

بررسی کارایی روش‌های مختلف خوشه‌بندی سلسله مراتبی در بازیابی تصویر به شیوه چند پرسشی

اسما شمسی گوشکی حسین نظام‌آبادی پور سعید سریزدی

دانشکده مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران

چکیده

بازیابی تصویر بر اساس محتوا (CBIR) شامل مجموعه‌ای از روش‌ها برای پردازش ویژگی‌های دیداری یک تصویر پرس‌وجو به منظور پیدا کردن تصاویر مشابه آن در یک پایگاه تصویر است. در بازیابی تصویر، بردار پرس‌وجو در فضای جستجو به نحوی حرکت می‌کند که از تصاویر نامرتب دور و به تصاویر مرتبط نزدیک شود. در روش‌های چندپرسشی، ابتدا بردار پرس‌وجو به چند بردار تبدیل شده سپس در فضای جستجو حرکت می‌کند. برای این منظور، تصاویر مرتبط خوشه‌بندی شده و مرکز هر خوشه به عنوان یک پرس‌وجوی جدید در نظر گرفته می‌شود. در این تحقیق روش‌های مختلف خوشه‌بندی سلسله مراتبی در بازیابی تصویر به شیوه چند پرسشی با هم مقایسه شده‌اند. در الگوریتم خوشه‌بندی سلسله مراتبی، برای ترکیب کردن دو خوشه باید میزان شباهت بین آن دو خوشه را محاسبه کرد. روش‌های متفاوت اندازه‌گیری شباهت بین دو خوشه منجر به الگوریتم‌های مختلفی شده است که در این تحقیق کارایی آن‌ها در بازیابی تصویر به شیوه چند پرسشی روی یک پایگاه تصویر شامل ۱۰۰۰۰ تصویر از ۸۲ گروه معنایی با هم مقایسه شده است. نتایج آزمایش‌ها برتری روش UPGMC را نشان می‌دهد.

کلمات کلیدی: بازیابی تصویر، بازخورد ربط، خوشه‌بندی، چند پرسشی، الگوریتم سلسله مراتبی.

۱- مقدمه

شباهت، وزن ویژگی‌های سطح پایین تصویر پرس‌وجو در خلال بازخورد ربط به منظور رسیدن به هدف کاربر تغییر می‌کند [۹، ۱۰] و هدف از بهبود بردار پرس‌وجو، حرکت این بردار در فضای جستجو به نحوی است که از تصاویر نامرتب دور و به تصاویر مرتبط نزدیک شود. این روش اولین بار در سامانه MARS در سال ۱۹۹۷ پیشنهاد شد [۱۱]. در جایجایی بردار پرس‌وجو به یک مکان جدید، فرض می‌شود که همه تصاویر مرتبط در کنار هم قرار دارند سپس میانگین همه نقاط مرتبط حساب شده و به عنوان بردار پرس‌وجوی جدید در نظر گرفته می‌شود. اما اگر تصاویر مرتبط در فضای ویژگی نزدیک هم نباشند، دیگر این روش چندان موثر نیست. به همین دلیل از روش‌های تبدیل بردار پرس‌وجو به چند بردار و جایجایی آن‌ها در فضای جستجو (چند پرسشی) استفاده می‌شود. در روش چند پرسشی، کاربر چندین تصویر را به عنوان تصاویر مرتبط انتخاب می‌کند، سپس سامانه، بردارهای ویژگی تصاویر مرتبط را خوشه‌بندی کرده و مرکز هر خوشه را به عنوان یک پرس‌وجوی جدید در نظر می‌گیرد.

در چند سال گذشته، روش‌ها و سیستم‌های متفاوتی برای بازیابی تصویر بر اساس محتوی^۱ ارائه شده است [۱]. در بازیابی بر اساس محتوی، تصاویر بر اساس ویژگی‌های سطح پایین همچون رنگ، شکل و بافت نمایه‌سازی می‌شوند [۲-۶]. از طرف دیگر، مطلوب کاربرد ویژگی‌های سطح بالای نهفته در تصاویر است. این موضوع باعث شده است که کارایی سیستم‌های بازیابی تصویر محدود شود. بنابراین، تحقیقات اخیر در این زمینه سعی بر تلفیق ویژگی‌های معنایی^۲ و ویژگی‌های سطح پایین^۳ برای بازیابی تصویر دارند [۷، ۸]. استفاده از بازخورد ربط به سیستم کمک می‌کند تا ارتباطی معنی‌دار بین ویژگی‌های سطح پایین و ویژگی‌های سطح بالا^۴ برقرار کند.

بهبود بردار پرس‌وجو^۵ و بهبود تابع شباهت^۶ از اهداف مهم استفاده از بازخورد ربط^۷ است که کمک شایانی به سرعت و دقت سامانه می‌کند. در بهبود تابع

مرحله، وزن مربوط به هر خوشه را جهت استفاده در مراحل بعدی محاسبه می‌کند؛ به اینصورت که وزن هر خوشه به صورت مجموع وزنداری از تصاویر مرتبط قبلی و تصاویر مرتبط مرحله کنونی در آن خوشه در نظر گرفته می‌شود. در این محاسبه، تصاویر مرتبط مرحله کنونی پر اهمیت‌تر فرض شده و وزن بیشتری می‌گیرند.

تحقیق [۱۱] روش MARS را بکار می‌گیرد و آن را با شیوه تبدیل بردار پرس‌وجو به چند بردار بهبود می‌دهد؛ در این تحقیق از روش خوشه‌بندی پیشنهادی در مقاله [۱۹] برای خوشه‌بندی تصاویر مرتبط استفاده می‌کند، بعد از اینکه تصاویر مرتبط خوشه‌بندی شدند، به منظور ترکیب نتایج از ترکیب بردارهای پرس‌وجو استفاده می‌کند و فاصله هر بردار از بردارهای پرس‌وجو به صورت ترکیب وزنداری از فاصله‌ها تعریف می‌شود، وزن هر خوشه برابر با تعداد تصاویر مرتبط در آن خوشه در نظر گرفته می‌شود.

در MARS ابتدا وزن‌های تابع شباهت تصحیح می‌شود، سپس برای بهبود بردار پرس‌وجو، از میانگین نقاط مرتبط استفاده می‌کند. در تحقیق [۲۰] از الگوریتم خوشه‌بندی سلسله مراتبی ward و در [۲۱] از الگوریتم خوشه‌بندی سلسله مراتبی single link برای خوشه‌بندی در روش چندپرسی استفاده شده است.

مرور تحقیقات نشان می‌دهد که بازیابی تصویر به شیوه‌های چندپرسی بر مبنای خوشه‌بندی انجام می‌شود. از سوی دیگر با توجه به ماهیت و ذات خوشه‌بندی سلسله مراتبی تجمعی، این نوع خوشه‌بندی به کرات در این نوع روشها مورد استفاده قرار می‌گیرد. طبقه‌بندی سلسله مراتبی تجمعی در ابتدا هر نمونه را به عنوان یک خوشه در نظر گرفته، سپس در هر مرحله با ادغام دو خوشه‌ای که بیشترین شباهت را به یکدیگر دارند ادامه می‌دهد. تاکنون روشهای زیادی برای تعیین شباهت خوشه‌ها و ادغام آنها در خوشه‌بندی سلسله مراتبی تجمعی ارائه شده است [۲۲]. یقیناً استفاده از روشهای مختلف شباهت در این نوع خوشه‌بندی، نتایج متفاوتی در بازیابی تصویر به همراه دارد. در این تحقیق به منظور یافتن مناسب‌ترین الگوریتم خوشه‌بندی سلسله مراتبی در بازیابی تصویر، روش‌های مختلف خوشه‌بندی سلسله مراتبی در بازیابی تصویر به شیوه چند پرسی با هم مقایسه می‌شوند. ادامه مقاله اینگونه سازماندهی شده است که در بخش ۲ مقدماتی در مورد سامانه بازیابی تصویر ارائه شده است. سپس در بخش ۳ روش چندپرسی و الگوریتم خوشه‌بندی سلسله مراتبی توضیح داده خواهند شد. در بخش ۴ نتایج مقایسه روش‌های مختلف خوشه‌بندی در بازیابی تصویر به شیوه چندپرسی آورده شده است و در بخش ۵ جمع‌بندی پرداخته شده است.

۲- مروری بر یک سامانه بازیابی تصویر

از اوایل ۱۹۹۰ بازیابی تصویر بر اساس محتوی، یک زمینه فعال برای تحقیقات قلمداد می‌شود. تا به حال سامانه‌های بازیابی تجاری و تحقیقاتی فراوانی ایجاد و معرفی شده است [۲۳-۳۲]. در همه این سامانه‌ها ویژگی‌های سطح پایین تصاویر بطور خودکار استخراج شده و تصاویر با آنها نمایه‌سازی می‌شوند. به طور کلی سامانه‌های بازیابی تصویر در دو مرحله عمل می‌کنند. در مرحله اول، ویژگی‌های دیداری تصاویر پایگاه به صورت خودکار استخراج شده و تصاویر با آنها نمایه‌سازی می‌شوند.

در مرحله دوم، پس از دریافت پرس‌وجوی کاربر، ویژگی‌های سطح پایین یا ویژگی‌های دیداری آن استخراج شده و پایگاه ویژگی‌های دیداری برای یافتن نزدیک‌ترین تصاویر به پرس‌وجو، جستجو می‌شود. در یک سامانه بازیابی تصویر، برای اجرای مرحله دوم به یک زیر سامانه برای پردازش تصویر پرس‌وجو و استخراج ویژگی‌های سطح پایین آن نیاز است. ساختار یک سامانه بازیابی تصویر

تحقیق [۱۲] از روش چندپرسی استفاده می‌کند و الگوریتم خوشه-Q را پیشنهاد می‌کند. در این الگوریتم، ابتدا یک تصویر پرس‌وجو از سوی کاربر به سامانه داده می‌شود؛ سپس سامانه k تصویری که فاصله کمتری از تصویر پرس‌وجو دارند را به کاربر نشان می‌دهد. بعد از اینکه کاربر تصاویر مرتبط را مشخص کرد، سامانه در مرحله اول برای خوشه‌بندی تصاویر مرتبط از خوشه‌بندی سلسله مراتبی استفاده می‌کند؛ سپس در مراحل بعد با استفاده از تابع خوشه‌بندی بیز [۱۳]، هر تصویر مرتبط را در یکی از دسته‌های قبلی قرار داده یا به یک دسته جدید نسبت می‌دهد. همچنین در هر مرحله خوشه‌هایی که شباهت بالایی به هم دارند با استفاده از روش Hotelling's T2 [۱۴] با هم ترکیب می‌شوند و در نهایت مرکز خوشه‌های باقیمانده به عنوان پرس‌وجوهای جدید در نظر گرفته می‌شوند. در تحقیق مذکور، فاصله هر بردار از بردارهای پرس‌وجو به صورت ترکیب وزنداری از فاصله‌ها تعریف می‌شود و وزن مربوط به هر خوشه برابر با مجموع رتبه تصاویر مرتبط در آن خوشه تعریف می‌شود.

روش k-بهترین^۱ [۱۵] نیز از شیوه چندپرسی برای بازیابی تصویر استفاده می‌کند. در این روش به جای اینکه بردارهای پرس‌وجو با هم ترکیب شوند و سپس پایگاه تصاویر جستجو شود، ابتدا پایگاه تصاویر بر حسب هر کدام از بردارهای پرس‌وجو جستجو شده و نتایج بازگشتی از آنها به دوصورت با هم ترکیب می‌شود. الف) k تصویر با رتبه بالاتر از لیست بازیابی شده تصاویر، انتخاب و به کاربر نشان داده می‌شوند. ب) k تصویر مرتبط به صورت اتفاقی از لیست بازیابی شده تصاویر، انتخاب و به کاربر نشان داده می‌شوند.

در تحقیق [۱۶] با استفاده از ترکیب روش‌های چندپرسی و بهبود تابع شباهت، دو روش را در بازخورد ربط ارائه کرده است. در روش CR^۱ در هر مرحله تصاویر مرتبط را خوشه‌بندی کرده و مرکز هر خوشه را به عنوان یک پرس‌وجوی جدید در نظر می‌گیرد؛ سپس با استفاده از طبقه‌بند k-همسایه نزدیکتر، تصاویر نامرتب را به مراکز خوشه‌های مرحله قبل دسته‌بندی می‌کند. در نهایت در هر خوشه با استفاده از تصاویر مرتبط و نامرتب مربوط به آن خوشه، عمل بهبود بردار پرس‌وجو را انجام می‌دهد به اینصورت که بردار پرس‌وجوی جدید مربوط به هر خوشه را به صورت میانگین تصاویر مرتبط منهای میانگین تصاویر نامرتب مربوط به آن خوشه در نظر می‌گیرد.

در روش CNR^۱، در مرحله اول بعد از اینکه کاربر تصاویر مرتبط و نامرتب را مشخص کرد، تصاویر مرتبط را خوشه‌بندی می‌کند سپس در سایر مراحل تمام تصاویر مرتبط و نامرتب را به مراکز این خوشه‌ها (خوشه‌های مرحله اول) با استفاده از طبقه‌بند k-همسایه نزدیکتر طبقه‌بندی می‌کند و با استفاده از تصاویر مرتبط و نامرتب در هر خوشه به بهبود بردار پرس‌وجو می‌پردازد. به اینصورت که در هر مرحله برای هر خوشه میانگین تصاویر مرتبط را به پرس‌وجوی قبلی اضافه و میانگین تصاویر نامرتب را از آن کم می‌کند. در تحقیق [۱۷] از روش چند پرسی در فضای ویژگی‌های متعامد استفاده کرده است.

در تحقیق [۱۸] از ناحیه‌بندی تصاویر و روش چندپرسی با خوشه‌بندی و فقی برای بازیابی تصویر استفاده می‌شود. در مرحله اول، بعد از اینکه تصاویر مرتبط توسط کاربر مشخص شدند، سامانه تصاویر مرتبط را ناحیه‌بندی کرده و ویژگی‌های سطح پایین هر ناحیه را استخراج می‌کند؛ سپس این نواحی را با استفاده از روش خوشه‌بندی سلسله مراتبی خوشه‌بندی کرده و برای هر خوشه یک وزن در نظر می‌گیرد. در مرحله اول، وزن هر خوشه را برابر با تعداد تصاویر مرتبط در آن خوشه در نظر گرفته و از روش Hotelling's T2 [۱۴] با استفاده از معیار ماهالونوبیس^{۱۲} استفاده کرده و خوشه‌های شبیه به هم را با هم ترکیب می‌کند. سپس مرکز هر خوشه را به عنوان یک پرس‌وجوی جدید در نظر می‌گیرد. در مراحل بعدی، با استفاده از الگوریتم Hotelling's T2 با معیار ماهالونوبیس و وزن‌های خوشه‌های مرحله قبل، هر تصویر مرتبط را به یکی از خوشه‌های قدیمی تخصیص داده و یا خودش را به عنوان یک خوشه‌ی جدید در نظر می‌گیرد. در هر

۳- بازیابی تصویر با روش چندپرسی

در این روش، ابتدا کاربر یک تصویر را به عنوان مثال به سامانه ارائه می‌کند. سامانه ویژگی‌های سطح پایین آن را استخراج کرده و بازیابی بر اساس بردار ویژگی آن تصویر شروع می‌شود. سامانه تصاویر بازیابی شده را به کاربر نشان داده و کاربر تصاویر مرتبط را مشخص می‌کند. در مرحله بعد، برای بازیابی از روش چندپرسی استفاده می‌شود. به این صورت که تصاویر مرتبط خوشه‌بندی شده و مرکز هر خوشه به عنوان یک پرس‌وجوی جدید در نظر گرفته می‌شود. سپس پایگاه داده با هر کدام از بردارهای پرس‌وجو جستجو شده و نتایج حاصل از آن‌ها با هم ترکیب می‌شوند. بعد از جستجوی پایگاه تصاویر با مرکز هر خوشه، برای هر تصویر پایگاه، کمترین فاصله از مراکز خوشه‌ها به عنوان میزان شباهت تصویر به تصویر پرس‌وجو در نظر گرفته شده و تصاویر پایگاه بر اساس این مقادیر مرتب شده و تصاویر با شباهت بیشتر به عنوان تصاویر مرتبط به کاربر نشان داده می‌شوند. این تعامل تا رسیدن به هدف کاربر ادامه می‌یابد. در این روش عموماً برای خوشه‌بندی تصاویر مرتبط از الگوریتم خوشه‌بندی سلسله مراتبی تجمعی استفاده می‌شود [۲۱، ۲۰، ۱۲].

۳-۱- خوشه‌بندی

خوشه‌بندی داده‌ها به ما اجازه می‌دهد تا به منظور درک بهتر، داده‌های مشابه و متفاوت را از هم جدا کنیم [۳۴]. در خوشه‌بندی، هر داده به عنوان یک نقطه در فضای ویژگی در نظر گرفته می‌شود. در محبت بازشناسی الگو، نواحی پیوسته‌ای از فضا که شامل نقاط با چگالی بالا هستند را یک خوشه می‌گویند. داده‌ها باید طوری خوشه‌بندی شوند که هر داده، قطعاً به یک خوشه تعلق داشته باشد. داده‌ها در یک خوشه به هم شبیه باشند و هیچ دو خوشه‌ای با هم اشتراک نداشته باشند [۲۲]. برای بیان مطلب به زبان ریاضی فرض کنید $X = \{x_1, x_2, \dots, x_N\}$ مجموعه داده‌هایی است که باید خوشه‌بندی شوند و R یک افراز شامل m خوشه از X است $(R = \{C_1, C_2, \dots, C_m\})$. در چنین شرایطی:

$$C_i \neq \phi \quad i = 1, 2, \dots, m \quad (1)$$

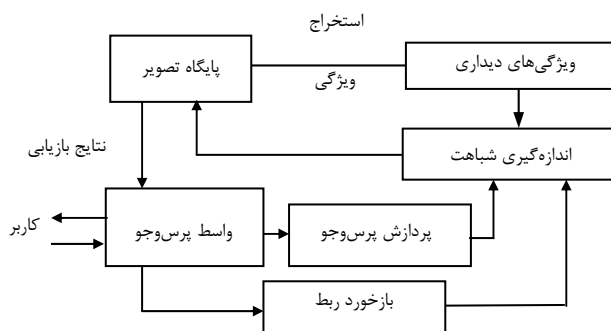
$$\bigcup_{i=1}^m C_i = X \quad i = 1, 2, \dots, m \quad (2)$$

$$C_i \cap C_j = \phi \quad i \neq j \quad i, j = 1, 2, \dots, m \quad (3)$$

۳-۲- خوشه‌بندی سلسله مراتبی تجمعی

در الگوریتم خوشه‌بندی سلسله مراتبی تجمعی [۲۲]، در مرحله اول، هر داده (در مساله مورد نظر بردار ویژگی مربوط به هر تصویر مرتبط به عنوان یک داده در خوشه‌بندی تجمعی است) به عنوان یک خوشه در نظر گرفته می‌شود (N تصویر و N خوشه). سپس در هر مرحله، شباهت بین هر دو خوشه محاسبه می‌شود و دو خوشه‌ای که به هم شبیه‌تر هستند با هم ترکیب شده و یک خوشه جدید را تشکیل می‌دهند. پس در مرحله دوم، $N-1$ خوشه بدست می‌آید و این روند ادامه پیدا می‌کند تا در مرحله N ام، یک خوشه که شامل همه تصاویر مرتبط است، تشکیل شود. حال برای روشن‌تر شدن مساله به یک مثال می‌پردازیم: فرض کنید $X = \{x_1, x_2, x_3, x_4, x_5\}$ مجموعه داده‌هایی است که باید خوشه‌بندی شوند. در شکل ۲ مراحل خوشه‌بندی نمایش داده شده است. در مرحله اول، ۵ خوشه

در شکل ۱ به تصویر کشیده شده است. در ادامه به طور مختصر بخش‌های مختلف این سامانه تشریح می‌شوند.



شکل ۱- ساختار سامانه پیشنهادی تعاملی بازیابی تصویر [۳۲]

پایگاه تصویر شامل تصاویری است که بازیابی در بین آن‌ها انجام می‌شود. نوع تصاویر پایگاه می‌تواند بر سایر بخش‌های سامانه بازیابی تصویر تاثیر بگذارد. به عنوان مثال سامانه‌ای که صرفاً برای تصاویر پزشکی خاص یا بازیابی لوگوهای تجاری استفاده می‌شود از نظر استخراج ویژگی با سامانه‌هایی که برای بازیابی از پایگاه‌های تصاویر عام به کار می‌روند، تفاوت بسیار دارند. در این مقاله فرض شده است که سامانه برای بازیابی از پایگاه‌های تصاویر عام مورد استفاده قرار می‌گیرد. از هر یک از تصاویر پایگاه تعدادی ویژگی دیداری استخراج می‌شود. به طور معمول در بازیابی تصویر لازم است که ویژگی‌های مناسبی برای توصیف خواص رنگ، شکل و بافت تصاویر استخراج شود. ویژگی‌های استخراج شده از تصاویر در پایگاه ویژگی‌های سطح پایین نگهداری می‌شوند.

این سامانه شامل یک رابط گرافیکی یا واسط پرس‌وجو است که با استفاده از آن با کاربر ارتباط برقرار کرده و با او در تعامل است. رابط گرافیکی باید دو عمل را انجام دهد. در وهله‌ی نخست باید اطلاعات کاربر از قبیل تصویر پرس‌وجو را در ورودی اخذ کرده و به سامانه تحویل دهد. سپس در مراحل بعدی باید تصاویر بازیابی شده توسط سامانه را به کاربر ارائه کرده و اطلاعات مربوط به نظر کاربر راجع به آن‌ها را اخذ و به سامانه تحویل نماید. در سامانه‌های بازیابی، روش‌های مختلفی برای ارائه پرس‌وجو وجود دارد که مهم‌ترین آن‌ها پرس‌وجو با ارائه مثال است.

زیر سامانه پردازش پرس‌وجو، ویژگی‌های مناسب را از تصویر پرس‌وجو استخراج می‌کند تا سامانه قادر به یافتن شبیه‌ترین تصاویر به آن از طریق مقایسه با پایگاه ویژگی‌های سطح پایین باشد. بنابراین این زیرسامانه، زیرسامانه استخراج ویژگی را به خدمت می‌گیرد تا اطلاعات دیداری لازم برای نمایه‌سازی تصویر پرس‌وجو را از آن استخراج کند. زیر سامانه اندازه‌گیری شباهت، شباهت بین بردار ویژگی تصویر پرس‌وجو و بردارهای ویژگی تصاویر پایگاه را محاسبه کرده و سامانه نزدیکترین تصاویر به تصویر پرس‌وجو را پیدا می‌کند. این تصاویر از طریق رابط گرافیکی به کاربر ارائه می‌شوند و با دریافت نظر کاربر راجع به آن‌ها، تعامل با کاربر در قالب بازخورد ربط و با استفاده از یادگیری کوتاه‌مدت تا رسیدن به تصاویر دلخواه او ادامه می‌یابد. در سامانه پیاده‌سازی شده برای یادگیری کوتاه‌مدت از روش تبدیل بردار پرس‌وجو به چند بردار و جایجایی آن‌ها در فضای جستجو استفاده شده است. زیر سامانه یادگیری کوتاه مدت، اطلاعات کاربر را در خصوص ارتباط تصاویر ارائه شده به پرس‌وجو، بررسی کرده و بر مبنای آن‌ها با استفاده از ابزارهای یادگیری ماشینی یا مجموعه‌ای از قواعد شهودی به تبدیل بردار پرس‌وجو به چند بردار و جایجایی آن‌ها در فضای جستجو به گونه‌ای که کاربر را گام به گام به تصاویر دلخواهش نزدیک کند، می‌پردازد.

قبل) تعریف می‌شود. از آنجا که در این رابطه، فاصله‌های بین خوشه جدید و زیرخوشه‌ها با ضریب ثابت با هم ترکیب شده‌اند، بنابراین زیرخوشه‌هایی که اعضای کمتری دارند وزن بیشتری و زیرخوشه‌هایی که اعضای بیشتری دارند وزن کمتری در تابع شباهت می‌گیرند. به همین دلیل این فاصله را فاصله وزن‌دار می‌گویند. در این الگوریتم‌ها، برای بدست آوردن فاصله بین اعضای دو خوشه (C_i, C_s) و $d(C_j, C_s)$ از فاصله اقلیدسی (رابطه (۶)) استفاده می‌شود.

$$d(C_j, C_s) = \left(\sum_{i=1}^L |C_{ji} - C_{si}|^2 \right)^{\frac{1}{2}} \quad (6)$$

که در این رابطه، C_j و C_s اعضای دو خوشه، L بیانگر طول بردار هر عضو و C_{si} و C_{ji} مولفه i ام از بردارهای C_j و C_s هستند.

اگر میانگین نسبی یا مرکز ثقل یک خوشه را $m_q = \frac{1}{n_q} \sum_{x \in C_q} x$ و فاصله بین دو خوشه را به صورت فاصله اقلیدسی بین مرکز ثقل دو خوشه یا $d_{qs} = \|m_q - m_s\|_2$ تعریف کنیم، الگوریتم‌های 15 UPGMC و 16 WPGMC بر مبنای این فاصله معرفی می‌شوند. روابط این الگوریتم‌ها در جدول ۲ آمده است. در نهایت، اگر فاصله بین دو خوشه C_j و C_i طبق رابطه (۷)، d'_{ij} تعریف شود.

$$d'_{ij}{}^2 = \frac{n_i n_j}{n_i + n_j} d_{ij}{}^2 \quad (7)$$

در اینصورت، الگوریتم Ward با تابع شباهت رابطه (۸) تعریف می‌شود.

$$d'_{sq} = \frac{n_i + n_s}{n_i + n_j + n_s} d'_{is} + \frac{n_j + n_s}{n_i + n_j + n_s} d'_{js} - \frac{n_s}{n_i + n_j + n_s} d'_{ij} \quad (8)$$

بنابراین، در این تحقیق، اثر ۷ نوع الگوریتم خوشه‌بندی تجمعی در یادگیری کوتاه مدت چندپرسشی در یک سیستم بازیابی تصویر بررسی خواهد شد.

۴- آزمایش‌ها و نتایج

در این مقاله از روش چندپرسشی با الگوریتم خوشه‌بندی سلسله مراتبی تجمعی در بازیابی تصویر استفاده شده است. همچنین برای تخمین تعداد مناسب خوشه‌ها از روش پیشنهادی در [۳۶] و به منظور اندازه‌گیری شباهت بین دو بردار ویژگی از تابع شباهت رابطه (۹) که به معیار χ^2 معروف است، استفاده شده است. انتخاب این معیار بر مبنای تحقیقاتی است که در این زمینه انجام شده است [۳۷].

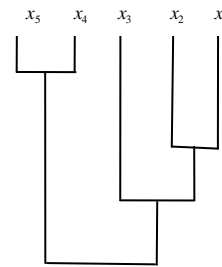
$$d(f_i, f_j) = \sum_i \left(\frac{f_{i,l} - f_{j,l}}{f_{i,l} + f_{j,l}} \right)^2 \quad (9)$$

که در این رابطه، L بیانگر طول بردار ویژگی و f_i و f_j به ترتیب بیانگر بردار ویژگی تصویر i ام و j ام هستند. $f_{i,l}$ بیانگر مولفه l ام از بردار ویژگی تصویر i ام است.

۴-۱- پایگاه تصویر

در این تحقیق از یک پایگاه تصویر عام استفاده شده است. برای تشکیل پایگاه از

وجود دارد که هر کدام شامل یک داده از X است. سپس فاصله بین هر جفت خوشه محاسبه شده و جفت خوشه‌ای که فاصله کمتری دارند (به عنوان مثال x_1 و x_2) با هم ترکیب می‌شوند. در مرحله سوم x_4 و x_5 با هم ترکیب می‌شوند (با فرض اینکه در این مرحله، این دو خوشه کمترین فاصله را از یکدیگر دارند). سپس x_3 به خوشه $\{x_1, x_2\}$ می‌پیوندد. و در نهایت در مرحله پنجم از خوشه‌بندی، خوشه‌های $\{x_1, x_2, x_3\}$ و $\{x_4, x_5\}$ با هم ترکیب شده و خوشه‌ای شامل همه مجموعه X را تشکیل می‌دهند.



شکل ۲- دندروگرام مربوط به خوشه‌بندی تجمعی: خوشه‌بندی داده‌ها در ۵ مرحله بوسیله الگوریتم خوشه‌بندی سلسله مراتبی تجمعی

لازم به ذکر است که در خوشه‌بندی تجمعی با آستانه‌زنی دندروگرام، آرایش خوشه‌ها و تعداد آنها مشخص می‌شود. در یک خوشه‌بندی تجمعی، فرض کنید خوشه C_q حاصل ترکیب دو خوشه C_j و C_i است. در این حالت، فاصله بین خوشه C_q و یک خوشه قدیمی، C_s ، تابعی به شکل رابطه (۴) است.

$$d(C_s, C_q) = f(d(C_i, C_s), d(C_j, C_s), d(C_i, C_j)) \quad (4)$$

در این رابطه، f متریکی جهت اندازه‌گیری عدم شباهت و d فاصله بین دو خوشه است. این تابع در مرجع [۳۵] به شکل رابطه (۵) معرفی شده است:

$$d(C_s, C_q) = a_i d(C_i, C_s) + a_j d(C_j, C_s) + b d(C_i, C_j) + c |d(C_i, C_s) - d(C_j, C_s)| \quad (5)$$

به ازای مقادیر مختلف a_i, a_j, b, c در این رابطه، الگوریتم‌های مختلفی برای اندازه‌گیری شباهت بین دو خوشه حاصل می‌شود. الگوریتم‌های single link و complete link که در جدول ۱ معرفی شده‌اند از مشهورترین این الگوریتم‌ها هستند. در الگوریتم single link، شباهت بین دو خوشه به صورت فاصله بین نزدیکترین اعضای دو خوشه تعریف می‌شود و دو خوشه‌ای که شباهت بیشتری با هم داشته باشند با هم ترکیب می‌شوند و در الگوریتم complete link، شباهت بین دو خوشه به صورت فاصله بین دورترین اعضای دو خوشه تعریف می‌شود.

اگر تعداد عناصر خوشه‌های C_j و C_i را با n_i, n_j نشان دهیم، دو الگوریتم 14 UPGMA و 13 WPGMA بر این اساس ارائه می‌شوند. (روابط این الگوریتم‌ها در جدول ۱ آمده است). در UPGMA شباهت بین دو خوشه به صورت میانگین غیروزنداری از فاصله بین خوشه جدید و زیر خوشه‌های قدیمی (مرحله قبل) تعریف می‌شود. از آنجا که در این رابطه، فاصله‌های بین خوشه جدید و زیرخوشه‌ها با ضریبی از تعداد اعضای تشکیل دهنده زیرخوشه‌ها با هم ترکیب شده‌اند، هر زیرخوشه به اندازه تعداد اعضایش در معیار شباهت نقش دارد و نه بیشتر، بنابراین این فاصله را غیر وزندار می‌نامند. اما در WPGMA شباهت بین دو خوشه به صورت میانگین وزنداری از فاصله بین خوشه جدید و زیرخوشه‌های قدیمی (مرحله

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \times \exp(j2\pi u_0 x) \quad (10)$$

$$G(u, v) = \exp\left(-\frac{1}{2}\left(\frac{(u-u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right) \quad (11)$$

که در این روابط u_0 ، فرکانس مرکزی موج سینوسی را در راستای محور x مشخص می‌کند. σ_x و σ_y به ترتیب انحراف معیار تابع پوش گوسی در راستای محورهای

$$x \text{ و } y \text{ و } \sigma_u = \frac{1}{2\pi\sigma_x} \text{ و } \sigma_v = \frac{1}{2\pi\sigma_y} \text{ به ترتیب انحراف معیار تابع گوسی در}$$

حوزه فرکانس در راستای محورهای u و v هستند.

موجک‌های گابور را می‌توان از طریق گسترش و چرخش موجک مادر $g(x, y)$ به صورت رابطه (۱۲) بدست آورد:

$$\begin{aligned} g_{m,n}(x, y) &= a^{-m} g(x', y') & a > 1, m = 0, 1, 2, \dots, s-1 & \quad (12) \\ y' &= a^{-m} (-x \sin\theta + y \cos\theta) & n = 0, 1, 2, \dots, k-1 & \\ x' &= a^{-m} (x \cos\theta + y \sin\theta) & \theta = \frac{n\pi}{k} & \end{aligned}$$

که در آن $g_{m,n}$ نسخه تغییر مقیاس و چرخش یافته‌ی فیلتر مادر $g(x, y)$ ، a پارامتر مقیاس، k تعداد کل جهت‌ها ۱۷، s تعداد کل مقیاس‌ها ۱۸ و m و n به ترتیب مشخص کننده مقیاس و جهت هر فیلتر می‌باشند. x' و y' نیز مختصات تبدیل یافته هستند. با استفاده از این فیلترها، از هر تصویر ورودی، تعدادی تصویر فیلتر شده استخراج می‌شود که هر کدام از این تصاویر، شامل تغییرات شدت روشنایی در یک محدوده باریک از فرکانس و جهت می‌باشد.

در این تحقیق، از فیلترهای گابور با ۵ مقیاس و ۶ جهت ($s=5$ و $k=6$)، در مجموع ۳۰ فیلتر، بر اساس توصیه‌ی استاندارد MPEG-7 استفاده شده است [۴۶].

هر تصویر از ۲۰ کانال فیلتر گابور عبور کرده و تصاویر فیلتر شده آن از رابطه (۱۳) محاسبه شدند.

$$I_{m,n}(x, y) = f^{-1}[I(u, v)G_{m,n}(u, v)] \quad (13)$$

که در آن $I_{m,n}(x, y)$ تصویر فیلتر شده با فیلتر در جهت m و مقیاس n ، $I(u, v)$ تبدیل فوریه تصویر و $G_{m,n}(u, v)$ تبدیل فوریه $g_{m,n}(x, y)$ است. از هر تصویر فیلتر شده دو پارامتر میانگین و انحراف معیار، مطابق رابطه (۱۴) به عنوان ویژگی استخراج شدند.

$$\begin{aligned} \mu_{m,n} &= \iint I_{m,n}(x, y) dx dy & (14) \\ \sigma_{m,n} &= \sqrt{\iint (I_{m,n}(x, y) - \mu_{m,n})^2 dx dy} \end{aligned}$$

که در آن $\mu_{m,n}$ و $\sigma_{m,n}$ به ترتیب میانگین و انحراف معیار تصویر فیلتر شده با فیلتر در جهت m و مقیاس n است. در نهایت، یک بردار ویژگی ۶۰ بعدی متناظر با تصویر ورودی مطابق رابطه (۱۵) ساخته می‌شود.

$$\bar{f} = [\mu_{00}, \sigma_{0,0}, \mu_{01}, \sigma_{0,1}, \dots, \mu_{5,6}, \sigma_{5,6}] \quad (15)$$

سه پایگاه تصویر مشهور که در بسیاری از تحقیقات به آن‌ها ارجاع شده، استفاده شده است: الف) پایگاه داده ۲۰۰۰۰ تصویری ImageCLEF 2007 photo [۳۸] (ب) پایگاه داده MSRC، شامل ۴۳۲۰ تصویر با ۳۳ گروه معنایی [۳۹] و (ج) پایگاه تصویر SIMPLisity شامل ۱۰۰۰ تصویر با ۱۰ گروه معنایی [۴۰]. از مجموع ۲۵۳۲۰ تصویر این سه پایگاه تصویر، تعداد ۱۰۰۰۰ تصویر انتخاب و پایگاه تصویر تشکیل شد. تمام تصاویر پایگاه، رنگی و درحوزه فشرده JPEG هستند. تصاویر انتخاب شده در ۸۲ گروه معنایی اصلی قرار دارند.

۲-۴- ویژگی‌های پایه

در این مقاله از سه نوع ویژگی برای نمایه‌سازی تصویر استفاده شده است. ویژگی‌های پیاده‌سازی شده عبارتند از: هیستوگرام رنگ به نمایندگی از ویژگی رنگ، هیستوگرام لبه به نمایندگی از ویژگی شکل و ویژگی گابور به نمایندگی از ویژگی بافت.

هیستوگرام رنگ

رنگ ویژگی‌ای است که اغلب محققان در بازیابی تصویر از آن استفاده کرده‌اند و در موارد متعدد هیستوگرام رنگ به عنوان اصلی‌ترین ویژگی در بازیابی مورد توجه بوده است. فوایدی همچون پایداری، موثر بودن، سادگی پیاده‌سازی، سادگی محاسباتی و حجم پایین بردار حاصل برای نمایه‌سازی، استفاده از هیستوگرام رنگ را توجیه می‌کند.

فضای رنگ RGB با سیستم بینایی انسان و روشی که انسان تصویر را درک می‌کند، همخوانی ندارد و برای نمایش تصویر توسط مانیتور و تلویزیون مناسب است. در این فضا تفاوت دو نقطه با توجه به مختصاتشان، متناسب با درک انسان نیست و فاصله اقلیدسی، تفاوت بین دو رنگ را به شکلی که برای انسان معنی‌دار باشد، بیان نمی‌کند. ما در این تحقیق از فضای رنگ HSV استفاده کرده‌ایم. این فضای رنگ در تحقیقات مشابه به دفعات استفاده شده است [۳۳، ۴۱، ۴۲].

در این تحقیق برای تشکیل هیستوگرام رنگ، فضای HSV بطور خطی چندی می‌شود. در این فضا بنا بر اهمیتی که مولفه H نسبت به سایر مولفه‌ها دارد، این مولفه به شش بازه و دو مولفه دیگر هر کدام به سه بازه چندی می‌شوند [۳۳، ۴۱، ۴۲] (در کل تعداد ۵۴ بازه به سطوح رنگ اختصاص داده شده است). علاوه بر آن ۶ بازه برای سطوح خاکستری در نظر گرفته می‌شود. از آنجا که در فضای رنگ HSV، پیکسل‌هایی که مقادیر S و V آن‌ها کمتر از ۰/۲ است رنگی به نظر نمی‌آیند و مثل پیکسل‌های با سطوح خاکستری هستند، این بخش از فضای رنگ به عنوان سطوح خاکستری در نظر گرفته شده و به ۶ سطح حول محور V به طور خطی و یکپارچه چندی می‌شود [۴۳]. بعد از چندی کردن رنگ نقاط یک تصویر، هیستوگرام رنگ آن با شمردن نقاطی که در هر بازه قرار می‌گیرند محاسبه می‌شود. برای حل مسئله تفاوت اندازه تصاویر، هیستوگرام رنگ به تعداد کل نقاط آن تصویر نرمالیزه می‌شود. بردار حاصل از نمایه‌سازی هر تصویر با استفاده از ویژگی رنگ، ۶۰ بعدی است.

ویژگی‌های گابور

فیلترهای گابور به علت خاصیت تفکیک پذیری خوب و چندگانه در حوزه‌های مکان و فرکانس به عنوان ابزاری قوی در تحلیل بافت، بکار می‌روند [۴۴، ۴۵]. تابع گابور دو بعدی و تبدیل فوریه آن طبق روابط (۱۰) و (۱۱) معرفی می‌شوند [۴۶، ۴۷].

هیستوگرام لبه

محاسبه می‌شوند. و در نهایت یکی از گراف‌های دقت بر حسب تعداد مراحل بازیابی، فراخوانی بر حسب تعداد مراحل بازیابی، گراف دقت_فراخوانی در یک مرحله خاص، معیار دقت بر حسب تعداد تصاویر بازیابی شده و معیار فراخوان بر حسب تعداد تصاویر بازیابی شده استفاده می‌شود. معیار دقت، بیانگر نسبت تصاویر مرتبط بازیابی شده به تصاویر بازیابی شده و معیار فراخوانی بیانگر نسبت تصاویر بازیابی شده به کل تصاویر مرتبط درون پایگاه است [۵۰].

در بعضی از تحقیقات از معیارهای دیگری بر پایه دقت و فراخوانی استفاده شده است: $[20]P(10)$ ، $[51]P(30)$ ، $[52]P(NR)$: دقت بعد از ۱۰، ۳۰ یا NR تصویر بالای بازیابی شده، که NR تعداد تصاویر مرتبط پایگاه را نشان می‌دهد. معیار میانگین رتبه‌های تصاویر مرتبط در [۱] استفاده شده است در این مقاله از گراف دقت بر حسب تعداد مراحل بازیابی شده و گراف دقت بر حسب تعداد تصاویر بازیابی شده به منظور ارزیابی سامانه استفاده شده است.

۴-۴- نتایج و مقایسه

برای مقایسه روش‌های مختلف خوشه‌بندی سلسه مراتبی، ۲۵ تصویر پرس‌وجو از ۲۵ گروه معنایی مختلف انتخاب و به سامانه ارائه شدند. تصاویر مورد پرس‌وجو در شکل ۳ نشان داده شده‌اند. از کاربر درخواست شد که هر نشست بازیابی را تا ۸ مرحله جلو برده و با سامانه تعامل کند. در هر مرحله تعداد تصاویر مرتبط با معنای کاربر ثبت می‌شود و بر مبنای تعداد تصاویر بازیابی شده مرتبط، برای هر تصویر گراف دقت بر حسب تعداد تکرارها بدست می‌آید. این عمل برای تمام ۲۵ تصویر پرس‌وجوی آزمون انجام شده و میانگین دقت بر حسب تکرار محاسبه و رسم می‌شود. به منظور بررسی اثر تعداد تصاویر بازیابی شده، آزمایش‌ها در تعداد مختلفی از تصاویر بازیابی شده (۱۰، ۲۰، ۳۰ و ۴۰) انجام و نتایج ارائه شدند.

در یک آزمایش، اثر وزن در الگوریتم‌های خوشه‌بندی سلسه مراتبی مورد ارزیابی قرار گرفت. هدف این بود که مشخص شود الگوریتم‌های تابع شباهت وزندار (WPGMC و WPGMA) بهتر عمل می‌کنند یا الگوریتم‌های تابع شباهت بدون وزن (UPGMC و UPGMA). به این منظور الگوریتم WPGMA با UPGMA و الگوریتم WPGMC با UPGMC مقایسه شدند. شکل ۴ گراف دقت بر حسب تعداد تکرار و گراف دقت بر حسب تعداد تصاویر بازیابی شده در تکرار هشتم را برای روش‌های WPGMA و UPGMA و شکل ۵ گراف دقت بر حسب تعداد تکرار و گراف دقت بر حسب تعداد تصاویر بازیابی شده در تکرار هشتم را برای روش‌های WPGMC و UPGMC نشان می‌دهند.

همانطور که در شکل‌های ۴-ب و ۵-ب مشاهده می‌شود، با افزایش تعداد تصاویر بازیابی شده که همان تصاویر شرکت کننده در خوشه‌بندی می‌باشند، الگوریتم‌های بدون وزن بهتر عمل می‌کنند. در واقع در این حالت، تعداد تصاویر موجود در هر خوشه اهمیت پیدا کرده و الگوریتم‌های بدون وزن که در آن‌ها هر خوشه به اندازه تعداد اعضایش در اندازه‌گیری شباهت بین دو خوشه تاثیر دارد، بهتر عمل می‌کنند. همچنین با توجه به شکل‌های ۴-الف و ۵-الف، با افزایش تعداد تکرارها، اختلاف بین دو الگوریتم مورد مقایسه زیاد می‌شود. با افزایش تعداد تکرار، به دلیل اینکه در هر مرحله تصاویر مرتبط مراحل قبل نیز در خوشه‌بندی شرکت می‌کنند، تعداد تصاویر و یا داده‌های شرکت کننده در خوشه‌بندی افزایش یافته و کارایی روش‌های بدون وزن نیز افزایش می‌یابد.

در الگوریتم‌های بدون وزن، اگر فاصله بین خوشه C_i از یکی از زیرخوشه‌های C_j که تعداد اعضای کمی دارد زیاد باشد، اثر این فاصله در محاسبه معیار شباهت بی‌تاثیر می‌شود و عملاً فاصله بین خوشه C_i با زیرخوشه‌ی با اعضای بیشتر در C_j تعیین کننده شباهت بین دو خوشه C_i و C_j است. این موضوع باعث می‌شود خوشه C_i با خوشه‌ای ترکیب شود که به زیرخوشه با تعداد اعضای بالاتر آن

لبه، یک ویژگی مهم در بیان محتوای تصویر است. در استاندارد MPEG-7 برای نمایش توزیع لبه در تصویر، از هیستوگرام لبه استفاده می‌شود. هیستوگرام توزیع لبه‌های محلی ۱۹ با ۸۰ کلاس در MPEG-7 به عنوان توصیف‌گر هیستوگرام لبه استاندارد شده است. برای بالا بردن کارایی بازیابی به توزیع لبه در کل تصویر نیاز است زیرا توزیع محلی لبه‌ها برای بیان ویژگی‌های توزیع لبه در کل تصویر کافی نیستند. بنابراین، دو نوع توزیع لبه دیگر نیز برای تصویر معرفی شده است: (الف) توزیع لبه سراسری (ب) توزیع لبه نیمه سراسری. برای بدست آوردن توزیع لبه محلی، تصویر به ۱۶ زیر تصویر تقسیم می‌شود.

در استاندارد MPEG-7، ۵ نوع لبه، شامل یک لبه بدون جهت و ۴ لبه جهت‌دار در جهت‌های افقی ۲۱، عمودی ۲۲، ۴۵ درجه و ۱۳۵ درجه برای تشکیل هیستوگرام لبه تعریف شده‌اند. برای استخراج ویژگی‌های لبه جهت‌دار و بدون جهت، هر تصویر به زیرتصویرهای مربعی ناهمپوشان تقسیم می‌شود. صرف نظر از اندازه تصویر، هر زیر تصویر به تعداد ثابتی بلوک تقسیم می‌شود. بنابراین، اندازه بلوک‌ها به اندازه تصویر بستگی دارد. اندازه هر بلوک باید ضریبی از ۲ باشد. در غیر این صورت از تعدادی از پیکسل‌ها صرف نظر می‌شود. اگر اندازه تصویر $P \times Q$ باشد، اندازه هر بلوک طبق رابطه (۱۶) بدست می‌آید:

$$x = \sqrt{\frac{p \times q}{desired_num_block}} \quad block_size = \left\lfloor \frac{x}{2} \right\rfloor \times 2 \quad (16)$$

عددی ثابت و متناسب با کیفیت $desired_num_block$ که در آن تصویر است.

برای استخراج ویژگی لبه از هر بلوک، هر بلوک به ۴ زیربلوک تقسیم می‌شود. سپس میانگین شدت روشنایی هر زیربلوک محاسبه شده و این مقادیر در ضرایب فیلترهای مربوط به هر لبه ضرب می‌شوند تا اندازه هر نوع لبه بدست آید. بنابراین، برای هر بلوک ۵ مقدار لبه بدست می‌آید. اگر بیشترین مقدار از یک حد آستانه تعریف شده بیشتر باشد آن لبه، لبه متناظر بلوک در نظر گرفته می‌شود. در بسیاری از مقالات حد آستانه ۱۱ و $desired_num_block$ برابر ۱۱۰۰ در نظر گرفته شده است. در هیستوگرام لبه برای هر زیر تصویر، ۵ کلاس وجود دارد. هر کلاس، متناظر با یک نوع لبه است و مقدار آن برابر با تعداد بلوک‌های شامل لبه متناظر در زیر تصویر مربوطه است. چون ۱۶ زیر تصویر داریم، بنابراین هیستوگرام لبه $16 \times 5 = 80$ کلاس دارد. برای چندی کردن هیستوگرام لبه، عناصر هر کلاس بر تعداد بلوک‌های زیر تصویر متناظر که شامل حداقل یک نوع لبه باشند، تقسیم می‌شود. بنابراین، مقادیر همه کلاس‌های هیستوگرام بین ۰ و ۱ خواهد بود.

هیستوگرام لبه سراسری، بیانگر توزیع لبه در تمام تصویر است و دارای ۵ عضو می‌باشد که هر عضو آن متناسب با یک نوع لبه و مقدار آن برابر با تعداد بلوک‌های شامل لبه متناظر در کل تصویر است. برای بدست آوردن هیستوگرام لبه نیمه سراسری ۲۳، هر تصویر به ۱۳ زیر تصویر تقسیم می‌شود. در هر زیر تصویر، توزیع لبه برای ۵ نوع مختلف لبه محاسبه می‌شود. بنابراین، هیستوگرام نیمه‌کلی تصویر، $13 \times 5 = 65$ عضو دارد. هیستوگرام نهایی، مجموع هیستوگرام‌های محلی، سراسری و نیمه سراسری است و ۱۵۰ کلاس دارد [۴۹].

۴-۳- معیار ارزیابی

رایج‌ترین معیارهای ارزیابی در بازیابی تصویر، معیارهای دقت و فراخوانی هستند. به منظور ارزیابی یک سامانه، در هر مرحله از بازیابی، معیارهای دقت و فراخوانی

در تکرار هشتم بیشتر از روش‌های دیگر است. در روش چند پرسشی، میانگین هر خوشه یا مرکز جرم آن به عنوان یک پرس و جوی جدید در نظر گرفته می‌شود. در الگوریتم‌های مبتنی بر مرکز جرم، خوشه‌هایی که مرکز جرم نزدیک به هم دارند با هم ترکیب می‌شوند. بنابراین در خوشه‌بندی نهایی، در هر خوشه، مرکز جرم زیر خوشه‌ها کمترین فاصله را از هم دارند. به همین دلیل مرکز جرم هر خوشه (پرس‌وجو) به مرکز جرم زیر خوشه‌هایش نزدیک است. و مرکز جرم‌های (پرس‌وجو) هر دو خوشه از هم دورند. بنابراین پرس‌وجوها یا مرکز جرم‌های خوشه‌ها به خوبی از هم تفکیک شده‌اند به عبارت دیگر تصاویر مرتبط بهتر خوشه‌بندی شده‌اند که این امر باعث افزایش کارایی الگوریتم‌های مبتنی بر مرکز جرم می‌شود.

۵- نتیجه‌گیری

در روش چندپرسشی، تصاویر مرتبط خوشه‌بندی شده و مرکز هر خوشه به عنوان یک پرس‌وجو در نظر گرفته می‌شود. در این تحقیق، روش‌های مختلف خوشه‌بندی سلسله مراتبی در بازیابی تصویر به شیوه چندپرسشی روی یک پایگاه تصویر شامل ۱۰۰۰۰ تصویر از ۸۲ گروه معنایی با هم مقایسه شدند. نتایج نشان دادند که با افزایش تعداد تصاویر شرکت کننده در خوشه‌بندی، روشهایی که در آنها تعداد اعضای هر خوشه در تعیین شباهت بین خوشه‌ها تاثیر دارد یا روشهایی که از آنها به عنوان روش‌های بدون وزن نام برده می‌شود، کارایی بهتری پیدا می‌کنند و دقت روش UPGMC در تکرار هشتم بیشتر از روش‌های دیگر است.

نزدیکتر باشد. بنابراین تعداد داده‌های نزدیک به هم در یک خوشه افزایش می‌یابد. به همین دلیل وقتی میانگین اعضای یک خوشه به عنوان یک پرس‌وجو در نظر گرفته می‌شود، پرس و جو به مرکز داده‌های با تراکم بالاتر نزدیکتر می‌باشد.

در الگوریتم‌های وزندار، چون زیر خوشه‌های C_q بدون در نظر گرفتن تعداد اعضایشان در اندازه‌گیری شباهت بین C_q و C_s تاثیر یکسان دارند، پراکندگی داده‌ها در یک خوشه بیشتر از الگوریتم‌های بدون وزن می‌باشد. بنابراین میانگین داده‌ها یا پرس‌وجو از مرکز داده‌های با تجمع بالاتر دور می‌شود که باعث پایین آمدن کارایی الگوریتم‌های وزندار نسبت به الگوریتم‌های بدون وزن در بازیابی تصویر به شیوه چندپرسشی می‌شود.

با افزایش تعداد تصاویر بازیابی شده و همینطور تعداد مراحل بازیابی، تعداد تصاویر مرتبطی که در خوشه‌بندی شرکت می‌کنند افزایش می‌یابد. هرچه تعداد تصاویر مرتبطی که در خوشه‌بندی شرکت می‌کنند بیشتر باشد تراکم داده‌ها در خوشه‌های خوشه‌بندی شده با الگوریتم‌های بدون وزن بالاتر رفته و کارایی این الگوریتم‌ها نسبت به الگوریتم‌های وزندار افزایش می‌یابد.

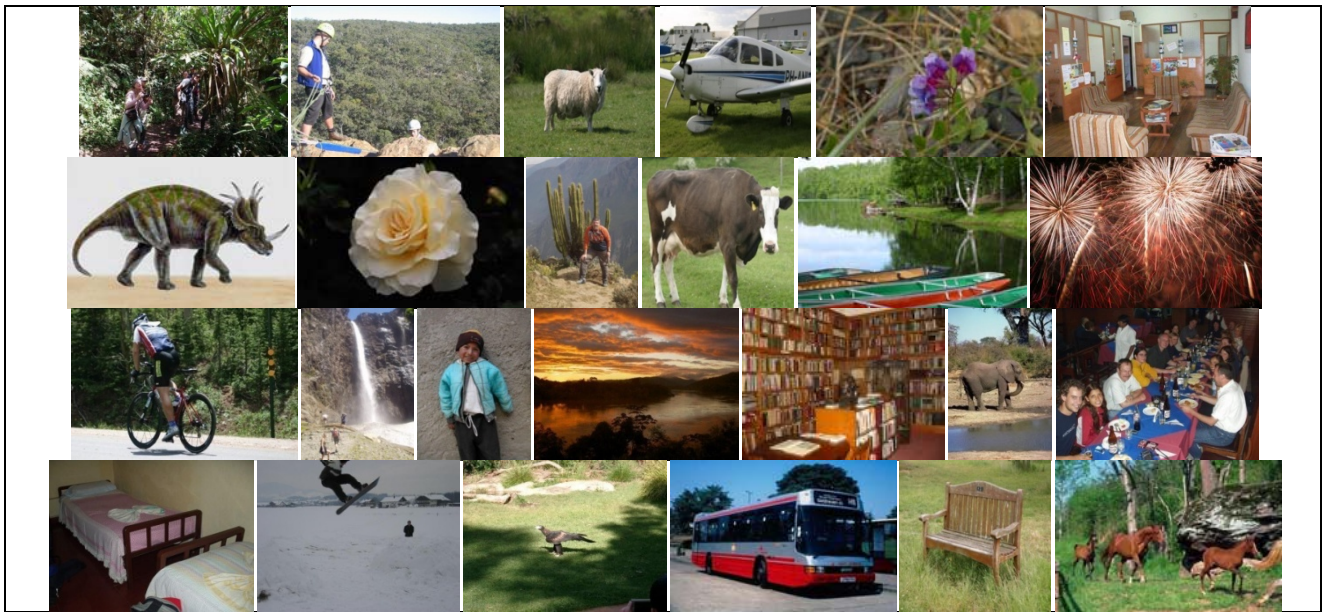
به منظور تعیین اثر میانگین و مرکز جرم در اندازه‌گیری شباهت بین دو خوشه، دو روش UPGMA و UPGMC با هم مقایسه شده‌اند. نتایج شکل ۶ نشان می‌دهند که در تعداد بازیابی‌های کم، روش UPGMA - محاسبه شباهت بین دو خوشه بر حسب میانگین فاصله بین اعضای دو خوشه - و در تعداد تصاویر بازیابی بالا، روش UPGMC - محاسبه شباهت بین دو خوشه بر اساس مرکز جرم دو خوشه - بهتر عمل می‌کنند. در نهایت گرافهای دقت بر حسب تعداد تکرار با ۴۰ و ۳۰ تصویر بازیابی شده در شکل ۷ و گراف دقت بر حسب تعداد تصاویر بازیابی شده در تکرار هشتم برای همه الگوریتم‌های خوشه‌بندی سلسله مراتبی در شکل ۸ نشان داده شده است. همانگونه که ملاحظه می‌شود، دقت روش UPGMC

جدول ۱- الگوریتم‌های single link, complete link, WPGMA و UPGMA

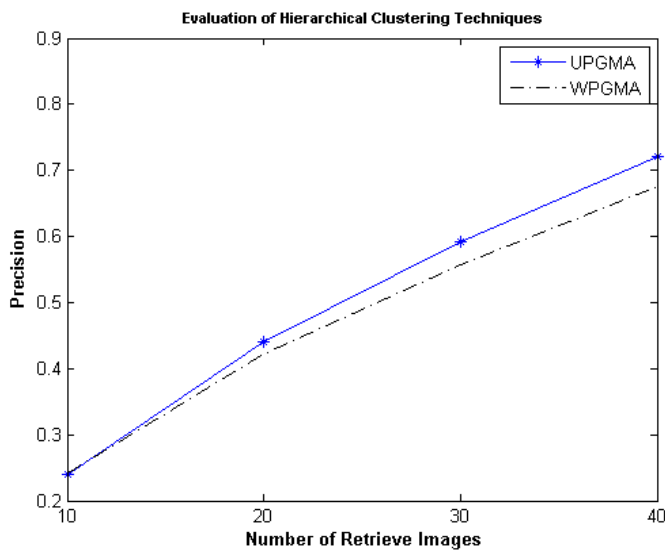
الگوریتم خوشه‌بندی	a_i	a_j	b	c	معیار شباهت
single link	$\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$d(C_i, C_q) = \min(d(C_i, C_s), d(C_j, C_s))$
complete link	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$d(C_s, C_q) = \max(d(C_i, C_s), d(C_j, C_s))$
WPGMA	$\frac{1}{2}$	$\frac{1}{2}$	0	0	$d(C_s, C_q) = \frac{1}{2}(d(C_i, C_s) + d(C_j, C_s))$
UPGMA	$\frac{n_i}{n_i + n_j}$	$\frac{n_j}{n_i + n_j}$	0	0	$d(C_s, C_q) = \frac{n_i}{n_i + n_j}d(C_i, C_s) + \frac{n_j}{n_i + n_j}d(C_j, C_s)$

جدول ۲- الگوریتم‌های UPGMC و WPGMC

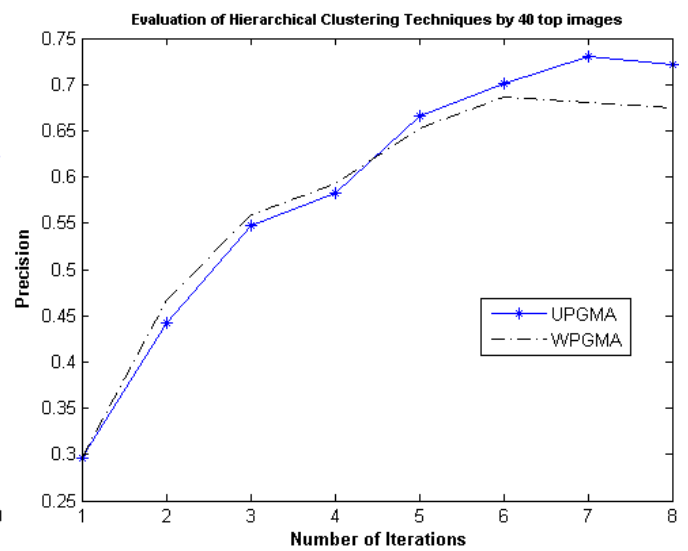
الگوریتم خوشه‌بندی	a_i	a_j	b	c	معیار شباهت
UPGMC	$\frac{n_i}{n_i + n_j}$	$\frac{n_j}{n_i + n_j}$	$\frac{n_i n_j}{(n_i + n_j)^2}$	0	$d_{sq} = \frac{n_i}{n_i + n_j}d_{is} + \frac{n_j}{n_i + n_j}d_{js} - \frac{n_i n_j}{(n_i + n_j)^2}d_{ij}$
WPGMC	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{4}$	0	$d_{sq} = \frac{1}{2}d_{is} + \frac{1}{2}d_{js} - \frac{1}{4}d_{ij}$



شکل ۳- تصاویر پرس‌وجو که در ارزیابی سامانه استفاده شده‌اند

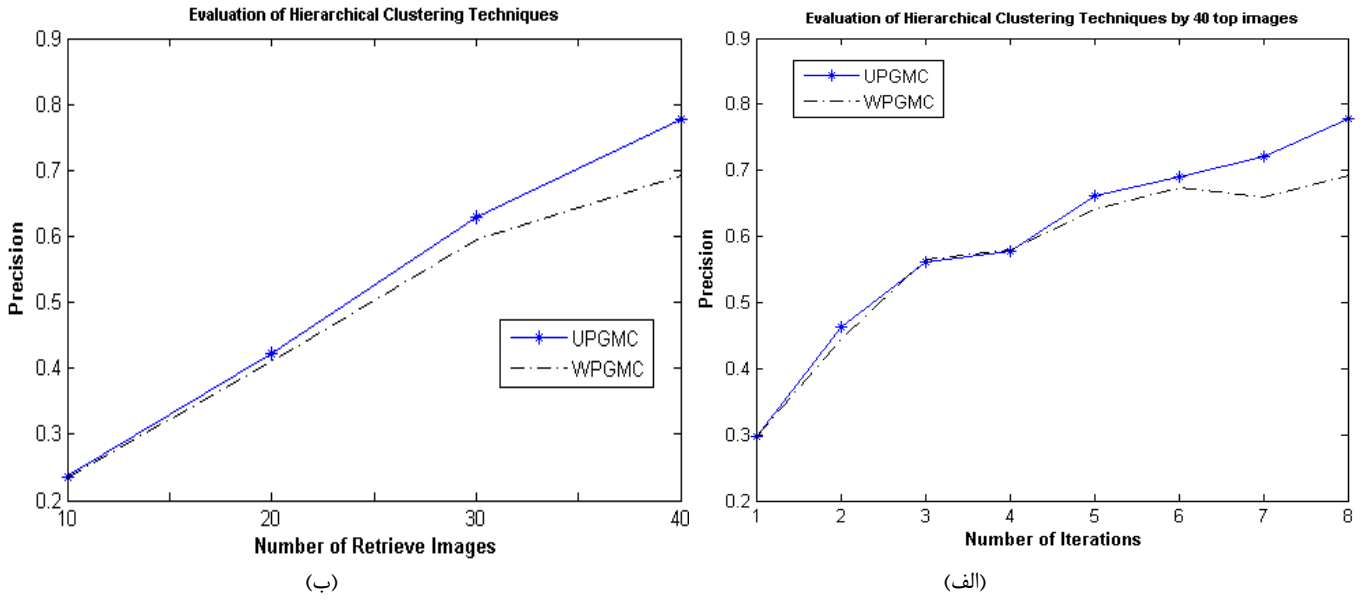


(ب)

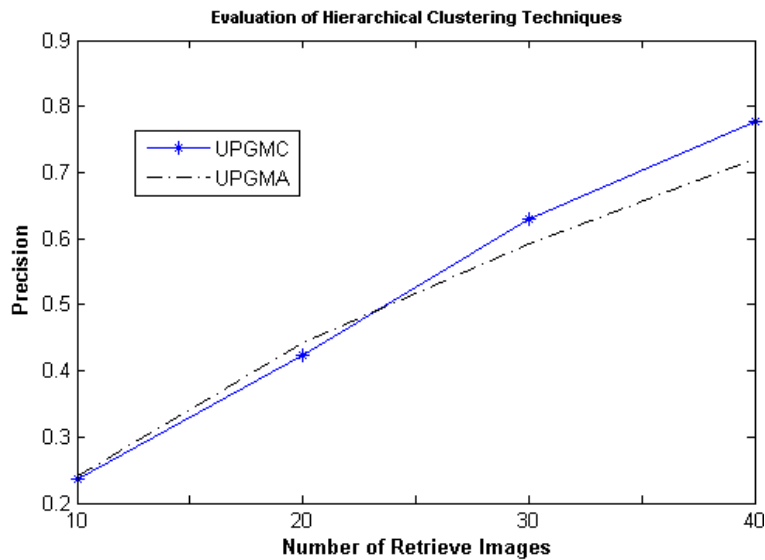


(الف)

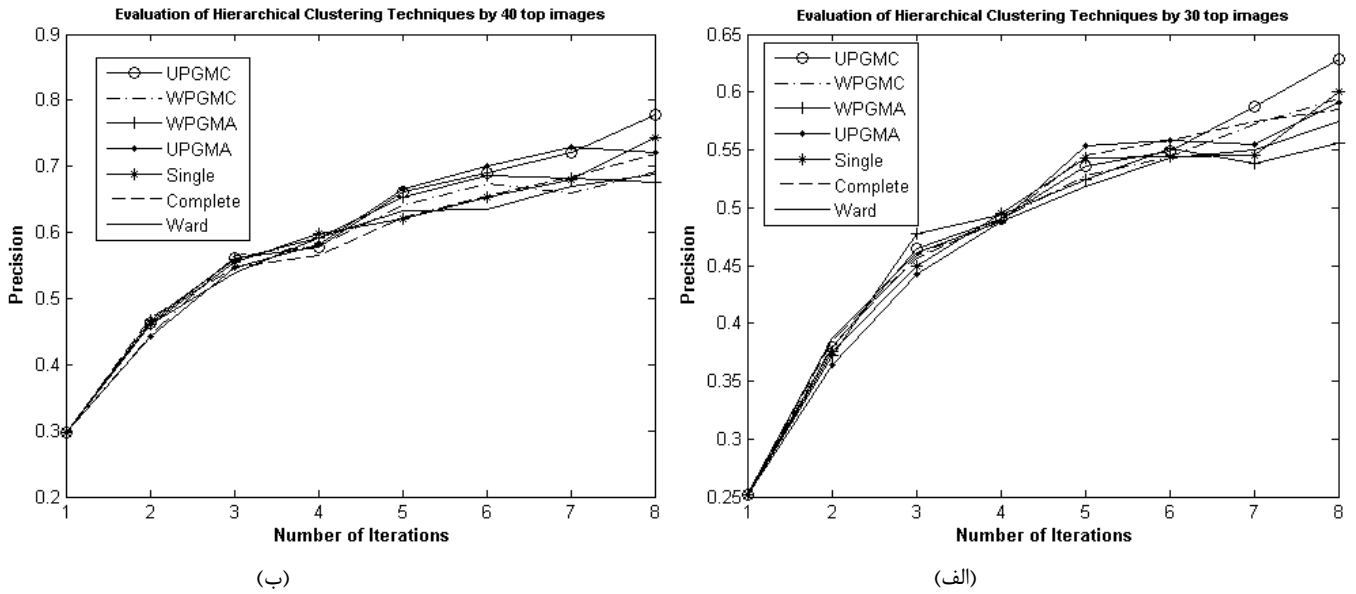
شکل ۴- الف) گراف دقت بر حسب تعداد تکرار برای الگوریتم‌های WPGMA و UPGMA (ب) گراف دقت بر حسب تعداد تصاویر مرتبط بازیابی شده در تکرار هشتم برای الگوریتم‌های WPGMA و UPGMA



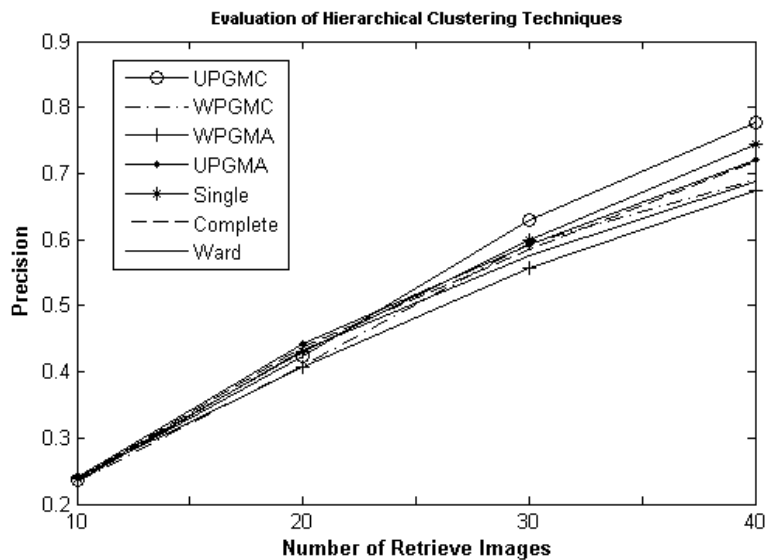
شکل ۵- الف) گراف دقت بر حسب تعداد تکرار برای الگوریتم‌های WPGMC و UPGMC (ب) گراف دقت بر حسب تعداد تصاویر مرتبط بازیابی شده در تکرار هشتم برای الگوریتم‌های WPGMC و UPGMC



شکل ۶- گراف دقت بر حسب تعداد تصاویر مرتبط بازیابی شده در تکرار هشتم برای الگوریتم‌های UPGMC و UPGMA در بازیابی تصویر به روش چندپرسشی



شکل ۷- گرافهای دقت بر حسب تعداد تکرار برای الگوریتم‌های خوشه‌بندی سلسله‌مراتبی با ۳۰ (شکل الف) و ۴۰ (شکل ب) تصویر بازیابی شده



شکل ۸- گراف دقت بر حسب تعداد تصاویر بازیابی شده در تکرار هشتم برای الگوریتم‌های خوشه‌بندی سلسله‌مراتبی

مراجع

- [14] R. A. Johnson, and D. W. Wichern, "Applied Multivariate Statistical Analysis," *Prentice-Hall*, New Jersey, 1998.
- [15] J. French, and X. Y. Jin, "An empirical investigation of the scalability of a multiple viewpoint CBIR system," *in: Proceedings of the CIVR*, Ireland, pp. 252–260, 2004.
- [16] N. Nhu-Van, B. Alain, J. Ogier, and S. Tabbone, "Clusters-Based Relevance Feedback for CBIR: A Combination of Query Movement and Query Expansion," *IEEE International Conference on Research, Innovation and Vision for the Future - RIVF*, 2010.
- [17] V. Khanh, C. Hao, and A. H. Kien, "Image retrieval in multipoint queries," *International Journal of Imaging Systems and Technology*, 18(2-3):170{181, 2008.
- [18] D. H. Kim, and S. H. Yu, "A new region filtering and region weighting approach to relevance feedback in content based image retrieval," *The Journal of Systems and Software*, 2007.
- [19] M. Charikar, C. Chekuri, T. Feder, and R. Motwani, "Incremental clustering and dynamic information retrieval," *Proc. Of ACM Symposium on theory of computing*, 1997.
- [20] J. Urban, and J. M. Jose, "Evidence Combination for Multi-Point Query Learning in Content-Based Image Retrieval," *Proceedings of the IEEE (ISMSE'04)*, 2004.
- [۲۱] ا. شمسی گوشکی، بازیابی تصاویر رنگی بر پایه محتوا با استفاده از روش‌های چند پرسشی و بازخورد ربط، پایان‌نامه کارشناسی‌ارشد، بخش مهندسی برق، دانشگاه شهید باهنر کرمان، ۱۳۸۹.
- [22] S. Theodoridis, and K. Koutroumbas, *Pattern Recognition*, Academic Press, ISBN: 0-21- 686140-4, 1999.
- [23] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The qbic project: Querying images by content, using color, texture, and shape," *in Storage and Retrieval for Image and Video Databases*, pp. 173–187, 1993.
- [24] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. F. Shu, "Virage image search engine: An open framework for image management," *in Storage and Retrieval for Image and Video Databases (SPIE)*, pp. 76–87, 1996.
- [25] Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," *International Journal of Computer Vision*, vol. 18, no. 3, pp. 233–254, 1996.
- [26] V. E. Ogle, and M. Stonebraker, "Chabot: Retrieval from a relational database of images," *IEEE Computer*, vol. 28, no. 9, pp. 40–48, 1995.
- [27] J. R. Smith, and S. F. Chang, "Visualseek: A fully automated content-based image query system," *in ACM Multimedia*, pp. 87–98, 1996.
- [1] Y. Liu, D. Zhang, G. Lu, and W. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, 40, pp. 262-282, 2007.
- [2] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the newAge," *ACM Comput. Surv.* 40, 2, 60 pages. 2008.
- [3] M. Banerjee, K. M. Kundu, and P. Maji, "Content-based image retrieval using visually significant point features," *Knowledge-Based Systems* Volume 160, Issue 23, 1 December, pp. 3323-3341, 2009.
- [4] T. Hurtut, Y. Gousseau, and F. Schmitt, "Adaptive image retrieval based on the spatial organization of colors," *Computer Vision and Image Understanding*, vol. 112, pp. 101–113, 2008.
- [5] S. W. Teng, and G. Lu, "Image indexing and retrieval based on vector quantization," *Pattern Recognition* 40, 3299–3316, 2007.
- [6] T. C. Lu, and C. C. Chang, "Color image retrieval technique based on color features and image bitmap," *Information Processing and Management* 43, 461–472, 2007.
- [7] T. Leña, P. Zuccarellob, G. Ayalaa, E. deVesb, and J. Domingoc, "Applying logistic regression to relevance feedback in image retrieval systems," *Pattern Recognition* 40, 2621 – 2632, 2007.
- [8] P. C. Cheng, B. C. Chien, H. R. Ke, and W. P. Yang, "A two-level relevance feedback mechanism for image retrieval," *Expert Systems with Applications*, 34, 2193–2200, 2008.
- [9] S. Newsam, B. Sumengen, and B. S. Manjunath, "Category-based Image retrieval," *in IEEE International Conference on Multimedia and Expo*, vol. 2, pp. 596- 599, 2001.
- [۱۰] ا. شمسی گوشکی، ح. نظام آبادی‌پور، س. سریزدی، و ا. کبیر، "روش جدید در بازخورد ربط برای بازیابی تصویر براساس محتوا،" پنزدهمین کنفرانس سالانه کامپیوتر، مرکز توسعه فناوری نیرو، تهران، ۱۳۸۸.
- [11] K. Porkaew, K. Chakrabarti, and S. Mehrotra, "Query refinement for multimedia similarity retrieval in MARS," *Proceedings of the ACM International, Multimedia Conference*, pp. 235–238, 1999.
- [12] D. H. Kim, C. W. Chung, and K. Barnard, "Relevance feedback using adaptive clustering for image similarity retrieval," *The Journal of Systems and Software*, vol. 78, pp. 9–23, 2005.
- [13] R. D. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley, New York, 2001.

- [43] M. Bertini, A. D. Bimbo, C. Torniai, C. Grana, R. Vezzani, and R. Cucchiara, "Sports Video Annotation Using Enhanced HSV Histograms in Multimedia Ontologies," *Image Analysis and Processing Workshops, 14th International Conference*, pp. 160-170, 2007.
- [44] N. Vo, P. N. Nguyen, and S. Orintara, "Texture image retrieval using complex directional filter bank," *IEEE*, pp. 5495-5498, 2006.
- [45] G. M. Haley, and B. S. Manjunath, "Rotation invariant texture classification using the modified Gabor filters," *IEEEICIP.95*, pp.262-265, 1995.
- [46] D. A. Clausi, and M. Jernigan, "Designing Gabor filters for optimal texture separability," *Elsevier, Pattern Recognition* 33, pp. 1835-1849, 2000.
- [47] W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no.12, pp. 1349-1380, 2000.
- [48] B. S. Manjunath, "Introduction to MPEG-7," *Wiley, New York*, 2002.
- [49] D. K. Park, Y. S. Jeon, and C. S. Won, "Efficient use of local edge histogram descriptor," *ACM workshops*, 1-58113-311-1, 2000.
- [50] H. Muller, W. Muller, D. M. Squire, S. M. Maillent, and T. Pun, "Performance evaluation in contentbased image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, pp. 593-601, 2001.
- [51] S. K. Saha, A. K. Das, and B. Chanda, "Image retrieval based on indexing and relevance feedback," *Pattern Recognition Letter*, 28, 357-366, 2007.
- [52] V. P. S. Rallabandi, and S. K. Sett, "Citations, Knowledge-based image retrieval system," *Knowledge-Based Systems* 21, 89-100, 2008.
- [28] J. R. Smith, and S. F. Chang, "Visually searching the web for content," *IEEE MultiMedia*, vol. 4, no. 3, pp. 12-20, 1997.
- [29] M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang, "Supporting similarity queries in mars," in *ACM Multimedia*, pp. 403-413, 1997.
- [30] C. Nastar, M. Mitschke, C. Meilhac, and N. Boujemaa, "Surfimage: A flexible content-based image retrieval system," in *ACM Multimedia*, pp. 339-344, 1998.
- [31] W. Y. Ma, and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," in *ICIP (1)*, pp. 568-571, 1997.
- [32] S. Ravela, and R. Manmatha, "Retrieving images by appearance," in *ICCV*, pp. 608-613, 1998.
- [33] H. Nezamabadi-pour, and E. Kabir, "Image retrieval using histograms of unicolor and bicolor blocks and directional changes in intensity gradient." *Pattern Recognition Letters*. vol. 25. no. 14. PP. 1547-1557, 2004.
- [34] M. R. Anderberg, *Cluster Analysis for Applications*, Academic Press, 1973.
- [35] G. N. Lance, and W. T. Williams, "A general theory of classificatory sorting strategies," *Computer Journal*, vol. 10, pp. 271-277, 1967.
- [36] S. Salvador, and P. Chan, "Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms," *Technical Report CS-2003-18*, Florida Institute of Technology, 2003.
- [37] ح. نظام‌آبادی‌پور، و ا. کبیر، "ارزیابی معیارهای عدم‌شبهت در بازیابی و طبقه‌بندی تصویر،" *مجله علمی و پژوهشی فنی مدرس*، شماره ۲۲، ص ۸۹ تا ۹۸، ۱۳۸۳.
- [38] P. Clough, M. Grubinger, A. Hanbury, and H. Muller, "Overview of the imageclef 2007 photographic retrieval task. In CLEF 2007 Workshop," *LNCS, in press*, Budapest, Hungary, 2008.
- [39] J. Winn, A. Criminisi, and T. Minka, "Object categorization by learned universal visual dictionary," *In ICCV*, volume 2, pp. 1800-1807, Beijing, China, Oct. 2005.
- [40] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLiCity: semantics-sensitive integrated matching for picture libraries," *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (9) 947-963, 2001.
- [۴۱] ح. نظام‌آبادی‌پور، ا. کبیر، و س. سریزدی، "بازیابی تصویر با استفاده از اطلاعات رنگ و شکل،" *دومین کنفرانس بینایی ماشین و پردازش تصویر ایران*، جلد دوم، ص ۲۲۶ تا ۲۳۲، ۱۳۸۱.
- [۴۲] ح. نظام‌آبادی‌پور، ا. کبیر، و س. سریزدی، "بازیابی تصویر براساس هم‌رخدادی رنگ‌ها در بلوک‌های لبه،" *هشتمین کنفرانس انجمن کامپیوتر ایران*، ص ۳۵۹ تا ۳۶۴، ۱۳۸۱.



اسما شمسی گوشکی کارشناسی خود را در مهندسی پزشکی در سال ۱۳۸۶ از دانشگاه شاهد تهران اخذ کرد. او سپس موفق به دریافت کارشناسی ارشد در رشته مهندسی برق - مخابرات از دانشگاه شهید باهنر کرمان در سال ۱۳۸۹ شد و هم‌اکنون در مقطع دکتری مهندسی برق دانشگاه صنعتی اصفهان مشغول به تحصیل است. زمینه‌های پژوهشی مورد علاقه او پردازش تصویر، بازشناسی الگو و رایانش نرم است. آدرس پست‌الکترونیکی ایشان عبارت است از:

asma.shamsi@ice.iut.ac.ir



حسین نظام‌آبادی‌پور کارشناسی و کارشناسی ارشد خود را در مهندسی برق - الکترونیک به ترتیب از دانشگاه شهید باهنر کرمان و دانشگاه تربیت مدرس در سال‌های ۱۳۷۷ و ۱۳۷۹ دریافت کرد. او سپس مدرک دوره دکتری خود را در مهندسی برق - الکترونیک از دانشگاه تربیت مدرس در سال

۱۳۸۳ اخذ کرد و در همان سال در سمت استادیاری در بخش مهندسی برق دانشگاه شهید باهنر کرمان مشغول به فعالیت شد. وی در حال حاضر در سمت دانشیاری در این دانشگاه مشغول به خدمت است. زمینه‌های پژوهشی مورد علاقه او الگوریتم‌های ابتکاری، پردازش تصویر، بازشناسی الگو و رایانش نرم است. آدرس پست‌الکترونیکی ایشان عبارت است از:

nezam@mail.uk.ac.ir



سعید سریزدی کارشناسی مهندسی الکترونیک و کارشناسی ارشد مهندسی مخابرات گرایش سیستم خود را به ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۶ از دانشگاه صنعتی اصفهان و مدارک D.E.A در زمینه پردازش سیگنال و دکتری در زمینه پردازش تصویر را از دانشگاه Rennes فرانسه به ترتیب در سال‌های ۱۳۷۳ و ۱۳۷۶ دریافت کرد. وی هم‌اکنون عضو هیات علمی دانشگاه شهید باهنر کرمان است. زمینه‌های پژوهشی مورد علاقه او مورفولوژی ریاضی، بازیابی تصویر، واترمارکینگ دیجیتال و کاربرد معادلات دیفرانسیل پاره‌ای در پردازش تصویر است. آدرس پست‌الکترونیکی ایشان عبارت است از:

saryazdi@mail.uk.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۹۰/۳/۵

تاریخ اصلاح: ۹۰/۱۲/۵

تاریخ قبول شدن: ۹۰/۱۲/۱۵

نویسنده مرتبط: دکتر حسین نظام‌آبادی‌پور، دانشکده مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران.

¹ Content Based Image Retrieval (CBIR).

² Semantic Features.

³ Low Level Features.

⁴ High Level Features.

⁵ Query Refinement.

⁶ Similarity Refinement.

⁷ Relevance Feedback.

⁸ Q_Cluster.

⁹ Top_K approach.

¹⁰ Clustering-Repeat.

¹¹ The Clustering-No-Repeat.

¹² Mahalanobis.

¹³ Weighted Pair Group Method Average.

¹⁴ Unweighted Pair Group Method Average.

¹⁵ Unweighted Pair Group Method Centroid.

¹⁶ Weighted Pair Group Method Centroid.

¹⁷ Orientations.

¹⁸ Scale.

¹⁹ Local Edge.

²⁰ Global Edge.

²¹ Horizontal Edge.

²² Vertical Edge.

²³ Semi_Global.