منطقی مناسب ترکیب کرده و فقط یک بردار که متناظر با تمام درخواست‌های منتشر شده است، به‌دست می‌آید. در مرحله بعدی این بردار به سامانه جستجو اعمال شده و حداقل فاصله این بردار را با بردارهای مجموعه پایگاه‌داده محاسبه کرده و کم‌ترین مقادیر به‌عنوان تصاویر مشابه در نظر گرفته می‌شود. برای اندازه‌گیری این فاصله از عملگر منطقی XOR و شمارش تعداد بیت‌های فعال بردار نهایی مؤلفه‌های باینری استفاده می‌شود.

در مجموع با این روش سه نوع عملگر جدید جستجو می‌توان معرفی کرد. اولین جستجو مبتنی بر ترکیب هر دو درخواست است. در حقیقت نتیجه بازیابی‌شده باید مرتبط به تمام مؤلفه‌های هر دو تصویر باشد. بدین منظور عملگر OR پیشنهاد شده است. عملگر دیگر جستجو برپایه داشتن مؤلفه‌های تصویر اول و حذف مؤلفه‌های تصویر دوم خواهد بود. در حقیقت عملگری باید باشد که مؤلفه‌هایی از تصویر اول که در تصویر دیگر وجود ندارد را در نظر بگیرد. عملگر بعدی برپایه داشتن مؤلفه‌هایی است که در هر دو تصویر مشترک باشد. جدول ۱ عملگرها و شیوه پیاده‌سازی آن‌ها را نشان می‌دهد.

جدول ۱: عملگرهای جستجوی پیشنهادی در روش ارائه‌شده

عملگر جستجو	نام عملگر	پیاده‌سازی
ترکیب کامل هر دو درخواست تصویری	$A + B$	$A \text{ or } B$
ترکیب مؤلفه‌هایی که فقط در یک درخواست تصویری وجود دارد	$A - B$	$A \text{ and } (A \text{ xor } B)$
ترکیب فقط مؤلفه‌هایی که در هر دو درخواست تصویری مشترک است	$A \cdot B$	$A \text{ and } B$

همان‌گونه که از جدول ۱ مشاهده می‌شود برای پیاده‌سازی انواع عملگرهای جستجوی پیشنهادی روابط ساده منطقی استفاده شده است. از مزایای این روش سادگی محاسبات و کاهش هزینه آن است در نتیجه به‌راحتی می‌توان سخت‌افزار مناسب و بسیار سریعی برای آن طراحی کرد. در این مقاله روش جستجوی منطبق بر ترکیب کامل هر دو درخواست پیاده‌سازی شده است.

جدول ۳: گروه‌های متفاوت تعریف‌شده در مجموعه پایگاه‌داده مورد آزمایش

Human	Beach	River	Sky	Building	Mountain	Trees
Bus	Dinosaur	Elephant	Lea	Flower	Horse	Food

مجموعه مشتمل بر ۶۱۶ تصویر به‌عنوان گروه «آموزش» در نظر گرفته‌شده و هر تصویر دارای یک بردار باینری مؤلفه ۱۴ تایی است. سپس در برنامه دیگری تصاویری را به‌عنوان درخواست مورد «آزمایش»، بررسی کرده و بر اساس ویژگی‌های سطح پایین استخراج شده از هر دو درخواست؛ یک مرحله بازیابی به‌صورت مستقل برای درخواست‌ها انجام شده و از ترکیب بردارهای مؤلفه تصاویر بازیابی شده، بردار مؤلفه هر دو تصویر محاسبه و ترکیب می‌شوند و یک درخواست ساده که شامل یک بردار مؤلفه است به‌دست می‌آید. سپس تعدادی از تصاویری که در پایگاه‌داده مشابه با این درخواست هستند بازیابی شده و اندازه‌گیری بر اساس تشابه با بردار باینری انجام می‌شود. در حقیقت بازیابی نهایی بر اساس بردار باینری مؤلفه‌های تصویر انجام می‌گردد. در نهایت معیارهای تعریف‌شده برای ارزیابی رویکرد پیشنهادی اندازه‌گیری می‌شوند.

همان‌گونه که در بخش‌های پیشین این مقاله اشاره شد برای استخراج ویژگی از ۴ روش متفاوت و ساده استفاده شده است. حال برای تعیین بردار مؤلفه‌ها برای درخواست‌های صادرشده در مرحله اول، این چهار روش متفاوت اجرا و نتیجه هرکدام بررسی می‌شوند. همچنین یک روش پایه نیز در نظر گرفته شده است. در این حالت خود کاربر به‌طور مستقیم برای درخواست‌های صادرشده در اولین مرحله، بردار مؤلفه تعریف می‌کند و ادامه مراحل بازیابی بر اساس همین نظر کاربر تکمیل می‌شود. دیگر روش‌های استخراج ویژگی با این روش پایه مقایسه شده‌اند؛ زیرا روش پایه بر اساس نظر کاربر می‌تواند تا حدودی به‌عنوان یک حالت بهینه در نظر گرفته شود. در جدول ۴ روش‌های ارائه شده بر اساس استخراج ویژگی به همراه طول بردار ویژگی به‌طور خلاصه بیان شده است. حداکثر طول بردار ویژگی مربوط به محاسبه هیستوگرام رنگ است که دارای ۳ کانال رنگ جداگانه می‌باشد و به‌طور یقین طول بردار آن در مقایسه با ویژگی‌های دیگر افزایش خواهد یافت.

جدول ۴: مشخصه الگوریتم‌های به‌کارگرفته‌شده در این شبیه‌سازی

الگوریتم	توضیحات	طول بردار ویژگی
الگوریتم ۱	هیستوگرام مقادیر شدت پیکسل‌های محلی تصویر	۸۰
الگوریتم ۲	هیستوگرام محلی لبه‌های افقی و عمودی تصویر	۳۲
الگوریتم ۳	هیستوگرام محلی بافت تصویر با عملگر LBP	۸۰
الگوریتم ۴	هیستوگرام رنگ محلی تصویر	۲۴۰

k تعداد تصاویر بازیابی‌شده را نشان می‌دهد. عبارت لگاریتمی یک کاهش‌دهنده هموار است که مکان نتیجه بازیابی‌شده را در رابطه تأثیر می‌دهد. در مجموع محاسبه آن بر اساس رابطه (۳) انجام می‌شود [۲۰].

$$DCCG_k = rel_1 + \sum_{i=2}^k \frac{rel_i}{\log_2(i)} \quad (3)$$

در مجموع بدون در نظر گرفتن نوع عملگر جستجو، می‌توان انواع عملگرهای منطقی که در این تحقیق استفاده شده را در جدول ۲ خلاصه کرد.

جدول ۲: عملگرهای منطقی استفاده‌شده در روش پیشنهادی

نوع عملیات	عملگر منطقی
Compute component binary vector	OR
Combination	Related to type of search
Similarity	XOR
MQUR	AND, OR

۵- نتایج تجربی

به‌منظور بررسی کارایی سامانه پیشنهادی از یک پایگاه‌داده شامل ۶۱۶ تصویر آموزش و ۱۸۴ تصویر آزمایش استفاده شده است. معمولاً پایگاه‌داده‌های استاندارد که وجود دارند و در بازیابی از آن‌ها استفاده می‌شود مبتنی بر یک درخواست ساده است و هر تصویر به یک گروه یا حداکثر به دو گروه مجزا مرتبط هستند. بنابراین برای مسئله مطرح‌شده در این مقاله مناسب نیستند. درواقع نیاز به معرفی یک پایگاه‌داده مناسب برای بررسی الگوریتم پیشنهادی الزامی است و به‌دنبال مجموعه‌ای می‌باشیم که تصویر بر اساس مؤلفه‌های آن تعریف شده و به‌صورت یک بردار باینری ذخیره شده باشد و هر تصویر می‌تواند به مؤلفه‌های متعددی منتسب شود. از این‌رو، برای انجام این تحقیق این مجموعه را تعریف نموده و به‌طور تقریبی بیش از ۹۰٪ تصاویر این مجموعه از پایگاه‌داده استاندارد Wang انتخاب شده‌اند. بقیه تصاویر نیز به‌صورت دستی از مجموعه‌های مختلف جمع‌آوری شده‌اند. به‌دلیل اینکه این تحقیق به‌دنبال بازیابی تصاویر مرتبط با درخواست‌های با نواحی و اشیاء‌های مجزا هست، پایگاه‌داده موردبررسی به‌صورت یک مجموعه تصویر با چند مؤلفه در نظر گرفته شده است. با این نوع تعریف درواقع هر تصویر می‌تواند بیش از یک مؤلفه داشته باشد. این ویژگی اجازه می‌دهد که قابلیت توانایی الگوریتم‌های پیشنهادشده در بازیابی تصاویر مرتبط به چند درخواست به‌طور همزمان، بهتر و دقیق‌تر ارزیابی شود. حداکثر تعداد مؤلفه‌های تصویر از پایگاه‌داده ۱۴ تا در نظر گرفته‌شده و به هر تصویر یک بردار مؤلفه باینری مناسب تخصیص داده شده است. در جدول ۳ مؤلفه‌های این پایگاه‌داده آورده شده است.

مرحله ۱: محاسبه بردار باینری مؤلفه‌های درخواست تصویری	
درخواست تصویری ۱	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 غذا اسب گل چمن قیل دایناسور آتهوس درخت کوه ساختمان آسمان رود ساحل انسان
درخواست تصویری ۲	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 غذا اسب گل چمن قیل دایناسور آتهوس درخت کوه ساختمان آسمان رود ساحل انسان
مرحله ۲: محاسبه بردار مؤلفه نهایی	
درخواست تصویری	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 غذا اسب گل چمن قیل دایناسور آتهوس درخت کوه ساختمان آسمان رود ساحل انسان
مرحله ۳ و ۴: انجام بازیابی بر اساس بردار باینری مؤلفه در مجموعه بردارهای باینری مجموعه آموزش - نمایش تصاویر خروجی	
بردار باینری ۱۰ تصویر خروجی	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0

شکل ۴: نتایج مرحله ۱ تا ۴ روش پیشنهادی شامل بردارهای محاسبه‌شده هر درخواست تصویری نشان‌دهنده در شکل ۵



شکل ۵: تعداد ۱۰ تصویر بازیابی‌شده بر مبنای الگوریتم پیشنهادی برای هر دو درخواست مشخص‌شده

دو درخواست تصویری دیگر نیز بازیابی شده است که نتایج هر مرحله روش پیشنهادی در شکل ۶ و تصاویر خروجی بازیابی شده در شکل ۷ به نمایش گذاشته شده‌اند. در شکل ۶ برای ارزیابی راحت‌تر در مرحله ۱ و ۲، بخش‌هایی که در بردار باینری به رنگ نارنجی هستند اشتباه محاسباتی توسط الگوریتم پیشنهادی را نشان می‌دهد. هر ۱۰ تصویر خروجی دارای یک بردار مؤلفه مستقل هستند که بردار محاسبه‌شده آن‌ها در شکل ۶ نشان داده شده است. در مرحله چهارم مؤلفه‌هایی که به‌طور صحیح محاسبه شده‌اند به‌صورت رنگ زرد مشخص شده‌اند.

برای ارزیابی بهتر الگوریتم پیشنهادی، نتایج با روش‌های ارائه‌شده در [۱۵] مقایسه شده‌اند. برای مقایسه کار، سه روش Joint-Avg، MQ-Max و MQ-Avg را بر روی پایگاه‌داده تعریف‌شده در این مقاله پیاده‌سازی کرده و نتایج نهایی بررسی و ارزیابی شده‌اند. برای پیاده‌سازی هر سه روش از نمایش BoW در نقاط کلیدی با توصیفگر بردار ۱۲۸ تایی SIFT استفاده شده و برای پیاده‌سازی آن از کتابخانه VL-Feat 0.9.20 [۲۱] استفاده شده است. در حقیقت روش‌های مرور شده بر اساس ترکیب‌های خطی جبری می‌باشد که آن‌ها را پیاده‌سازی نموده و با نوع ترکیب منطقی ارائه‌شده در این پژوهش مقایسه شده‌اند. نتایج ارزیابی در جدول‌ها و شکل‌هایی که در ادامه آورده شده‌اند، نمایش داده شده است. در مجموعه پایگاه‌داده آزمایش ۱۸۴ تصویر مختلف وجود دارد که در مجموع تعداد جفت درخواست‌های بررسی‌شده برابر با ۱۶۸۳۶ است. برای راحتی مقادیر DCG را بر مقادیر حداکثرشان تقسیم نموده و به‌صورت مقادیر نرمال‌شده نمایش داده شده‌اند. داده‌های جدول ۵ با تعداد بازیابی $k = 15$ اندازه‌گیری شده است.

جدول ۵: مقایسه میان DCG در الگوریتم‌های مختلف

Technique	DCG
Basic Method	۰/۷۲
Algorithm 1	۰/۵۳
Algorithm 2	۰/۳۸
Algorithm 3	۰/۳۸
Algorithm 4	۰/۳۸
Joint-Avg [15]	۰/۰۲
MQ-Max [15]	۰/۱۴
MQ-Avg [15]	۰/۰۱

دو نمونه اجرای بازیابی برای درخواست‌های تصویری متفاوت همراه با نتایج هر مرحله با تعداد ۱۰ تصویر خروجی به‌ترتیب در شکل‌های ۷-۴ نشان داده شده است. همان‌گونه که از خروجی مشخص است در صورتی که جفت درخواست دارای مؤلفه‌های منطبق بر هم باشند تصاویر بازیابی شده در خروجی دارای بالاترین کارایی خواهند بود. در شکل ۴ هر چهار مرحله برای بازیابی دو درخواست تصویری شکل ۵ نشان داده شده است. در این مثال اجرا شده تمام بردارهای مؤلفه‌های تصاویر خروجی در مرحله چهارم فقط دارای مؤلفه فعال «گل» هستند و مؤلفه‌هایی که صحیح محاسبه شده‌اند با رنگ زرد مشخص شده‌اند و درستی کار را بیان می‌نمایند.

مقایسه دیگری که در این تحقیق انجام شده است، متوسط زمان اجرای برنامه برای هر جفت درخواست در الگوریتم‌های موردبررسی می‌باشد که مجموع نتایج در جدول ۶ نشان داده شده است. نتایج، نشان‌دهنده زمان اجرای بهتر برای الگوریتم پیشنهادی در این کار است. تمام برنامه‌ها بر روی پردازنده core i7 اینتل اجرا شده‌اند.

جدول ۶: متوسط زمان اجرای الگوریتم‌های پیشنهادی برای بازیابی بر اساس یک جفت درخواست تصویری

الگوریتم	زمان اجرا (میلی ثانیه)
الگوریتم ۱	۲۶/۵
الگوریتم ۲	۳۵/۶
الگوریتم ۳	۴۵۰
الگوریتم ۴	۳۶/۷
[15] Joint-Avg	۶۳۶
[15] MQ-Max	۶۲۸
[15] MQ-Avg	۲۶۷۰

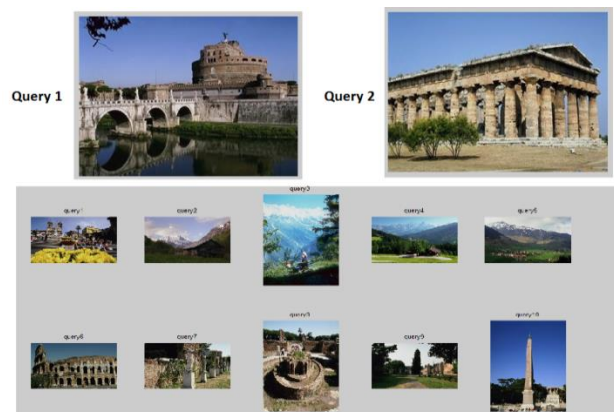
یکی از مزایای روش پیشنهادی، به‌کارگیری بردار مؤلفه باینری است که برای نمایش هر تصویر از آن‌ها استفاده شده است. با این بردار سرعت جستجو در پایگاه‌داده بسیار سریع‌تر خواهد شد. علاوه بر این امتیاز، این نوع تعریف سبب طراحی عملگرهای جستجوی متفاوت و متناسب با برنامه‌های کاربردی متنوع می‌شود که بسیار ساده و راحت نیز قابل پیاده‌سازی هستند درحالی‌که تعریف و پیاده‌سازی این نوع عملگرهای جستجو در شیوه‌های دیگر می‌تواند بسیار پیچیده و هزینه‌بر باشد. یکی دیگر از امتیازات این پژوهش، ایجاد حاشیه‌نویسی به‌طور خودکار برای تصویر ورودی است که در آینده بیش‌تر می‌توان به آن پرداخت.

یکی دیگر از مزایای این رویکرد استفاده بهینه‌تر از فضای حافظه است. با فرض یک پایگاه‌داده با n تصویر، بردار مؤلفه با m عنصر که ترکیبی از مقادیر صفر و یک است، می‌تواند یک ماتریس اسپارس ایجاد نماید؛ زیرا هر تصویر به‌طور معمول کم‌تر از m تا مؤلفه فعال دارد. پس تعداد صفرها می‌تواند زیاد باشد و از دیگر نکات مهم این ماتریس فقط وجود داده‌های صفر و یک است. بنابراین در مقایسه با ماتریس اسپارس مرسوم فضای کم‌تری نیاز دارد. البته اگر هر بردار باینری با حداکثر تعدادی بایت ذخیره شوند شرایط هنوز بهتر خواهد شد. برای مثال در این پیاده‌سازی طول بردار ۱۴ تایی است که می‌توان دو بایت برای آن در نظر گرفت. در این شرایط صرفه‌جویی حافظه بسیار بیش‌تر خواهد شد و دیگر نیازی به ذخیره‌سازی ماتریس اسپارس نیست. این مزیت امکان به‌کارگیری این روش را در سامانه‌های تعبیه شده^{۱۵} فراهم می‌کند.

در مجموع، مسئله‌ای که می‌توان در این روش به آن اشاره کرد، اجرای بازیابی مرحله اول برای محاسبه بردار مؤلفه تصویر است که

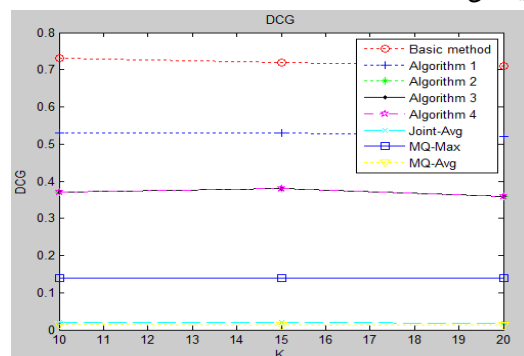
مرحله ۱: محاسبه بردار باینری مؤلفه‌های درخواست تصویری																																																																																																																																																																									
درخواست تصویری ۱	1 0 0 1 1 1 1 0 0 0 0 0 0 0 0 غذا اسب گل چمن قیل دایناسورتیوس درخت کوه ساختمان آسمان رود ساحل انسان																																																																																																																																																																								
درخواست تصویری ۲	0 0 0 1 0 0 1 0 0 1 1 1 1 1 0 غذا اسب گل چمن قیل دایناسورتیوس درخت کوه ساختمان آسمان رود ساحل انسان																																																																																																																																																																								
مرحله ۲: محاسبه بردار مؤلفه نهایی																																																																																																																																																																									
درخواست تصویری	1 0 0 1 1 1 1 0 0 1 1 1 1 1 0 غذا اسب گل چمن قیل دایناسورتیوس درخت کوه ساختمان آسمان رود ساحل انسان																																																																																																																																																																								
مرحله ۳ و ۴: انجام بازیابی بر اساس بردار باینری مؤلفه در مجموعه بردارهای باینری مجموعه آموزش-نمایش تصاویر خروجی																																																																																																																																																																									
بردار باینری ۱۰ تصویر خروجی	<table border="1"> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	1	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	1	0	0	0	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0
1	0	0	1	1	0	1	0	0	0	0	1	0	0																																																																																																																																																												
0	0	0	1	1	1	1	0	0	0	1	0	0	0																																																																																																																																																												
1	0	0	1	0	1	1	0	0	0	1	0	0	0																																																																																																																																																												
0	0	0	1	1	1	1	0	0	0	1	0	0	0																																																																																																																																																												
0	0	0	1	1	1	1	0	0	0	1	0	0	0																																																																																																																																																												
1	0	0	1	1	0	1	0	0	0	0	0	0	0																																																																																																																																																												
0	0	0	1	1	0	1	0	0	0	1	0	0	0																																																																																																																																																												
1	0	0	1	1	0	1	0	0	0	0	0	0	0																																																																																																																																																												
0	0	0	1	1	0	1	0	0	0	1	0	0	0																																																																																																																																																												
1	0	0	1	1	0	1	0	0	0	0	0	0	0																																																																																																																																																												
0	0	0	1	1	0	1	0	0	0	1	0	0	0																																																																																																																																																												
1	0	0	1	1	0	1	0	0	0	0	0	0	0																																																																																																																																																												

شکل ۶: نتایج مرحله ۱ تا ۴ روش پیشنهادی شامل بردارهای محاسبه‌شده هر درخواست تصویری نشان‌داده‌شده در شکل ۷



شکل ۷: خروجی ۱۰ تصویر بازیابی شده بر پایه دو درخواست ارائه‌شده در الگوریتم پیشنهادی

محاسبه DCG برای سه مقدار مختلف بازیابی ۱۰، ۱۵ و ۲۰ تصویر در شکل ۸ نشان داده شده است. نتایج برای هر سه تعداد بازیابی تقریباً یکسان است.



شکل ۸: مقادیر مختلف DCG برای بازیابی به ازای ۱۰، ۱۵ و ۲۰ تصویر

یادگیری می‌تواند بسیار مؤثر باشد. همچنین در آینده برای تسریع عملیات می‌توان بازیابی مرحله اول را بر اساس پردازنده‌های موازی چند هسته‌ای و محاسبات ابری انجام داد که می‌توان افزایش سرعت را پیش‌بینی کرد.

مراجع

- [1] P. Liu, K. Jia and Z. Wang, "An effective image retrieval method based on color and texture combined features," *third International conference on Intelligent Information Hiding and Multimedia Signal Processing, IHHMSP*, pp. 169-172, 2007.
- [2] M. H. Saad, H. I. Saleh, H. Konbar and M. Ashour, "Image retrieval based on integration between YCbCr color histogram and shape feature," *the 7th International Computer Engineering Conference (ICENCO)*, pp. 97-102, 2011.
- [3] A. W. Smeulders, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, 2000.
- [4] B. Wang, X. Zhang and N. Li, "Relevance feedback technique for content-based image retrieval using neural network learning," *IEEE International Conference on Machine Learning Cybernetics*, pp. 3692-3696, 2006.
- [۵] اسما شمسی گوشکی، سعید سریزدی، حسین نظام آبادی پور و محمدشهرام معین، «روشی جدید در بازخورد ربط برای بازیابی تصویر بر اساس محتوا به شیوه چندپرسشی»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۰، شماره ۲، صفحه ۵۲-۶۱، ۱۳۸۹.
- [6] X. Jin and J. C. French, "Improving image retrieval effectiveness via multiple queries," *Multimedia tools and applications*, vol. 26, no.2, pp.221-245, 2005.
- [7] J. Assfalg, A. Del Bimbo and P. Pala, "Using multiple examples for content-based image retrieval," *IEEE International Conference on Multimedia and Expo, ICME*, pp. 335-338, 2000.
- [8] S. M. M. Tahaghoghi, J. A. Thom and H. E. Williams, "Multiple example queries in content-based image retrieval," *String processing and information Retrieval .Springer Berlin Heidelberg*, pp. 227-241, 2002.
- [9] K. Hsiao, J. Calder and A. O. Hero, "Pareto-Depth for multiple-query image retrieval," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 583-594, 2015.
- [10] J. Deng, O. Russakovsky, J. Krause, M. Bernstein, A. Berg and L. Fei-Fei, "Scalable multi-label annotation," *Proceedings of the SIGCHI conference on Human Factors in Computing Systems, ACM*, 2014.
- [11] J. Tang and S. Acton, "An image retrieval algorithm using multiple query image," *7th International Symposium on Signal Processing and Its Applications*, pp. 193-196, 2003.
- [12] R. Arandjelovic and A. Zisserman, "Three things everyone should know to improve object retrieval," *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 2911-2918, 2012.
- [13] O. Chum, J. Philbin, J. Sivic, M. Isard and A. Zisserman, "Total recall: Automatic query expansion with a generative feature model for object retrieval," *IEEE 11th International Conference on Computer Vision, ICCV*, pp. 1-8, 2007.

می‌تواند بار محاسباتی به سامانه تحمیل نماید. از کارهایی که در آینده باید به آن پرداخته شود، استفاده از راه‌حل‌های جایگزینی و حذف این مرحله از بازیابی است.

با بررسی نتایج معیار اندازه‌گیری DCG در الگوریتم‌های پیشنهادی و مقایسه آن با کارهای پیشین، بالاترین مقدار به روش پایه تعلق دارد که چنین انتظاری نیز منطقی می‌باشد زیرا مبتنی بر نظر کاربر است. در ادامه الگوریتم ۱ بهترین نتیجه را در تعداد بازیابی‌های مختلف با زمان محاسباتی کم‌تری نشان می‌دهد. در مجموع نتایج به‌دست‌آمده از الگوریتم‌های مختلف استخراج ویژگی منطبق بر ۴ مرحله پیشنهادی در این مقاله، کارایی و سرعت اجرای بهتری را به هر سه روش Joint-Avg، MQ-Max و MQ-Avg نشان می‌دهند. از دلایل این نتایج می‌توان به کارا بودن روش ترکیب تصاویر بر اساس محتوا اشاره نمود و همچنین زمان اجرای بهتر آن ناشی از به‌کارگیری عملگرهای سریع منطقی است که تأثیر به‌سزایی در کاهش زمان اجرا دارند. برای بهبود دقت نتایج و افزایش معیار DCG می‌توان از دیگر روش‌های برجسته در استخراج ویژگی تصاویر استفاده نمود. هر چه میزان محاسبه بردارهای باینری مؤلفه درخواست‌های تصویر ورودی دقیق‌تر باشد نتایج نهایی به انتظار کاربر نزدیک‌تر خواهند بود. همان‌گونه که در بالا در شکل ۵ و ۴ مشاهده شد، برای درخواست‌های تصویری «گل» چون هر دو بردار باینری تخمین شده دقیق بوده‌اند دقت نتیجه بازیابی ۱۰۰٪ است. برای نمونه در دو شکل دیگر ۷ و ۶ که در بالا نشان داده شد، خطای ناشی در محاسبه بردارهای باینری مؤلفه برای هر درخواست نتیجه مستقیمی در بازیابی نهایی داشته است. از این‌رو برای کاهش این خطا، بهره‌بردن از ترکیب چند توصیف‌گر برای استخراج ویژگی می‌تواند نتایج دقیق‌تری را ارائه نماید.

۶- نتیجه‌گیری

در این مقاله یک چارچوب جدید برای ترکیب محتویات تصاویر درخواستی با تکیه بر محاسبات ساده بر پایه عملگرهای منطقی پیشنهاد شد. این نوع ترکیب می‌تواند برای بازیابی چند درخواست تصویری با مفاهیم متفاوت اعمال شود. در این روش بردار مؤلفه‌های باینری برای تصاویر ورودی مورد آزمایش به‌طور خودکار محاسبه می‌شوند. سپس بازیابی بر اساس آن انجام می‌شود. در این روش دو مرحله بازیابی انجام می‌شود. این روش در مقایسه با روش‌های ارائه‌شده قبلی برای بازیابی‌های مبتنی بر چند تصویر، به‌دلیل به‌کارگیری عملگرهای منطقی پردازش بسیار ساده‌تر و کم‌حجم‌تری را پیشنهاد می‌کند و در نتیجه سرعت اجرا بهتر خواهد بود. یکی دیگر از امتیازات این روش ایده بردارهای مؤلفه‌های تصویر است که نتایج بازیابی بر اساس آن‌ها برای چند درخواست بسیار مناسب‌تر خواهد بود. از کارهای آینده، اجرای این ایده بر روی پایگاه‌داده بزرگ‌تر می‌باشد و همچنین برای محاسبات بردارهای مؤلفه‌های تصویر روش‌های

- Conference on Computer Vision, ICCV, pp. 2544-2551, 2013.
- [18] B. Xu, J. Bu, C. Chen, D. Cai, X. He, W. Liu and J. Luo, "Efficient manifold ranking for image retrieval," *34th International ACM SIGIR Conference on Research and Development Information Retrieval*, pp. 525-534, 2011.
- [19] D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, "Local binary patterns and its application to facial image analysis: A survey," *IEEE Transactions on Systems, man, and cybernetics, Part C: Applications and Reviews*, vol. 41, no. 6, pp. 765-781, 2011.
- [20] P. Ravikumar, A. Tewari and E. Yang, "On nDCG consistency of listwise ranking methods," *International Conference on Artificial Intelligence and Statistics*, pp. 618-626, 2011.
- [21] <http://www.vlfeat.org/>
- [۱۴] رضا خدایی، محمدعلی بالافر و سیدناصر رضوی، «اثربخشی بسط پرس و جو مبتنی بر خوشه‌بندی اسناد شبه بازخورد با الگوریتم K-NN»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۶، شماره ۱، صفحه ۱۴۳-۱۵۱، ۱۳۹۵.
- [15] R. Arandjelovic and A. Zisserman, "Multiple queries for large scale specific object retrieval," *British Machine Vision Conference, BMVC*, pp. 1-11, 2012.
- [16] J. Sivic and A. Zisserman, "Video-google: A text retrieval approach to object matching in videos," *9th IEEE International Conference on Computer Vision*, pp. 1470-1477, 2003.
- [17] B. Fernando and T. Tuytelaars, "Mining multiple queries for image retrieval: on-the-fly learning of an object-specific mid-level representation," *IEEE International*

زیرنویس‌ها

- ¹ Query
- ² Text-based Image Retrieval
- ³ Content-based Image Retrieval
- ⁴ Relevance Feedback
- ⁵ Multiple-Query
- ⁶ Binary component vector
- ⁷ Query Expansion
- ⁸ Bag of Words
- ⁹ Pareto front
- ¹⁰ Efficient manifold ranking
- ¹¹ Discounted Cumulative Gain
- ¹² Precision
- ¹³ Recall
- ¹⁴ Multiple-Query Unique Relevance
- ¹⁵ Embedded systems