مرور مؤثّر نتایج جستجوی تصاویر با تلخیص بصری و متنّوع از طریق خوشه‌بندی

فاطمه علمدار و محمدرضا کیوان‌پور
دانشگاه الزهراء

چکیده:

با رشد بی‌سابقهٔ تولید تصاویر دیجیتال و استفاده از منابع چندرسانه‌ای، نیاز به جستجوی تصاویر و مطالب، افزایش یافته است. پردازش نظام‌مند این اطّلاعات پیش‌نیازی اساسی برای تحلیل، سازمان‌دهی و مدیریت مؤثّر آن محسوب می‌شود. از طرفی مجموعهٔ عظیمی از تصاویر بر روی وب در دسترس عموم قرار گرفته‌اند و بسیاری از موتورهای جستجو، امکان جستجوی تصاویر وب را بر مبنای کلمات کلیدی مهیا کرده‌اند. برای یافتن تصویر مطابق با نیاز و خواست افراد توسط موتورهای جستجوی تصاویر، چالش‌هایی همچون نارسابودن کلمهٔ پرس‌وجو، تعداد زیاد تصاویر نامرتبط با جستجوی انجام شده، تعداد زیاد تصاویر برگشتی و نبودن تلخیص، وقت‌گیربودن مرور تمامی تصاویر و عدم تنوع وجود دارد. خوشه‌بندی نتایج جستجوی تصاویر می‌تواند راه حل مؤثّری برای این مشکلات باشد.

در این پژوهش، چند الگوریتم برای خوشه‌بندی نتایج تصاویر موتورهای جستجو پیشنهاد شده است. تلخیص ایجاد شده از خوشه‌بندی، به کاربر این امکان را می‌دهد که تصاویر را به‌راحتی مرور کرده و حدود و محتوای کلّی تمامی تصاویر بازگشتی را در زمانی کوتاه و تعداد کمی کلیک به‌دست آورد. با خوشه‌بندی، مجموعهٔ متنوعی از تصاویر که بسیاری از تفاسیر ممکن کلمهٔ کلیدی را دربر دارد، ارائه می‌شود و خوشه‌های ایجاد شده، علاوه‌بر نمایش‌دادن تنوع‌های ناشی از ابهامات، تنوع بصری را نیز پوشش می‌دهند. بر اساس آزمایش‌های صورت گرفته، این رویکرد پیشنهادی باعث بهبود در نتایج خوشه‌بندی تصاویر می‌شود.

واژگان کلیدی: خوشه‌بندی تصاویر، الگوریتم Folding، استخراج ویژگی، تنوع بصری، مرور مؤثّر، موتور جستجوی تصاویر.

۱- مقدّمه

با توجّه به پیشرفت سریع در سخت‌افزار و نرم‌افزار، وب جهان‌گستر[1] به‌عنوان یک شیوهٔ انتشار برخط[2] به یک منبع چندرسانه‌ای در حال رشد تبدیل شده است و مجموعهٔ عظیمی از تصاویر بر روی وب در دسترس عموم قرار گرفته‌اند. ممکن است به نظر برسد که وب جهان‌گستر با تعداد بسیار وسیع و دائماً در حال رشد تصاویر دیجیتال، برای پیداکردن تصویری که مطابق با نیاز و خواست افراد است، حقیقتاً کمک کننده است؛ امّا این نظر تا حدودی غلط است. درعوض اغلب کاربران با تعداد زیاد تصاویر دردسترس سردرگم می‌شوند.

به‌منظور کمک به کاربر که تصویر مورد نظر خود را بر روی وب بیابد، مسألهٔ بازیابی تصاویر وب، سال‌ها مورد مطالعه قرار گرفته و تعدادی از راهبردها پیشنهاد شده‌اند(*Liu, 2004*). بسیاری از موتورهای جستجوی تجاری، با توسعهٔ فناوری‌هایی، این امکان را به کاربر می‌دهند که میلیون‌ها تصویر وب را بر مبنای کلمات کلیدی جستجو کنند.

بیشتر موتورهای جستجوی تصاویر موجود مانند Google image search، Microsoft Bing image search، Yahoo! Image search، از متن همبسته با تصاویر (فراداده‌ها[3]) برای جستجو استفاده می‌کنند؛ بنابراین برای جستجوی تصاویر از روش‌های جستجوی مبتنی بر متن استفاده می‌شود. نتایج جستجو معمولاً به‌صورت فهرستی مرتب شده نمایش داده می‌شود که بازتاب‌دهندهٔ شباهت فراداده‌های تصاویر به پرس‌وجوی متنی[4] است. درنتیجه

[1] - World Wide Web
[2] - On-line publication mechanism
[3] - Meta data
[4] - Textual query

هرقدر که فراداده‌ها بهتر بتوانند محتوای تصویر را نشان دهند، بازدهی بازیابی بهتر خواهد بود.

جستجوی تصویر در موتورهای جستجوی تصاویر برای کاربر مشکلاتی را ایجاد می‌کند، از جمله:

نارسابودن کلمهٔ پرس‌وجو: با مشاهدهٔ یک تصویر، مفهوم، مقصود و زیبایی آن به‌سادگی درک می‌شود و یک تصویر می‌تواند اطّلاعاتی را به بیننده القا کند که کلمات قادر به بیان آن نیستند(*van Leuken, 2009*). به‌دلیل همین طبیعت غنی و توان‌گر محتوای تصاویر و معنای محدودی که در ساختار پرس‌وجو مبتنی بر کلمهٔ کلیدی وجود دارد، اغلب برای کاربران مشکل است که اطّلاعات مورد نیاز خود را به‌دقّت در قالب یک کلمهٔ پرس‌وجو فرموله کنند(*van Zwol, 2008*).

تعداد زیاد تصاویر نامرتبط: کاربر ممکن است با تعداد زیاد تصاویر تکراری که حتی بعضی از آنها چندان مرتبط با مورد جستجو نیستند، مواجه شود (*Zhang, 2005*). مرتبط‌نبودن تصاویر ممکن است به‌دلیل ابهام کلمهٔ مورد جستجو باشد (*Fergus, 2005*) و یا اینکه فراداده‌های تصاویر، نشان‌دهندهٔ آن تصاویر نباشند. نتایج مرتبط‌تر به‌طورمعمول در اوایل نتایج جستجو قرار دارند و برای موضوعاتی که رایج‌ترند و بیشتر جستجو می‌شوند، تصاویر مرتبط بیشتری برگردانده می‌شود. چنین تجربیاتی در جستجو، باعث ناامیدی کاربران از دستیابی به تصویر موردنظر می‌شود.

نبود تلخیص[1]: درصورتی‌که کاربر بخواهد حدود و محتوای کلی تمامی تصاویر بازگشتی را به‌دست آورد، باید بر روی تمامی صفحات، کلیک کند و یا scroll صفحه را پایین بکشد تا نگاهی کلی به همهٔ تصاویر بیاندازد. حتی با مشاهده تمامی تصاویر، هنوز هم دریافت سریع محتوای تعداد زیادی از تصاویر (به‌طور معمول در حدود ۱۰۰۰ تصویر) برای یک کاربر عادی آسان نیست (*Wang, 2010*).

وقت‌گیر بودن: کاربران موتورهای جستجو باید فهرست مرتب شده را برای یافتن تصویر موردنظر بگردند. این کار امری وقت‌گیر است؛ زیرا نتایج ممکن است شامل موضوعات[2] مختلف باشند و این موضوعات به‌طور نامشخّصی با هم ترکیب شده باشند. این موقعیت می‌تواند بدتر شود، زمانی‌که یک موضوع غالب بوده و سراسر فهرست را بپوشاند؛ امّا چیزی نباشد که کاربر به دنبال آن است (*Cai, 2004*).

عدم تنوع[3]: کلمهٔ پرس‌وجو ممکن است مبهم باشد و رتبه‌بندی نتایج به‌طور مناسب جهات مختلف این ابهام را تحت پوشش قرار ندهد. در مواردی که کلمهٔ پرس‌وجو مبهم نیست، این امکان وجود دارد که رتبه‌بندی حاصل فاقد تنوع بصری باشد (*van Leuken, 2009*).

تنوع مجموعهٔ تصاویر، بستگی به ابهام کلمهٔ پرس‌وجو دارد. این ابهام شامل دو نوع «ابهام مفهوم کلمه[4]» و«ابهام مختص نوع[5]» است (*van Zwol, 2008*). برای درک نوع ابهام، فرض کنید که کاربر "apple" را به عنوان کلمهٔ پرس‌وجو وارد کرده است. این کلمهٔ کلیدی می‌تواند هم مفهوم «میوه سیب» و هم «شرکت سیب» داشته باشد. به این نوع ابهام، «ابهام مفهوم کلمه» گفته می‌شود. زمانی که کلمهٔ پرس‌وجو به "apple company" تصحیح شود، نوع دیگری از ابهام ظاهر می‌شود. به طور ایده‌آل نتایج جستجو، هنوز هم متنوع و گوناگون هستند و شامل نمونه‌هایی از محصولات مختلف شرکت، لوگوی آن و ... می‌شوند. به این نوع ابهام، «ابهام مختص نوع» می‌گویند. اگر کلمهٔ پرس‌وجو مبهم نباشد، نوع دیگری از تنوع وجود دارد؛ همچون تنوع بصری که این نوع تنوع از فراداده‌های همبسته‌شده با یک تصویر به‌دست نمی‌آید.

توجّه به ویژگی‌های بصری و خوشه‌بندی نتایج جستجوی تصاویر بر مبنای این ویژگی‌ها، می‌تواند راه حلی برای مشکلات ذکر شده باشد. ویژگی‌های سطح پایین در یک تصویر، معرّف کلیات تصویرند و اشیا یا مفاهیم موجود در تصویر را توصیف نمی‌کنند، مانند ویژگی‌های رنگ، بافت، شکل و موقعیت مکانی[6]. در این راه حل، پس از اتمام جستجوی تصاویر بر مبنای کلمهٔ پرس‌وجو، در گام بعد، نتایج برطبق ویژگی‌های سطح پایین خوشه‌بندی می‌شوند. برای سازمان‌دهی به نحوهٔ نمایش نتایج جستجوی تصاویر خوشه‌بندی شده، نمایندهٔ[7] هر خوشه به کاربر نشان داده می‌شود و کاربر بنابر نیاز و علاقهٔ یکی از این نماینده‌ها را انتخاب کرده و تصاویر مربوط به آن خوشه را مشاهده می‌کند. تلخیص ایجاد شده از خوشه‌بندی، به کاربر این امکان را می‌دهد که تصاویر را به‌راحتی مرور کرده[8] و حدود و محتوای کلی تمامی تصاویر بازگشتی را در زمانی کوتاه و تعداد کمی کلیک به‌دست آورد. با خوشه‌بندی، مجموعهٔ متنوعی از تصاویر که بسیاری از تفاسیر ممکن کلمهٔ کلیدی

[1] - Summarization
[2] - Topics

[3] - Diversity
[4] - Word sense ambiguity
[5] - Type specific ambiguity
[6] - Spatial location
[7] - Representative
[8] - Browse

را دربر دارند، ارائه می‌شود و خوشه‌های ایجاد شده، علاوه‌بر نمایش دادن تنوع‌های ناشی از ابهامات، تنوع بصری را نیز پوشش می‌دهند. خوشه‌بندی علاوه‌بر راحت‌ترکردن کار کاربر، باعث بازیابی مؤثّرتر تصاویر می‌شود.

در این پژوهش روشی برای خوشه‌بندی نتایج تصاویر موتورهای جستجو پیشنهاد شده است که با استفاده از پیش‌پردازش و استخراج ویژگی‌های مؤثر و چند الگوریتم تعمیم داده شده، انجام گرفته است. بر اساس آزمایش‌های صورت گرفته، این روش باعث بهبود در نتایج خوشه‌بندی تصاویر می‌شود.

ادامهٔ مقاله بدین صورت است: در بخش دو پیشینهٔ تحقیق بیان می‌گردد. تعدادی از الگوریتم‌های خوشه‌بندی تصاویر وب در بخش سه مطرح می‌شوند. بخش چهار سیستم پیشنهادی را تشریح می‌کند. آزمایش‌ها و نتایج در بخش پنجم آورده شده است و درنهایت در بخش شش نتیجه‌گیری ارائه می‌شود.

۲- پیشینهٔ تحقیق

در این بخش به تعدادی از کارهای انجام شده در زمینهٔ خوشه‌بندی تصاویر وب پرداخته می‌شود.

Group-I *(Jing, 2006)* تنها از متن احاطه‌کنندهٔ[1] تصویر برای گروه‌بندی نتایج جستجوی تصاویر استفاده می‌کند. این روش با استفاده از تحلیل متن، ابتدا خوشه‌های معنایی مرتبط با کلمهٔ پرس‌وجو را شناسایی کرده و سپس خوشه‌بندی معنایی را با انتساب تصاویر متناظر با هر خوشه، انجام می‌دهد. ویژگی این روش ایجاد گروه‌ها با اسامی معنادار می‌باشد؛ امّا مضامین و ویژگی‌های بصری را برای دسته‌بندی درنظر نمی‌گیرد.

Gao و همکاران، روشی برای خوشه‌بندی تصاویر وب با ترکیب سازگار ویژگی‌های سطح پایین و متن احاطه کنندهٔ تصویر در *(Gao, 2005)* پیشنهاد کردند. این روش به‌عنوان یک مسألهٔ بهینه‌سازی چند منظورهٔ محدود تنظیم و فرموله شده است. این روش بیشتر بر روی برچسب‌های همبسته‌شده[2] به تصاویر توجّه دارد.

Li و همکارانش در *(Li, 2005)*، یک الگوریتم خوشه‌بندی پیوندی برای کار با فضای ویژگی نامتجانس[3] مطرح کرده است. این روش برای متون استخراج شده از صفحات وب، رویکرد خوشه‌بندی مشترک[4] و برای ویژگی‌های سطح پایین، رویکرد خوشه‌بندی یک‌مسیره[5] را اتخاذ می‌کند و نتایج خوشه‌بندی رویکردها به‌منظور گروه‌بندی تصاویر موتور جستجو، ترکیب می‌شوند. در این پژوهش تنها از دو ویژگی بصری استفاده شده است و پایگاه داده استفاده شده محدود به ۶ مورد می‌باشد.

در *(Cai, 2004)* یک رویکرد خوشه‌بندی سلسله-مراتبی ارائه شده است که نتایج جستجوی تصاویر وب را با توجه به تحلیل بصری، متنی و لینک‌ها گروه‌بندی می‌کند. هدف این روش خوشه‌بندی نتایج جستجوی کلمات مبهم به منظور تسهیل مرور تصاویر توسط کاربر است. این رویکرد، بیشتر بر روی خوشه‌بندی نتایج موارد جستجوشدهٔ مبهم تمرکز کرده است و یک اشکال بارز آن عدم توجّه به میزان ارتباط تصاویر به کلمهٔ جستجو شده در هنگام خوشه‌بندی است.

در *(Deselaers, 2003)* برای بهبود راحتی کارکردن کاربر[6] با موتورهای جستجوی تصویر که مبتنی بَر متن کار می‌کنند، ترکیبی از روش‌های بینایی ماشین و داده‌کاوی ارائه شده است. در روش ذکر شده در این مرجع، نیاز است که تعداد خوشه‌ها تعیین شود و نتایج خوشه‌بندی تصاویر نتیجه شده از جستجو به صورت کمی و مقایسه‌ای ارائه نشده است.

رویکرد «نزدیک‌ترین همسایه‌های مشترک[7] (SNN)» در *(Moëllic, 2008)* برای خوشه‌بندی تصاویر وب استفاده شده است که هم ویژگی‌های بصری و هم ویژگی‌های متنی را در نظر می‌گیرد. SNN یک روش خوشه‌بندی مبتنی بر چگالی است، امّا تضمین تنوع نتایج آن به‌سادگی امکان‌پذیر نیست. این روش متکّی به پارامتری سراسری برای تخمین چگالی است و توانایی کمی در ایجاد خوشه‌هایی با فشردگی[8] متفاوت دارد. این روش نیاز به توان پردازشی بالایی برای تولید خوشه‌بندی نهایی دارد.

Jia و همکاران، یک الگوریتم خوشه‌بندی بر مبنای انتشار وابستگی[9] ارائه کرده و با استفاده از ویژگی‌های بصری سعی در یافتن نمونه‌های تصاویر و سازمان‌دهی نتایج جستجوی تصاویر دارد *(Jia, 2008)*.

TeBIC[10] یک روش خوشه‌بندی نتایج جستجوی تصاویر بر مبنای اطلاعات متنی است *(Wang, 2009)*. هدف این روش، ایجاد خوشه‌ها با اسامی معنادار و غلبه بر هزینهٔ

[1] - Surrounding text
[2] - Associated tags
[3] - Inhomogeneous
[4] - Co-clustering
[5] - One-way
[6] - User-freindliness
[7] - Shared nearest neighbors
[8] - Compactness
[9] - Affinity propagation
[10] - Text Based Image Clustering

محاسباتی بالای خوشه‌بندی بر مبنای ویژگی‌های بصری عنوان شده است. اشکال اصلی این روش، این فرض است که تمامی اطلاعات موردنیاز برای خوشه‌بندی از متن قابل استخراج است و ویژگی‌های بصری، اطّلاعات اضافه‌تری دربرندارند.

مسألهٔ تنوع بصری در نتایج جستجوی تصاویر در (van Leuken, 2009) بررسی شده است. این پژوهش یک معیار شباهت وزن‌دار پویا پیشنهاد کرده است و با استفاده از این معیار و چند ویژگی بصری و سه روش خوشه‌بندی ارائه شده، سعی در تلخیص مؤثر و متنوع دارد. در این روش به‌منظور سازمان‌دهی نمایش نتایج جستجوی تصویر، نمایندهٔ خوشه‌ها به کاربر نشان داده می‌شود و با انتخاب نمایندهٔ مورد نظر توسط کاربر، تصاویر خوشه قابل مشاهده هستند. تنوع بصری نتایج در این روش تضمین شده است.

«گردش تصادفی جذاب پویا[1]» (Wang, 2010)، یک روش تلخیص بصری مبتنی بر تصویر برای نمایش نتایج جستجوی تصاویر می‌باشد. در این روش تنوع خوب[2] با استفاده از یک شمای تنظیم وزن پویا، تضمین شده است و میزان ارتباط و شباهت بین تصاویر با یک روش ارزیابی شباهت بصری که به‌صورت محلی مقیاس‌دهی شده، بررسی شده است. این رویکرد با استفاده از یک ساختار سلسله-مراتبی، یک نمایش تعاملی مؤثر به کاربر ارائه می‌دهد، تا تصاویر توسط وی راحت‌تر مرور شوند.

در سیستم ارائه شده در این پژوهش، از چند ویژگی بصری برای محسوب‌کردن جهات بصری متفاوت تصاویر استفاده شده است. چند الگوریتم برای خوشه‌بندی پیشنهاد شده است که مبتنی‌بر الگوریتم Folding (van Leuken, 2009) هستند. این الگوریتم ایدهٔ ساده و نتایج قابل قبولی دارد. در الگوریتم‌های پیشنهادی سعی شده با انتخاب مؤثرتر نماینده‌های خوشه‌ها، فازی‌کردن الگوریتم، وزن، تکرار و بهره‌بردن از مزایای الگوریتم‌های سلسله‌مراتبی عملکرد نسبت به الگوریتم Folding بهبود یابد. در این الگوریتم‌ها نیاز به مشخص ساختن تعداد خوشه‌ها نیست و به میزان ارتباط تصاویر به کلمهٔ جستجو شده در هنگام خوشه‌بندی نیز توجه شده است. پایگاه دادهٔ تصاویر این پژوهش، شامل تصاویری است که با جستجوی ۶۵ عنوان کلمهٔ «پرس‌وجو» در موتورهای جستجوی رایج جمع‌آوری شده‌است که این عناوین شامل کلمات مبهم و غیرمبهم می‌باشد.

۳- الگوریتم‌های خوشه‌بندی تصاویر وب

به‌دلیل مقایسه‌ها و ارجاعات در بخش‌های بعدی به سه الگوریتم خوشه‌بندی (van Leuken, 2009) Folding، Maxmin، Reciprocal election، در این بخش به‌طور اختصار به این الگوریتم‌ها پرداخته می‌شود. قبل از توصیف الگوریتم‌ها علامت‌گذاری‌های به‌کاربرده شده در این الگوریتم‌ها و الگوریتم‌های پیشنهادی، بیان می‌گردند:

I: مجموعه‌ای از تصاویر نتیجه شده از جستجو که شامل N تصویر است.

L: فهرست مرتب تصاویر I یعنی: $L = L_1, L_2, ..., L_N$.

C: یک خوشه‌بندی و افرازبندی از I است که در آن همهٔ تصاویر به K خوشه تقسیم می شوند $C_1, C_2, ..., C_K$ که $C_k \cap C_l = 0$ برای همهٔ $l, k \in K$ و $\bigcup_{k=1}^{K} C_k = I$. n_k تعداد تصاویر در خوشهٔ C_k است، بنابراین $\sum_{k=1}^{K} n_k = N$.

R: مجموعهٔ نماینده‌های همهٔ خوشه‌هاست که R_k نمایندهٔ خوشهٔ C_k است.

۳-۱- الگوریتم Folding

الگوریتم Folding به رتبه‌بندی اولیه و اصلی نتایج جستجوی تصویر هنگام انجام خوشه‌بندی، اهمیت می‌دهد و به تصاویر با مرتبهٔ بالاتر احتمال بالاتری نسبت می‌دهد که نمایندهٔ خوشه باشند. مرحلهٔ اوّل این راهبرد انتخاب تصاویر نماینده است. بعد از مشخّص‌شدن نماینده‌ها، خوشه‌ها در اطراف هر نمایندهٔ خوشه با استفاده از قانون نزدیک‌ترین همسایه شکل می‌گیرند.

برای انتخاب نماینده‌ها، تصویر اوّل رتبه‌بندی یعنی L_1، همیشه به‌عنوان نماینده انتخاب می‌شود. در هنگام پیمودن فهرست مرتب شده به سمت پایین، هر تصویر با مجموعه‌ای از نماینده‌های پیشین مقایسه می‌شود. وقتی که یک تصویر به قدر کافی نسبت به همهٔ نمایندگان انتخاب شده در R بی‌شباهت باشد، به فهرست نمایندگان اضافه می‌شود. این پارامتر مقایسه به‌عنوان متوسط فاصله‌ای که همهٔ تصاویر در I تا «تصویر میانگین[3]» دارند، تعریف می‌شود. در این روش، تصویر میانگین، تصویری با کوچک‌ترین متوسط فاصله تا تمامی دیگر تصاویر در نظر گرفته شده است (van Leuken, 2009).

۳-۲- الگوریتم Maxmin

الگوریتم Maxmin رتبه‌بندی اوّلیه و اصلی تصاویر را در نظر نمی‌گیرد و اوّلین نماینده R_1 به‌صورت تصادفی انتخاب می‌گردد. دومین نماینده R_2، تصویری است که بزرگترین

[1] - Dynamic absorbing random walk
[2] - A good diversity
[3] - Average image

فاصله را تا R_1 دارد. برای دیگر نماینده‌ها، تصویری انتخاب می‌شود که بزرگترین حدّاقل فاصله[1] را تا تمامی دیگر نماینده‌های انتخاب شده داشته باشد. پس از انتخاب نماینده‌ها، شکل‌گیری خوشه‌ها با استفاده از قانون نزدیک‌ترین همسایه انجام می‌شود *(van Leuken, 2009)*.

۳-۳- الگوریتم Reciprocal election

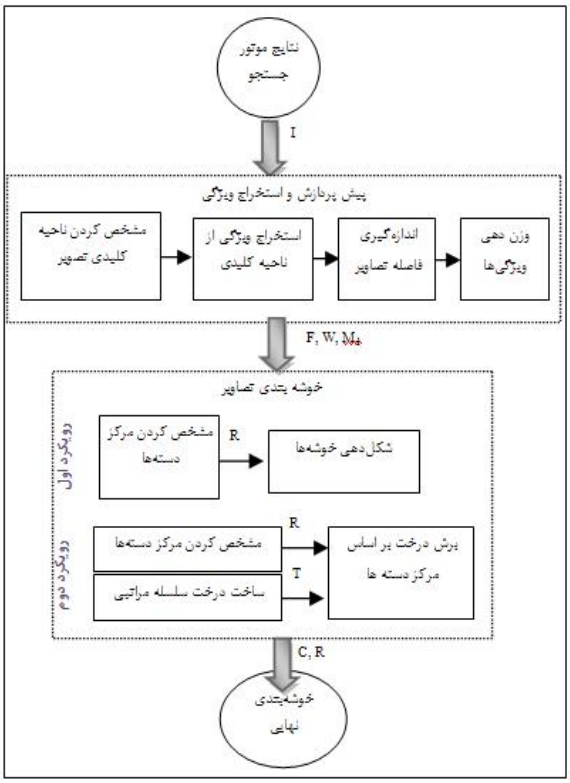
برخلاف الگوریتم Folding راهبرد Reciprocal election فرآیندهای انتخاب نمایندگان خوشه و تشکیل خوشه در میان یکدیگر انجام می‌شود. ایدهٔ اصلی این روش این است که هر تصویر در I تصمیم می‌گیرد که به‌وسیلهٔ کدام تصویر، بهتر بازنماینده[2] می‌شود. فرآیند رأی‌گیری برای هر تصویر مبتنی‌بر محاسبهٔ رتبه‌های متقابل[3] در رتبه‌بندی‌های I است. پس از تعیین رأی‌های تمامی تصاویر، تصویر با بیشترین تعداد رأی، به‌عنوان اوّلین نمایندهٔ R_1 انتخاب می‌گردد. و سپس خوشهٔ R_1 با افزودن آن تصاویری که R_1 را در m تصویر بالای فهرست رتبه‌بندی خود دارند، شکل می‌گیرد. سپس اعضا و نمایندهٔ آن خوشه از فهرست کاندیدای نمایندگی حذف خواهند شد و فرآیند تا جایی ادامه پیدا می‌کند که هر تصویر یا به‌عنوان نماینده یا عضوی از یک خوشه انتخاب شده باشد *(van Leuken, 2009)*.

۴- سیستم پیشنهادی

(شکل ۱) نشان‌دهندهٔ ساختار «سیستم خوشه‌بندی نتایج جستجو» پیشنهادی است. ورودی سیستم تصاویری هستند که با جستجوی یک کلمهٔ پرس‌وجو در موتور جستجو نتیجه شده‌اند. سیستم خوشه‌بندی پس از طی مراحلی، تصاویر خوشه‌بندی‌شده را به کاربر ارائه می‌دهد. کاربر می‌تواند با انتخاب مرکز خوشه‌ها، تمامی تصاویر متعلق به هر خوشه را مشاهده کند و تصویر مورد نظر خود را بیابد. این سیستم شامل دو زیرسیستم می‌باشد. ابتدا در زیرسیستم اوّل، بر روی تصاویر پیش‌پردازش‌های لازم انجام شده و ویژگی‌های آن‌ها استخراج می‌شوند. زیرسیستم دوم، خوشه‌بندی تصاویر را انجام می‌دهد. زیرسیستم «پیش‌پردازش و استخراج ویژگی» شامل مراحل مشخّص‌کردن ناحیهٔ کلیدی تصویر، استخراج ویژگی از ناحیهٔ کلیدی، اندازه‌گیری فاصلهٔ تصاویر و وزن‌دهی ویژگی‌ها است. ورودی این زیرسیستم، تصاویر منتج شده از موتورجستجو I است و خروجی آن بُردار ویژگی تمامی تصاویر F، وزن‌ها W و ماتریس‌های فاصله M_d می‌باشد. زیرسیستم «خوشه‌بندی تصاویر» قابل انجام با دو رویکرد کلی است که رویکرد اوّل به‌طورکلّی شامل دو مرحلهٔ مشخّص‌کردن مرکز دسته‌ها R و شکل‌دهی خوشه‌ها است. رویکرد دوم که برمبنای درخت‌های سلسله‌مراتبی می‌باشد دارای سه مرحلهٔ کلّی مشخّص‌کردن مرکز دسته‌ها R، ساخت درخت سلسله‌مراتبی T و بُرش درخت بر اساس مرکز دسته است. خوشه‌بندی C و مراکز خوشه‌ها R خروجی این زیرسیستم هستند.

الگوریتم‌های پیشنهاد شده در این پژوهش مبتنی بر الگوریتم Folding هستند و در سیستم پیشنهادی سعی شده است مراحل پیش‌پردازش و خوشه‌بندی مؤثّرتر انجام شود. در مرحلهٔ پیش‌پردازش به‌جای استخراج ویژگی از کل تصویر، ویژگی‌ها از زیرتصویری استخراج می‌شوند که از لحاظ بصری بیشتر مورد توجّه انسان است. و همچنین به ویژگی‌ها وزنی نسبت داده می‌شود تا ویژگی‌های توصیف‌کننده‌تر تأثیر بیشتری داشته باشند. در مرحلهٔ خوشه‌بندی، الگوریتم‌هایی پیشنهاد شده که در این الگوریتم‌ها سعی شده است با انتخاب مؤثّرتر نماینده‌های خوشه‌ها، فازی کردن الگوریتم، وزن، تکرار و بهره‌بردن از مزایای الگوریتم‌های سلسله‌مراتبی عملکرد نسبت به الگوریتم Folding بهبود یابد.

در ادامه مراحل زیرسیستم‌ها با جزئیات بیشتری شرح داده می‌شوند.



(شکل ۱): ساختار «سیستم خوشه‌بندی نتایج جستجو پیشنهادی»

[1] - Largest minimum distance
[2] - Best represented
[3] - Reciprocal

۴-۱- مشخّص‌کردن ناحیهٔ کلیدی تصویر

برای استخراج ویژگی، می‌توان ناحیه یا نواحی از تصویر را در نظر گرفت که از لحاظ انسانی مورد توجّه بصری[1] هستند. آنالیز توجّه بصری، به‌خوبی برای اهداف بینایی ماشین مورد ارزیابی قرار گرفته و مدل‌هایی برای محاسبهٔ توجّه بصری توسعه یافته‌اند (*Ma, 2003; Lamming, 1991; Niebur, 1998; Milanse, 1995; Baluja, 1997; Itti, 1998; Ahmad, 1991*). در (*Ma, 2003*)، به جای بررسی شیوهٔ ادراک انسان توسط الگوریتم‌ها، یک چارچوب آنالیز توجّه بصری برای تصاویر معرفی شده است؛ که این روش ساده و مؤثّر بوده و نیاز به محاسبات پیچیده و زمان بالا ندارد. این چارچوب از هر تصویر سه سطح «توجّه» استخراج می‌کند: چشم‌انداز مورد توجّه[2]، سطوح مورد توجّه[3] و نقاط مورد توجّه[4]. «چشم‌انداز مورد توجّه» با استخراج زیرتصویری که دارای مهم‌ترین اطّلاعات موجود در تصویر است، دقّت استخراج ویژگی را به‌طور مؤثّری تسریع می‌کند. درحقیقت، «چشم‌انداز مورد توجّه» قسمت اصلی و کلیدی تصویر با برقراری توازن میان ترکیب اجزا[5] و اطّلاعات مفید تصویر است. در این پژوهش برای استخراج ناحیهٔ کلیدی تصویر از چارچوب کلی «چشم‌انداز مورد توجّه» (*Ma, 2003*) استفاده شده است.

یکی از مراحل استخراج این ناحیه، کوانتیزه‌کردن رنگ است، که الگوریتم F-PSO-GA (ترکیبی از الگوریتم‌های PSO و GA به‌همراه FCM) بدین منظور پیشنهاد شده است (برای مشاهدهٔ جزییات این روش کوانتیزاسیون، به (*Alamdar, 2010*) مراجعه شود). این الگوریتم از مزایای هر سه الگوریتم بهره می‌برد؛ مسأله اساسی PSO این است که الگوریتم به‌طور زودرس به نقطه‌ای پایدار همگرا می‌شود، که لزوماً بیشینه نیست. برای جلوگیری از این واقعه به‌روزرسانی موقعیت از طریق سازوکار پیوندی GA انجام می‌شود و میزان شایستگی کروموزم‌ها با استفاده از تابع هدف فازی محاسبه می‌گردد. این الگوریتم باعث کوانتیزاسیون بهتر با خطای کمتر می‌شود (*Alamdar, 2010*). (شکل ۲) یک تصویر و ناحیهٔ کلیدی مشخّص‌شده در آن نشان می‌دهد.

(الف) (ب)

(شکل ۲): (الف) تصویر اصلی، (ب) ناحیهٔ کلیدی استخراج شده از تصویر (الف)

[1] - Visual attention
[2] - Attended view
[3] - Attended areas
[4] - Attended points
[5] - Composition

۴-۲- استخراج ویژگی از ناحیهٔ کلیدی تصویر

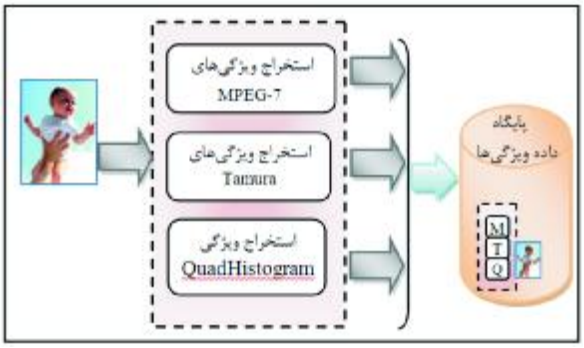
پس از مشخّص‌شدن ناحیهٔ کلیدی تصاویر، باید ویژگی‌های تصاویر از این ناحیه استخراج می‌شوند (شکل ۳). در آزمایش‌های انجام شده در این پروژه، از هر تصویر هفت ویژگی استخراج شده‌اند تا خصوصیات مختلف تصویر را از نظر رنگ، بافت و موقعیت مکانی استخراج کنند. سعی شده است ویژگی‌هایی انتخاب شوند که در کارهای پیشین به مراتب بیشتر استفاده شده‌اند و نتایج بهتری داشته‌اند، ضمن اینکه استخراج آنها سرعت مناسبی داشته باشد و تا حدّی فشرده باشد. ویژگی‌های سیستم پیشنهادی، ترکیبی از سه ویژگی Mpeg-7 (CLD[6]، SCD[7]، EHD[8]) (*Manjunath, 2001*)، سه ویژگی بافت Tamura (درشتی و ریزی[9]، شدّت‌نور[10]، راستا[11]) (*Tamura, 1978*) و یک ویژگی پیشنهادی به‌نام QuadHistogram استفاده شده است (جزئیات مربوط به این ویژگی در (*Alamdar, 2011*) آمده است). در ویژگی پیشنهادی QuadHistogram تجزیهٔ درخت چهارتایی[12] بر روی تصاویر اعمال شده تا بلوک‌های همگن[13] و در اندازه‌های مختلف مشخّص شوند. سپس هیستوگرام رنگ و پیچیدگی برای هر سطح از بلوک (بلوک‌های هم‌اندازه) استخراج می‌شود. برای محاسبهٔ تطابق دو تصویر در این ویژگی از فاصلهٔ ذیل استفاده شده است:

$$d_{QH}(Q,I) = \alpha(\alpha_1(\sum_{h=1}^{L} Wh_b \sum_{k=1}^{K} |\, h_b^I(k) - h_b^Q(k)\,|) + \beta_1(\sum_{k=1}^{K} |\, h^I(k) - h^Q(k)\,|)) + \beta(\sum_{k=1}^{L} Ws_b \,|\, S_b^I - S_b^Q\,|) \quad (1)$$

I و Q تصاویر مورد مقایسه هستند. h_b هیستوگرام bامین سطح از بلوک‌های درخت چهارتایی است ($b = 1,2,...,L$)، که L آخرین سطح است؛ به‌عنوان مثال تصویری با اندازهٔ ۱۲۸x۱۲۸دارای هشت سطح بلوک می‌باشد(۱۲۸=2^{8-1}) و سطح دوم بلوک این تصویر، شامل بلوک‌هایی با اندازه ۶۴x۶۴ است و حداکثر تعداد این بلوک‌ها می‌تواند ۴ بلوک باشد. h_b هیستوگرام رنگ استخراج شده برای بلوک‌های همگن سطح bام است. S_b تعداد بلوک‌های همگن همین سطح (پیچیدگی) می‌باشد. Wh'_b وزن نرمال شدهٔ h_b است. K تعداد بین‌های هیستوگرام است و h هیستوگرام رنگ سراسری است. Ws_b وزن S_b می‌باشد. α_1, β_1 نسبت ترکیب h_b با h و α, β نسبت ترکیب

[6] - Color Layout Descriptor
[7] - Scalable Color Descriptor
[8] - Edge histogram descriptor
[9] - Coarseness
[10] - Contrast
[11] - Directionality
[12] - Quadtree decomposition
[13] - Homogenous blocks

هیستوگرام با پیچیدگی است. این ویژگی در مقایسه با هیستوگرام رنگ عملکرد قابل توجهی از خود نشان می‌دهد *(Alamdar, 2011)*.



(شکل ۳): استخراج ویژگی از ناحیهٔ کلیدی تصویر

۴-۳- اندازه‌گیری فاصلهٔ تصاویر

در این مرحله باید معیاری برای اندازه‌گیری فاصله و شباهت مشخص شود که فاصله یا شباهت بین دو تصویر را بر مبنای ویژگی‌های استخراج شده، کمّی‌سازی کند. یکی از مؤلفه‌های کلیدی در الگوریتم‌های خوشه‌بندی، نحوهٔ ترکیب ویژگی‌ها است که درحقیقت همان اندازه‌گیری شباهت و فاصلهٔ بین اشیاست *(van Leuken, 2009)*. در خوشه‌بندی و بازیابی تصویر، استفادهٔ همزمان از ویژگی‌های مختلف برای محاسبهٔ شباهت یا فاصلهٔ بین تصاویر، معمول است. هر یک از این ویژگی‌ها نمایان‌گر جهات مختلف[1] تصویر هستند و هریک بازنمایی[2] مخصوص خود (به‌عنوان مثال یک عدد[3]، یک بُردار، یک هیستوگرام) و روش تطابق[4] متناظر (به‌عنوان مثال فاصله اقلیدسی، نرم L_1) دارند. این ویژگی‌ها ممکن است در حدود[5] و توزیع[6] متفاوتی باشند، که لازم است همهٔ آن‌ها به منظور استفاده در الگوریتم خوشه‌بندی، در یک مقدار واحد تجمیع شوند.

در اینجا از یک «استراتژی رتبه‌بندی پویا[7]» استفاده شده است که در *(van Leuken, 2009)* پیشنهاد شده است. در این استراتژی، ویژگی‌های مختلف توسط واریانس فواصل (نرمال شده) همهٔ تصاویر، وزن‌دهی می‌شود. فاصلهٔ تصویر بر طبق یک ویژگی بر واریانس آن ویژگی تقسیم می‌شود. این امر باعث می‌شود که فواصل تصاویر بر طبق ویژگی‌های مختلف که در محدوده‌های مختلف قرار دارند، در یک محدودهٔ مشابه قرار گیرند و به ویژگی‌هایی که تفکیک‌کننده‌های[8] خوبی هستند، وزن بزرگ‌تری نسبت داده می‌شود. فاصلهٔ بین تصویر i و j این‌چنین محاسبه می‌شود:

$$d_{ij} = \frac{1}{f}\sum_{k=0}^{f}\frac{1}{\sigma_k^2}d_k(i,j) \qquad (۲)$$

که f تعداد ویژگی‌ها است، $d_k(i,j)$ فاصلهٔ بین تصویر i و j بر طبق ویژگی kاُم است؛ برای محاسبه فاصله در ویژگی‌های EHD و SCD از نُرم L1 استفاده می‌شود، ویژگی CLD فاصلهٔ منحصربه‌فرد دارد *(Manjunath, 2001)*. ویژگی‌های تامورا از فاصلهٔ اقلیدسی استفاده می‌کنند *(Tamura, 1978)* و فاصله در ویژگی QuadHistogram طبق رابطهٔ ۱ محاسبه می‌شود. σ_k^2 واریانس تمامی فواصل تصاویر بر طبق ویژگی kاُم در مجموعهٔ تصاویر است.

۴-۴- وزن‌دهی ویژگی‌ها

در این مرحله، ویژگی‌های استخراج‌شده وزن‌دهی می‌شوند. هر ویژگی استخراج شده می‌تواند یک «درجهٔ اهمیت[9]» داشته باشد که وزن ویژگی نامیده می‌شود *(Wang, 2004)*. انتساب وزن به ویژگی‌ها، توسعه‌ای بر «انتخاب ویژگی»[10] است. مقادیر وزن در انتخاب ویژگی فقط یک یا صفر هستند، ولی در «وزن‌دهی ویژگی‌ها»، وزن‌ها در بازهٔ [۱،۰] هستند. در این پروژه وزن‌دهی ویژگی‌ها بر مبنای روشی صورت می‌گیرد که در *(Wang, 2004)* پیشنهاد شده است. این وزن‌دهی به‌وسیلهٔ یادگیری وزن ویژگی‌ها بر مبنای شباهت بین نمونه‌ها[11] با روش نزول گرادیان انجام می‌شود. این یادگیری با حدّاقل‌کردن تابع ارزیابی[12] $E(w)$ انجام می‌شود و $E(w)$ با رابطهٔ ۳ تعریف می‌شود:

$$E(w) = \frac{2}{n(n-1)}\sum_i\sum_{j\neq i}\frac{1}{2}(\rho_{ij}^{(w)}(1-\rho_{ij}^{(1)}) + \rho_{ij}^{(1)}(1-\rho_{ij}^{(w)})) \qquad (۳)$$

در این رابطه $\rho_{ij}^{(w)}$ مقدار اندازهٔ شباهت وزن‌دار، $\rho_{ij}^{(1)}$ مقدار اندازهٔ شباهت بدون وزن و n نشان‌دهندهٔ تعداد نمونه‌ها (تصاویر) است. برای حدّاقل‌کردن رابطهٔ فوق از روش نزول گرادیان استفاده می‌شود.

۴-۵- رویکرد اوّل خوشه‌بندی پیشنهادی

این رویکرد کلّی شامل دو مرحله است که مرحلهٔ اول آن مشخّص‌کردن مرکز یا همان نمایندهٔ خوشه‌هاست. این مراکز به روشی مشابه روش انتخاب نمایندهٔ الگوریتم Folding، انتخاب می‌شوند. پس از مشخّص‌شدن مراکز، در مرحلهٔ دوم

[1] - Different aspects
[2] - Representation
[3] - A scalar
[4] - Matching method
[5] - Range
[6] - Distribution
[7] - Dynamic ranking strategy
[8] - Discriminators
[9] - Importance degree
[10] - Feature selection
[11] - Samples
[12] - Evaluation function

خوشه‌ها شکل می‌گیرند. در اینجا الگوریتم خوشه‌بندی پیشنهادی Fuzzy Folding معرفی می‌شود که تعمیمی بر الگوریتم Folding *(van Leuken, 2009)* است و طی این دو مرحله خوشه‌بندی را انجام می‌دهد. در ادامه ابتدا روش‌های مختلف مشخّص‌کردن مرکز دسته‌ها بیان می‌شود و سپس الگوریتم پیشنهادی Fuzzy Folding تشریح می‌گردد.

۴-۵-۱- مشخّص‌کردن مرکز دسته‌ها

برای مشخّص‌کردن مراکز، مانند روش انتخاب نماینده در الگوریتم Folding (RS1) عمل می‌شود. تصویر اوّل رتبه‌بندی یعنی L_1 همیشه به عنوان نماینده انتخاب می‌شود. در هنگام پیمودن فهرست مرتب‌شده به سمت پایین، هر تصویر با مجموعه‌ای از نماینده‌ها که پیش از این انتخاب شده‌اند مقایسه می‌شود. وقتی که یک تصویر مثل i به‌قدر کافی نسبت به همهٔ نمایندگان انتخاب شده در R بی‌شباهت باشد، به R اضافه می‌شود، یعنی:

$$\forall j \in R \quad d_{ij} > \varepsilon \qquad (۴)$$

می‌توان مراکز دسته‌ها را مبتنی بر تصویر میانگین مشخّص کرد. در این روش (RS2) مانند روش پایه عمل می‌شود با این تفاوت که تصویر میانگین تصویری در نظر گرفته می‌شود که ویژگی‌های آن میانگین ویژگی‌های تمامی تصاویر است، یعنی:

$$x\_avg_s = \sum_{i=1}^{N} x_{is}, \quad s = 1,..., f \qquad (۵)$$

که x_{is} نشان‌دهندهٔ sامین ویژگی تصویر i است.

در این مقاله روش «مشخّص‌کردن مراکز دسته کاهشی مبتنی بر تصویر میانگین» برای تعیین مراکز پیشنهاد شده است. در این روش (RS3) همانند روش مشخّص‌کردن مرکز دسته مبتنی بر تصویر میانگین عمل می‌شود و علاوه‌بر آن به این موضوع نیز توجّه شده است که تصاویر در فهرست مرتب شده به‌گونه‌ای قرار گرفته‌اند که نتایج مرتبط‌تر به‌طور معمول در اوایل نتایج جستجو قرار دارند و هرچه که به انتهای فهرست برویم از این ارتباط بیشتر کاسته می‌شود. بدین منظور ε از ابتدا تا انتها یکسان نیست و با پیمودن فهرست به سمت پایین مقدار آن بزرگتر می‌شود تا از انتهای فهرست نماینده‌های کمتری انتخاب شود. اگر c یک مقدار ثابت از پیش تعیین شده باشد، این تغییر مقدار به‌صورت رابطهٔ ۶ قابل بیان است.

$$\varepsilon_{new} = \varepsilon_{old} + c\,\varepsilon_{old} \qquad (۶)$$

الگوریتم مشخص کردن مرکز دسته‌ها در (شکل ۴) نشان داده شده است.

Algorithm Representatives Selection
Input: Ranked list L of I and distance matrix m_d
Output: Representatives R

1: Let the image L_1 be the first representative R_1
2: **for** Each image L_i **do**
3: **if** $d(L_i, R_k) > \varepsilon$ for all representatives R_k **then**
4: add L_i to the set of representatives R
5: **end for**

(شکل ۴): شبه کُد مشخّص‌کردن مرکز دسته‌ها

۴-۵-۲- الگوریتم Fuzzy Folding

در این الگوریتم پیشنهادی، نمایندگان خوشه به همان صورتی که در بخش قبل ذکر شد، مشخّص می‌شوند. بعد از مشخّص‌شدن نماینده‌ها، خوشه‌ها در اطراف هر نمایندهٔ خوشه با استفاده درجهٔ عضویت فازی شکل می‌گیرند. یعنی هر تصویر در L به خوشه‌ای نسبت داده می‌شود که درجهٔ عضویت بیشتری داشته باشد. در این روش یک درجه عضویت کلّی پیشنهاد شده است که برای محاسبهٔ آن، ابتدا درجهٔ عضویت هر تصویر بر طبق هریک از ویژگی‌های استخراج شده (u_s) و بر طبق تمامی ویژگی‌های استخراج شده (u_d) محاسبه می‌شود و سپس درجهٔ عضویت کلّی (u_{ij}) بر اساس درجهٔ عضویت‌های ویژگی‌ها شکل می‌گیرد. اگر $m \in [1,\infty)$ میزان فازی‌بودن باشد، درجهٔ عضویت طبق روابط ۷ و ۸ محاسبه می‌شود.

$$u_{ij} = \left(\sum_{s=1}^{f} w_s\, u_s(i,j) + u_d(i,j)\right)/2 \qquad (۷)$$

$$u_s(i,j) = \frac{(d_s(i,j))^{-2/(m-1)}}{\sum_{l=1}^{K}(d_s(l,j))^{-2/(m-1)}},$$

$$u_d(i,j) = \frac{(d_{ij})^{-2/(m-1)}}{\sum_{l=1}^{K}(d_{lj})^{-2/(m-1)}}, \qquad (۸)$$

$$\text{for} \quad s = 1,\ldots,f, \quad i,j = 1,\ldots,N \quad \text{and} \quad j \in R$$

که در روابط فوق، f نشان‌دهندهٔ تعداد ویژگی‌ها و K تعداد نماینده‌های انتخاب شده در مرحلهٔ قبل است، یعنی همان تعداد خوشه‌ها. $d_s(i,j)$ فاصلهٔ بین تصویر i و j بر طبق ویژگی sام است و d_{ij} فاصلهٔ بین تصویر i و j بر طبق تمامی ویژگی‌ها طبق معادلهٔ ۱ است. w_s وزن نرمال شدهٔ متناسب با هر ویژگی است که اگر برای هر ویژگی وزنی درنظر گرفته نشود، مقدار تمامی وزن‌ها $1/f$ قرار داده می‌شود.

پس از محاسبهٔ درجهٔ عضویت کلّی هر تصویر تا نماینده‌ها، تصویر به خوشه‌ای نسبت داده می‌شود که این مقدار برای نمایندهٔ آن حداکثر باشد یعنی:

$$j = \underset{1 \le j \le K}{\operatorname{argmax}}(u_{ij}) \qquad (۹)$$

اگر تصویر j ، kامین نماینده در R باشد، C_k خوشهٔ این تصویر است و مرکز آن هم R_k می‌باشد. مراحل الگوریتم در (شکل ۵) نشان داده شده است.

Algorithm Fuzzy Folding
Input: Ranked list L of I and distance matrix m_d
Output: Clustering C

1: R Select by Representatives Selection
2: **for** Each image $L_i \notin R$ do
3: Find representative R_k that L_i has the highest membership to
4: Assign L_i to the cluster R_k
5: **end for**

(شکل ۵): شبه کد الگوریتم Fuzzy Folding

۴-۶- رویکرد دوم خوشه‌بندی پیشنهادی

این رویکرد کلّی برمبنای درخت‌های سلسله‌مراتبی می‌باشد که دارای سه مرحلهٔ کلی مشخّص‌کردن مرکز دسته‌ها، ساخت درخت سلسله‌مراتبی و برش درخت بر اساس مرکز دسته است. در ادامه الگوریتم پیشنهادی Folding Tree شرح داده می‌شود.

۴-۶-۱- الگوریتم Folding Tree

الگوریتم پیشنهادی Folding Tree بر پایهٔ خوشه‌بندی سلسله‌مراتبی شکل می‌گیرد. خوشه‌بندی سلسله‌مراتبی به‌طور متوالی با ادغام خوشه‌های کوچک‌تر در خوشه‌های بزرگ‌تر و یا با تقسیم خوشه‌های بزرگ‌تر، پیش می‌رود. نتیجهٔ الگوریتم یک درخت از خوشه‌هاست، که نمودار درختی[1] نامیده می شود؛ که این نمودار ارتباط خوشه‌ها را با هم نشان می‌دهد. با قطع‌کردن نمودار درختی در یک سطح دلخواه، خوشه‌بندی موردنظر به‌دست می‌آید(*Halkidi, 2001*).

یکی از مسائل مطرح در خوشه‌بندی سلسله‌مراتبی، مشخّص‌کردن تعداد مناسب خوشه‌هاست. به‌طور عمومی برای تعیین این تعداد از تحلیل آماری دسته‌های تولید شده و فضای بین آن‌ها استفاده می‌شود(*Milligan, 1985; Salvador, 2004; Boberg, 1993*).

در فرآیند خوشه‌بندی در الگوریتم Folding Tree، ابتدا درخت خوشه‌بندی براساس یکی از روش‌های خوشه‌بندی سلسله‌مراتبی ساخته می‌شود. پس از آن باید درخت طی یک روال مناسب برش زده شود. بدین منظور مراکز دسته‌ها طبق روش مطرح شده در بخش ۴-۵-۱ انتخاب می‌شوند و طبق مکان قرار گرفتن این مراکز در درخت، برش به‌گونه‌ای انجام می‌شود که جزئی‌ترین دسته‌بندی ممکن تولید شود. از برگ‌ها که شامل تمامی تصاویر است شروع کرده و طی ادغام نزدیک‌ترین زیردرخت‌ها، به‌گونه‌ای در درخت بالا می‌رود که دو نمایندهٔ مختلف با هم در یک خوشه قرار نگیرند. بدین منظور پیمایش از برگ‌ها به‌صورت سطری به سمت گرهٔ ریشه انجام می‌شود و اگر برگی با برگی که مرکز یکی از خوشه‌ها است، ادغام شد، خوشهٔ این برگ مشخّص شده و برابر با خوشهٔ مرکز مذکور می‌گردد. در سطوح بعدی، خوشهٔ هر زیردرخت قبل از ادغام با زیردرخت دیگری بررسی می‌شود، اگر خوشهٔ هر دو یکسان باشد و یا فقط خوشهٔ یکی از زیردرخت‌ها مشخّص شده باشد، دو زیردرخت با هم ادغام می‌شوند و خوشهٔ آن نیز تعیین می‌گردد، این امر در (شکل ۶ الف) مشخّص شده است. ولی اگر خوشهٔ هر دو زیردرخت نامشخص باشد، تعیین خوشه به سطوح بعد ارجاع داده می‌شود. اگر خوشهٔ هر دو زیردرخت مشخص و متفاوت باشد، این دو زیردرخت قابل ادغام نیستند و درخت در این قسمت برش‌خورده و این دو خوشه نیز بیش از این قابل گسترش نیستند و این روال تا ریشهٔ درخت ادامه پیدا می‌کند.

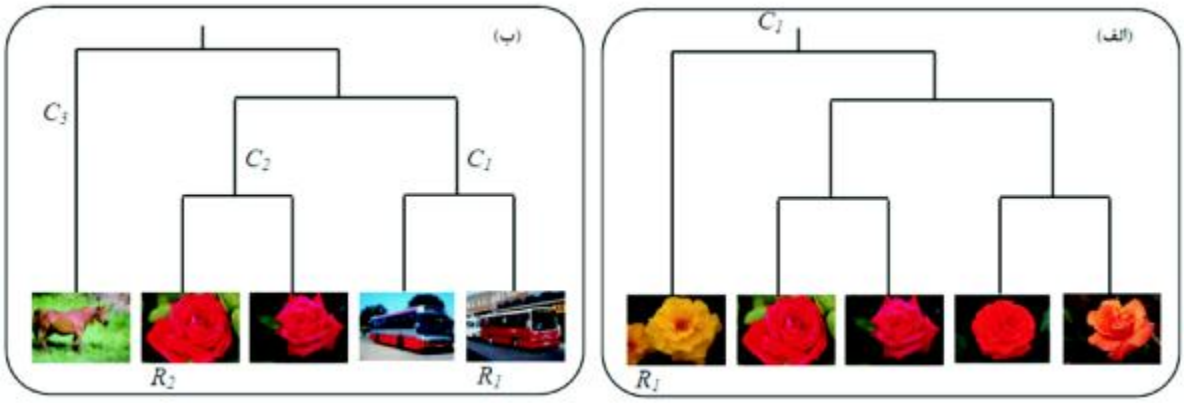
یکی از مواردی که ممکن است در پیمایش درخت مشکل ایجاد کند، زمانی است که خوشهٔ یک برگ یا زیردرخت تعیین نشده باشد و در گام بعدی باید با قسمتی از درخت ادغام شود که قبل از آن برش خورده است؛ به‌عنوان مثال زمانی که باید با دو زیردرخت که قابل ادغام نبوده و هر کدام خوشهٔ متفاوتی را مشخص می‌کنند، ادغام شود که امر امکان‌ناپذیر است. برای حلّ این مشکل این زیردرخت یک خوشهٔ جدید را تشکیل می‌دهد که نماینده آن می‌تواند آن عنصری از خوشه باشد که میانگین فاصله‌اش تا دیگر اعضای خوشه حدّاقل باشد (رابطه ۱۰).

$$R'_k = \arg\min_{1<j<n_k} \left(\sum_{i=1}^{n_k} d_{ij} / n_k\right), \quad for\ k = 1,\ldots,K\ and\ i,j \in C_k \qquad (۱۰)$$

این مسأله در ش(کل ۶ب) نشان داده شده است که اگر R_1 و R_2 مراکز خوشه باشند، دو خوشهٔ C_1 و C_2 تشکیل شده و ادغام متوقف می‌شود. زیردرخت بعدی که خوشه‌اش نامعین است قابل ادغام با آن‌ها نیست و درنتیجه خوشهٔ C_3 جدید تشکیل می‌گردد.

الگوریتم شرح داده شده در (شکل ۷) نشان داده شده است.

[1] - Dendrogram



(شکل ۶): (الف) ادغام زیردرخت‌ها و شکل‌گیری خوشه، (ب) شکل‌گیری خوشهٔ جدید

Algorithm Folding Tree

Input: Ranked list L of I and distance matrix m_d

Output: Clustering C

1: R Select by Representatives Selection
2: Construct Hierarchical Tree T
3: **for** Each R_i, Assign R_i to the cluster of R_i, **end for**
4: **for** Each two closest Subtree T_i, T_j do
5: Merge T_i, T_j to T_k
6: **if** T_i was cut
7: **if** T_j was assigned to a cluster **then** T_j, T_k are cut
8: **else** T_j was assigned to a new cluster with new representative, T_j, T_k are cut
9: **else if** T_i, T_j were assigned to clusters **then** T_i, T_j, T_k are cut
10: **else if** T_i is assigned to a cluster and T_j is not assigned to a cluster **then**
11: Assign T_j, T_k to the cluster of T_i
12: **do** lines 6-11 for T_j
13: **end for**

(شکل ۷): شبه کُد الگوریتم Folding Tree

۵- آزمایش‌ها و نتایج

در این بخش، ابتدا پایگاه داده تصویری مورد استفاده معرفی شده است و سپس به معیارهای ارزیابی برای اندازه‌گیری کارایی رویکردها و الگوریتم‌های پیشنهادی پرداخته می‌شود. در انتها روش آزمون و نتایج حاصل از آنها بیان خواهند شد.

۵-۱- پایگاه داده تصاویر

برای ایجاد پایگاه دادهٔ مورد استفاده جهت آزمون روش‌های پیشنهادی از رویکرد مطرح در (*Liu, 2004; van Leuken, 2009; van Zwol, 2008; Wang, 2010; Cai, 2004; Jing, 2008; Tian, 2010*) استفاده شده است. مطابق با این رویکرد تصاویر به‌دست آمده از موتوهای جستجوی مختلف، مبنای تشکیل این پایگاه داده را شکل می‌دهد.

پایگاه دادهٔ تصاویر این پژوهش، شامل تصاویری است که با جستجوی ۶۵ عنوان کلمهٔ «پرس‌وجو» در موتورهای جستجوی تصویر Google و Yahoo جمع‌آوری شده‌اند. برای هر عنوان حداکثر هزار تصویر اوّل دانلود شده‌اند. برخی تصاویر ممکن است به علّت فیلترینگ و یا نبود تصویر اصلی بر روی سرور مورد نظر غیر قابل دانلود باشند، درنتیجه تعداد تصاویر هر عنوان ممکن است کمتر از تعداد ذکر شده باشد. موتور جستجوی Yahoo در اکثر موارد ۴۸ صفحه نمایش می‌دهد، که این تعداد صفحات ممکن است با توجّه به کلمهٔ پرس‌وجو کمتر باشد. حداکثر تعداد صفحات تصویری که موتور جستجوی Google نشان می‌دهد پنجاه است که در اکثر موارد تعداد این صفحات کمتر است. در موتور جستجوی Google تعداد تصاویر نمایش داده شده در هر صفحه بیست عدد است و در موتور جستجوی Yahoo این تعداد ۲۸ می‌باشد (در صفحهٔ ۴۸ تعداد تصاویر ۲۰ عدد است).

سعی شده است که کلمات پرس‌وجویی انتخاب شوند که بیشتر مورد جستجو قرار می‌گیرند و عناوین متفاوت و متنوعی را شامل شوند. کلمات جستجو شده هم شامل کلمات ذاتاً مبهم است چه «ابهام مفهوم کلمه» وچه «ابهام مختص نوع» و هم شامل کلمات غیرمبهم. این امر سبب می‌شود که کارایی سیستم در هر دو حالت مورد ارزیابی قرار گیرد.

در (جدول ۱) برخی از کلمات جستجو شده به همراه نمونه‌هایی از تصاویر آن آورده شده‌اند. به بعضی از مشخّصات مربوط به این کلمات پرس‌وجو نیز اشاره شده است، مانند معانی و بعضی از انواع معنایی و بصری، موتور جستجوی مورد استفاده و تعداد دسته‌های تعیین شده توسط کاربر در مرحلهٔ ایجاد حقیقت پایه.

۵-۲- معیارهای ارزیابی

در خوشه‌بندی یک مجموعه داده، هیچ کلاس از پیش تعیین شده‌ای و هیچ نمونه‌ای که بتواند اعتبار خوشه‌های شکل‌گرفته توسط الگوریتم‌های خوشه‌بندی را نشان دهد، وجود ندارد. بنابراین برای مقایسهٔ الگوریتم‌های مختلف خوشه‌بندی، وجود چند معیار اعتبارسنجی لازم است. در کل سه معیار بنیادین برای بررسی صحت اعتبار خوشه وجود دارد: معیارهای خارجی[1]، معیارهای داخلی[2]، معیارهای نسبی[3] (*Gan, 2007*). در راهبرد معیارهای خارجی، نتایج الگوریتم خوشه‌بندی بر مبنای یک دسته‌بندی از پیش معین شده ارزیابی می‌شود، که این دسته‌بندی بر مجموعه داده اعمال شده و ساختار ذاتی این مجموعه را باز تاب می‌دهد. درحقیقت هدف این معیار ارزیابی دو خوشه‌بندی متفاوت با هم است. در اینجا دو معیار خارجی Fowlkes-Mallows (*Fowlkes, 1983*) و Variation of information (*Meilă, 2007*) شرح داده می‌شوند که در (*van Leuken, 2009; Wang, 2010*) برای مقایسه نتایج حاصل از خوشه‌بندی تصاویر موتورهای جستجو مورد استفاده قرار گرفته‌اند.

Fowlkes-Mallows(FM): اگر I یک مجموعهٔ تصویر و C و C' دو خوشه‌بندی باشند، همهٔ جفت تصاویر ممکن بر مبنای I به چهار دستهٔ قابل تقسیم‌اند: N_{11}: هر دو تصویر در هر دو خوشه‌بندی C و C'، در یک خوشه هستند. N_{00}: هر دو تصویر در هر دو خوشه‌بندی C و C'، در خوشه‌های متفاوت قرار دارند. N_{10}: دو تصویر در خوشه‌بندی C در یک خوشه هستند اما در خوشه‌بندی C' در خوشه‌های متفاوت قرار دارند. N_{01}: دو تصویر در خوشه‌بندی C' در یک خوشه هستند امّا در خوشه‌بندی C در خوشه‌های متفاوت قرار دارند.

مقدار بالای شاخص FM، نشان‌دهندهٔ شباهت دو خوشه است و بر اساس دو معیار نامتقارن ذیل شکل می‌گیرد.

$$W_I(C,C') = \frac{N_{11}}{N_{11}+N_{01}}$$
$$W_{II}(C,C') = \frac{N_{11}}{N_{11}+N_{10}} \qquad (۱۱)$$

این شاخص میانگین هندسی این دو معیار است که آن را معیاری متقارن می‌سازد.

$$FM(C,C') = \sqrt{W_I(C,C')W_{II}(C,C')} \qquad (۱۲)$$

Variation of information(VI): این معیار مقایسه، بر مبنای ساختاری است که «جدول هم‌رخدادی[4]» یا «ماتریس آشفتگی[5]» نامیده می‌شود. «جدول هم‌رخدادی» دو خوشه‌بندی، یک ماتریس $K \cdot K'$ است که اُمین kk' مؤلّفهٔ، تعداد نقاط مشترک در خوشه‌های C_k از C و $C_{k'}$ از C' است. این معیار مقایسهٔ خوشه‌بندی، معیار تغییرپذیری اطّلاعات[6] $VI(C,C')$ است و بر مبنای مفهوم آنتروپی شرطی

[1] -External criteria
[2] -Internal criteria
[3] -Relative criteria
[4] - Contingency table
[5] - Confusion matrix
[6] - Variation of information

می‌باشد(Fowlkes, 1983). اگر C یک خوشه‌بندی باشد، احتمال اینکه یک تصویر که به‌طور تصادفی انتخاب شده متعلق به خوشهٔ C_k با اندازهٔ n_k باشد، برابر است با

$$P(k)=\frac{n_k}{N} \qquad (۱۳)$$

این رابطه متغیری تصادفی را تعریف می‌کند که K مقدار می‌گیرد. «عدم قطعیت[1]» خوشه‌ای که تصویر به آن متعلق است برابر با آنتروپی متغیر تصادفی رابطهٔ ۱۴ است.

$$H(C)=-\sum_{k=1}^{K}P(k)\log P(k) \qquad (۱۴)$$

اطّلاعات متقابل[2] $I(C,C')$، یعنی اطلاعاتی که یک خوشه‌بندی درمورد یک خوشه‌بندی دیگر دارد، به‌طور مشابه قابل تعریف است. ابتدا، احتمال اینکه یک تصویر که به‌طور تصادفی انتخاب شده متعلق به خوشهٔ C_k در خوشه‌بندی C و متعلق به خوشهٔ $C'_{k'}$ در خوشه‌بندی C' باشد ، برابر است با

$$P(k,k')=\frac{|C_k\cap C'_{k'}|}{N} \qquad (۱۵)$$

سپس، اطّلاعات متقابل $I(C,C')$ با مجموع آنتروپی‌های تمامی جفت خوشه‌های ممکن تعریف می‌شود.

$$I(C,C')=\sum_{k=1}^{K}\sum_{k'=1}^{K'}P(k,k')\log\frac{P(k,k')}{P(k)P'(k')} \qquad (۱۶)$$

ضریب اطّلاعات متقابل، می‌تواند به‌عنوان کاهش‌دهندهٔ عدم قطعیت از یک خوشه‌بندی به خوشه‌بندی دیگر در نظر گرفته شود. تغییرپذیری اطّلاعات به‌صورت رابطهٔ ذیل نوشته می‌شود.

$$VI(C,C')=[H(C)-I(C,C')]+[H(C')-I(C',C)] \qquad (۱۷)$$

تغییرپذیری اطّلاعات بر روی ارتباط بین یک نقطه[3] و خوشهٔ آن متمرکز می‌شود. این معیار، تفاوت در این ارتباط را که بر روی تمامی نقاط میانگین گرفته شده است، بین دو خوشه‌بندی اندازه می‌گیرد. اگر تغییرپذیری اطلاعات کم باشد، نشان‌دهندهٔ شباهت دو خوشه‌بندی است.

۵-۳- آزمون‌ها

برای ارزیابی کارایی رویکردها و الگوریتم‌های پیشنهادی همانند روش مطرح شده در (*van Leuken, 2009; Wang, 2010*) عمل می‌شود که در این روش ارزیابی با مقایسه نتایج الگوریتم‌ها با نتایج حاصل از دسته‌بندی تصاویر توسط کاربر با استفاده از معیارهای FM و VI انجام می‌گیرد. در این بخش به چگونگی ساخت این دسته‌بندی از پیش تعیین شده و نتایج آزمون روش‌های پیشنهادی پرداخته می‌شود. همچنین روش‌ها پیشنهادی با خود و دیگر روش‌ها مقایسه می‌شوند.

۵-۳-۱- ایجاد حقیقت پایه[4]

برای اینکه بتوان نتایج و کارایی سیستم را مورد ارزیابی قرار داد، باید یک «حقیقت پایه» داشت یعنی یک دسته‌بندی از پیش تعیین شده که بتوان آن دسته‌بندی را با نتایج خوشه‌بندی سیستم مقایسه کرد. بدین منظور از شش کاربر خواسته شد که این ۶۵ عنوان را بر مبنای ویژگی‌های بصری گروه‌بندی کنند. این شش کاربر به‌صورت مستقل و بدون دانستن و یا دیدن جواب سیستم، ارزشیابی خود را انجام دادند. فرآیند دسته‌بندی توسط کاربر به‌صورت ذیل انجام گرفت:

۱. مشاهدهٔ تمامی تصاویر: ابتدا کلّ پنجاه تصویر اوّل به کاربر نشان داده می‌شود و کاربر می‌تواند تمامی تصاویر را به سرعت بررسی و مرور کند. این امر به کاربر کمک می‌کند تا درک و تجسّمی کلّی از این تصاویر به‌دست آورد و ایدهٔ چگونگی دسته‌بندی و تعداد دسته‌ها در ذهن او نقش ببندد.

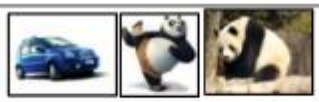
۲. شکل‌دهی دسته‌ها: در این مرحله کاربر به تعداد دسته‌های مورد نظر پوشهٔ جدید ساخته و با قرار دادن تصاویر در پوشهٔ مربوطه گروه‌بندی تصاویر انجام می‌شود. در طول این پروسه، ممکن است کاربر به تصویری برسد که به نظر او به هیچ یک از دسته‌های ایجاد شده متعلق نیست؛ در این حالت کاربر می‌تواند گروه جدیدی به گروه‌های موجود بیافزاید و همچنین تصاویر نامرتبط در یک پوشه قرار می‌گیرند.

۳. مشخّص‌کردن نمایندهٔ دسته: پس از شکل‌گیری کلّ دسته‌ها از کاربر خواسته می‌شود که در هر دسته یکی از تصاویر را به‌عنوان نمایندهٔ دستهٔ جاری انتخاب کند.

بیست تصویر اوّل نتیجه شده از جستجوی کلمه "mouse" و حقیقت پایه ایجادشده توسط دسته‌بندی کاربر، در (شکل ۸) نشان داده شده است. همچنین در ستون آخر (جدول ۱)، تعداد دسته‌های تعیین شده توسط کاربر برای بعضی از کلمات جستجوشده آورده شده است.

[1] - Uncertainty
[2] - Mutual information
[3] -Point
[4] - Ground truth establishment

(جدول ۱): بعضی از کلمات جستجو شده به همراه انواع بصری و معنایی آنها

ردیف	کلمهٔ پرس-وجو	تصاویر نمونه	معانی	بعضی از انواع معنایی و بصری	موتور جستجو	تعداد دسته‌های تعیین شده توسط کاربر
۱	apple		سیب	میوهٔ سیب، آرم شرکت، محصولات متنوع شرکت	Google	۱۰
۲	beetle		سوسک	انواع سوسک، نوعی ماشین، شخصیت کارتونی beetle man	Yahoo!	۱۱
۳	glasses		عینک	انواع عینک، لیوان‌ها و جام‌ها	Yahoo!	۹
۴	Honda		هوندا	انواع اتومبیل و موتور هوندا و آرم شرکت	Google	۱۰
۵	jaguar		پلنگ خالدار	انواع پلنگ، اتومبیل جگوار و داخل و اجزای آن و آرم شرکت	Google	۶
۶	Mercury		جیوه، سیاره عطارد، تیر	سیاره عطارد، نوعی اتومبیل، نوعی اسکیت، الههٔ یونانی	Google	۷
۷	mouse		موش	انواع موش واقعی و کارتونی، موشواره رایانه	Google	۸
۸	panda		پاندا	انواع پاندای واقعی و کارتونی، نوعی اتومبیل	Yahoo!	۵
۹	polo		چوگان	ورزش چوگان، نوعی اتومبیل، نوعی تیشرت، آرم	Google	۵
۱۰	swan		قو	انواع قوی سفید و سیاه و نقاشی آن، نوعی کفش	Yahoo!	۱۲

(شکل ۸): ۲۰ تصویر اوّل نتیجه شده از جستجوی کلمه mouse و حقیقت پایه ایجادشده توسط دسته‌بندی کاربر

۵-۳-۲- نتایج آزمون

پس از مشخّص‌شدن «حقیقت پایه»، نتایج بر مبنای معیارهای Fowlkes-Mallows (FM) و Variation of Information (VI) مقایسه می‌شوند. افزایش نسبی امتیاز FM و کاهش نسبی امتیاز VI نیز محاسبه می‌گردد. افزایش نسبی روش i در مقایسه با روش j برای FM به‌صورت رابطهٔ ۱۸ قابل اندازه‌گیری است.

$$\text{Relative increase of FM} = \frac{FM_i - FM_j}{FM_j} \qquad (۱۸)$$

و کاهش نسبی روش i در مقایسه با روش j برای VI برابر است با

$$\text{Relative reduction of VI} = \frac{VI_j - VI_i}{VI_j} \qquad (۱۹)$$

در تمامی مقایسه‌ها بر مبنای این دو پارامتر، نتایج روش جاری با الگوریتم Folding اوّلیه انجام شده است. نتایج حاصل از اعمال روش‌ها بر روی پنجاه تصویر اوّل برای تمامی کلمات پرس‌وجو و بر پایهٔ حقیقت پایه ایجاد شده می‌باشند، که این مقایسهٔ کمّی همانند روش ذکر شده در (*van Leuken, 2009; Wang, 2010*) می‌باشد.

۵-۳-۲-۱- بررسی تحلیلی تأثیر روش انتخاب مرکز دسته و وزن‌دهی ویژگی‌ها بر نتایج

در اینجا تأثیر روش انتخاب مرکز دسته و وزن‌دهی بر نتایج الگوریتم‌های مختلف به تفکیک بیان می‌گردد.

- **بررسی نتایج الگوریتم Folding**

در این بخش الگوریتم‌های مختلف Folding (F) با یکدیگر مقایسه می‌شوند. در (جدول ۲) مقایسهٔ الگوریتم F با نحوهٔ متفاوت مشخّص‌کردن مرکز دسته‌ها و اعمال وزن‌دهی آورده شده است. همان‌طور که ذکر شد، در F1 تصویر میانگین، تصویری در نظر گرفته می‌شود که کوچک‌ترین متوسط فاصله تا تمامی دیگر تصاویر دارد. F2 از روش دوم تعیین مرکز (RS2) بهره می‌برد یعنی تصویر میانگین، تصویری است که ویژگی‌های آن میانگین ویژگی‌های تمامی تصاویر است. F3 روش سوم را به کار می‌برد؛ یعنی روش مشخّص‌کردن مرکز دستهٔ کاهشی مبتنی بر تصویر میانگین. همان‌گونه که در این جدول مشخص است F3 از عملکرد بهتری برخوردار است که به دلیل اهمّیت بیشتر به تصاویر با اولویت بالاتر می‌باشد. در (جدول ۲) نتایج اعمال وزن‌دهی نیز نشان داده شده است که این وزن‌دهی باعث بهبود بیشتر F2 و F3 در هر دو معیار ارزیابی شده است، که میزان بهبود با معیارهای نسبی در مقایسه با الگوریتم F1 مشخّص است. نتایج بهترین روش در این جدول، Bold مشخّص شده‌اند. این نتایج به‌صورت نمودار در (اشکال ۹ الف و ب) ترسیم شده‌اند که میزان تأثیر روش انتخاب مراکز و وزن‌دهی را بر روی F نشان می‌دهد. همان‌گونه که در (جدول ۲) و (اشکال ۹ الف و ب) مشخّص است دو روش انتخاب مرکز RS2 و RS3 هم با اعمال وزن‌دهی و هم بدون اعمال آن، عملکرد بهتری داشته و نشان‌دهندهٔ کارا بودن این دو روش پیشنهادی در انتخاب مرکز می‌باشد.

(جدول ۲): نتایج الگوریتم Folding براساس روش انتخاب مرکز دسته و وزن‌دهی ویژگی

weight	Representative selection	Algorithm	FM	VI	Relative increase of FM	Relative reduction of VI
without weight	RS1	F1	0.3463	3.0184	-	-
	RS2	F2	0.3622	2.9471	4.604%	2.362%
	RS3	F4	0.3656	2.9385	5.571%	2.646%
with weight	RS1	W-F1	0.3460	3.0179	-	-
	RS2	W-F2	0.3632	2.9447	4.893%	2.444%
	RS3	W-F4	**0.3674**	**2.9299**	6.115%	2.933%

- **بررسی نتایج الگوریتم پیشنهادی Fuzzy Folding**

در این بخش نتایج سیستم با استفاده از الگوریتم پیشنهادی Fuzzy Folding (FF) بررسی می‌شوند. نتایج حاصل از این الگوریتم با روش‌های مختلف تعیین مرکز خوشه‌ها و همچنین اعمال وزن‌دهی بر روی آن‌ها در مقایسه با الگوریتم Folding اوّلیه در (جدول ۳) نشان داده شده است. همانگونه که مشخّص است الگوریتم‌های FF علاوه‌بر اینکه نسبت به الگوریتم Folding اولیه بهبود داشته‌اند، در هر دو معیار FM و VI نسبت به نسخه‌های غیرفازی خود، عملکرد بهتری داشته‌اند. این مطلب در مورد نسخه‌های وزن‌دار فازی در مقایسه با نسخه‌های وزن‌دار غیرفازی نیز صادق است. در این نتایج FF2 و FF3 به مانند نسخه‌های غیرفازی خود عملکرد خوبی از خود نشان داده‌اند. (اشکال ۹ ج و د) این نتایج را نمایش می‌دهد که نمایش‌گر برتری FF3 و WFF3 بر دیگر روش‌های فازی پیشنهادی در هر دو معیار FM و VI می‌باشد.

- **بررسی نتایج الگوریتم پیشنهادی Folding Tree**

نتایج حاصل از استفاده از الگوریتم Folding Tree (FT) برای خوشه‌بندی تصاویر در (جدول ۴) عنوان شده است. این نتایج با روش‌های مختلف تعیین مرکز خوشه‌ها و همچنین اعمال وزن‌دهی بر روی آن‌ها در مقایسه با الگوریتم F1 کسب شده‌اند. این نتایج به‌صورت نمودار میله‌ای نیز در (اشکال ۹ه و و) نشان داده شده است. واضح است که الگوریتم‌های FT2 و FT3 علاوه بر اینکه نسبت به الگوریتم Folding اوّلیه بهبود داشته‌اند، در هر دو معیار FM و VI نسبت به نسخه‌های اوّلیه و حتّی فازی خود،

عملکرد بهتری داشته‌اند. این مطلب در مورد نسخه‌های وزن‌دار این الگوریتم‌ها نیز صحیح می‌باشد.

- **مقایسهٔ نتایج الگوریتم Folding و الگوریتم‌های پیشنهادی**

در اینجا نتایج حاصل از الگوریتم‌های پیشنهادی با الگوریتم پایه Folding و الگوریتم‌های Maxmin (MM) و Reciprocal election (RE) مقایسه می‌گردند. نتایج مربوطه در (جدول ۵) آمده است. این نتایج بر روی پایگاه داده ذکر شده و با اعمال زیرسیستم پیش‌پردازش و استخراج ویژگی پیشنهادی حاصل شده‌اند. در کلیهٔ روش‌ها به‌جز RE که مستقل از روش تعیین مراکز است، مراکز با سه روش RS1، RS2، RS3 مشخّص شده و مقایسه شده‌اند. برای مقایسهٔ راحت‌تر، نتایج به‌صورت نمودار در (اشکال ۱۰الف و ب) آورده شده است. نتایج حاکی از آن است که الگوریتم W-FT3 بهترین نتایج را در هر دو معیار دارد. دومین الگوریتم از نظر نتایج، الگوریتم FT3 بر طبق معیار FM و الگوریتم WFF3 بر مبنای معیار VI می‌باشد. در بین روش‌های F، MM و RE، الگوریتم W-F3 بهترین نتیجه را دارد که این نتایج در جدول به‌صورت Bold مشخص شده‌اند.

۶- نتیجه‌گیری

در این پژوهش، چالش‌های فراروی کاربر برای جستجوی تصویر در موتورهای جستجوی تصاویر بیان شد و سعی شد رویکرد و سیستمی ارائه شود که نیاز کاربر را به صورت کاراتری برآورد.

در این پژوهش، چند الگوریتم برای خوشه‌بندی نتایج تصاویر موتورهای جستجو پیشنهاد شد، که تلخیص بصری و متنوع ایجاد شده توسط آن، به کاربر این امکان را می‌دهد که تصاویر را به راحتی مرور کرده و حدود و محتوای کلی تمامی تصاویر بازگشتی را در زمانی کوتاه و تعداد کمی کلیک بدست آورد. بر اساس ارزیابی‌های انجام شده، رویکرد پیشنهادی باعث بهبود در نتایج خوشه‌بندی تصاویر نتیجه شده از جستجو می‌شود.

(جدول ۳): نتایج الگوریتم Fuzzy Folding

براساس روش انتخاب مرکز دسته و وزن‌دهی ویژگی

weight	Representative selection	Algorithm	FM	VI	Relative increase of FM	Relative reduction of VI
without weight	RS1	FF1	0.3493	2.9899	0.866%	0.944%
	RS2	FF2	0.36515	2.9352	5.455%	2.759%
	RS3	FF3	0.3672	2.9340	6.059%	2.797%
with weight	RS1	W-FF1	0.3491	2.9903	0.809%	0.9310%
	RS2	W-FF2	0.36523	2.9337	5.480%	2.806%
	RS3	W-FF3	**0.3683**	**2.9250**	6.358%	3.096%

(جدول ۴): نتایج الگوریتم Folding Tree

براساس روش انتخاب مرکز دسته و وزن‌دهی ویژگی

weight	Representative selection	Algorithm	FM	VI	Relative increase of FM	Relative reduction of VI
without weight	RS1	FT1	0.356576	2.9903	2.967%	0.931%
	RS2	FT2	0.3702	2.9302	6.918%	2.922%
	RS3	FT4	0.3717	2.9335	7.351%	3.098%
with weight	RS1	W-FT1	0.356584	2.9903	2.970%	0.931%
	RS2	W-FT2	0.3710	2.9249	7.351%	2.815%
	RS3	W-FT4	**0.3730**	**2.9227**	7.734%	3.171%

(جدول ۵): مقایسهٔ الگوریتم‌های F، MM، RE

با الگوریتم‌های پیشنهادی

Algorithm	weight	Representative selection	FM	VI
Folding	without weight	RS1	0.3463	3.0184
		RS2	0.3622	2.9471
		RS3	0.3656	2.9385
	with weight	RS1	0.3460	3.0179
		RS2	0.3632	2.9447
		RS3	**0.3674**	**2.9299**
Fuzzy Folding	without weight	RS1	0.3493	2.9899
		RS2	0.36515	2.9352
		RS3	0.3672	2.9340
	with weight	RS1	0.3491	2.9903
		RS2	0.36523	2.9337
		RS3	0.3683	2.9250
Folding Tree	without weight	RS1	0.356576	2.9903
		RS2	0.3702	2.9302
		RS3	0.3717	2.9335
	with weight	RS1	0.356584	2.9903
		RS2	0.3710	2.9249
		RS3	**0.3730**	**2.9227**
Maxmin	without weight	RS1	0.3243	3.0640
		RS2	0.3334	3.0411
		RS3	0.3361	2.9999
	with weight	RS1	0.3272	3.0400
		RS2	0.3419	3.0181
		RS3	0.3345	3.0331
Reciprocal Election	without weight	-	0.3346	3.1112
	with weight	-	0.3348	3.1090

۷- مراجع

Ahmad, S., 1991. "VISIT: A neural model of covert attention". Advances in Neural Information Processing Systems, 4:420-427.

Alamdar, F., bahmani, Z., Haratizadeh, S., 2010. "Color Quantization with Clustering By F-PSO-GA". In Proceeding of IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS), 233-238.

Alamdar, F., Keyvanpour, M. R., 2011. "A New Color Feature Extraction Method Based on QuadHistogram". Procedia Environmental Science, 10(2011):777-783.

Baluja, S., Pomerleau, D.A., 1997. "Expectation-based selective attention dor visual monitoring and control of a robot vehicle". Robotics and Autonomous System, 22(3-4):329-344.

Boberg, J., Salakoski, T., 1993. "General formulation and evaluation of agglomerative clustering methods with metric and non-metric distances". Pattern Recognition, 26(9):1395-1406.

Cai, D., He, X., Li, Z., Ma, W. Y., Wen, J. R., 2004. "Hierarchical clustering of WWW image search results using visual, textual and link information". In Proceeding of ACM multimedia, 952–959.

Deselaers, T., Keysers, D., Ney, H., 2003. "Clustering Visually Similar Images to Improve Image Search

Manjunath, B. S., Ohm, J., Vasudevan V. V., Yamada, A., 2001. "Color and Texture Descriptors". IEEE Transactions on Circuits and Systems for Video Technology, 11(6):703-715.

Meilă, M., 2007. "Comparing clusterings: an information based distance". Journal of Multivariate Analysis, 98(5) 873-895.
Milanese, R., Gil S., Pun, T., 1995. "Attentive Mechanism for dynamic and static scene analysis". Optical Engineering, 34(8):2428-2434.

Milligan, G., Cooper, M., 1985. "An examination of procedures for determining the number of clusters in a data set". Psychometrika, 50(2):159-179.

Moëllic, P. A., Haugeard, J. E., Pitel, G., 2008. "Image clustering based on a shared nearest neighbors approach for tagged collections". In Proceeding of CIVR, 269-278.

Niebur, E., Koch, C., 1998. "Computational architectures for attention," R. Parasuraman (Ed.), The attentive brain, Cambridge, MA: MIT Press, 163-186.

Salvador, S., Chan, P., 2004. "Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms". In Proceeding of 16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 576–584.

Tamura, H., Mori, S., Yamawaki, T., 1978. "Textural features corresponding visual perception". IEEE Trans. Syst. Man Cybern., 8(6):460–473.

Tian, X., Tao, D., Hua, X. S., Wu, X., 2010. "Active Reranking for Web Image Search". IEEE Transaction on Image Processing, 19(3): 805-820.

van Leuken, R. H., Garcia, L., Olivares, X., 2009. "Visual Diversification of Image Search Results". In Proceedings of the 18th international conference on World wide web, 341-350.

van Zwol, R., Murdock, V., Garcia, L., Ramirez, G., 2008. "Diversifying image search with user generated content". In Proceedings of the International ACM Conference on Multimedia Information Retrieval, 67-74.

Wang, H., Missura, O., Gärtner T., Wrobel, S., 2009. "Context-Based Clustering of Image Search Results". Lecture Notes in Computer Science, 5803:153-160.

Wang, J., Jia, L., Hua, X. S., 2010. "Interactive browsing via diversified visual summarization for image search results". Multimedia Systems, 17(5): 379-391.

Wang, X., Wang, Y., Wang, L., 2004. "Improving fuzzy c-means clustering based on feature-weight learning". Pattern Recognition Letters, 25(10):1123-1132.

Zhang, B., Li, H., Liu, Y., Ji, L., Xi, W., Fan, W., Chen, Z., Ma, W. Y., 2005. "Improving web search results using affinity graph". In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, 504-511.

Engines". In Proceeding of Informatiktage 2003 der Gesellschaft für Informatik.

Fergus, R., Fei-Fei, L., Perona, P., Zisserman, A., 2005. "Learning object categories from Google's image search". In Proceeding of Tenth IEEE International Conference on Computer Vision (ICCV), 1816-1823,

Fowlkes E., Mallows, C., 1983. "A method for comparing two hierarchical clusterings". Journal of the American Statistical Association, 78(383)553-569.

Gan, G., Ma, Ch., Wu, J., 2007. "Data Clustering Theory, Algorithms, and Applications". ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

Gao, B., Liu, T. Y., Qin, T., Zheng, X., Cheng, Q., Ma, W. Y., 2005. "Web image clustering by consistent utilization of visual features and surrounding texts". In Proceeding of ACM multimedia, 112-121.

Halkidi, M., Baistakis, Y., Vazirgiannis, M., 2001. "On clustering validation techniques". Journal of Intelligent Information Systems, 17(2-3):107-145.

Itti, L., Koch, C., Niebur, E., 1998. "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis". IEEE Trans. on Pattern Analysis and Machine Intelligence, 20(11):1254-1259.

Jia, Y., Wang, J., Zhang, C., Hua, X. S., 2008. "Finding image exemplars using fast sparse affinity propagation". In Proceeding of ACM multimedia, 639-642.

Jing, F., Wang, C., Yao, Y., Deng, K., Zhang, L., Ma, W., 2006. "IGroup: A Web Image Search Engine with Semantic Clustering of Search Results". In Proceedings of the 14th annual ACM international conference on Multimedia, 497-498.

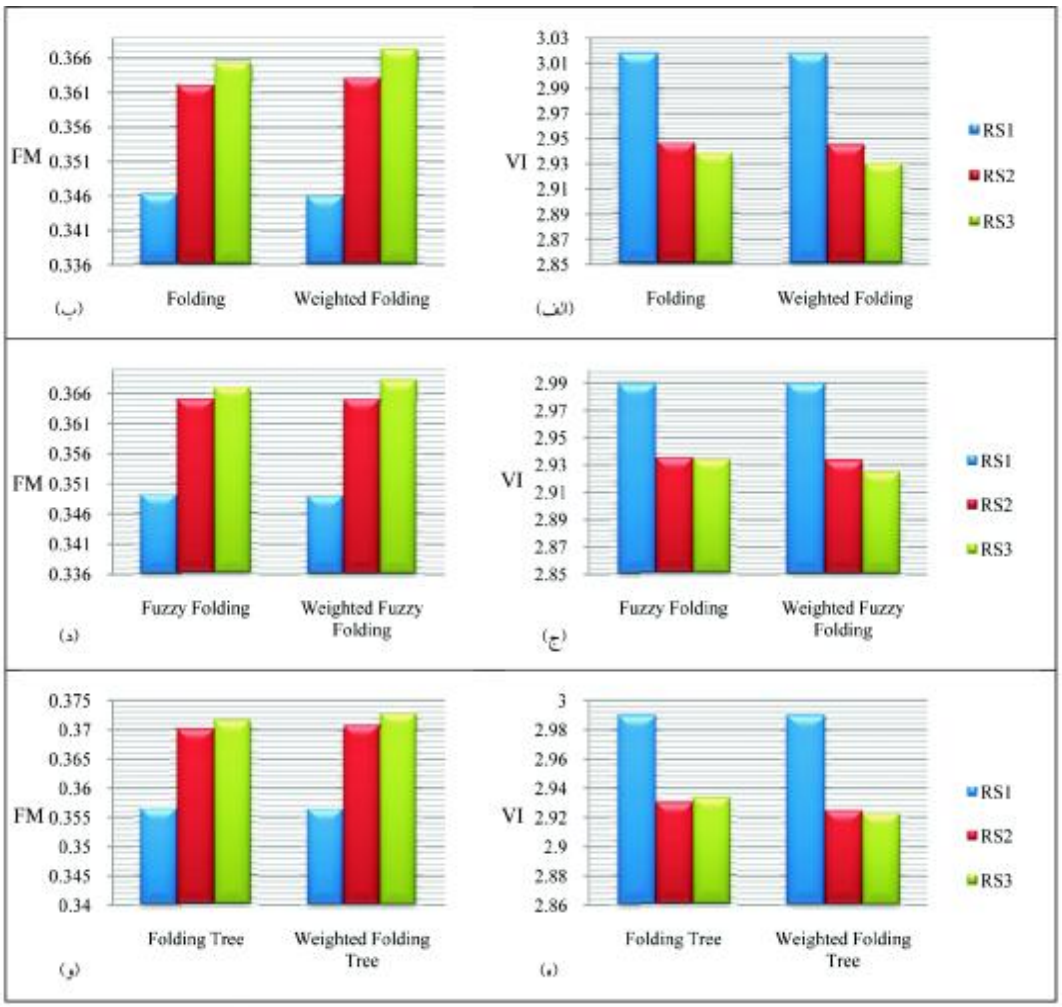
Jing, Y., Baluja, S., 2008. "VisualRank: Applying PageRamk to Large-Scale Image Search". IEEE Transaction on Pattern Analysis and Machine Intelligence, 30(11): 1877-1890.

Lamming, D., 1991. "Contrast Sensitivity". Chapter 5. In: Cronly-Dillon, J., Vision and Visual Dysfunction, Vol. 5, London: Macmillan Press.

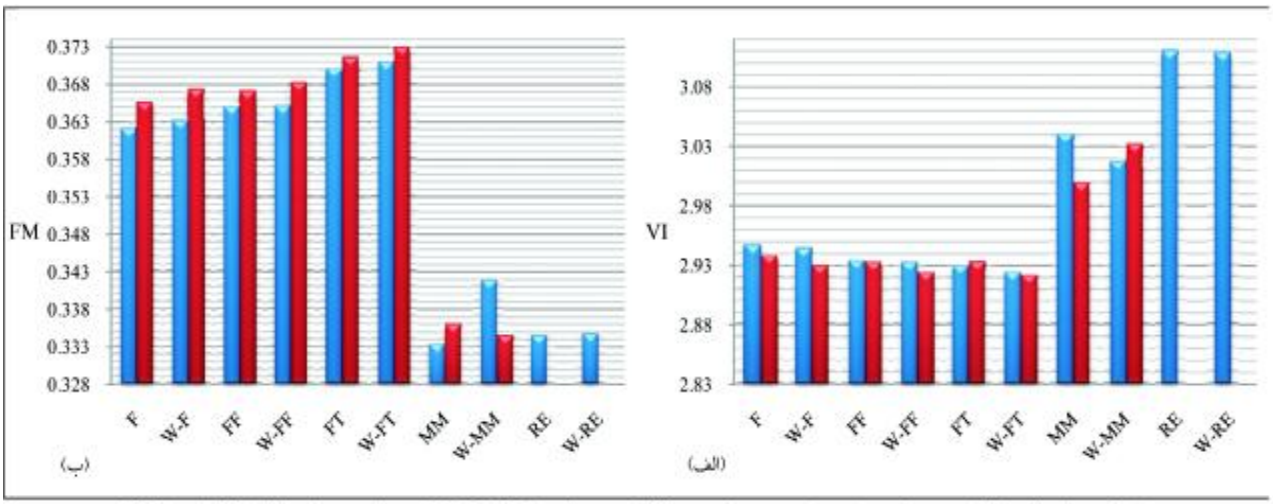
Li, Z., Xu, G., Li, M., Ma, W. Yi., Zhang, H. J., 2005. "Grouping WWW Image Search Results by Novel Inhomogeneous Clustering Method". In Proceeding of the 11th International Multimedia Modeling Conference, 255-261.

Liu, H., Xie, X., Tang, X., Li, Z., Ma, W. Y., 2004. "Effective browsing of web image search results". In Proceeding of Multimedia information retrieval, 84-90.

Ma, Y., Zhang, H., 2003. "Contrast-based image attention analysis by using fuzzy growing". In Proceedings of the eleventh ACM international conference on Multimedia, 374-381.



(شکل ۹): (الف، ب)نتایج الگوریتم‌های F و W-F برمبنای معیار VI و FM، (ج، د) نتایج الگوریتم‌های FF و W-FF برمبنای معیار VI و FM، (ه، و) نتایج الگوریتم‌های FT و W-FT برمبنای معیار VI و FM.



(شکل ۱۰): نتایج الگوریتم‌های پیشنهادی در مقایسه با الگوریتم‌های F، MM، RE بر طبق معیار (الف) VI، (ب) FM

فاطمه علمدار مدرک کارشناسی خود را در رشتهٔ مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۸۷ از دانشگاه شهید باهنر کرمان اخذ نمود. وی تحصیلات خود را در مقطع کارشناسی ارشد کامپیوتر گرایش هوش مصنوعی در دانشگاه الزهراء (س) در سال ۱۳۹۰ به پایان رسانده است. زمینه‌های پژوهشی مورد علاقهٔ او پردازش تصویر، بازشناسی الگو، الگوریتم‌های تکاملی و شبکهٔ عصبی است.

نشانی رایانامک ایشان عبارتست از:
fatemeh.alamdar@student.alzahra.ac.ir

محمدرضا کیوان‌پور مدرک کارشناسی خود را در سال ۱۳۷۶ در رشتهٔ مهندسی کامپیوتر گرایش نرم‌افزار از دانشگاه علم و صنعت ایران، و مدارک کارشناسی ارشد و دکترای خود را نیز به‌ترتیب در سال‌های ۱۳۷۹ و ۱۳۸۶ در رشتهٔ مهندسی کامپیوتر گرایش نرم‌افزار از دانشگاه تربیت مدرس اخذ نمود. وی از سال ۱۳۸۷ تاکنون عضو هیأت علمی دانشکدهٔ فنی و مهندسی دانشگاه الزهرا(س) است. زمینه‌های تحقیقاتی مورد علاقهٔ او عبارتند از: پایگاه داده، یادگیری ماشین، داده کاوی، پردازش ویدئو و تصویر.

نشانی رایانامک ایشان عبارتست از:
keyvanpour@alzahra.ac.ir