

# روشی برای بازخورد ربط براساس بهبود تابع

## شباهت در بازیابی تصویر بر اساس محتوا

اسما شمسی گوشکی،<sup>۱</sup> احسین نظام آبادی پور،<sup>۲</sup> سعید سریزدی و<sup>۳</sup> احسان الله کبیر  
<sup>۱</sup> و <sup>۲</sup> دانشکده مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران  
<sup>۳</sup> دانشکده مهندسی برق، دانشگاه تربیت مدرس، تهران، ایران

### چکیده

در سامانه‌های بازیابی تصویر بر اساس محتوا، ویژگی‌های دیداری تصاویر پایگاه استخراج شده و پایگاه ویژگی‌های دیداری برای یافتن نزدیک‌ترین تصاویر به تصویر پرس‌وجو، جستجو می‌شود. کاربر با سامانه تا رسیدن به هدف نهایی خود تعامل کرده و سامانه از اطلاعات کاربر و ابزار بازخورد ربط به منظور بالا بردن کارایی بازیابی استفاده می‌کند. در این تحقیق، هر تصویر با یک بردار ویژگی شامل سه نوع ویژگی و مجموع ۲۷۰ مؤلفه نمایه می‌شود. برای اندازه‌گیری شباهت بین دو تصویر از یک معیار وزن دار استفاده شده که در آن برای هر نوع و هر مؤلفه ویژگی یک وزن در نظر گرفته شده است. این وزن‌ها در طول بازیابی با استفاده از اطلاعات کاربر و روش پیشنهادی بهبود می‌یابند. در این روش، وزن هر نوع ویژگی با توجه به رتبه تصاویر مرتبط در بازیابی بر اساس آن نوع ویژگی تنظیم و برای تصحیح وزن هر مؤلفه ویژگی از میانگین و انحراف معیار آن مؤلفه روی تصاویر مرتبط و نامرتبط استفاده می‌شود. روش پیشنهادی روی یک پایگاه تصویر شامل ده هزار تصویر از ۸۲ گروه معنایی متفاوت آزموده شده است. نتایج آزمایش‌ها، برتری روش پیشنهادی به روش‌های موجود را نشان می‌دهد.

واژگان کلیدی: بازخورد ربط، بازیابی تصویر، بهبود بردار پرس‌وجو، بهبود تابع شباهت.

### ۱- مقدمه

در سال‌های اخیر رشد فناوری رایانه، اهمیت فراوان اطلاعات چندرسانه‌ای و وجود آرشیوهای بزرگ دیجیتال، تلاش بسیاری از محققان را برای ایجاد ابزارهای مناسب بازیابی تصاویر به خود جلب کرده است. از اوایل ۱۹۹۰، بازیابی تصویر بر اساس محتوا، یک زمینه فعال برای تحقیقات قلمداد می‌شود (Liu, et al., 2007). در سامانه‌های بازیابی تصویر، ابتدا ویژگی‌های سطح پایین<sup>۱</sup> تصاویر پایگاه به صورت خودکار استخراج شده و در پایگاه ویژگی‌های دیداری ذخیره می‌شوند؛ سپس بعد از دریافت تصویر پرس‌وجوی کاربر، ویژگی‌های سطح پایین آن استخراج شده و پایگاه ویژگی‌های دیداری برای یافتن نزدیک‌ترین تصاویر به تصویر پرس‌وجو، جستجو می‌شود.

در سامانه‌های بازیابی تصویر بر اساس محتوا، سامانه، پایگاه داده را بر اساس ویژگی‌های سطح پایین جستجو کرده و تصاویری را که ویژگی‌های سطح پایین نزدیک به تصویر پرس‌وجو دارند به کاربر ارائه می‌کند؛ در حالی که کاربر به دنبال جستجوی معنایی تصاویر است و تصاویر ارائه شده توسط سامانه، ممکن است دارای معنای دلخواه کاربر نباشند. بنابراین ویژگی‌های سطح پایین برای ارائه ویژگی‌های معنایی تصاویر کافی نیستند. از سوی دیگر، سلیقه افراد مختلف با یکدیگر متفاوت می‌باشد و ممکن است افراد مختلف معانی متفاوتی از یک تصویر برداشت کنند (نظام آبادی پور، ۱۳۸۳). برای رفع این محدودیت‌ها، لازم است عمل بازیابی با تعامل کاربر همراه باشد. در این تعامل سعی می‌شود پرس‌وجوی دلخواه کاربر مدل شود و با تنظیم ضریب ویژگی‌ها، کاربر به تصاویر دلخواهش برسد.

<sup>۱</sup> Low Level Features

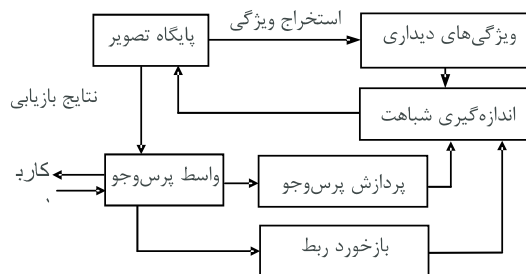
شده و در بخش سوم، روش پیشنهادی معرفی شده است. در بخش چهارم نتایج عملکرد سامانه بازیابی تصویر پیشنهادی و مقایسه با روش‌های استاندارد آورده شده است و در نهایت، بخش پنجم به جمع‌بندی روش پیشنهادی می‌پردازد.

## ۲- مروری بر روش‌های یادگیری کوتاه‌مدت

در یادگیری کوتاه‌مدت، سامانه 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 تصحیح بردار پرس‌وجو و بهبود تابع شباهت می‌پردازد؛ به‌گونه‌ای که کاربر را گام به گام به تصاویر دلخواهش نزدیک کند. الگوریتم‌های مبتنی بر بردار پرس‌وجو و تابع شباهت فقط از تصاویر بازخوردشده توسط کاربر استفاده می‌کنند؛ درحالی‌که روش‌های یادگیری ماشین روی تصاویر کل پایگاه اعمال شده و از نظر زمانی پرهزینه‌ترند. در روش‌های بهبود تابع شباهت، هدف، استفاده از تصاویر مرتبط و نامرتب به‌منظور تنظیم وزن مؤلفه‌های بردار ویژگی است. به این منظور در تحقیقات مختلف از اطلاعات متفاوتی مثل میانگین یا واریانس هر مؤلفه روی تصاویر مرتبط و نامرتب، یا فاصله هر مؤلفه از بردارهای ویژگی تصاویر بازخوردشده با بردار ویژگی تصویر پرس‌وجو در روابط مختلف استفاده شده است.

روش‌های یادگیری کوتاه‌مدت به دو دسته کلی الگوریتم‌های مبتنی بر یادگیری ماشین و الگوریتم‌های مبتنی بر بردار پرس‌وجو و تابع شباهت تقسیم می‌شوند. روش‌های مبتنی بر یادگیری ماشین<sup>۱</sup> به دو دسته با نظارت<sup>۲</sup> و بدون نظارت<sup>۳</sup> تقسیم می‌شوند. روش‌های یادگیری با نظارت، مثل ماشین بردار پشتیبان (SVM)<sup>۴</sup> (Wei, et al., 2009; Seo, 2007; Chen, et al., 2001) Fuzzy (Barrett, et al., 2009)، طبقه‌بند بیس<sup>۵</sup> (Wu & Lu, 2010)، شبکه‌های عصبی<sup>۶</sup> (Modagheh, et al., 2010) و درخت تصمیم<sup>۷</sup> (Qian, et al., 2003 MacArthur, et al., 2000)، به‌طور معمول برای یادگیری مفاهیم سطح بالا از

ساختار یک سامانه بازیابی تصویر در شکل (۱-۱) نشان داده شده است. در این سامانه، پایگاه تصویر شامل تصاویری است که کاربر برای یافتن تصاویر دلخواهش، آنها را جستجو می‌کند. در پایگاه ویژگی‌های دیداری یا سطح پایین، ویژگی‌های دیداری استخراج شده از تصاویر نگهداری می‌شوند (سریزدی و نظام آبادی پور، ۱۳۸۴؛ Nezamabadi-pour & Kabir, 2005; Pour & Saryazdi, 2005). 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. زیرسامانه پردازش پرس‌وجو، ویژگی‌های مناسب را از تصویر پرس‌وجو استخراج کرده و زیرسامانه اندازه‌گیری شباهت، شباهت بین بردار ویژگی تصویر پرس‌وجو و بردارهای ویژگی تصاویر پایگاه را محاسبه می‌کند؛ سپس سامانه نزدیک‌ترین تصاویر به تصویر پرس‌وجو را پیدا می‌کند. این تصاویر از طریق رابط گرافیکی به کاربر ارائه شده و با دریافت نظر کاربر راجع به آنها، تعامل با کاربر تا رسیدن به تصاویر دلخواه او ادامه می‌یابد. از این فرآیند که به بازخورد ربط معروف است، به‌منظور یادگیری کوتاه‌مدت استفاده می‌شود.



شکل ۱-۱: ساختار سامانه تعاملی بازیابی تصویر (Nezamabadi-pour & Kabir, 2004).

در این تحقیق، روشی جدید در بازخورد ربط پیشنهاد شده است که در آن تابع شباهت با استفاده از اطلاعات تصاویر مرتبط و نامرتب تصحیح می‌شود. در روش پیشنهادی، برای تصحیح وزن هر مؤلفه ویژگی از اطلاعات مربوط به میانگین و انحراف معیار آن مؤلفه روی تصاویر مرتبط و نامرتب استفاده می‌شود و وزن هر نوع ویژگی با توجه به رتبه تصاویر مرتبط در بازیابی بر اساس آن نوع ویژگی تنظیم می‌شود.

ادامه مقاله این‌گونه سازمان‌دهی شده است که در بخش دوم، مروری بر کارهای مرتبط با موضوع تحقیق انجام

<sup>1</sup> Machine Learning.

<sup>2</sup> Supervised Learning.

<sup>3</sup> Unsupervised Learning.

<sup>4</sup> Support Vector Mechine (Svm).

<sup>5</sup> Bayesian Classifier.

<sup>6</sup> Artificial Neural Network.

<sup>7</sup> Decision Tree.

الگوریتم‌های مبتنی بر بردار پرس‌وجو و تابع شباهت به دو دسته بهبود بردار پرس‌وجو و بهبود تابع شباهت تقسیم می‌شوند. هدف از بهبود بردار پرس‌وجو، حرکت بردار پرس‌وجو در فضای جستجو است؛ به نحوی که از تصاویر نامرتب دور و به تصاویر مرتبط نزدیک شود. این روش اولین بار در سال ۱۹۹۷ در سامانه MARS پیشنهاد شد (Porkaew & Chakrabarti, 1999). در این روش، میانگین ویژگی‌های همه تصاویر مرتبط حساب شده و به عنوان بردار پرس‌وجوی جدید در نظر گرفته می‌شوند.

در تحقیق (Rocchio, 1971) روش روچیو طبق رابطه ۱ معرفی شده است. در این روش از تصاویر مرتبط و نامرتب برای بهبود بردار پرس‌وجو استفاده شده است؛ به این صورت که در هر مرحله، بردار پرس‌وجو به تصاویر مرتبط نزدیک‌تر و از تصاویر نامرتب دور می‌شود.

$$F_q^{new} = F_q + \frac{\alpha}{|Q^+|} \sum_{F^+ \in Q^+} F^+ - \frac{1-\alpha}{|Q^-|} \sum_{F^- \in Q^-} F^- \quad (1)$$

که در آن  $F_q^{new}$  و  $F_q$  به ترتیب بردار پرس‌وجوی جدید و پرس‌وجو در مرحله قبل بوده و  $Q^+$  و  $Q^-$  به ترتیب مجموعه تصاویر مرتبط و نامرتب را نشان می‌دهند.  $F^+$  و  $F^-$  به ترتیب بیان‌گر بردار ویژگی تصاویر مرتبط و نامرتب انتخاب‌شده توسط کاربر بوده و  $\alpha$  یک ضریب قابل تنظیم است.

در بهبود تابع شباهت، وزن ویژگی‌های سطح پایین بردار پرس‌وجو در خلال بازخورد ربط به منظور رسیدن به هدف کاربر تغییر می‌کند. در تحقیق (Doulamis, et al., 1999)، طبق رابطه ۲ وزن هر یک از مؤلفه‌های ویژگی در تابع شباهت بر حسب فاصله بردار ویژگی تصویر پرس‌وجو با بردارهای ویژگی تصاویر مرتبط تخمین زده می‌شود.

$$h_i = 1 / \sum_{j=1}^m (F_{q,i} - F_{j,i}^+)^2 \quad (2)$$

که در آن  $h_i$  وزن مؤلفه  $i$ ام و  $F_{q,i}$  و  $F_{j,i}^+$  به ترتیب مؤلفه  $i$ ام از بردار ویژگی پرس‌وجو و تصویر مرتبط  $j$ ام بوده و  $m$  تعداد تصاویر مرتبط بازخورد شده را نشان می‌دهد. در مراجع (Zhuang, et al., 2001; Rui, et al., 1998) وزن هر یک از مؤلفه‌های ویژگی متناسب با عکس انحراف معیار آن مؤلفه روی مجموعه تصاویر مرتبط در نظر گرفته می‌شود. در مرجع (Schettini, et al., 1999)، برای

روی ویژگی‌های سطح پایین به کار می‌روند. در یادگیری کوتاه‌مدت از SVM به صورت متمایزکننده داده‌ها در یک و دو دسته استفاده شده است. روش SVM از یک ابر صفحه برای تفکیک داده‌ها استفاده می‌کند. در SVM با قابلیت جداسازی دو دسته، سامانه تصاویر پایگاه داده را به دو دسته مرتبط و نامرتب تقسیم می‌کند (Setia, et al., 2005). در SVM یک‌دسته‌ای، سامانه فقط تصاویر مرتبط پایگاه داده را جدا می‌کند (Chen, et al., 2001).

مقایسه عملکرد SVM با یک و دو دسته در فرآیند بازخورد ربط به منظور بازیابی تصویر، در تحقیق (Setia, et al., 2005) آمده است. نتایج نشان می‌دهند که در تعداد مراحل بازیابی‌های کم، SVM با یک دسته مناسب‌تر است در حالی که با بالا رفتن تعداد دفعات بازیابی، SVM با دو دسته بهتر عمل می‌کند.

در مرجع (Ferreira, et al., 2011) از GP<sup>۱</sup> در بازخورد ربط استفاده شده است. در تحقیق دیگری، بازخورد ربط به صورت یک مسأله طبقه‌بندی بیز مبتنی بر بازخورد کاربر در نظر گرفته شده است که تصاویر را به دو دسته مرتبط و نامرتب تقسیم می‌کند. نتایج نشان می‌دهد که این روش بهتر از سامانه MARS<sup>۲</sup> عمل می‌کند (Duan, et al., 2005).

روش‌های یادگیری بدون نظارت، بدون دخالت کاربر و به صورت خودکار، پایگاه تصاویر را خوشه‌بندی می‌کنند. هدف از خوشه‌بندی تصاویر<sup>۳</sup>، دسته‌بندی تصاویر است به طوری که شباهت درون خوشه‌ای زیاد و شباهت بین خوشه‌ای کم باشد. روش‌های خوشه‌بندی k-میانگین، MAP<sup>۴</sup>، ML<sup>۵</sup>، EM<sup>۶</sup>، D-EM<sup>۷</sup>، LVQ<sup>۸</sup>، TS-SOM<sup>۹</sup>، Ncut<sup>۱۰</sup>، LPC<sup>۱۱</sup>، OPF<sup>۱۲</sup> (da Silva, et al., 2010) و CLUE<sup>۱۳</sup> جزو روش‌های یادگیری بدون نظارت هستند. در تحقیق (Liu, et al., 2007) نشان داده شده است که روش خوشه‌بندی LPC از Ncut و این روش از روش k-میانگین، کارایی بهتری در بازیابی تصویر دارند.

<sup>1</sup> Genetic Programming

<sup>2</sup> Multimedia Analysis And Retrieval System (Mars).

<sup>3</sup> Image Clustering.

<sup>4</sup> Maximum-A-Posteriori (Map).

<sup>5</sup> Maximum-Likelihood (MI).

<sup>6</sup> Expectation Maximization (Em).

<sup>7</sup> Discriminant Em (D-Em).

<sup>8</sup> Learning Vector Quantization (Lvsq).

<sup>9</sup> Self-Organizing Map Tree-Structured (Ts-Som).

<sup>10</sup> Normalized Cut (Ncut).

<sup>11</sup> Locality Preserving Clustering (Lpc).

<sup>12</sup> Optimum-Path Forest (Opf).

<sup>13</sup> Cluster-Based Retrieval (Clue).

تنظیم ضرایب مربوط به هر مؤلفه ویژگی، از تصاویر مرتبط و نامرتب استفاده شده است. به این صورت که وزن هر مؤلفه ویژگی را با مقایسه واریانس آن مؤلفه روی تصاویر مرتبط و واریانس روی تصاویر مرتبط و نامرتب، تغییر می‌دهد. مرجع (Das & Ray, 2005) نیز از اطلاعات تصاویر نامرتب استفاده کرده است و وزن هر یک از مؤلفه‌های ویژگی را طبق رابطه ۳ می‌آورد.

$$h_i = \frac{1 + \sigma_-^i}{1 + \sigma_+^i} \quad (3)$$

که در این رابطه  $\sigma_+^i$  و  $\sigma_-^i$  به ترتیب بیان‌گر انحراف معیار مؤلفه  $i$ ام روی بردارهای ویژگی تصاویر مرتبط و تصاویر نامرتب با پرس‌وجو هستند.

در مراجع (Zhuang, et al., 2001; Rui, et al., 1998)، برای تعیین وزن مربوط به تأثیر هر یک از مؤلفه‌های ویژگی در تابع شباهت، از تصاویر مرتبط استفاده شده است. بازیابی تصویر با استفاده از وزن‌های از پیش تعیین شده شروع شده و بر این اساس تصویری به کاربر ارائه می‌شود. پس از اعلام نظر کاربر، برای تنظیم وزن هر یک از مؤلفه‌های ویژگی، بازیابی بر اساس تک‌تک مؤلفه‌های ویژگی انجام می‌شود. فصل مشترک تصویری که توسط کاربر به‌عنوان تصاویر مرتبط مشخص شده‌اند با تصاویر برتر بازیابی شده بر اساس هر یک از مؤلفه‌های ویژگی تعیین می‌شود. با این کار مؤلفه ویژگی مؤثرتر در بازیابی مشخص شده و وزن آن در بازیابی مرحله بعدی افزایش می‌یابد. در مرجع (Cheng, et al., 2008) نیز وزن هر یک از مؤلفه‌های ویژگی با استفاده از اطلاعات کسب‌شده از کاربر و تصاویر بازیابی‌شده بر اساس آن مؤلفه ویژگی تنظیم می‌شود.

در مرجع (Ciocca & Schettini, 2001) برای هر یک از مؤلفه‌های ویژگی، میانگین فواصل آن مؤلفه از تصویر پرس‌وجو را یک‌بار با همان مؤلفه از تصاویر مرتبط و بار دیگر با همان مؤلفه از تصاویر مرتبط و نامرتب محاسبه می‌کند. چنانچه یک مؤلفه از ویژگی در بازیابی موثر باشد، باید میانگین فواصل آن مؤلفه از تصاویر مرتبط کمتر از میانگین فواصل آن مؤلفه از مجموعه تصاویر مرتبط و نامرتب باشد. با در نظر گرفتن فرض مذکور، وزن مربوط به هر یک از مؤلفه‌های ویژگی افزایش یا کاهش می‌یابد. در مرجع (Newsam, et al., 2001)، میانگین فواصل بین تصاویر مرتبط بر اساس هر نوع ویژگی محاسبه می‌شود و ضریب

مربوط به تأثیر نوع ویژگی، متناسب با عکس این مقدار در نظر گرفته می‌شود.

در مرجع (Guldogan & Gabbouj, 2009) وزن هر مؤلفه ویژگی در قالب بازخورد ربط و طبق رابطه ۴، با استفاده از اطلاعات تصاویر مرتبط و نامرتب تصحیح می‌شود.

$$w_f = 1 - \frac{\alpha_f^- + \alpha_f^+}{2} \quad (4)$$

$$\alpha_f^+ = \frac{\Delta}{2 \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} d(F_i, F_j)} \quad (5)$$

$$\alpha_f^- = |\delta(x, y)| = \left| \frac{\sum_{i=1}^{N_x} (f_{xi} - \mu_x)(f_{yi} - \mu_y)}{N_x N_y \delta_x \delta_y} \right| \quad (6)$$

که در این روابط  $\Delta$  واریانس تصاویر مرتبط،  $N$  تعداد تصاویر مرتبط و  $d$  فاصله اقلیدسی بین تصاویر مرتبط را نشان می‌دهد.  $f_{xi}$  و  $f_{yi}$  آلمین عنصر و  $N_x$  و  $N_y$  تعداد عناصر خوشه‌های  $x$  و  $y$  را نشان می‌دهند.  $\mu_x$ ،  $\mu_y$ ،  $\delta_x$ ،  $\delta_y$  به ترتیب نشان دهنده میانگین و انحراف معیار خوشه‌های  $x$  و  $y$  می‌باشند. فرض می‌شود تعداد عناصر خوشه‌های  $x$  و  $y$  با هم برابر است.

نویسندگان این مقاله در مرجع (شمسی گوشکی و همکاران، ۱۳۸۸) روشی جدید در بازخورد ربط پیشنهاد داده‌اند که در آن وزن‌های مربوط به انواع ویژگی و مؤلفه‌های ویژگی با استفاده از اطلاعات تصاویر مرتبط و نامرتب تصحیح می‌شوند.

### ۳- روش پیشنهادی برای بهبود تابع شباهت

در این تحقیق، روشی به‌منظور بهبود تابع شباهت با استفاده از اطلاعات تصاویر مرتبط و نامرتب ارائه شده است. ابتدا لازم است اشاره شود که در سامانه پیشنهادشده، ویژگی‌های عام سطح پایین از تصاویر پایگاه استخراج می‌شوند. برای این منظور، فرض کنید تعداد  $N$  تصویر در پایگاه تصویر وجود دارد  $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$ . متناظر با تصویر  $X_i$  بردار ویژگی  $F_i$  وجود دارد که شامل ویژگی‌های سطح پایین مرتبط با آن است؛ بنابراین، پایگاه ویژگی شامل  $N$  بردار ویژگی به قرار  $F = \{F_1, F_2, \dots, F_N\}$  است. بردار ویژگی  $F_i$ ، ترکیبی از چند بردار ویژگی ناهمگون است. به‌عنوان مثال فرض کنید از یک تصویر،  $K$  نوع مختلف ویژگی (ویژگی نوع  $k$  شامل

ویژگی مذکور، متناسب با عکس مجموع رتبه تصاویر مرتبط در نظر گرفته می‌شود. برای توضیح بیشتر از زبان ریاضیات استفاده می‌کنیم. فرض کنید در یک مرحله از بازیابی، مجموعه تصاویر  $Q^+ = \{X_1^+, X_2^+, \dots, X_m^+\}$ ، به‌عنوان تصاویر مرتبط انتخاب شوند. به‌منظور تعیین وزن ویژگی نوع  $k$ ، برای هر یک از تصاویر پایگاه و تصویر پرس‌وجو، بردار ویژگی فقط شامل مؤلفه‌های ویژگی نوع  $k$  در نظر گرفته شده و تصاویر پایگاه بر اساس شباهت نسبت به تصویر پرس‌وجو مرتب می‌شوند. پس از آن، رتبه هر یک از تصاویر مجموعه  $Q^+$ ، در بازیابی مذکور مشخص شده و وزن ویژگی نوع  $k$ ، از رابطه ۸ و هنجارسازی شده آن از رابطه ۹ به‌دست می‌آید.

$$w^k = \frac{1}{\sum_{i=1}^m \text{rank}^k(X_i^+)} \quad (8)$$

$$w_n^k = \frac{w^k}{\sum_{k=1,2,\dots,K} w^k} \quad (9)$$

که در آن،  $\text{rank}^k(X_i^+)$  بیان‌گر رتبه تصاویر مرتبط  $X_i^+$  در فهرست بازیابی بر اساس ویژگی نوع  $k$  و  $w_n^k$  وزن هنجارسازی شده مربوط به آن ویژگی است.

### ۳-۲- روش پیشنهادی برای تصحیح وزن‌های مؤلفه‌های ویژگی ( $h$ )

همان‌گونه که ذکر شد، در شروع بازیابی، وزن تمام مؤلفه‌ها برابر یک در نظر گرفته می‌شود. در مراحل بعدی، وزن‌های مربوط به هر یک از مؤلفه‌های ویژگی، بر اساس روش پیشنهادی تنظیم می‌شوند. در این روش برای تنظیم وزن‌های مؤلفه‌های ویژگی که یک روش مبتنی بر قانون است از اطلاعات واریانس و میانگین هر مؤلفه روی تصاویر مرتبط و نامرتب استفاده شده است (شمسی گوشکی، ۱۳۸۹). برای ارائه قواعد، وزن‌های مربوط به مؤلفه‌های ویژگی مختلف کنار یکدیگر قرار گرفته و بردار وزن، برای کل ویژگی‌های تصویر تعریف می‌شود. به‌عبارتی دیگر داریم  $h = [h^1, h^2, \dots, h^K]$  یا به‌طریق دقیق

$$h = [h_1^1, h_2^1, \dots, h_{L_1}^1, h_1^2, h_2^2, \dots, h_{L_2}^2, \dots, h_1^K, h_2^K, \dots, h_{L_K}^K]$$

قواعد پیشنهادی در روابط ۱۰ تا ۱۲ ارائه شده‌اند. برای رسیدن به این روابط از این ایده استفاده شده است که وقتی واریانس یک مؤلفه روی تصاویر مرتبط کمتر از واریانس آن روی تصویر نامرتب باشد، آن مؤلفه در بازیابی مفید بوده و

$L_k$  مؤلفه ویژگی است و  $(k=1,2,\dots,K)$  استخراج می‌شود و ویژگی‌های استخراج شده در قالب یک بردار ویژگی به یکدیگر الحاق می‌شوند، یعنی این‌که  $F_i = [f_i^1, f_i^2, \dots, f_i^K]$  که در آن،  $f_i^k$  بیان‌گر ویژگی نوع  $k$  است و  $k=1,2,\dots,K$  (نظام آبادی پور و کبیر، ۱۳۸۳; Nezamabadi-Pour & Kabir, 2004). در روش‌های بهبود تابع شباهت، تابع شباهت در قالب رابطه ۷ ارائه می‌شود.

$$d(F_i, F_j) = w_n^1 \sum_l h_l^1 \left( \frac{f_{i,l}^1 - f_{j,l}^1}{f_{i,l}^1 + f_{j,l}^1} \right)^2 + \dots \quad (7)$$

$$+ w_n^K \sum_l h_l^K \left( \frac{f_{i,l}^K - f_{j,l}^K}{f_{i,l}^K + f_{j,l}^K} \right)^2$$

که در این رابطه  $F_i, F_j$  و  $d$  به‌ترتیب بیان‌گر بردارهای ویژگی دو تصویر  $i$  و  $j$  و معیار عدم شباهت بین این دو بردار ویژگی هستند.  $f_{i,l}^k$  مؤلفه  $l$  از ویژگی نوع  $k$  استخراج شده از تصویر  $i$  است و  $L_k$  طول بردار ویژگی نوع  $k$  را نشان می‌دهد.  $w_n^k$  بیان‌گر وزن هنجارسازی شده مربوط به ویژگی نوع  $k$  و  $h_l^k$  وزن مربوط به مؤلفه  $l$  از ویژگی نوع  $k$  است. در روش‌های بهبود تابع شباهت، این وزن‌ها با توجه به تصاویر مرتبط و نامرتب تنظیم می‌شوند.

در بعضی از تحقیقات، تنها مجموعه تصاویر مرتبط و در بعضی دیگر هم مجموعه تصاویر مرتبط و هم مجموعه تصاویر نامرتب برای بهبود تابع شباهت به‌کار می‌روند. در سامانه پیشنهادی، تابع شباهت با استفاده از اطلاعات تصاویر مرتبط و نامرتب تصحیح می‌شود. در اولین مرحله از بازیابی، بعد از آنکه ویژگی‌های تصویر پرس‌وجو استخراج شد، وزن‌های مربوط به تأثیر هر یک از انواع ویژگی ( $w$ ) و مؤلفه‌های ویژگی ( $h$ ) برابر یک در نظر گرفته می‌شوند. در این روش، برای تصحیح وزن‌های نوع ویژگی ( $w$ ) از اطلاعات مجموعه تصاویر مرتبط و برای تصحیح وزن‌های مؤلفه‌های ویژگی از هر دو مجموعه تصاویر مرتبط و نامرتب استفاده می‌شود.

### ۳-۱- روش پیشنهادی برای تصحیح وزن‌های نوع ویژگی ( $w$ )

این روش، در هر مرحله، ابتدا بازیابی با وزن‌های مرحله قبل انجام شده و تصاویر مرتبط توسط کاربر مشخص می‌شوند؛ سپس برای تنظیم وزن هر نوع ویژگی، بازیابی تصویر فقط بر اساس آن نوع ویژگی انجام شده و رتبه تصاویر مرتبط، در این بازیابی محاسبه می‌شوند؛ سپس وزن مربوط به نوع

#### ۴- آزمایش‌ها، نتایج و مقایسه

در این تحقیق، ساختار سامانه‌ی بازیابی تصویر ارائه‌شده در شکل (۱-۱) پیاده‌سازی شده است. در ادامه، اجزای سامانه‌ی پیاده‌سازی‌شده، ارائه شده و نتایج الگوریتم پیشنهادی با روش‌های شناخته‌شده در این زمینه، مقایسه می‌شود.

#### ۴-۱- پایگاه تصویر

در سامانه‌ی پیاده‌سازی‌شده از یک پایگاه تصویر عام استفاده شده‌است. برای تشکیل پایگاه از سه پایگاه تصویر مشهور که در بسیاری از تحقیقات به آنها ارجاع شده، استفاده شده است: الف) پایگاه داده بیست‌هزار تصویری Image CLEF photo 2007 (Grubinger, et al., 2008) ب) پایگاه داده MSRC، شامل ۴۳۲۰ تصویر با ۳۳ گروه معنایی (Winn, et al., 2005) و ج) پایگاه تصویر Simplicity، شامل هزار تصویر با ده گروه معنایی (Wang, et al., 2001). از مجموع ۲۵۳۲۰ تصویر این سه پایگاه تصویر، تعداد ده‌هزار تصویر انتخاب و پایگاه تصویر تشکیل شد. تمام تصاویر پایگاه، رنگی و در حوزه فشرده JPEG هستند. تصاویر انتخاب‌شده در ۸۲ گروه معنایی اصلی قرار دارند. تمام تصاویر متعلق به یک گروه معنایی به‌طورقطع در یک ویژگی معنایی با یکدیگر اشتراک دارند که به‌عنوان گروه معنایی اصلی برای آن تصویر در نظر گرفته شده است.

#### ۴-۲- ویژگی‌های پایه

از خصوصیات سامانه‌ی ارائه‌شده این است که چندان به نوع ویژگی استفاده‌شده در آن حساس نیست. از این رو سعی شده است ویژگی‌ها هر چه عام‌تر بوده و تعداد مؤلفه‌های آنها زیاد باشند. در این مطالعه از سه نوع ویژگی رنگ، بافت و شکل در استاندارد MPEG7 استفاده شده است. هیستوگرام رنگ هیستوگرام راستای لبه‌ها به‌ترتیب به‌عنوان ویژگی‌های رنگ و شکل انتخاب و برای استخراج ویژگی بافت از فیلترهای گابور استفاده شده است. ویژگی‌های استفاده‌شده در سامانه در ادامه توصیف شده‌اند. طول بردارهای ویژگی رنگ، لبه و بافت به‌ترتیب برابر ۶۰ و ۱۵۰ و ۶۰ است و در مجموع هر یک از تصاویر با برداری به اندازه ۲۷۰ بعد نمایه شده‌اند.

#### • هیستوگرام رنگ

برای نمایه‌سازی رنگ در تصاویر، از هیستوگرام رنگ در فضای HSV استفاده شده است. در این تحقیق،

وزن آن افزایش می‌یابد (رابطه ۱۱) و در غیر این صورت کاهش می‌یابد (رابطه ۱۰) که این مهم با چندضابطه‌ای کردن قوانین به دست آمده است؛ در عین حال، در هر دو حالت، هرچه میانگین یک مؤلفه روی تصاویر مرتبط بیشتر از میانگین آن روی تصاویر نامرتب باشد، مؤلفه در بازیابی موثرتر بوده و باید وزن آن افزایش یابد که برای رسیدن به این منظور از توابع نمایی بر حسب اختلاف میانگین مؤلفه مذکور روی تصاویر مرتبط و نامرتب استفاده شده است. در رابطه ۱۰، رابطه نمایی استفاده شده، باعث می‌شود تا در حالتی که واریانس یک مؤلفه روی تصاویر مرتبط بیشتر از واریانس آن روی تصویر نامرتب بوده و وزن ملزم به کاهش است، اگر میانگین آن مؤلفه روی تصاویر مرتبط بیشتر از میانگین آن روی تصاویر نامرتب باشد، وزن آن مؤلفه کمتر کاهش یابد. و در رابطه ۱۱ هنگامی که واریانس یک مؤلفه روی تصاویر مرتبط کمتر از واریانس آن روی تصویر نامرتب بوده و وزن ملزم به افزایش است، تابع نمایی موجب می‌شود که اگر میانگین آن مؤلفه روی تصاویر مرتبط بیشتر از میانگین آن روی تصاویر نامرتب باشد، وزن آن مؤلفه بیشتر افزایش یابد.

$$\text{If } (\sigma_+^l > \sigma_-^l) \text{ and } (\sigma_+^l \neq 0) \text{ Then} \quad (10)$$

$$h_l(\text{new}) = h_l(\text{old}) - e^{-(\mu_+ - \mu_-) - 1} \times \frac{\min(\sigma_-^l)}{\sigma_-^l} \times h_l(\text{old})$$

$$\text{If } (\sigma_+^l \leq \sigma_-^l) \text{ and } (\sigma_+^l \neq 0) \text{ Then} \quad (11)$$

$$h_l(\text{new}) = h_l(\text{old}) + e^{(\mu_+ - \mu_-) - 1} \times \frac{\min(\sigma_+^l)}{\sigma_+^l} \times h_l(\text{old})$$

$$\text{If } (\sigma_+^l = 0) \text{ Then } h_l(\text{new}) = 2 \times h_l(\text{old}) \quad (12)$$

که در آن  $l = 1, 2, \dots, L_F$  ،  $L_F = L_1 + L_2 + \dots + L_K$

$L_F$  طول بردار ویژگی هر تصویر است. واریانس مؤلفه  $l$  ام روی تصاویر مرتبط و نامرتب را به‌ترتیب با  $\sigma_+^l$  و  $\sigma_-^l$  نشان می‌دهیم و  $\mu_+^l$  و  $\mu_-^l$  به‌ترتیب بیان‌گر میانگین مؤلفه  $l$  ام روی تصاویر مرتبط و نامرتب می‌باشند.  $h_l(\text{new})$  وزن تصحیح‌شده و  $h_l(\text{old})$  وزن مؤلفه  $l$  ام در مرحله قبلی است.

بازیابی به توزیع لبه در کل تصویر نیاز است؛ زیرا توزیع محلی لبه‌ها برای بیان ویژگی‌های توزیع لبه در کل<sup>۵</sup> تصویر کافی نیستند. بنابراین، دو نوع توزیع لبه<sup>۶</sup> دیگر نیز برای تصویر معرفی شده است: الف) توزیع لبه<sup>۷</sup> سراسری (ب) توزیع لبه<sup>۸</sup> نیمه سراسری.

در استاندارد MPEG-7، پنج نوع لبه، شامل یک لبه بدون جهت و چهار لبه جهت‌دار در جهت‌های افقی، عمودی، ۴۵ درجه و ۱۳۵ درجه برای تشکیل هیستوگرام لبه تعریف شده‌اند. طول بردار ویژگی هیستوگرام لبه (ترکیب ویژگی‌های سراسری و نیمه‌سراسری) ۱۵۰ بعد است. برای مشاهده جزئیات مربوط به استخراج هیستوگرام لبه به مرجع (Park, et al., 2000) مراجعه کنید.

#### ۴-۳- معیار عدم شباهت و ترکیب ویژگی‌ها

نتایج آزمایش‌ها در مرجع (نظام آبادی پور و کبیر، ۱۳۸۴) نشان می‌دهد که معیار  $\chi^2$  (رابطه ۱۳) کارایی قابل قبولی را در بازیابی برای هر سه ویژگی هیستوگرام رنگ، لبه و ویژگی‌های گابور دارد.

$$d(F_i, F_j) = \sum_l \left( \frac{f_{i,l} - f_{j,l}}{f_{i,l} + f_{j,l}} \right)^2 \quad (13)$$

که در این رابطه  $F_i$ ،  $F_j$  و  $d$  به ترتیب بیان‌گر بردارهای ویژگی دو تصویر  $i$  و  $j$  و معیار عدم شباهت بین این دو بردار ویژگی هستند.  $f_{i,l}$  مؤلفه  $l$  از بردار ویژگی تصویر  $i$ ام است و  $L_F$  طول بردار ویژگی را نشان می‌دهد.

بنابراین، این معیار به‌عنوان معیار عدم شباهت در ترکیب ویژگی‌ها انتخاب می‌شود. اما از آنجا که اعمال معیار عدم شباهت  $\chi^2$  به ویژگی‌های مختلف، نتایجی در محدوده دینامیکی متفاوت تولید می‌کند، ترکیب ویژگی‌ها در یک تابع عدم شباهت، باید با هنجارسازی کردن فواصل ویژگی‌های متفاوت همراه باشد. در روش هنجارسازی کردن خطی، باید فاصله تمام تصاویر درون پایگاه بر اساس یک ویژگی خاص از یکدیگر محاسبه شود. سپس، مقدار بیشینه و کمینه این فواصل تعیین شود. در نهایت فاصله هنجارسازی شده با کاهش فاصله دو تصویر از مقدار کمینه فواصل و تقسیم بر تفاضل مقادیر بیشینه و کمینه فواصل در محدوده ۰ و ۱ هنجارسازی می‌شود. بنابراین، این روش تحت تأثیر مقادیر کمینه و بیشینه فواصل تصاویر در پایگاه

هیستوگرام رنگ پیشنهادی با شصت بازه از تصاویر استخراج و تمام پایگاه تصویر بر این مینا نمایه‌سازی شده است. برای محاسبه هیستوگرام رنگ پیشنهادی، هر یک از بازه‌های H، S و V به ترتیب به ۶، ۳ و ۳ بازه به‌طور خطی و یکنواخت چندی شده‌اند<sup>۱</sup> (در کل، تعداد ۵۴ بازه به سطوح رنگی و شش بازه به سطوح خاکستری اختصاص یافته است). 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 به‌عنوان سطوح خاکستری در نظر گرفته شده و به شش سطح حول محور V به‌طور خطی و یکنواخت چندی شده است. در نهایت هیستوگرام رنگ با شصت طبقه محاسبه می‌شود.

#### • ویژگی‌های گابور

ویژگی‌های گابور به‌علت خاصیت تفکیک‌پذیری خوب و چندانکه در حوزه‌های مکان و فرکانس به‌عنوان ابزاری قوی در تحلیل بافت استفاده می‌شوند. در این تحقیق در مجموع از سی فیلتر گابور در پنج مقیاس<sup>۲</sup> و شش جهت<sup>۳</sup> استفاده شده است. فرکانس مرکزی حداکثر،  $2\sqrt{128}$  سیکل بر پهنا<sup>۴</sup> تصویر و فرکانس مرکزی حداقل،  $2\sqrt{16}$  سیکل بر پهنا<sup>۴</sup> تصویر انتخاب شده‌اند. بنابراین، مقدار ضریب مقیاس، دو خواهد بود. این فیلترها در ابعاد  $256 \times 256$  پیکسل طراحی شده‌اند. برای نمایه‌سازی تصویر، در مرحله پیش‌پردازش، از تصویر رنگی، تصویر خاکستری ساخته شده و این تصویر به ابعاد  $256 \times 256$  نرمالیزه می‌شود. سپس زیر-تصویرهای فیلتر شده، محاسبه می‌شوند. از هر زیرتصویر، دو مشخصه میانگین و انحراف معیار به‌عنوان ویژگی استخراج می‌شود. این ویژگی بر مبنای استاندارد MPEG-7 انتخاب شده است (Manjunath, et al., 2002). بردار نمایه ویژگی‌های گابور شصت بعدی است.

#### • هیستوگرام راستای لبه

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<sup>۴</sup> با هشتاد کلاس در MPEG-7 به‌عنوان توصیف‌گر هیستوگرام لبه استاندارد شده است. برای بالابردن کارایی

<sup>1</sup> Quantize

<sup>2</sup> Scale

<sup>3</sup> Orientation

<sup>4</sup> Local Edge

<sup>5</sup> Global Edge

$$w_n^k = \frac{w^k}{\sum_{k \in \{c, e, g\}} w^k} \quad (17)$$

بنابراین، تابع شباهت نهایی طبق رابطه ۱۸ توصیف می‌شود. این تابع در اولین مرحله از بازبایی از وزن‌های رابطه ۱۷ و در سایر مراحل از وزن‌های رابطه ۹ استفاده می‌کند.

$$d(F_i, F_j) = w_n^c \sum_l^{Lc} h_l^c \left( \frac{f_{i,l}^c - f_{j,l}^c}{f_{i,l}^c + f_{j,l}^c} \right)^2 + \quad (18)$$

$$w_n^e \sum_l^{Le} h_l^e \left( \frac{f_{i,l}^e - f_{j,l}^e}{f_{i,l}^e + f_{j,l}^e} \right)^2 +$$

$$w_n^g \sum_l^{Lg} h_l^g \left( \frac{f_{i,l}^g - f_{j,l}^g}{f_{i,l}^g + f_{j,l}^g} \right)^2$$

که در این رابطه  $Lc$ ،  $Le$  و  $Lg$  به ترتیب بیان‌گر طول بردارهای هیستوگرام رنگ، هیستوگرام راستای لبه و ویژگی‌های گابور هستند.

#### ۴-۴- نتایج و مقایسه با سایر روش‌ها

در یادگیری کوتاه‌مدت، سه رهیافت (دو رهیافت پایه و یک رهیافت پیشنهادی) پیاده‌سازی شده است. رهیافت پایه اول، در تصحیح بردار پرس‌وجو از روش روچو (Rocchio, 1971) و در تصحیح تابع شباهت از روش Rui (Rui, et al., 1998) استفاده می‌کند که برای سادگی آن را "RUI STL" می‌نامیم و رهیافت پایه دوم، در تصحیح بردار پرس‌وجو از روش روچو (Rocchio, 1971) و در تصحیح تابع شباهت از روش پیشنهادی در مرجع (Guldogan & Gabbouj, 2009) استفاده می‌کند که برای سادگی آن را "GUL STL" می‌نامیم. رهیافت پیشنهادی، تصحیح بردار پرس‌وجو را با استفاده از روش روچو (Rocchio, 1971) انجام داده و برای تصحیح تابع شباهت از روش پیشنهادی استفاده می‌کند که آن را به‌اختصار "The Proposed STL" می‌نامیم. درحقیقت در مقایسه انجام‌شده در این بخش از آنجا که روش تصحیح بردار پرس‌وجو در سه رهیافت یکسان است، روش‌های مذکور در تصحیح تابع شباهت مقایسه می‌شوند.

از دیگر ابتکارات پیشنهادشده برای تصحیح روش یادگیری کوتاه‌مدت، حذف تصاویری است که توسط کاربر در تکرارهای قبلی به‌عنوان تصویر نامرتب اعلام شده است.

قرار گرفته و ممکن است منجر به نتایج رضایت‌بخش نشود. از سوی دیگر در هنجارسازی کردن گوسی، فاصله هنجارسازی‌شده با کاهش فاصله دو تصویر از مقدار میانگین کل فواصل در پایگاه و تقسیم بر ضریبی از انحراف معیار فواصل (به‌طور معمول ضریب ۳) در محدوده ۱- و ۱ هنجارسازی می‌شود.

در این تحقیق از پایگاه داده نسبتاً بزرگی استفاده می‌شود. بنابراین به دست آوردن فاصله هر تصویر از تصاویر دیگر به‌منظور هنجارسازی کردن آنها، نیازمند زمان و حافظه بالایی است. به همین دلیل برای هنجارسازی کردن از یک روش ابداعی استفاده شده است. در این روش به جای هنجارسازی کردن فواصل و سپس جمع وزن‌دار آن‌ها با وزن‌های مساوی، از جمع وزن‌دار فواصل با وزن‌های غیرمساوی استفاده می‌شود. درعمل، ضرایب هنجارسازی کردن به ضرایب ترکیب فواصل منتقل شده و به این صورت هنجارسازی کردن انجام می‌شود. به عبارتی، ما استفاده از رابطه ۱۴ را برای ترکیب فواصل پیشنهاد می‌کنیم.

$$d(F_i, F_j) = \sum_{k \in \{e, c, g\}} w_n^k d^k(f_i^k, f_j^k) \quad (14)$$

که در آن  $c$ ،  $e$  و  $g$  به ترتیب بیان‌گر ویژگی‌های هیستوگرام رنگ، هیستوگرام راستای لبه و گابور هستند.  $f_i^k$  بیان‌گر ویژگی نوع  $k$  از تصویر  $i$ ام و  $d^k$ ، معیار شباهت تصاویر بر اساس ویژگی نوع  $k$  است. همان‌گونه که ملاحظه می‌شود، فواصل  $d^k$  بدون تغییر باقی مانده و هنجارسازی نشده است و فقط از یک ضریب هنجارسازی کردن  $w_n^k$  استفاده می‌شود.

برای محاسبه  $w_n^k$ ، ابتدا میانگین کل فواصل برای هر ویژگی با معیار شباهت مورد نظر محاسبه می‌شود (رابطه ۱۵).

$$\mu^k = \frac{2}{N(N-1)} \left( \sum_{i=1}^N \sum_{j=1}^N d^k(f_i^k, f_j^k) \right) \quad k \in \{c, e, g\} \quad (15)$$

سپس، وزن‌های ترکیب‌کننده اولیه در تابع شباهت استفاده از رابطه ۱۴ و هنجارسازی‌شده آنها از رابطه ۱۷ محاسبه می‌شوند.

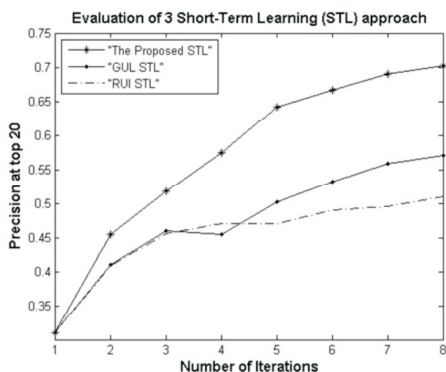
$$w^k = \frac{1}{\mu^k} \quad (16)$$



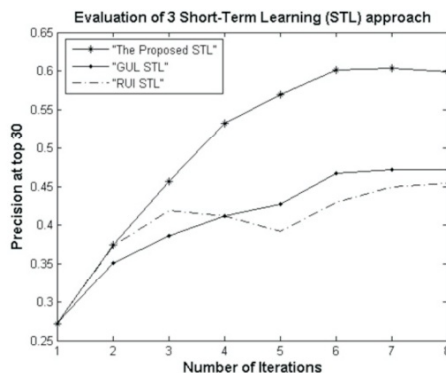
دقت روش پیشنهادی در تکرار هشتم بیش از روش‌های مورد مقایسه است. بیان مذکور در نتایج دیداری نشان داده شده در شکل‌های (۴-۴) تا (۹-۴) نیز قابل مشاهده است. این موضوع نشان‌دهنده آن است که رهیافت پیشنهادی بهتر توانسته است ارتباط بین ویژگی‌های دیداری و معانی سطح بالا را برقرار کند.

(جدول ۴-۱): دقت روش‌های پیاده‌سازی شده در مرحله هشتم با ۲۰، ۳۰ و ۴۰ تصویر بازیابی شده در هر مرحله

تعداد تصاویر بازیابی شده در هر مرحله	۲۰	۳۰	۴۰
دقت روش The Proposed STL	٪۷۰	٪۶۰	٪۵۴
دقت روش GUL STL	٪۵۷	٪۴۷	٪۴۲
دقت روش RUI STL	٪۵۱	٪۴۵	٪۴۵
افزایش دقت روش پیشنهادی نسبت به روش GUL STL	٪۱۳	٪۱۳	٪۱۲
افزایش دقت روش پیشنهادی نسبت به روش RUI STL	٪۱۹	٪۱۵	٪۹



(شکل ۴-۱): گراف دقت بر حسب تعداد تکرار برای سه رهیافت پیاده‌سازی شده با بیست تصویر بازیابی شده در هر مرحله

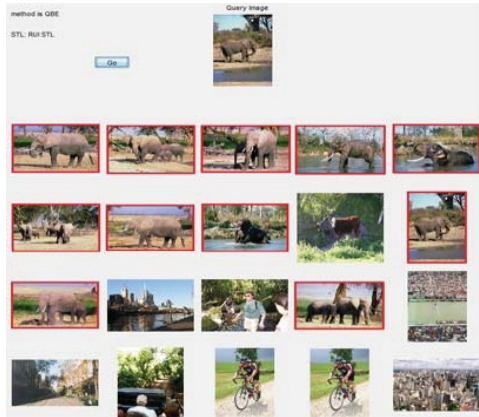


(شکل ۴-۲): گراف دقت بر حسب تعداد تکرار برای سه رهیافت پیاده‌سازی شده با سی تصویر بازیابی شده در هر مرحله

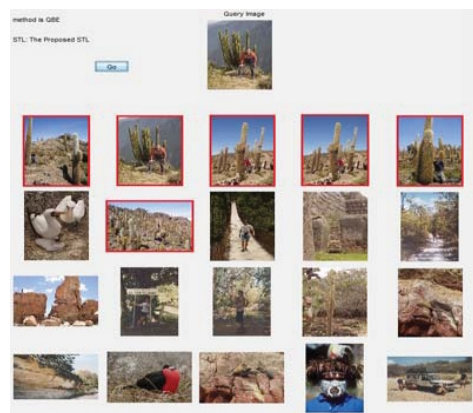
ممکن است بعضی از تصاویر نامرتبلی که پیش از این توسط کاربر در نشست فعلی معرفی شده‌اند، در تکرارهای بعدی الگوریتم در رتبه‌های بالایی از نظر شباهت قرار گرفته و دوباره در فهرست کاربر قرار گیرند. برای رفع این مشکل یک حافظه موقت در هر نشست بازیابی در نظر گرفته می‌شود و تصاویری که توسط کاربر به‌عنوان تصاویر نامرتبلی اعلام می‌شوند، در این حافظه نگه‌داری شده و در ادامه نشست به کاربر ارائه نخواهند شد؛ هرچند که در فهرست بازیابی قرار داشته باشند. لازم به ذکر است که این ایده در هر سه رهیافت یادگیری کوتاه‌مدت پیاده‌سازی شده مورد استفاده قرار گرفته است.

برای مقایسه روش‌های یادگیری کوتاه‌مدت، ۲۵ تصویر پرس‌وجو از ۲۵ گروه معنایی مختلف انتخاب و به سامانه ارائه شد. از کاربر درخواست شد که در هر نشست، بازیابی را تا هشت مرحله جلو برده و با سامانه تعامل کند. در هر مرحله، سامانه تصاویر بازیابی شده را به کاربر نشان داده و کاربر نظر خود را مبنی بر مرتبط‌بودن یا نامرتبلی‌بودن تصاویر اعمال می‌کند. بر مبنای تعداد تصاویر بازیابی شده مرتبط، برای هر تصویر، گراف دقت بر حسب تعداد تکرارها به دست می‌آید. این عمل برای تمام ۲۵ تصویر پرس‌وجوی آزمون انجام شده و میانگین دقت بر حسب تکرار محاسبه و رسم می‌شود. لازم به ذکر است که تصاویر پرس‌وجو برای مقایسه هر سه رهیافت یادگیری کوتاه‌مدت یکسان هستند. گراف دقت بر حسب فراخوانی برای تعداد متفاوتی از تصاویر بازیابی شده در هر مرحله، رسم شده است. شکل‌های ۲، ۳ و ۴ گراف‌های دقت بر حسب تعداد تکرار را برای این سه رهیافت به ترتیب با ۲۰، ۳۰ و ۴۰ تصویر بازیابی شده در هر مرحله به تصویر کشیده‌اند. همچنین، نتایج دیداری برای بازیابی بر اساس پرس‌وجویی که روش پیشنهادی در آن بالاترین دقت را دارد (گروه معنایی فیل) برای هر سه رهیافت یادگیری کوتاه‌مدت، در شکل‌های ۵ تا ۷ و برای پرس‌وجویی که روش پیشنهادی در آن کمترین دقت را دارد (گروه معنایی درخت کاکتوس) برای هر سه رهیافت، در شکل‌های ۸ تا ۱۰ آورده شده است. این نتایج در تکرار هشتم و برای حالت بیست تصویر بازیابی شده در هر مرحله ارائه شده‌اند. همین‌طور در جدول (۴-۱) نیز، دقت روش‌های پیاده‌سازی شده در تکرار هشتم نشان داده شده است.

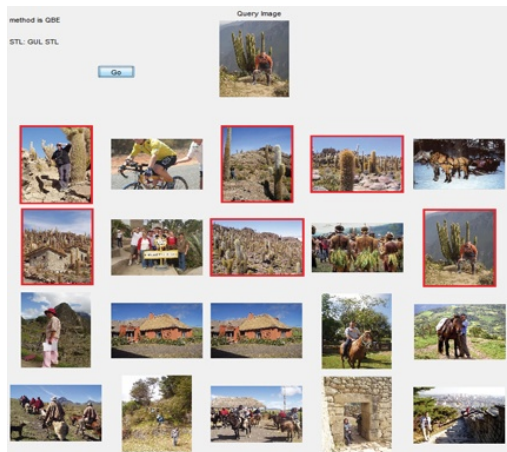
همان‌گونه که ملاحظه می‌شود، در هر گراف با افزایش تعداد مراحل بازیابی، دقت سامانه در بازیابی تصاویر مرتبط بیشتر می‌شود. همچنین با توجه به جدول (۴-۱)،



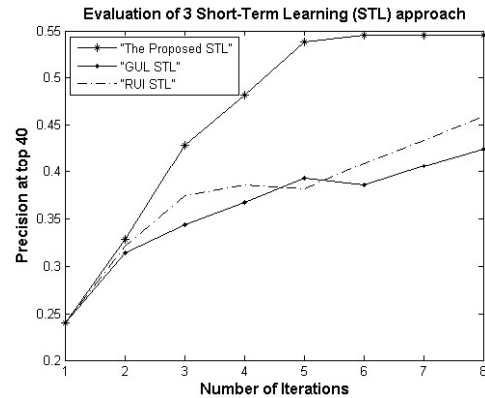
(شکل ۴-۶): تصاویر بازیابی شده با روش "RUI STL" برای پرس و جو با گروه معنایی "فیل"، یازده تصویر مرتبط بازیابی شده است



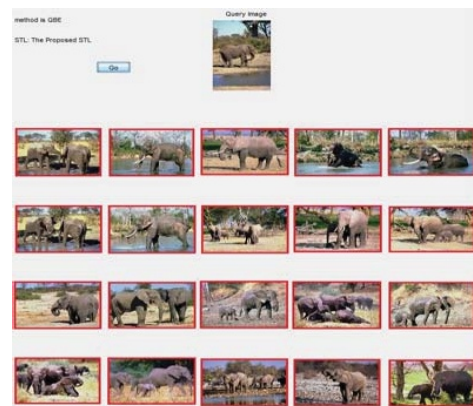
(شکل ۴-۷): تصاویر بازیابی شده با روش "The Proposed STL" برای پرس و جو با گروه معنایی "درخت کاکتوس"، شش تصویر مرتبط بازیابی شده است



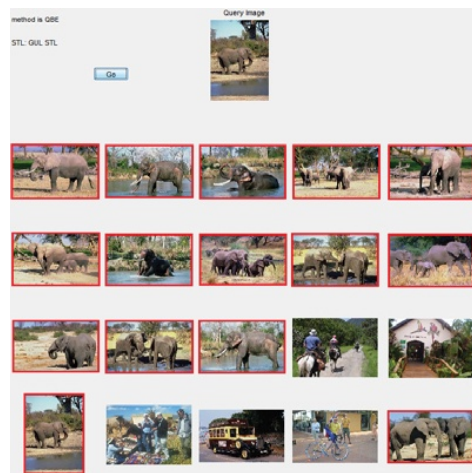
(شکل ۴-۸): تصاویر بازیابی شده با روش "GUL STL" برای پرس و جو با گروه معنایی "درخت کاکتوس"، شش تصویر مرتبط بازیابی شده است



(شکل ۴-۳): گراف دقت بر حسب تعداد تکرار برای سه رهیافت پیاده سازی شده با چهل تصویر بازیابی شده در هر مرحله



(شکل ۴-۴): تصاویر بازیابی شده با روش "The Proposed STL" برای پرس و جو با گروه معنایی "فیل". بیست تصویر مرتبط بازیابی شده است



(شکل ۴-۵): تصاویر بازیابی شده با روش "GUL STL" برای پرس و جو با گروه معنایی "فیل". پانزده تصویر مرتبط بازیابی شده است

## مراجع

سرزیدی، س و نظام آبادی پور، ح، ۱۳۸۴. "روش جدیدی برای نمایه سازی تصاویر رنگی مبتنی بر فشرده سازی و مورفولوژی ریاضی". نشریه علمی پژوهشی انجمن کامپیوتر ایران، علوم و مهندسی کامپیوتر، مجلد ۳، شماره ۳ (الف)، صفحات ۸۱-۹۴.

شمسی گوشکی، ا، ۱۳۸۹. "بازیابی تصاویر رنگی بر پایه محتوا با استفاده از روش های چند پرسشی و بازخورد ربط". پایان نامه کارشناسی ارشد، دانشگاه شهید باهنر کرمان، بخش مهندسی برق.

شمسی گوشکی، ا، نظام آبادی پور، ح، سرزیدی، س و کبیر، ا، "روشنی جدید در بازخورد ربط برای بازیابی تصویر بر اساس محتوا"، پانزدهمین کنفرانس سالانه انجمن کامپیوتر ایران، ۱۳۸۸.

نظام آبادی پور، ح، ۱۳۸۳. "کاربرد ویژگی های وابسته به نوع تصویر در یک سامانه بازیابی تصویر رنگی". پایان نامه دکتری، دانشگاه تربیت مدرس، بخش مهندسی برق.

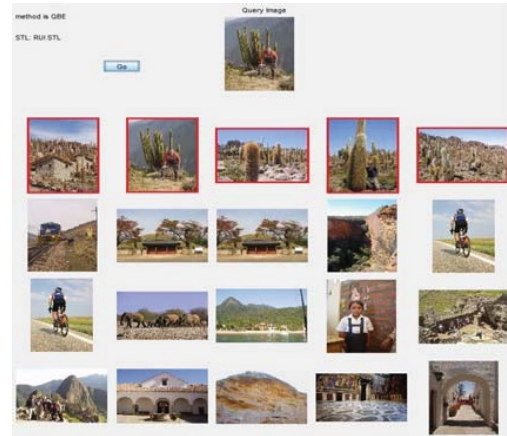
نظام آبادی پور، ح و کبیر، ا، ۱۳۸۴. "ارزیابی معیارهای عدم شباهت در بازیابی و طبقه بندی تصویر". مجله علمی پژوهشی فنی مدرس، شماره ۲۲، صفحات ۸۹-۹۸.

نظام آبادی پور، ح و کبیر، ا، ۱۳۸۳. "ترکیب ویژگی های سطح پایین برای طبقه بندی معنایی تصاویر". نشریه علمی پژوهشی انجمن کامپیوتر ایران، علوم و مهندسی کامپیوتر، ۱۳۸۳، مجلد ۲، شماره ۱ و ۳ (الف)، صفحات ۳۷-۴۶.

Barrett, S., Chang, R. & Qi, X., 2009. A fuzzy combined learning approach to content-based image retrieval. In IEEE International Conference on Multimedia and Expo. IEEE, pp. 838-841.

Chen, Y., Zhou, X.S. & Huang, T.S., 2001. One-class SVM for learning in image retrieval. In International Conference on Image Processing, 2001. IEEE, pp. 34-37.

Cheng, P.-C. et al., 2008. A two-level relevance feedback mechanism for image retrieval. Expert Systems with Applications, 34(3), pp.2193-2200.



(شکل ۴-۹): تصاویر بازیابی شده با روش "RUI STL" برای پرس و جو با گروه معنایی "درخت کاکتوس"، پنج تصویر مرتبط بازیابی شده است

## ۵- جمع بندی

در سامانه های بازیابی تصویر بر اساس محتوا از ابزار بازخورد ربط مثل بهبود بردار پرس و جو و بهبود تابع شباهت به منظور بالابردن کارایی سامانه استفاده می شود. در بهبود بردار پرس و جو، به طور معمول از تصاویر مرتبط و نامرتب برای تصحیح بردار پرس و جو استفاده می شود و در بهبود تابع شباهت، وزن های مؤلفه های ویژگی در خلال تعامل با کاربر تصحیح می شوند. در این تحقیق، روشی جدید در بازخورد ربط پیشنهاد شده که در آن تابع شباهت با استفاده از اطلاعات تصاویر مرتبط و نامرتب تصحیح شده است. در روش پیشنهادی برای تصحیح وزن های نوع ویژگی ها از اطلاعات مجموعه تصاویر مرتبط و برای تصحیح وزن های مؤلفه های ویژگی از هر دو مجموعه تصاویر مرتبط و نامرتب استفاده شده است. نتایج نشان می دهند که دقت روش پیشنهادی بیش از روش های مورد مقایسه است. این موضوع نشان دهنده آن است که رهیافت پیشنهادی بهتر توانسته است ارتباط بین ویژگی های دیداری و معانی سطح بالا را برقرار کند.

## سپاس گذاری

این تحقیق با حمایت مالی مرکز تحقیقات مخابرات ایران در چارچوب پروژه "ارائه روش های نوین بازیابی تصاویر رنگی به شیوه مالتی مودال" انجام شده است.

- Nezamabadi-pour, H. & Kabir, E., 2003. Image retrieval using block-based color histogram and local distribution of intensity gradient. In 2nd Int. Symposium on Telecommunication. Isfahan, Iran, pp. 176–180.
- Nezamabadi-pour, H. & Kabir, E., 2004. Image retrieval using histograms of uni-color and bi-color blocks and directional changes in intensity gradient. *Pattern Recognition Letters*, 25(14), pp.1547–1557.
- Nezamabadi-Pour, H. & Saryazdi, S., 2005. Object-based image indexing and retrieval in DCT domain using clustering techniques. In *Proceedings of World Academy of Science Engineering and Technology*. pp. 207–210.
- Park, D.K., Jeon, Y.S. & Won, C.S., 2000. Efficient use of local edge histogram descriptor. In *workshops on Multimedia Proceedings of the ACM*. ACM, pp. 51–54.
- Porkaew, K. & Chakrabarti, K., 1999. Query refinement for multimedia similarity retrieval in MARS. In *seventh ACM international conference on Multimedia (Part 1)*. ACM, pp. 235–238.
- Qian, F. et al., 2003. Alternating feature spaces in relevance feedback. *Multimedia Tools and Applications*, 21(1), pp.35–54.
- Rocchio, J.J., 1971. Relevance feedback in information retrieval. *SMART Retrieval System Experiments in Automatic Document Processing*, pp.313 – 323.
- Rui, Y. et al., 1998. Relevance feedback: a power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(5), pp.644–655.
- Schettini, R., Ciocca, G. & Gagliardi, I., 1999. Content-based color image retrieval with relevance feedback. In *International Conference on Image Processing*. Kobe: IEEE, pp. 75–79.
- Seo, K.-K., 2007. An application of one-class support vector machines in content-based image retrieval. *Expert Systems with Applications*, 33(2), pp.491–498.
- Setia, L., Ick, J. & Burkhardt, H., 2005. SVM-based Relevance Feedback in Image Retrieval using Invariant Feature Histograms. In *Workshop on Machine Vision Applications*. Japan, pp. 542–545.
- Da Silva, A.T., Xavier, A. & Magalhães, L.P., 2010. A new CBIR approach based on relevance feedback and optimum-path forest classification. *WSCG*, 18(1-3), pp.73–80.
- Wang, J.Z., Li, J. & Wiederhold, G., 2001. SIMPLicity: Semantics-sensitive integrated matching for
- Ciocca, G. & Schettini, R., 2001. Content-based similarity retrieval of trademarks using relevance feedback. *Pattern Recognition*, 34(8), pp.1639–1655.
- Das, G. & Ray, S., 2005. A compact feature representation and image indexing in content-based image retrieval. In *Image and Vision Computing*. New Zealand: Citeseer, pp. 387 – 391.
- Doulamis, A.D. et al., 1999. Interactive content-based retrieval in video databases using fuzzy classification and relevance feedback. In *IEEE International Conference on Multimedia Computing and Systems*. IEEE, pp. 954–958.
- Duan, L. et al., 2005. Adaptive relevance feedback based on Bayesian inference for image retrieval. *Signal Processing*, 85(2), pp.395–399.
- Ferreira, C.D. et al., 2011. Relevance feedback based on genetic programming for image retrieval. *Pattern Recognition Letters*, 32(1), pp.27–37.
- Grubinger, M. et al., 2008. Overview of the ImageCLEFphoto 2007 photographic retrieval task. In *Advances in Multilingual and Multimodal Information Retrieval*. Springer, pp. 433–444.
- Guldogan, E. & Gabbouj, M., 2009. Dynamic feature weights with relevance feedback in content-based image retrieval. In *24th International Symposium on Computer and Information Sciences*. IEEE, pp. 56–59.
- Liu, Y. et al., 2007. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1), pp.262–282.
- MacArthur, S.D., Brodley, C.E. & Shyu, C.-R., 2000. Relevance feedback decision trees in content-based image retrieval. In *IEEE Workshop on Content-based Access of Image and Video Libraries*. IEEE, pp. 68–72.
- Manjunath, B.S., Salembier, P. & Sikora, T., 2002. *Introduction to MPEG-7: multimedia content description interface*, John Wiley & Sons.
- Modagheh, H. et al., 2010. Learning of Relevance Feedback Using a Novel Kernel Based Neural Network. *Australian Journal of Basic and Applied Sciences*, 4(2), pp.171–186.
- Newsam, S., Sumengen, B. & Manjunath, B.S., 2001. Category-based image retrieval. In *International Conference on Image Processing*. IEEE, pp. 596–599.
- Nezamabadi-Pour, H. & Kabir, E., 2004. Combining low level features for semantic image classification. *The CSI Journal on Computer Science and Engineering*, 2(1-3 (a)), pp.37 – 46.

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. زمینه‌های پژوهشی مورد علاقه ایشان الگوریتم‌های ابتکاری، پردازش تصویر، بازشناسی الگو و رایانش نرم است.

نشانی رایانامه ایشان عبارت است از :

nezam@mail.uk.ac.ir



**سعید سریزدی** کارشناسی مهندسی الکترونیک و کارشناسی ارشد مهندسی مخابرات گرایش سیستم خود را به‌ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۶ از دانشگاه صنعتی اصفهان و مدرک D.E.A

در زمینه پردازش سیگنال و دکتری در زمینه پردازش تصویر از دانشگاه رن فرانسه را به‌ترتیب در سال‌های ۱۳۷۳ و ۱۳۷۶ دریافت کرد. وی هم‌اکنون عضو هیأت 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.

نشانی رایانامه ایشان عبارت است از :

saryzadi@mail.uk.ac.ir



**احسان‌اله کبیر** کارشناسی ارشد پیوسته خود را در رشته مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستم‌های الکترونیک از دانشگاه اسکس در انگلستان،

به‌ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۹ دریافت کرد. او هم‌اکنون در بخش مهندسی برق دانشگاه تربیت مدرس مشغول تحقیق و تدریس است. زمینه‌های پژوهشی مورد علاقه او بازشناسی الگو به‌ویژه بازشناسی متون چاپی و دست‌نویس و همچنین بینایی ماشین است.

نشانی رایانامه ایشان عبارت است از :

kabir@modares.ac.ir

picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence., 23(9), pp.947-963.

Wei, L., Yang, Y. & Nishikawa, R.M., 2009. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. Pattern recognition, 42(6), pp.1126-1132.

Winn, J., Criminisi, A. & Minka, T., 2005. Object categorization by learned universal visual dictionary. In Tenth IEEE International Conference on Computer Vision. IEEE, pp. 1800-1807.

Wu, J. & Lu, M., 2010. Asymmetric bayesian learning for image retrieval with relevance feedback. In 16th international conference on Advances in Multimedia Modeling. Verlag Berlin: Springer, pp. 650-655.

Zhuang, Y., Li, Q. & Lau, R.W.H., 2001. Web-based image retrieval: a hybrid approach. In Computer Graphics International. Hong Kong: IEEE, pp. 62-69.



**اسما شمسی** گوشکی کارشناسی خود را در رشته مهندسی پزشکی در سال ۱۳۸۶ از دانشگاه شاهد تهران اخذ کرد. سپس موفق به دریافت کارشناسی ارشد در رشته مهندسی برق-مخابرات از دانشگاه

شهید باهنر کرمان در سال ۱۳۸۹ شد و هم‌اکنون در مقطع دکترای مهندسی برق دانشگاه صنعتی اصفهان مشغول به تحصیل است. زمینه‌های تحقیقاتی مورد علاقه ایشان، پردازش تصویر، بازشناسی الگو و رایانش نرم است. نشانی رایانامه ایشان عبارت است از :

asma.shamsi@ice.iut.ac.ir



**حسین نظام آبادی‌پور** کارشناسی و کارشناسی ارشد خود را در رشته مهندسی برق-الکترونیک به‌ترتیب از دانشگاه شهید باهنر کرمان و دانشگاه تربیت مدرس در سال‌های ۱۳۷۷

و ۱۳۷۹ دریافت کرد. او سپس دوره دکترای خود را در