



## بکارگیری یک مدل گراف دولایه به منظور بازیابی تصاویر با استفاده از محتوی و مفاهیم سطح بالای آنها

دکتر محسن سریانی  
soryani@iust.ac.ir

دکتر محمود فتحی  
mahfathy@iust.ac.ir

مهدی بدیع فرخند  
mahdi\_badie@comp.iust.ac.ir

### چکیده

الگوریتم قبلی و نیز تغییر در نحوه استخراج و تفسیر اطلاعات درون لایه‌ها، کارایی مدل گراف دولایه به طور قابل ملاحظه‌ای بهبود یافته است.

**واژه‌های کلیدی:** پردازش تصویر، بازیابی تصویر، گراف، بازخورد مرتبط

### 1. مقدمه

با افزایش سریع استفاده از اطلاعات چندرسانه‌ای مانند متن، تصویر، صدا و ویدئو تقاضا برای ذخیره، جستجو و بازیابی این اطلاعات به شدت افزایش یافته است. در این بین بازیابی تصاویر مبتنی بر محتوی با توجه به کاربرد وسیع آن از دیرباز مورد توجه پژوهشگران بوده است [1-2]. یک سیستم بازیابی تصویر مبتنی بر محتوی از ویژگیهای سطح پایین مانند رنگ، شکل و بافت جهت شاخص‌گذاری تصاویر استفاده می‌کند. اگر چه بازیابی تصویر مبتنی بر محتوی در برخی از کاربردها نتایج خوبی داشته است اما مشکلات و هزینه‌هایی مثل قطعه بندی تصویر، استخراج ویژگیهای تصویری مناسب و منطبق با ادراک کاربر و ... روبرو

در این مقاله یک سیستم بازیابی تصویر مبتنی بر یک مدل گراف دولایه پیاده‌سازی و با انجام اصلاحات بهبود یافته است. این مدل این قابلیت را دارد که علاوه بر اطلاعات بازخوردی در طول یک جلسه بازیابی<sup>1</sup>، از اطلاعات بازخوردی در سایر پرس‌وجوهای دیگر نیز استفاده کند. به همین علت علاوه بر کارایی بسیار خوب در بازیابی کوتاه‌مدت، در بازیابی بلندمدت نیز کارایی خوبی دارد. لایه اول گراف، همبستگی بین تصاویر را از نقطه‌نظر شباهتهای بصری نشان می‌دهد که به کمک یک سیستم بازیابی تصویر مبتنی بر محتوی<sup>2</sup> (CBIR) ایجاد می‌شود. لایه دوم نماینده شباهتهای معنایی بین تصاویر است که بوسیله یک الگوریتم یادگیری مفاهیم و روابط معنایی بین تصاویر را، از طریق بازخوردهای کاربران استخراج می‌کند. همچنین بر مبنای این مدل گراف دولایه، روش تحلیل اتصالات درون گراف برپایه ایده انتشار شباهت از طریق اتصالات بصری و معنایی، به منظور بازیابی تصاویر به کار گرفته می‌شود. در این مقاله با ارائه یک الگوریتم جدید یادگیری بر پایه

<sup>1</sup> - Retrieval session

<sup>2</sup> - Content-based image retrieval

است. همچنین وجود فاصله معنایی<sup>1</sup> بین مفاهیم سطح بالا و ویژگیهای سطح پایین و نیز مسأله ذهنی بودن ادراک انسانها، کارایی روشهای مبتنی بر محتوی را کاهش داده است [3]. از این رو با توجه به نتیجه خیلی از کاربردها، تحقیقات در زمینه بازیابی تصویر به سمت بازیابی معنایی تصویر و استفاده از بازخورد مرتبط<sup>2</sup> تغییر کرده است. روشهای زیادی تاکنون با استفاده از بازخورد مرتبط پیشنهاد شده اند [4-7]. در [5] یک متد احتمالی یادگیری ویژگیهای مرتبط<sup>3</sup> (PFRL) معرفی شد که به طور خودکار با استفاده از بازخوردهای کاربر ویژگیهای مرتبط را تشخیص می‌داد. در [7] یک شیوه بازخورد مرتبط جدید بر پایه SVM یک کلاسه معرفی شده است که مفاهیم سطح بالای کاربر را به دست آورده و از آن در پرس‌وجوهای آینده استفاده می‌کند. در [8] یک روش یادگیری معنایی جدید برای بازیابی تصویر مبتنی بر محتوی با عنوان AHP<sup>4</sup> (فرآیند سلسله‌مراتبی تحلیلی) ارائه شده است و طبق یک مکانیزم سلسله‌مراتبی، یک بردار معنایی تشکیل می‌دهد که دارای مقادیر برانزنگی معانی مختلف تصویر داده شده می‌باشد. اغلب روشهای بازخوردی، به غیر از تعداد معدودی [9-11]، مکانیزمی برای یادگیری و ذخیره سازی اطلاعات بازخوردی برای استفاده مجدد از آنها در پرس‌وجوهای بعدی ندارند. اگر یک جلسه بازیابی به صورت ارائه یک پرس‌وجو توسط کاربر و مجموعه تمام بازخوردهای کاربر بعد از همان پرس‌وجو، تعریف شود، در این صورت اکثر روشها فقط قادرند نتیجه بازیابی را در خلال یک جلسه بازیابی بهبود بخشند (کارایی کوتاه‌مدت) بدون اینکه این بهبود و نتایج حاصل از آن در جلسات بعدی بازیابی مورد استفاده قرار گیرد (کارایی بلندمدت). آن دسته از مکانیزمهای یادگیری هم، که وجود دارد، اغلب نیازمند مدل‌های محاسباتی پیچیده [9 و 11] و یا استفاده از یک سری کلمات کلیدی [10] می‌باشند. مدل گراف دولایه [12]، با بهره‌گیری از ویژگیهای محتوایی تصویری در لایه اول و مفاهیم سطح بالای آنها

در لایه دوم سعی دارد تا حدودی مشکلات ذکر شده روشهای بازخوردی را برطرف کند. سازمان مقاله بدین‌گونه است که در بخش دوم، مدل گراف دولایه تشریح می‌شود. بخش سوم به فرآیند بازیابی تصویر با استفاده از تحلیل اتصالات و انتشار شباهت می‌پردازد. بهینه‌سازیهای مدل گراف دولایه در بخش چهارم معرفی می‌شود. در بخش پنجم نتایج آزمایشات و نهایتاً در بخش ششم نتیجه‌گیری ارائه خواهد شد.

## 2. مدل گراف دولایه [12]

این مدل شامل دو لایه می‌باشد. یک لایه بصری<sup>5</sup> و یک لایه معنایی<sup>6</sup>. در هر لایه یک گراف بدون جهت وجود دارد و هر گره در این دو گراف معرف یک تصویر در پایگاه داده تصویری می‌باشد و اتصالات بین گره‌ها وجود نوعی شباهت و همبستگی بین تصاویر را نشان می‌دهد و وزن هر یال شدت این شباهت را نشان می‌دهد (شکل 1). گره‌های هر گراف در هر لایه دقیقاً متناظر با گره‌های گراف دیگر می‌باشند. یعنی به ازاء هر گره در لایه بصری یک گره متناظر با آن در لایه معنایی وجود دارد. اما اتصالات بین گره‌ها در دو لایه متفاوت است. یک اتصال در لایه معنایی (اتصال معنایی) نمایانگر میزان شباهت دو تصویر از نظر معنایی می‌باشد، در حالیکه یک اتصال در لایه بصری (اتصال بصری)، شباهت بین دو تصویر را از نقطه نظر ویژگیهای سطح پائین تصویری نشان می‌دهد. بازیابی تصویر بر اساس این گراف دو لایه صورت می‌گیرد و نحوه توزیع اتصالات و وزن آنها کارایی بازیابی را تعیین می‌کند.

<sup>1</sup> - Semantic gap

<sup>2</sup> - Relevance feedback

<sup>3</sup> - Probabilistic feature relevance feedback

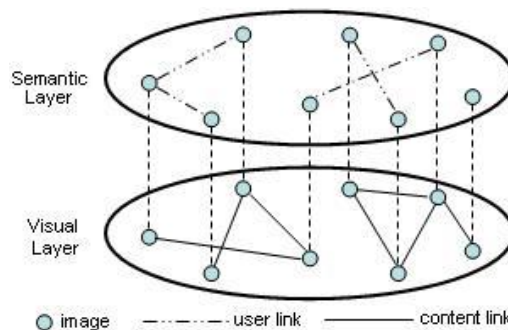
<sup>4</sup> - Analytic hierarchical process

<sup>5</sup> - Visual Layer

<sup>6</sup> - Semantic layer

وجود ندارد، آنگاه یک اتصال معنایی بین گره‌های متناظر با آنها در لایه‌های گراف ایجاد شود و مقدار اولیه آن برابر یک قرار داده شود. اگر هم اتصالی بین آن دو وجود دارد، وزن آن را یک واحد افزایش می‌دهیم.

(3) برای هر  $I_i \in I_N$  اگر اتصال معنایی بین  $I_i$  و  $I_s$  وجود دارد، مقدار وزن آن تقسیم بر 4 می‌شود و اگر حاصل این تقسیم از یک کوچکتر شود، آن اتصال حذف می‌شود.



شکل 1 - مدل گراف دو لایه [12]

اتصالات معنایی برای پردازش پرس‌وجوهایی که بعداً به سیستم وارد می‌شوند به کار می‌آید. به این ترتیب با این الگوریتم یادگیری روشی برای جمع‌آوری، ذخیره‌سازی و استفاده از اطلاعات بازخوردی کاربران، ایجاد می‌شود که می‌تواند در بلندمدت و برای همه پرس‌وجوها مورد استفاده قرار گیرد.

### 3. فرآیند بازیابی تصویر با استفاده از تحلیل اتصالات و انتشار شباهت<sup>1</sup>

اگر تصویر پرس‌وجو از تصاویر پایگاه داده نباشد قبل از اینکه فرایند بازیابی شروع شود، ابتدا تصویر جدید به پایگاه داده اضافه شده و یک گره نیز در هر کدام از لایه‌ها به گراف اضافه می‌شود و تمامی اتصالات بصری بین تصویر جدید و هر کدام از تصاویر پایگاه داده محاسبه و شکل می‌گیرد. بنابراین پرس‌وجو می‌تواند از یک گره (گره متناظر با تصویر پرس‌وجو) در گراف دو لایه آغاز می‌شود. روش بازیابی بر پایه ایده انتشار شباهت است. شباهت (با تصویر پرس و جو) می‌تواند از یک گره به تمامی گره‌های مرتبط با آن از طریق اتصالات بصری و معنایی انتشار یابد.

برای پردازش پرس‌وجو ابتدا یک مقدار اولیه شباهت در گره نمایانگر تصویر پرس‌وجو فرض می‌شود و سپس اجازه داده می‌شود که این شباهت از طریق اتصالات بصری و معنایی جریان

اتصالات بصری قبل از اینکه سیستم مورد استفاده قرار گیرد، محاسبه و شکل می‌گیرد. وزن این اتصالات برابر با مقدار شباهت دو تصویر تنظیم می‌شود. طراحی لایه بصری در مدل گراف دو لایه دارای انعطاف زیادی است چرا که هر نوع سیستم CBIR که بتواند ویژگی‌های سطح پایین تصاویر را استخراج و مقایسه کند، می‌تواند در این لایه به کار گرفته شود. اما از آنجا که کاربران انسانی، قضاوت مشابه بودن تصاویر را بیشتر از نقطه نظر ذهنی و معنایی انجام می‌دهند، در فرآیند بازیابی تصویر اتصالات معنایی اهمیت بیشتری دارند. برای کسب مفاهیم تصویر که کار پیچیده‌ای می‌باشد، سیستم به طور مستقیم با کاربر در تعامل است و با استفاده از یک الگوریتم یادگیری و اخذ اطلاعات وی، اتصالات معنایی را تشکیل می‌دهد. در [12] الگوریتمی معرفی شد که ایده آن، ساده و قابل درک است. هنگامی که کاربر یک تصویر نمونه را به عنوان تصویر پرس‌وجو ارائه می‌دهد و تصاویر بازیابی شده را به دو گروه مرتبط و نامرتب تقسیم می‌کند، آنگاه اتصالات معنایی بین تصویر پرس‌وجو و هر یک از تصاویر مرتبط ایجاد و یا تقویت می‌شود و اگر اتصالی بین تصویر پرس‌وجو و هر یک از تصاویر نامرتب وجود داشته باشد، تضعیف و یا اینکه حذف می‌شود. این الگوریتم در زیر آمده است:

(1) تصویر پرس و جو را  $I_S$ ، مجموعه تصاویر مرتبط را  $I_R$  و مجموعه تصاویر نامرتب را  $I_N$  می‌نامیم.

(2) برای هر  $I_i \in I_R$ ، اگر اتصالی معنایی بین  $I_i$  و  $I_S$

<sup>1</sup> -Similarity propagation

یابد و مقدار انتشار یافته آن برابر با حاصلضرب وزن اتصالات در مقدار شباهت می‌باشد. نحوه توزیع شباهت بین گره‌ها متناسب با میزان ارتباط تصاویر پایگاه داده با تصویر پرس‌وجو است. فرض کنید هر یک از لایه های بصری و معنایی به ترتیب توسط ماتریسهای مجاورت،  $M_c$ ،  $M_u$  نشان داده شوند. هر درایه غیر قطری  $m_{i,j}$  در این ماتریسها مقدار وزن اتصال بین تصویر  $I_i$ ،  $I_j$  را دارد و همه درایه‌های قطری برابر صفر است. همچنین اگر اتصالی بین  $I_i$ ،  $I_j$  وجود نداشته باشد،  $m_{i,j}$  برابر صفر است. فرایند انتشار در مراحل گسسته  $t=0, 1, \dots, N$  انجام می‌شود. بردار  $R(t)$  به صورت بردار شباهت در مرحله  $t$  ام تعریف می‌شود و عناصر این بردار،  $r_i(t)$ ، میزان شباهت تصویر  $I_i$  را با تصویر پرس‌وجو در مرحله  $t$ ام نشان می‌دهد. در ابتدا با توجه به  $t=0$  عناصر بردار  $R(0)$  محاسبه می‌شوند. آن دسته از عناصر این بردار که متناظر با تصاویری هستند که با تصویر پرس‌وجو اتصال دارند برابر یک و سایر عناصر بردار صفر می‌شوند. از آنجائیکه در این مدل دو لایه وجود دارد، فرایند انتشار در بین دو لایه به صورت یک شیوه تناوبی<sup>1</sup> (یعنی یک مرحله در لایه بصری و مرحله دیگر در لایه معنایی) طبق رابطه زیر انجام می‌شود.

$$R(2t+1)=[a_u M_u + (1-a_u)I]R(2t)$$

$$R(2t+2)=[a_c M_c + (1-a_c)I]R(2t+1) \quad t=0,1, \dots, N$$

در رابطه بالا  $I$  یک ماتریس واحد<sup>2</sup> است و  $a_u$ ،  $a_c$  پارامترهایی در بازه  $[0,1]$  هستند که میزان جریان شباهت در بین اتصالات را کنترل می‌کنند. معمولاً  $a_u$  به مراتب بزرگتر از  $a_c$  انتخاب می‌شود. این مساله باعث می‌شود که در یک شرایط یکسان (از نظر وزن اتصالات در هر دو لایه) شباهت در اتصالات معنایی به مراتب جریان بیشتری نسبت به اتصالات بصری داشته باشد. تعداد مراحل انتشار،  $N$ ، معمولاً به صورت تجربی طوری انتخاب می‌شود که بعد از طی آن تعداد مرحله، توزیع شباهتها به یک

حالت پایدار برسد. بردار  $R(t)$  بعد از  $N$  مرحله نتیجه نهایی شباهت هر تصویر با تصویر پرس‌وجو را در خود خواهد داشت. بعد از نمایش سری اول نتایج بازیابی، کاربر می‌تواند نتایج را با تعیین نمونه‌های مرتبط تصحیح کند و به سیستم بازخورد دهد. پس از آن اتصالات معنایی با استفاده از یک الگوریتم یادگیری، بهنگام می‌شود و سپس شباهت هر یک از تصاویر پایگاه داده با تصویر پرس‌وجو طبق قاعده انتشار شباهت و بر مبنای این ایده که تصویری که در مجاورت یک تصویر مرتبط و در عین حال دور از تصاویر نامرتب باشد به احتمال قوی یک نتیجه خوب برای بازیابی خواهد بود، دوباره محاسبه می‌شود. مجاورت یک تصویر با تصویر دیگر یعنی اینکه بتوان از طریق یک یا چند اتصال (چه اتصالات معنایی و چه اتصالات ویزوالی) از یک تصویر به تصویر دیگر رسید. الگوریتم پیشنهادی فرایند بازیابی تصویر در [12] به صورت زیر می‌باشد:

(1) مجموعه تصاویر مرتبط در بین نتایج بازیابی را  $I_R$  و تصاویر نامرتب را با  $I_N$  نشان می‌دهیم.

(2) الگوریتم یادگیری را به کار می‌گیریم و اتصالات معنایی را بهنگام می‌کنیم.

(3) بردار شباهت اولیه  $R_R(0)$  را طوری تشکیل می‌دهیم که اگر  $I_i \in I_R$  باشد آنگاه  $r_i(0)=1$  و در غیر این صورت  $r_i(0)=0$ . قاعده انتشار شباهت را بر روی بردار  $R_R(0)$  به تعداد  $N = M_R$  مرحله بکار می‌بریم و بردار  $R_R(M_R)$  را به دست می‌آوریم.

(4) بردار شباهت اولیه  $R_N(0)$  را طوری تشکیل می‌دهیم که اگر  $I_i \in I_N$  باشد. آنگاه  $r_i(0)=1$  و در غیر این صورت  $r_i(0)=0$ . قاعده انتشار شباهت را بر روی بردار  $R_N(0)$  به تعداد  $N = M_N$  مرحله به کار می‌بریم و بردار  $R_N(M_N)$  را به دست می‌آوریم.

$$R^* = R_R(M_R) - R_N(M_N) \quad (5)$$

(6) عناصر  $R^*$  را به صورت نزولی مرتب می‌کنیم و تعداد  $K$  تصویر متناظر با  $K$  مقدار اول بردار  $R^*$  را به عنوان نتیجه

<sup>1</sup> - Interleaved manner

<sup>2</sup> - Identity Matrix

- (1) تصویر پرس وجو را  $I_s$  ، مجموعه تصاویر مرتبط را  $I_R$  و مجموعه تصاویر نامرتبط را  $I_N$  می‌نامیم.
- (2) برای هر  $I_i \in I_R$  ، یک واحد به وزن اتصال بین  $I_i$  ،  $I_s$  اضافه شود.
- (3) برای هر  $I_i \in I_N$  ، یک واحد از وزن اتصال بین  $I_i$  ،  $I_s$  کسر شود.

الگوریتم پیشنهادی بالا فرض می‌کند که اتصال بین گره‌ها همیشه وجود دارد، اما مقدار وزن آن می‌تواند مثبت، منفی و یا صفر باشد. مقدار مثبت نمایانگر شبیه بودن، مقدار منفی به معنی نامشابه بودن و مقدار صفر به این معنی است که قضاوتی در مورد دو تصویر نمی‌توان داشت. عمده تغییر الگوریتم جدید در مقایسه با الگوریتم یادگیری قبلی، ایجاد اتصالات با وزن منفی در لایه معنایی است که این مقادیر منفی به خوبی می‌تواند در فرآیند انتشار شباهت که در بخش 4-1 توضیح داده شد تاثیرگذار باشد.

#### 4-2. تغییر در نحوه استخراج و تفسیر اطلاعات از لایه معنایی

در مدل گراف دو لایه برای پردازش تصویر پرس‌وجو ابتدا یک مقدار اولیه شباهت در گره نمایانگر تصویر پرس‌وجو فرض می‌شود و سپس اجازه داده می‌شود که این شباهت از طریق اتصالات بصری و معنایی جریان یابد و مقدار انتشار یافته آن برابر با حاصلضرب وزن اتصالات در مقدار شباهت می‌باشد. اما واضح است که تمامی اتصالات مخصوصاً در لایه معنایی، جدا از مقدار وزنشان دارای اهمیت یکسانی نمی‌باشند و در فرآیند انتشار شباهت باید نقش متفاوتی داشته باشند. به عنوان مثال در پردازش یک تصویر پرس‌وجو اتصالاتی که یک طرف آن مستقیماً به خود تصویر پرس‌وجو وصل است قطعاً از اهمیت بیشتری در مقایسه با سایر اتصالات قرار دارند در حالیکه به این نکته در فرآیند انتشار شباهت توجهی نشده است. به همین خاطر در بهینه‌سازی این مرحله سعی شد تا حد امکان این مسأله در نظر گرفته شود. اساس

الگوریتم بالا قاعده انتشار شباهت را با استفاده از تصاویر مرتبط و نامرتبط، انجام می‌دهد. در پایان مرحله سوم  $R_R(M_R)$  شباهت تصاویر پایگاه داده با تصاویر مرتبط را مشخص می‌کند و در پایان مرحله چهارم  $R_N(M_N)$  شباهت تصاویر پایگاه داده با تصاویر نامرتبط را مشخص می‌کند و  $R^*$  که در مرحله پنجم به دست می‌آید، یک تخمین مناسب و کاملاً منطقی از تصاویر مشابه با تصویر پرس‌وجو را ارائه می‌کند.

#### 4. بهینه‌سازی مدل گراف دو لایه

##### 4-1. ارائه یک الگوریتم جدید یادگیری

ایده اصلی الگوریتم یادگیری اولیه که در [12] معرفی شده است بر پایه تقسیم کردن تصاویر به دو گروه تصاویر مرتبط و نامرتبط، توسط کاربر بود. در این الگوریتم به ازاء هر تصویر بازیابی شده‌ای که به عنوان تصویر مرتبط شناخته می‌شد تغییری در لایه معنایی به صورت ایجاد و یا تقویت یک اتصال بین دو گره، اعمال می‌شد. از اینرو تصاویر مرتبط بر یادگیری سیستم تاثیر مستقیم داشتند و از طرف دیگر به ازاء هر تصویر بازیابی شده‌ای که به عنوان تصویر نامرتبط شناخته می‌شد تغییری در لایه معنایی به صورت تضعیف و یا حذف یک اتصال بین دو گره اعمال می‌شد. اما در بیشتر مواقع ممکن است بین تصویر پرس‌وجو و تصویر نامرتبط در لایه معنایی اتصالاتی وجود نداشته باشد و لذا در این حالت الگوریتم یادگیری هیچ تغییری در لایه معنایی ایجاد نمی‌کند. یعنی الگوریتم از این تصویر نامرتبط بازیابی شده نمی‌تواند هیچگونه اطلاعاتی کسب و ذخیره کند و در واقع هیچ فرقی بین این نوع از تصویر نامرتبط بازیابی شده و تصاویری که اصلاً بازیابی نشده‌اند، نمی‌باشد. این همان ایده تغییر الگوریتم یادگیری بود که ما را بر آن داشت برای آن چاره‌اندیشی کنیم. برای حل این مشکل الگوریتم یادگیری طوری تغییر کرد که تصاویر نامرتبط به مانند تصاویر مرتبط همیشه در لایه معنایی تاثیرگذار باشند. از اینرو الگوریتم جدید یادگیری به

	Precision Ratio	Recall Ratio
Color Histogram	0.2444	0.0385
Energy calculation with wavelet transform	0.3472	0.0558
Image blocking & Color averaging	0.5556	0.1099

جدول 1- مقایسه مقادیر دقت و فراخوانی سه روش CBIR

در هر سری، آزمایشات بر روی 12 فایل تصویری پرس‌وجوی مختلف و برای هر تصویر پرس‌وجو تا 10 مرحله بازخوردهای کاربر اخذ شد. نتایج نهایی از میانگین‌گیری بر روی نتایج تمامی فایل‌های پرس‌وجو به دست آمد. پیاده‌سازی پروژه در محیط نرم افزار MATLAB-7 انجام شد و برای انجام آزمایشات، پایگاه داده دانشگاه واشنگتن از آدرس اینترنتی زیر دانلود گردید:

<http://www.cs.washington.edu/research/imagetatabase/groundtruth>

معمولاً برای ارزیابی سامانه بازیابی تصویر از دو معیار فراخوانی<sup>2</sup> و دقت<sup>3</sup> استفاده می‌شود. نرخ دقت نسبت تعداد تصاویر مرتبط بازیابی شده به تعداد کل تصاویر بازیابی شده و نرخ کارایی نسبت تعداد تصاویر مرتبط بازیابی شده به تعداد کل تصاویر مرتبط با تصویر پرس‌وجو در داخل پایگاه داده تصویری، را مشخص می‌کند. برای سهولت کار تعداد تصاویر بازیابی شده برای هر تصویر پرس‌وجو، به میزان تعداد تصاویر مشابه با آن در پایگاه داده تصویری انتخاب شد. لذا در این حالت دو معیار فراخوانی و دقت برابر می‌شوند که اصطلاحاً به آن نرخ Accuracy می‌گوییم. شکل 2 نمودار مقایسه‌ای Accuracy را در حالت بازیابی کوتاه‌مدت، برای مدل اولیه گراف دولایه به همراه هر دو مرحله بهینه‌سازیهای صورت‌گرفته بر روی آن نشان می‌دهد. و شکل 3 همین مقایسه را برای بازیابی بلندمدت نشان می‌دهد. همانگونه که مشاهده می‌کنید و در بخش 4-2 نیز عنوان شد بهینه‌سازی سطح دوم در بازیابی کوتاه‌مدت تأثیری ندارد.

تغییر ایجاد شده به این صورت است که اتصالاتی که مستقیماً به خود تصویر پرس‌وجو متصل هستند از اهمیت بیشتری برخوردار باشند. لذا برای اینگونه اتصالات در فرآیند انتشار شباهت به طور کامل و بدون هیچ گونه تغییری از مقدار وزن خالص این اتصالات استفاده شد و برای سایر اتصالاتی که دسترسی به آنها از گره تصویر پرس‌وجو، حداقل بعد از دو یا چند گام بیشتر امکانپذیر است، از علامت وزن آنها طبق رابطه زیر استفاده می‌شود.

$$w_{i,j} = \begin{cases} 1 & \text{if } w_{i,j} \geq 1 \\ 0 & \text{if } w_{i,j} = 0 \\ -1 & \text{if } w_{i,j} \leq -1 \end{cases}$$

البته واضح است که این بهینه‌سازی هیچ تأثیری در نتایج بازیابی کوتاه‌مدت نخواهد داشت چرا که در این نوع بازیابی فقط اتصالاتی شرکت دارند که مستقیماً به تصویر پرس‌وجو متصل هستند.

## 5. نتایج آزمایشات

برای استخراج ویژگیهای سطح پایین و ایجاد اتصالات بین گره‌ها در لایه بصری مدل گراف دولایه سه روش شناخته شده بازیابی مبتنی بر محتوی پیاده‌سازی شد. اولین روش بازیابی تصاویر مبتنی بر رنگ با استفاده از فناوری تطبیق هیستوگرام‌های تصویر در فضای رنگی HSV است. در روش دوم انرژی تصویر با استفاده از تبدیل موجک<sup>1</sup> هرگونه محاسبه می‌شود و روش سوم بلاک‌بندی تصویر و میانگین‌گیری مقدار رنگ هر بلاک می‌باشد. جدول 1 نتایج حاصل از مقایسه سه روش CBIR پیاده‌سازی شده را نشان می‌دهد که برتری روش سوم در مقایسه با دو روش دیگر مشخص است. به همین خاطر، این روش برای بکارگیری در لایه بصری مدل گراف دولایه، انتخاب گردید.

<sup>2</sup> - Recall ratio

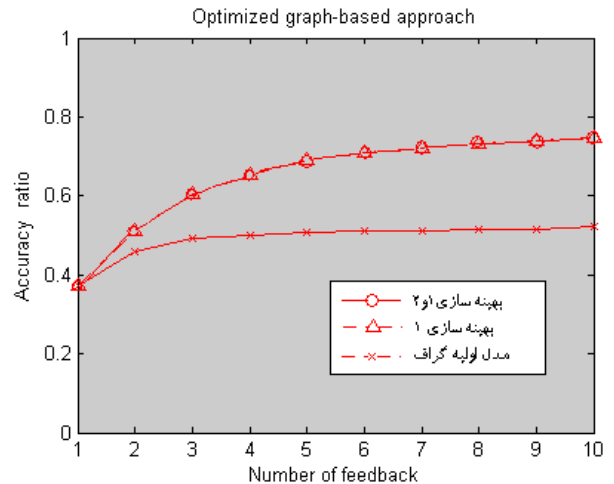
<sup>3</sup> - Precision ratio

<sup>1</sup> - Wavelet transform

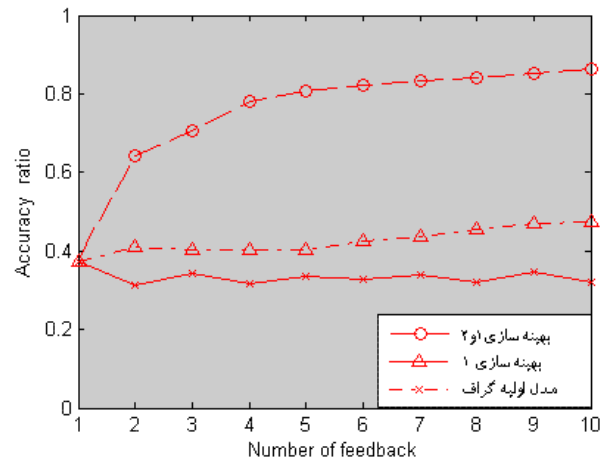
توجه به اهمیت مفاهیم سطح بالای تصویر و نقش بازخوردهای مرتبط در بازیابی تصویر، بهینه‌سازیهایی صورت گرفته در مدل گراف دولایه بیشتر مربوط به لایه معنایی می‌باشد. نتایج این تحقیق علاوه بر نشان دادن قابلیت و پتانسیل بالای مدل گراف دولایه برای یک سیستم بازیابی تصویر، به خوبی تأثیر بهینه‌سازی اول و ارائه الگوریتم جدید یادگیری را در هر دو حالت بازیابی کوتاه‌مدت و بلندمدت نشان داد. همچنین تأثیر بهینه‌سازی دوم در بازیابی بلندمدت از دیگر نتایجی است که به خوبی در این تحقیق مشخص شد. علیرغم کارایی بسیار خوب مدل گراف دولایه و بهینه‌سازیهایی صورت گرفته، این مدل هنوز از پتانسیل بالایی برای بهبود و توسعه برخوردار است. بهبود نحوه ذخیره و بازیابی اتصالات، تقویت لایه بصری با به کارگیری روشهای CBIR کارا تر و نیز شناسایی و حذف اتصالات نویزی، افزودن یک لایه متنی یا لایه کلاسترینگ به منظور افزایش قابلیت‌های مدل و بالاخره قابلیت به کارگیری این مدل در سایر زمینه‌ها، مثل بازیابی اطلاعات و یا مسائل کلاسه‌بندی از جمله کارهایی است که می‌توان در ادامه این کار انجام داد.

## مراجع

- [1]. C. Faloutsos, M. Flocker, W. Niblack, D. Petkovic, W. Equitz, R. Baeber, "Efficient and effective querying by image content", Technical Report RJ 9453 (83074), IBM Research Report, 1993.
- [2]. W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC project: querying images by content using color, texture, and shape", in Proc. Of storage and retrieval for image and video database, pp. 173-187, USA, 1993.
- [3]. Y.Rui, T.S.Huang, M.Ortega, and S.Mehrotra, "Relevance Feedback : A Power Tool in Interactive Content-Based Image Retrieval", IEEE Tran. Circuits and Systems for Video Tech, 8(5), pp. 644-655, Sept. 1998.
- [4]. J. Peng, B. Banerjee, D. Heisterkamp, "Kernel index for relevance feedback retrieval in large image



شکل 2 - مقایسه نرخ Accuracy بعد از اعمال بهینه‌سازی اول و دوم در مدل گراف دولایه برای بازیابی کوتاه‌مدت



شکل 3 - مقایسه نرخ Accuracy بعد از اعمال بهینه‌سازی اول و دوم در مدل گراف دولایه برای بازیابی بلندمدت

## 6. نتیجه‌گیری

در این مقاله یک سیستم بازیابی تصویر تعاملی مبتنی بر یک مدل گراف دولایه پیاده‌سازی و اصلاح شد. مدل گراف دولایه با بهره‌گیری از دو لایه مستقل بصری و معنایی هم از اطلاعات محتوایی تصاویر و هم از مفاهیم سطح بالای آنها استفاده می‌کند. در این مقاله سه روش مختلف بازیابی تصویر مبتنی بر محتوی به منظور به کارگیری در لایه بصری و همچنین الگوریتمهای یادگیری و فرآیند بازیابی جهت استفاده در لایه معنایی پیاده‌سازی شد. با

database", in proceeding of the 9<sup>th</sup> international conference on neural information processing, Singapore, 2002.

- [5]. J. Peng, B. Bhanu, S. Qing, "Probabilistic feature relevance learning for content based image retrieval" *Compute Vision and image understanding* 75(1/2): 150-164, 1999.
- [6]. Y. Rui, T. S. Huang, and S. Mehrotra, "Content-Based Image Retrieval with relevance feedback in MARS," *proc. IEEE Int. Conf. on Image Processing, Santa Barbara*, pp. 815-818, 1997.
- [7]. I. Gondra, D. R. Heisterkamp, J. Peng, "Improving image retrieval performance by inter-query learning with one-class support vector machines", *Neural Compute & Applic*, 13: 130-139, 2004.
- [8]. S. C. Cheng, T. C. Chou, C. L. Yang, H. Y. Chang, "A semantic learning for content-based image retrieval using analytical hierarchy process", *Expert systems with Application*, pp. 495-505, 2005.
- [9]. C. S. Lee, Y. W. Ma, H. J. Zhang, "Information Embedding Based on User's Relevance Feedback for Image Retrieval ", *Technical Report, HP Labs*, 1998.
- [10]. Y. Lu, et al. "A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems", *Proc. of ACM Multimedia*, pp. 31-38, 2000.
- [11]. T. P. Minka et al., "Interactive Learning with a 'society of models'", *Proc. of IEEE CVPR*, pp. 447-452, 1996.
- [12]. Y. Zhuang, J. Yang, Q. Li, Y. Pan, "A graphic-theoretic model for Incremental Relevance Feedback in Image Retrieval", *Proc. Of 2002 Int'l Conf. on Image Processing, New York, Sep*, 2002.