

Vol. 12/ No. 47/Spring 2023

Research Article

High-Scale Image Clustering with Semantic Cues Modeling and Spatial Simulation

Mahdi Jalali, Associate Professor^{1*} 

¹Department of Electrical Engineering, Naghadeh Branch, Islamic Azad University, Naghadeh, Iran. Microwave and Antenna Research Center, Urmia Branch, Islamic Azad University, Urmia, Iran, mahdi.jalali@iau.ac.ir

Correspondence

Mahdi Jalali, Associate Professor of Department of Electrical Engineering, Naghadeh Branch, Islamic Azad University, Naghadeh, Iran. Microwave and Antenna Research Center, Urmia Branch, Islamic Azad University, Urmia, Iran, mahdi.jalali@iau.ac.ir

Received: 28 September 2022

Revised: 5 November 2022

Accepted: 20 November 2022

Abstract

In recent years, image annotation is one of the active research topics. This paper suggests a semi-supervised cooperative clustering method for annotating images. In order to achieve the highest efficiency, the clustering results of six systems with different color space and similarity criteria are cooperatively combined with the majority vote. When the number of votes for an image is low, relevant feedback is used to annotate it. Nowadays, the linear similarity criterion is mostly used to determine the similarity between images, but the nonlinear models can have much better performance due to their proximity to the human vision system, for this purpose, the KMRBF nonlinear similarity criterion is used to simulate vision. Humans and improvement of recovery results are suggested. Experiments on the Corel image database and satellite images show that the proposed method has good performance. According to the results obtained in the satellite image database, the YIQ color space has a higher accuracy (82.5%). Also, the three color spaces CIE Lab, HSV and YIQ have higher efficiency, because in these color spaces, luminance is separated from chrominance and these color spaces are closer to the human vision system.

Keywords: Similarity Measure, Texture, Image Retrieval System, Clustering, Feature Extraction.

Highlights

- Combining system classification results with different color space and similarity criteria with multiple features.
- keeping the image vector modes constant while sampling at the greatest distance from the center of gravity.
- Extraction of Multiple Hybrid Features with spatial modeling.

Citation: M. Jalali, "High-Scale Image Clustering with Semantic Cues Modeling and Spatial Simulation," *Journal of Southern Communication Engineering*, vol. 12, no. 47, pp. 61–70, 2023, doi: 10.30495/jce.2022.1968473.1173, (in Persian)

خوشه‌بندی تصاویر مقیاس بالا با مدل‌سازی نشانه‌های معنایی و شبیه‌سازی مکانی

مهدی جلالی*^۱ ID

چکیده:

در سال‌های اخیر حاشیه‌نویسی تصاویر یکی از موضوعات تحقیقاتی فعال است. در این مقاله برای حاشیه‌نویسی تصاویر، تکنیک خوشه‌بندی تعاونی نیمه نظارت‌شده پیشنهاد می‌شود. روش‌های خوشه‌بندی به دلیل عدم نیاز به حاشیه‌نویسی بسیار مورد توجه هستند. برای دستیابی به بالاترین کارایی، نتایج خوشه‌بندی شش سیستم با فضای رنگ و معیار شباهت متفاوت با رأی اکثریت به صورت تعاونی، باهم ترکیب می‌شوند. در شرایطی که تعداد رأی‌ها برای یک تصویر کم باشد از بازخورد مرتبط برای حاشیه‌نویسی آن استفاده می‌شود. امروزه بیشتر از معیار شباهت خطی برای تعیین شباهت بین تصاویر استفاده می‌شود، ولی مدل‌های غیرخطی به دلیل نزدیکی به سیستم بینایی انسان می‌توانند کارایی بسیار بهتری داشته باشند، بدین منظور معیار شباهت غیرخطی KMRBF برای شبیه‌سازی بینایی انسان و بهبود نتایج بازیابی پیشنهاد می‌شود. آزمایش‌ها روی پایگاه داده تصاویر کورل و تصاویر ماهواره‌ای نشان می‌دهند که روش پیشنهادی دارای کارایی مناسبی است. با توجه به نتایج به دست آمده در پایگاه داده تصاویر ماهواره‌ای فضای رنگ YIQ دارای دقت بالاتری (به مقدار ۸۲/۵ درصد) است. همچنین سه فضای رنگ CIELab, HSV و YIQ دارای کارایی بالاتری هستند، چون در این فضاهای رنگی لومینانس از کرومینانس جدا بوده و این فضاهای رنگی به سیستم بینایی انسان نزدیک‌تر هستند.

کلید واژه‌ها: معیار شباهت، بافت، سیستم بازیابی تصویر، خوشه‌بندی، استخراج ویژگی.

اگرچه برق، واحد نقده، دانشگاه آزاد اسلامی، نقده، ایران
مرکز تحقیقات میکروویو و آنتن، واحد ارومیه، دانشگاه آزاد
اسلامی، ارومیه، ایران،
mahdi.jalali@iau.ac.ir

نویسنده مسئول

*مهدی جلالی، استادیار، گروه برق، واحد نقده، دانشگاه آزاد
اسلامی، نقده، ایران
مرکز تحقیقات میکروویو و آنتن، واحد ارومیه، دانشگاه آزاد
اسلامی، ارومیه، ایران،
mahdi.jalali@iau.ac.ir

تاریخ دریافت: ۶ مهر ۱۴۰۱

تاریخ بازنگری: ۱۴ آبان ۱۴۰۱

تاریخ پذیرش: ۲۹ آبان ۱۴۰۱

<https://doi.org/10.30495/jce.2022.1968473.1173>

۱-مقدمه

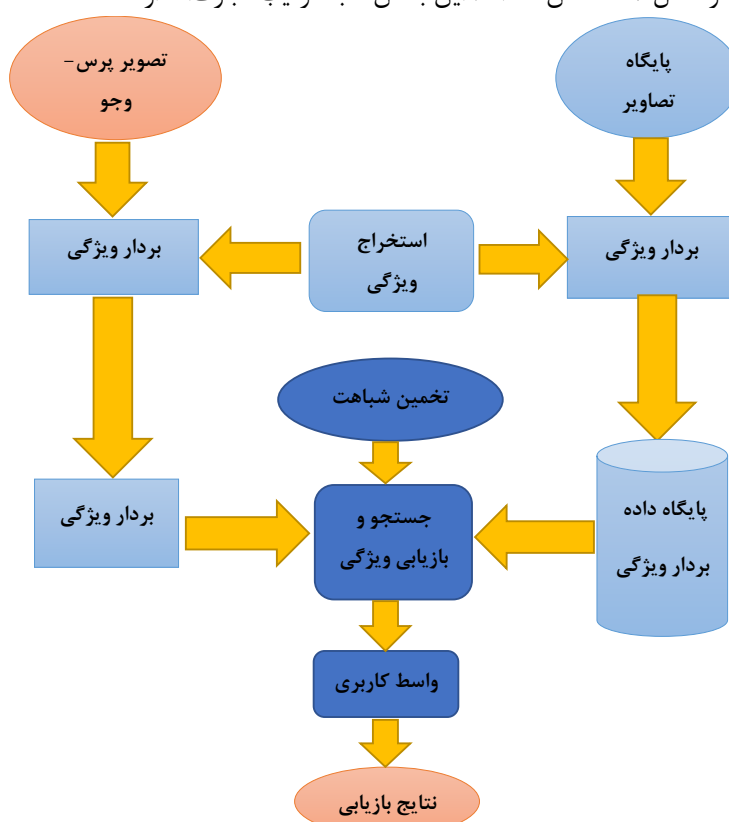
در ابتدای دهه نود با افزایش حجم تصاویر موجود در پایگاه‌های داده نظیر شبکه جهانی اینترنت و نیز برای غلبه بر مشکلات سیستم‌های مبتنی بر متن، توسعه سیستم‌های بازیابی تصویر مبتنی بر محتوا پیشنهاد شد که وظیفه استخراج تصاویر به صورت خودکار و با استفاده از مفاهیم دیداری نظیر رنگ، بافت و طرح‌بندی تصویر را بر عهده داشتند [۱-۳]. سیستم‌های اولیه برای بازیابی تصویر از کاربر درخواست می‌کرد تا یک یا چند ویژگی دیداری را انتخاب و محدوده‌ای برای مقدار آن‌ها تعیین کند، سپس عمل بازیابی بر اساس این اطلاعات صورت می‌گرفت. در سیستم‌های پیشرفته‌تر، کاربر قادر بود تا درجه اهمیت ویژگی‌های انتخاب‌شده را تعیین کند. به دلیل ناتوانی در جلب رضایت کاربر، در سیستم‌های نسل بعد قابلیت دریافت یک تصویر به‌عنوان الگوی بازیابی فراهم شد. با این امکان، کاربر در کنار قابلیت‌های ذکر شده می‌توانست تصویر یا شمای تصویر موردنظر خود را به سیستم وارد کرده و تصاویر مشابه با آن را از پایگاه داده استخراج نماید. در سیستم‌های پیشرفته

بعدی، کاربر قادر شد تا میزان اهمیت هر یک از این ویژگی‌ها را نیز تعیین نماید؛ همچنین قابلیت بهبود نتایج جستجو با استفاده از بازخورد کاربر نیز به این سیستم‌ها اضافه شد [۵،۴].

در این مقاله بعد از معرفی سیستم بازیابی تصاویر و ساختار و زیر بخش‌های آن به همراه الگوریتم پیشنهادی، فرمول‌های معیار شباهت ارائه و نتایج به‌دست‌آمده با نمودارها و جدول باهم مقایسه شده و بهترین سیستم‌ها از نظر دقت بازیابی معرفی می‌شود.

۲- ساختار کلی سیستم‌های بازیابی تصاویر

سیستم‌های بازیابی تصویر از ویژگی‌های تصویری برای دسته‌بندی و بازیابی تصاویر استفاده می‌کنند. در این سیستم‌ها پایگاه داده تصویری وجود دارد که شامل کلیه تصاویر قابل بازیابی برای کاربر است. کاربر برای استخراج تصاویر از این بانک، خصوصیات موردنظر خود را به سیستم ارائه می‌دهد و سیستم با استفاده از این اطلاعات، بانک تصاویر خود را جستجو می‌کند تا تصاویر مطلوب را استخراج و به کاربر نمایش دهد. بر همین اساس، یک سیستم بازیابی تصویر از پنج بخش اصلی تشکیل می‌شود که این بخش‌ها در شکل ۱ مشخص شده‌اند. این بخش‌ها به ترتیب عبارت‌اند از:



شکل ۱: سیستم بازیابی تصویر

Figure 1. Image retrieval system

- پایگاه داده تصویر: که شامل تصاویر خام اولیه است که قابل بازیابی هستند.
- استخراج ویژگی^۱: که وظیفه استخراج ویژگی‌های تصاویر را بر عهده دارد.
- پایگاه داده ویژگی: که خصوصیت‌های استخراج‌شده از تصاویر پایگاه داده در آن ذخیره می‌شوند.
- بخش جستجو و بازیابی^۲ که تصاویر پایگاه داده را جستجو می‌کند و بر اساس میزان شباهت آن‌ها به تصویر پرس‌وجو، شبیه‌ترین تصاویر را بازیابی و به‌منظور نمایش، به واسط کاربری می‌فرستد.

¹ Feature Extraction

² Search & Retrieval

- واسط کاربری^۱: که تصویر پرس‌وجوی را از کاربر دریافت کرده و تصاویر استخراج‌شده را برای وی نمایش می‌دهد.
- گورز^۲ در [۵] روش‌های بازیابی تصاویر را بر اساس کاربردهای مختلف در سه دسته طبقه‌بندی کرده است:
 ۱. جستجو بر اساس شباهت تصاویر^۳، هدف از این روش به دست آوردن تصاویر موردنظر کاربر از بین تصاویر موجود در یک گالری تصاویر است که در مراحل پی‌درپی و با استفاده از بازخورد کاربر صورت می‌گیرد.
 ۲. جستجو به هدف یافتن تصویر خاص^۴: هدف از این روش به دست آوردن تصاویر مشابه با یک تصویر مشخص از پایگاه داده تصاویر است. منظور از تصویر مشابه، تصویری است که (بخشی از آن) با تصویر موردنظر کاربر یکسان باشد و یا (بخشی از) یکشی در هر دو تصویر یکسان وجود داشته باشد.
 ۳. جستجو بر اساس دسته‌بندی^۵: هدف از این روش به دست آوردن تصویری است که متعلق به دسته یا کلاس بخصوصی باشد، مانند تصاویر آرایه‌گری منزل و یا تصاویر رادیولوژی.
- وی همچنین پایگاه داده تصاویر را بر اساس چگونگی میزان تغییر در ویژگی‌ها در دو دسته قرار داده است [۵]:
 ۱. دامنه محدود، شامل تغییرات محدود در ویژگی‌های دیداری تصاویر (مانند پایگاه داده تصاویر مربوط به پزشکی).
 ۲. دامنه وسیع، شامل تغییرات وسیع و غیرقابل‌پیش‌بینی در تصاویر (مانند پایگاه داده تصاویر موجود در اینترنت).
- نمونه‌ای از سیستم‌های بازیابی تصویر که به‌تازگی توسعه داده‌شده‌اند می‌توان سیستم بازیابی تصویر توسعه داده‌شده توسط دانشگاه کلمبیا تحت نام VisualSEEK [۶]، EFF^۶ [۷]، سیستم دانشگاه آمستردام به نام PicToSeek [۸] و سیستم دانشگاه جنوا به نام Viper [۹] را نام برد که در این مطالعه به برخی از آن‌ها اشاره شده است. همچنین برخی از ویژگی‌های QBIC^۷ [۱۰] که توسط شرکت آی‌بی‌ام توسعه داده‌شده است نیز توضیح داده می‌شود. سیستم QBIC جزو مهم‌ترین سیستم‌های بازیابی به شمار می‌رود، چراکه سایر سیستم‌های بازیابی از آن به‌عنوان مدل اولیه توسعه سیستم بازیابی تصویر استفاده کرده‌اند. برای بازیابی تصویر باید از مطلوبات کاربر مطلع شد. کاربر می‌تواند ویژگی‌ها و درجه اهمیت آن‌ها را برای تصاویر مطلوب خود به سیستم وارد کند و یا اینکه با استفاده از تصویری موسوم به تصویر پرس‌وجو^۸ مشخصات تصاویر مطلوب خود را اعلام کند. از روش دوم در اکثر سیستم‌های بازیابی تصویر نظیر PicToSeek، QBIC، EFF، VisualSEEK و Viper استفاده شده است. در اکثر سیستم‌ها برای جستجوی تصاویر و پیدا کردن تصاویر مشابه از جستجوی نزدیک‌ترین همسایه^۹ استفاده می‌شود که بنا به تعریف، این روش جستجو به دنبال نقطه‌ای از فضای ویژگی است که فاصله آن با بردار ویژگی تصویر پرس‌وجو کمینه باشد [۱۱]. روش‌های متنوعی برای به دست آوردن این مجموعه تصاویر عنوان شده است که اکثر آن‌ها از شیوه‌های اندیس‌گذاری تصاویر و قسمت‌بندی فضای ویژگی استفاده می‌کند. بدین ترتیب نزدیک‌ترین تصویر به تصویر مطلوب کاربر استخراج می‌شود که k می‌تواند عددی از پیش تعیین‌شده باشد و یا کاربر آن را مقداردهی کند. در VisualSEEK به کاربر این امکان داده می‌شود تا سایر تصاویر استخراج‌شده که از نظر سیستم از این فاصله همسایگی دورتر است را نیز مشاهده کند.

۳- معیار ارزیابی سیستم‌های بازیابی اطلاعات

برای ارزیابی سیستم‌های بازیابی محتوا که شامل بازیابی تصویر نیز می‌گردد، غالباً از دو مفهوم فراخوانی و دقت استفاده می‌شود. شکل ۲ شمایی از بخش‌بندی تصاویر موجود در پایگاه داده در زمان بازیابی که بر اساس یک پرس‌وجوی معین است را نشان می‌دهد. در این شکل تصاویر پایگاه داده به چهار بخش مجزا تقسیم می‌شوند:

۱. تصاویر نامرتبط

¹ User Interface

² Gevers

³ Search by Association

⁴ Target Search

⁵ Category Search

⁶ Effective Feature Fractional

⁷ Query by image and video content

⁸ Query

⁹ Nearest Neighborhood Search

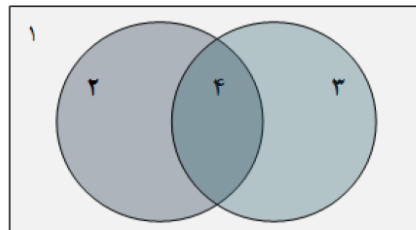
۲. تصاویر بازیابی شده

۳. تصاویر مرتبط

۴. تصاویر مرتبط بازیابی شده

بر این اساس، فراخوانی برابر تعداد تصاویر بازیابی شده مرتبط با پرس و جو تقسیم بر تعداد کل تصاویر مرتبط با پرس و جو در پایگاه داده تعریف می‌شود؛ در حالی که دقت برابر تعداد تصاویر بازیابی شده مرتبط با پرس و جو تقسیم بر تعداد کل تصاویر بازیابی شده است. در واقع فراخوانی معرف میزان حساسیت بازیابی و دقت معرف میزان کامل بودن آن است [۱۲]. واضح است که مقدار به دست آمده عددی در فاصله [۰,۱] است.

اگر نقطه برش را عددی در نظر بگیریم که برابر با ماکزیمم تعداد تصاویر خروجی به کاربر پس از مرحله جستجو باشد، با تغییر این نقطه می‌توان دقت فراخوانی‌های متفاوت نسبت به آن را به دست آورد. به گونه‌ای که با بالا بردن نقطه برش فراخوانی افزایش یافته و در مقابل تعداد تصاویر نامرتب نیز افزایش می‌یابد که خود موجب کاهش دقت خواهد شد. بر همین اساس گرافی تحت نام گراف فراخوانی دقت تعریف می‌شود که معرف تغییر دقت با توجه به تغییرات تعداد فراخوانی در نقاط برش متفاوت است [۱۳]. از این گراف برای ارزیابی سیستم بازیابی به دست آمده استفاده می‌گردد.



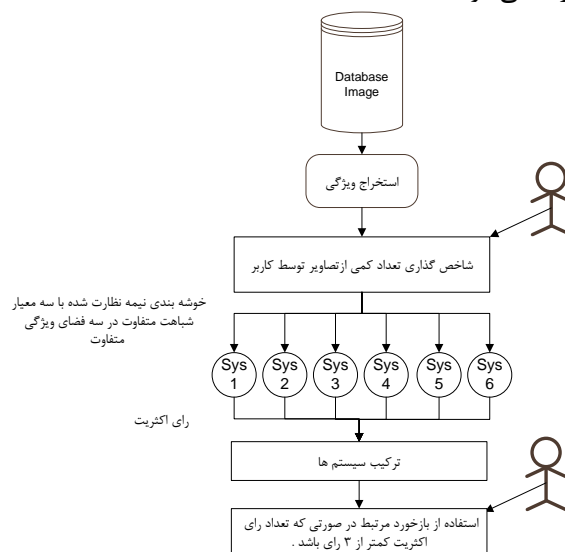
شکل ۲: نتایج بازیابی: ۱ - تصاویر نامرتب، ۲ - تصاویر بازیابی شده، ۳ - تصاویر مرتبط، ۴ - تصاویر بازیابی شده مرتبط [۱].
Figure 2. Retrieval results: 1 - Unrelated images, 2 - Retrieval images, 3 - Related images, 4 - Related retrieval images

۳-۱- روش پیشنهادی

الگوریتم k -means به دلیل پیچیدگی زمانی $O(n)$ ، برای خوشه‌بندی^۱ پایگاه داده‌های بزرگ تصاویر بسیار مناسب است ولی از مشکلات الگوریتم‌های k -means، همگرایی به بهینه محلی است که ناشی از انتخاب نامناسب مراکز اولیه است. همچنین در بازیابی تصویر در فضاهای رنگ متفاوت و با استفاده از معیارهای شباهت متفاوت خوشه‌بندی‌های متفاوتی حاصل می‌شود. در صورت ترکیب نتایج خوشه‌بندی می‌توان به نتایج بهتری دست یافت. مشکل اصلی برای ترکیب نتایج در خوشه‌بندی نامشخص بودن برچسب خوشه‌ها است. برای مثال در یک بار خوشه‌بندی خوشه اول به صورت تصادفی خوشه گل است و در بار دیگر خوشه اتوبوس است و نمی‌توان نتایج خوشه‌بندی با معیارهای متفاوت را باهم ترکیب کرد. برای حل این مشکلات از روش خوشه‌بندی نیمه نظارت شده K -means استفاده می‌شود. بدین منظور روش پیشنهادی ارائه شده و در شکل ۳ نشان داده شده است. ابتدا از تعداد کمی تصاویر بازخورد شده توسط کاربر برای تعیین مراکز خوشه ابتدایی استفاده می‌شود. در این حالت چون هر خوشه با یک مفهوم حاشیه‌نویسی می‌شود، می‌توان نتایج خوشه‌بندی متفاوت را باهم ترکیب کرد. برای دستیابی به بالاترین کارایی به صورت تجربی سه فضای رنگ و سه معیار شباهت که الگوریتم خوشه‌بندی روی آن‌ها بهترین کارایی را دارند، انتخاب می‌شوند. با استفاده از این فضاهای رنگ و معیارهای شباهت، نه سیستم بازیابی ایجاد می‌شود که از این بین شش سیستم که بهترین کارایی را دارند، انتخاب می‌شوند و آن‌ها به صورت خوشه‌بندی تعاونی (با رأی‌گیری اکثریت) باهم ترکیب می‌شوند. در برخی مواقع رأی‌گیری اکثریت ما را به چندین تصمیم متفاوت با رأی‌های مساوی یا تعداد رأی‌های کم هدایت می‌کند. برای افزایش دقت سیستم در شرایطی که تعداد رأی‌ها در مورد یک تصویر بیش از نصف تعداد رأی‌دهنده‌ها (سه رأی) است، به آن رأی اعتماد می‌شود. ولی اگر تعداد رأی‌های یکسان کمتر از نصف رأی‌دهنده‌ها باشد از بازخورد مرتبط برای تعیین خوشه مورد نظر استفاده می‌شود. به دلیل نقش مهمی که معیار شباهت در کارایی سیستم بازیابی تصویر دارد،

¹ Clustering

ابتدا یک بررسی روی انواع معیار شباهت معایب و مزایای آن‌ها انجام می‌شود و سپس یک معیار شباهت پیشنهادی جدید برای رفع برخی از این مشکلات ارائه می‌شود.



شکل ۳: الگوریتم پیشنهادی
Figure 3. Proposed algorithm

۳-۲- معیار شباهت

تابع فاصله یا اندازه‌گیری عدم شباهت مسئله اصلی خیلی از مدل‌ها و الگوریتم‌ها در یادگیری ماشین و شناسایی آماری الگو و بینایی ماشین است. بخش بسیار مهم سیستم بازیابی تصویر نیز تعیین معیار شباهت بین تصاویر است. همچنین در فرایند محاوره تحلیل نقش کاربر در درک شباهت تصاویر نیز بسیار با اهمیت است [۱۴]. شباهت یک تعریف پایه‌ای برای خوشه‌بندی است. اندازه‌گیری شباهت بین دو الگو در فضای ویژگی، اساس کار اکثر روش‌های خوشه‌بندی است. کارایی خیلی از روش‌های خوشه‌بندی و کلاس‌بندی (RBF^۱ و K-means) بستگی به تعریف مناسب تابع فاصله دارد. در عوض پیش تعریف تابع فاصله بر مبنای دانش قبلی درباره کاربرد مورد استفاده بسیار خوش‌آیند برای یادگیری تابع فاصله مناسب است و ممکن است با یک انتخاب اولیه شروع شود و بر مبنای اطلاعات نظارت‌شده موجود درباره کاربرد ادامه یابد.

معیارهای شباهت غیرخطی مثل روش‌های بر مبنای هسته دارای کارایی بهتری نسبت به روش‌های خطی مثل فاصله اقلیدسی و مینکوسکی و ماهالانوبیس^۲ و ... دارند. چون روش‌های بر مبنای هسته به سیستم بینایی انسان نزدیک‌تر هستند و در سیستم‌های بازیابی تصویر به خوبی عمل می‌کنند. تابع هسته فاصله بین نقاط داده را با نگاشت ضمنی به فضای با ابعاد بالاتر که داده‌ها به صورت خطی از هم قابل جداسازی باشند اندازه‌گیری می‌کند.

مجموعه داده X با d بعد در فضای حقیقی R^d مفروض است. تابع نگاشت غیرخطی از فضای ورودی به فضای ویژگی با ابعاد بالای H بدین صورت است [۱۵].

$$\psi: R^d \rightarrow H, \vec{X}_1 \rightarrow \psi(\vec{X}_1) \quad (1)$$

که،

$$\vec{X}_1 = [X_{i,1}, X_{i,2}, \dots, X_{i,d}]^T, \psi(\vec{X}_1) = [\psi_1(\vec{X}_1), \dots, \psi_H(\vec{X}_1)]^T \quad (2)$$

^۱ Radial Basis Function

^۲ Mahalanobis

با اعمال نگاشت ضرب داخلی \vec{x}_1, \vec{x}_j به $\Psi^T(\vec{x}_1), \Psi(\vec{x}_j)$ تبدیل می‌شود. ضرب داخلی $\Psi^T(\vec{x}_1), \Psi(\vec{x}_j)$ در فضای تبدیل می‌تواند به‌عنوان تابع هسته در $k(\vec{x}_1, \vec{x}_j)$ فضای ورودی محاسبه شود. تابع هسته k مربع ضرب داخلی بردار \vec{x}_1 و \vec{x}_j در R^d است. بنابراین اندازه فاصله بین دو الگوی \vec{x}_1 و \vec{x}_j در فضای هسته به‌صورت زیر است،

$$\begin{aligned} \left\| \psi(\vec{X}_1) - \psi(\vec{X}_j) \right\|^2 &= \left(\psi(\vec{X}_1) - \psi(\vec{X}_j) \right)^T \left(\psi(\vec{X}_1) - \psi(\vec{X}_j) \right) \\ &= \psi^T(\vec{X}_1) \cdot \psi(\vec{X}_1) - 2\psi^T(\vec{X}_1) \psi(\vec{X}_j) + \psi(\vec{X}_j) \cdot \psi(\vec{X}_j) = K(\vec{X}_1, \vec{X}_1) - 2K(\vec{X}_1, \vec{X}_j) + K(\vec{X}_j, \vec{X}_j) \end{aligned} \quad (3)$$

در میان انواع توابع هسته استفاده‌شده در مقالات تابع هسته گوسی که به RBF معروف است بسیار مشهور بوده و به‌صورت زیر نمایش داده می‌شود،

$$K(\vec{X}_1, \vec{X}_j) = \exp \left(- \frac{\left\| \vec{X}_1 - \vec{X}_j \right\|^2}{2\sigma^2} \right) \quad (4)$$

که $\sigma > 0$ است. در هسته گوسی واضح است که $k(\vec{x}_1, \vec{x}_1) = 1$ است [۱۵] بنابراین در رابطه (۳) داریم،

$$\left\| \psi(\vec{X}_1) - \psi(\vec{X}_j) \right\|^2 = 2 \left(1 - K(\vec{X}_1, \vec{X}_j) \right) \quad (5)$$

همان‌طور که پیش‌تر اشاره شد اکثر الگوریتم‌های خوشه‌بندی از معیارهای خطی برای تعیین شباهت بین الگوها استفاده می‌کنند. در این مقاله از رویکرد غیرخطی برای شبیه‌سازی درک انسان استفاده می‌شود که اجازه می‌دهد به‌طور کارا شکاف معنایی بین ویژگی‌های سطح پایین و مفاهیم معنایی سطح بالا موجود در درک انسان را کاهش دهد. معیار شباهت هسته گوسی ماهالانوبیس بر مبنای رابطه ۴ را به‌صورت رابطه ۵ تعریف می‌شود که برای اندازه‌گیری تابع شباهت برای الگوریتم خوشه‌بندی استفاده می‌شود و مقدار σ را به‌طور تجربی برابر ۱/۵ تنظیم می‌شود [۱۶].

$$D(\vec{Y}, \vec{X}) = 2 \left(1 - e^{-\frac{(\vec{Y}-\vec{X})^T S^{-1} (\vec{Y}-\vec{X})}{2\sigma^2}} \right) \quad (6)$$

در رابطه فوق ماتریس S کوواریانس بین X و Y است که به‌صورت زیر تعریف می‌شود،

$$S = E \left\{ (X - M)(X - M)^T \right\} = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \dots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1}^2 & \sigma_{n1}^2 & \dots & \sigma_{nn}^2 \end{bmatrix} \quad (7)$$

ماتریس M میانگین است و به‌صورت زیر تعریف می‌شود،

$$M = E \{ X \} \quad (8)$$

۴- بررسی و ارزیابی نتایج سیستم پیشنهادی

پایگاه داده مورد استفاده شامل ۱۰۰۰ تصویر از مجموعه کورل^۱ در ده گروه (خوشه‌ها به ترتیب آفریقا، ساحل، اتوبوس، بنای تاریخی، دایناسور، فیل، گل، اسب، کوهستان و غذا) در ابعاد 384×256 یا 256×384 است و ۲۴۰۰ تصویر ماهواره‌ای گرفته‌شده از شهر تهران که شامل پنج مفهوم در ابعاد 90×90 در قالب JPEG^۲ است که در شکل ۴ نشان داده‌شده است.

¹ corel

² Joint Photographic Experts Group

ویژگی‌های سطح پایین مورد استفاده شامل نه ویژگی میانگین، واریانس، کجی‌نمایی از فضای رنگی CIELab به‌عنوان ویژگی رنگ است. همچنین شش ویژگی انرژی، آنتروپی، کنتراست، واریانس، همبستگی و گشتاور تفاضلی معکوس به‌عنوان ویژگی بافت استخراج شده اند [۱۷].

در اولین قسمت برای بررسی تأثیر فضای رنگ در کارایی سیستم بازیابی تصویر، دقت خوشه‌بندی در فضای رنگ متفاوت بررسی شده است. دقت خوشه‌بندی در جدول ۱ نشان داده شده است. در پایگاه داده تصاویر ماهواره‌ای فضای رنگ YIQ دارای دقت بالاتری است. در کل کورل در سه فضای رنگ CIELab, HSV و YIQ دارای کارایی بالاتری هستند چون در این فضاهای رنگی لومینانس از کرومینانس جدا هستند و در واقع این فضاهای رنگی به سیستم بینایی انسان نزدیک‌تر است.



شکل ۴: نمونه‌ای از تصاویر ماهواره‌ای مورد استفاده
Figure 4. An example of satellite images used

جدول ۱: تأثیر فضای رنگ در دقت خوشه‌بندی

Table 1. Effect of color space on clustering accuracy

CIELab	HSV	YIQ	YCbCr	RGB	درصد خوشه‌بندی
۷۵/۸	۷۴/۷	۷۴/۴	۶۲/۴	۶۰	پایگاه داده کورل
۸۰/۶	۸۰/۱	۸۲/۵	۶۶/۳	۶۳/۱	تصاویر ماهواره‌ای

در شکل ۵ نمودار محدوده-فراخوانی مربوط به این دوازده معیار شباهت ارائه شده است. همان‌طور که مشاهده می‌شود معیار شباهت پیشنهادی به دلیل نگاشت فضای ویژگی به فضایی با ابعاد بالاتر توانسته فاصله الگوها را به خوبی اندازه‌گیری کند و دارای بهترین کارایی است و بعد از آن معیارها شباهت کنبرا^۱، ماهالانوبیس، جفری^۲ قرار دارند. معیارهای اینر^۳ و L_∞ دارای کمترین کارایی هستند چون معیارهای شباهت کنبرا، ماهالانوبیس و جفری به نوعی در فرمول‌های خود نوعی عمل نرمال‌سازی را انجام می‌دهند و کارایی بسیار بالاتری نسبت به بقیه معیارها دارند.

برای ارزیابی روش‌های پیشنهادی معمولاً از نمودارهای دقت-فراخوانی^۴ و یا نمودار محدوده-فراخوانی^۵ و یا نمودار تعداد تصاویر بازیابی-فراخوانی^۶ استفاده می‌شود. معیار محدوده-فراخوانی یکی از رایج‌ترین معیارهای ارزیابی در بازیابی تصویر است. در اینجا معیار فراخوانی بیانگر نسبت تعداد تصاویر مرتبط بازیابی شده به کل تصاویر مرتبط درون پایگاه است:

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{total of relevant images in the collection}} \quad (9)$$

در حقیقت فراخوانی، اندازه‌گیری توانایی بازیابی تصاویر مرتبط از پایگاه داده است.

¹ Conberra

² Jeffrey

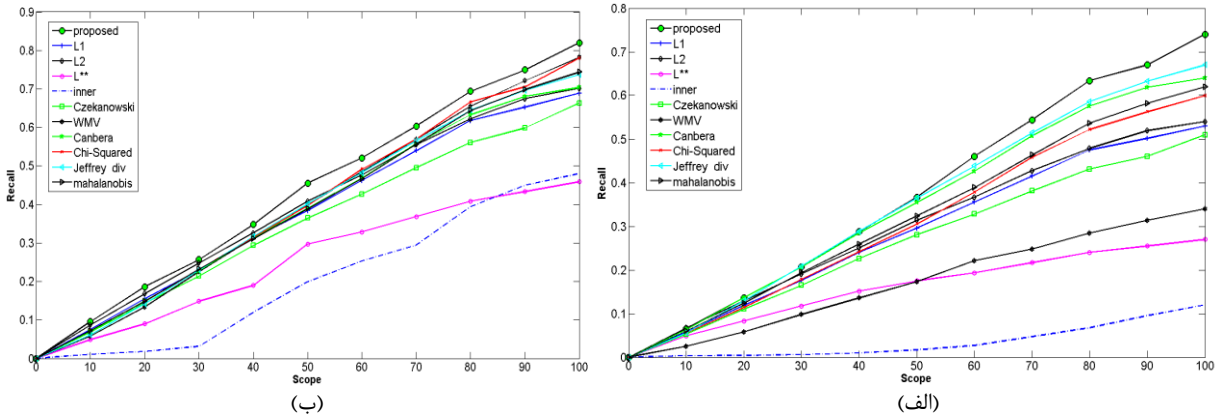
³ inner

⁴ Precision-Recall

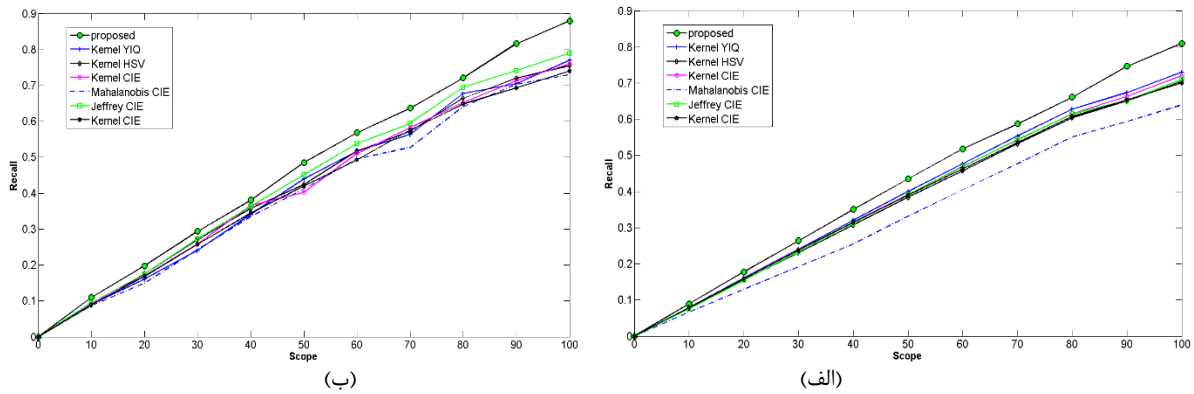
⁵ Recall- Scope

⁶ Recall- No. of retrieved images

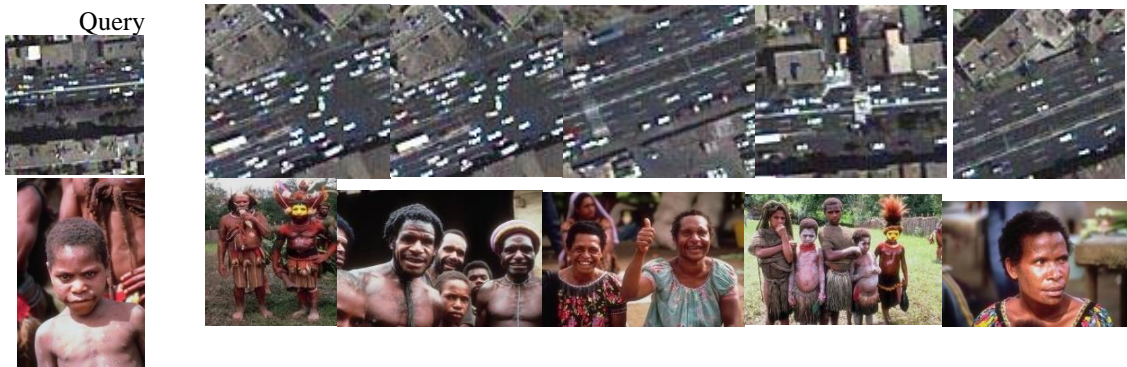
برای تعیین کارایی روش ترکیب روش‌های خوشه‌بندی به صورت خوشه‌بندی تعاونی در شکل ۶ با نمودار محدوده- فراخوانی شش سیستم را با سیستم ترکیبی ارائه شده است. همان‌طور که ملاحظه می‌کنید با ترکیب روش‌ها کارایی سیستم افزایش یافته است. همچنین با ارائه پرس‌وجو روی دو پایگاه داده در شکل ۷ ملاحظه می‌شود که سیستم دارای کارایی بسیار مناسبی است.



شکل ۵: مقایسه معیار شباهت پیشنهادی با ۱۱ معیار شباهت دیگر (الف) پایگاه داده کورل (ب) پایگاه داده تصاویر ماهواره‌ای
 Figure 5. Comparison of the proposed similarity measure with 11 other similarity measures a) Corel database b) Satellite image database



شکل ۶: مقایسه معیارهای شباهت با بهترین فضاها رنگ با سیستم ترکیبی پیشنهادی (الف) پایگاه داده کورل (ب) پایگاه داده تصاویر ماهواره‌ای
 Figure 6. Comparison of similarity measures with the best color spaces with the proposed hybrid system a) Corel database b) Satellite image database



شکل ۷: نتایج بازیابی دو پرس‌وجو از پایگاه داده تصاویر ماهواره‌ای و کورل و پنج تصویر برگردانده شده توسط روش پیشنهادی
 Figure 7. Retrieving results of two queries from the satellite image database and Corel and five images returned by the proposed method

۵- نتیجه‌گیری

سیستم بازیابی تصویر بر مبنای محتوا شامل سه بخش اصلی استخراج ویژگی سطح پایین و اندازه‌گیری شباهت و کاهش فاصله معنایی است. برای اینکه این سیستم به‌خوبی کار کند باید همه بخش‌های آن به‌خوبی کار کند و به سیستم بینایی و درک انسان نزدیک باشد، بدین منظور در این مقاله از معیار شباهت هسته مالهالانوبیس برای اندازه‌گیری شباهت استفاده شده است. همچنین با ترکیب نتایج خوشه‌بندی، سیستم‌هایی با ویژگی‌های استخراج‌شده در فضای رنگ متفاوت و معیار شباهت متفاوت می‌توان به نتایج بهتری دست یافت. نتایج آزمایش‌ها نیز کارایی خوب روش پیشنهادی را نشان می‌دهند. برای بازیابی، کاربر موظف است تا تصویر پرس‌وجو را به سیستم وارد کند. سیستم ویژگی‌های این تصویر را با ویژگی کلیه تصاویر موجود در پایگاه داده مقایسه می‌کند. در انتها تصاویری که در مرحله جستجو و تعیین شباهت بیشترین امتیاز را کسب کرده‌اند، به کاربر نمایش داده می‌شوند. برای بررسی این سیستم، نتایج آن با نتایج به‌دست‌آمده از بررسی سیستم بازیابی تصویر SIMPLicity مقایسه شده است. پایگاه داده مورد استفاده، مشتق شده از پایگاه داده کورل و تصاویر ماهواره‌ای است که بر اساس شباهت مفهومی، تصاویر در ده رده طبقه‌بندی شده‌اند. پس از بررسی، مشاهده شد که سیستم پیشنهادی در تمامی رده‌ها از سیستم SIMPLicity قوی‌تر است. در روش استخراجی ویژگی‌های شکل اگر در ابتدا پس‌زمینه توسط بعضی از روش‌ها مثل الگوریتم بیشینه‌سازی امید ریاضی حذف شود و فقط از تصویر شیء موجود در تصویر باقی بماند و سپس استخراج ویژگی شکل یا همان هیستوگرام جهت لبه‌های مشابه اعمال شود در این صورت ویژگی‌های استخراجی شکل نیز در افزایش دقت بازیابی تصاویر اهمیت بسزایی خواهند داشت.

مراجع

- [1] F. Cao , J. Liang and B. Liang, “ A new initialization method for categorical data clustering,” *Expert Systems with Applications*, vol. 36,no.7, pp. 10223–10228, 2009, doi:10.1016/j.eswa.2009.01.060.
- [2] A. Amato and V. Lecce, “ A knowledge based approach for a fast image retrieval system,” *Image and Vision Computing*, vol. 26,no.11, pp. 1466–1480, 2008, doi:10.1016/j.imavis.2008.01.005.
- [3] Y. Chen, J. Z. Wang and R. Krovetz, “ Cluster-Based Retrieval of Images by Unsupervised Learning,” in *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1187-1201, Aug. 2005, doi: 10.1109/TIP.2005.849770.
- [4] R. Zhang and Zh.Zhang, “ BALAS:Empirical Bayesian learning in the relevance feedback for image retrieval,” *Image and Vision Computing*, vol. 24,no.3, pp.211–223, 2006, doi:10.1016/j.imavis.2005.11.004.
- [5] T.h. Gevers and A.W.M. Smeulders, “ Image Search Engines An Overview,” University of Amsterdam 1098 SJ Amsterdam, The Netherlands, 2003.
- [6] Y. Rui, T. S. Huang and S. Mehrotra, “Content-based image retrieval with relevance feedback in MARS,” *Proceedings of International Conference on Image Processing*, 1997, pp. 815-818 vol.2, doi: 10.1109/ICIP.1997.638621.
- [7] J.A. Hartigan and M.A.Wong, “Algorithm AS136: A k-means Clustering algorithm,” *Applied Statistic*, vol.28,no.1, pp. 100-108. 1979, doi:10.2307/2346830.
- [8] C. Singh, and K.P. Kaur, “A fast and efficient image retrieval system based on color and texture features,” *Journal of Visual Communication and Image Representation*,vol. 41, pp.225-238. 2016, doi:10.1016/j.jvcir.2016.10.002.
- [9] J.M. Ho et al. “A novel content based image retrieval system using kmeans/knn with feature extraction,” *Computer Science and Information System*, vol.9,no.4,pp.1645-1661,2012,doi:10.1109/ICSAI.2012.6223128.
- [10] J. Mao and A.K. Jain, “Texture Classification and Segmentation using Multi-Resolution Simultaneous

- Autoregressive Models, ” *Pattern Recognition*, vol.25, no.2, pp.173-188, 2010, doi:10.1016/0031-3203(92)90099-5.
- [11] Y. Chen, J. Z. Wang and R. Krovetz, “CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning,” in *IEEE Transactions on Image Processing*, vol.14, no.8, pp. 1187 - 1201, 2005, doi: 10.1109/TIP.2005.849770.
- [12] M. Jalali “Classification Percentage Enhancement of Segmentation Indexed Image based on Clustering Algorithm,” *International journal of engineering & technology sciences* ,vol.1, no. 1, pp. 1-4, 2014.
- [13] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” *Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, pp. 1–14, 2017, doi:10.48550/arXiv.1609.02907.
- [14] R. Zhang and Z. Zhang, “BALAS: Empirical Bayesian learning in the relevance feedback for image retrieval,” *Image and Vision Computing*, vol.24, no.3, pp. 211–223, 2006, doi:10.1016/j.imavis.2005.11.004.
- [15] J. Smith, “Color for Image Retrieval”, *Image Databases: Search and Retrieval of Digital Imagery*, John Wiley & Sons, New York, pp.285-311, 2001.
- [16] A. Amato and V.D. Lecce, “A knowledge based approach for a fast image retrieval system,” *Image and Vision Computing*, vol. 26, no. 11, 2008, doi:10.1016/j.imavis.2008.01.005.
- [17] F. Hajiani, N. Parhizgar, and A. Keshavarz, “Hyperspectral Image Classification Using Low-Rank Representation and Spectral-Spatial Information, ” *Journal of Southern Communication Engineering*, vol.11, no.43, pp.27-38. 2022, (in Persian).

COPYRIGHTS

©2023 by the authors. Published by the Islamic Azad University Bushehr Branch. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0>

