روش ترکیب یادگیری تقویتی و الگوریتم ژنتیک برای جستجوی کمتر در بازیابی تصویر

نویسندگان: امین عابدینی- aminabedunu.ai@gmail.com
عفت جبارپور- ejabarpoor@yahoo.com

چکیده: الگوریتم جستجو در بازیابی تصاویر تمایل به تمرکز بر روی تصاویر به‌طور انحصاری دارد . به‌طور ضمنی، سیستم اکتشاف کاربران، از فضای تصویر و حذف بالقوه به دست می‌آید. بنابراین، در سال‌های اخیر، علاقه به توسعه الگوریتم‌های اکتشافی و بهره‌برداری برای جستجو در تصویر افزایش‌یافته است. در این مقاله ما یک سیستم بازیابی تصویر از ترکیب یادگیری تقویتی و الگوریتم ژنتیک به همراه یک رابط کاربر طراحی کرده‌ایم. در این سیستم، یادگیری تقویتی به کاربران اجازه می‌دهد تا بین اکتشاف و بهره‌برداری سویچ کنند و الگوریتم ژنتیک این سویچ و نتایج حاصل‌شده از یادگیری تقویتی را بهینه می‌کند .

کلیدواژه: الگوریتم ژنتیک ، بازیابی تصویر ، بازیابی تصویر بر اساس محتوا ، یادگیری تقویتی، *CBIR*

۱ – مقدمه

در سال‌های اخیر، تکنیک‌های بازیابی تصویر در متا داده‌ها، مانند حاشیه‌نویسی متنی و یا برچسب‌های مشخص‌شده توسط کاربر، استاندارد بازیابی از مجموعه تصویر بزرگ، مانند جستجوی تصویر گوگل، پیشرفت خوبی داشته‌اند. این رویکرد باکیفیت و اندازه خوبی در متا داده، در حال رشد بود که انفجاری در مجموعه تصویر، توسط کاربران فیس بوک ایجاد شد، به‌طور متوسط ۳۵۰ میلیون عکس در روز در طول سه‌ماهه چهارم سال ۲۰۱۲ ارسال شد و همچنین نمایش مشخصات عمومی تصاویر به‌طور متوسط ۵۵ میلیون بود. به همین دلیل معلوم شد که برچسب زدن تصاویر جدید با سرعت‌بالا و کارآمد بودن، [۱۲] همیشه ممکن نیست. ثانیاً، برچسب زدن تصاویر باکیفیت بالا به‌صورت آنی سخت است، هنوز هم در بسیاری از موارد جستجوی تصویر به درخواست کاربر مشکل‌ساز است.

ممکن است تعریف جستجو برای کاربر آسان باشد مثلاً کاربر به دنبال یک تصویر از یک گربه باشد، اما چگونه می‌توانیم مشخص کرد که گربه یک سایه بسیار خاص دارد و اندوهی در چشم آن است. یک‌راه حل برای این وضعیت وجود دارد که در آن کاربر قادر به مشخص کردن محتوای موردنیاز از طریق تگ یا سایر خواص تصویر می‌باشد، این بر اساس بازیابی محتوا تصویر (CBIR) است.

آزمایش اول از CBIR در تاریخ ۱۹۹۲ [۸] انجام شد و پس‌ازآن بسیاری از سیستم‌های بازیابی تصویر CBIR توسعه‌یافته‌اند [۴، ۱۰، ۱۵، ۱۷]. آزمایش‌های اولیه نشان می‌دهد که روند بازیابی تصویر را می‌توان تا حد زیادی با استفاده از بازخورد [۲۰] از طریق درگیر کاربر با حلقه جستجو و استفاده از دانش خود در فرایند جستجو تکراری [۳، ۱۱، ۱۸] بهبود داد.

استفاده از استراتژی‌های اکتشاف / بهره‌برداری‌های مختلف در بازخورد، یکی دیگر از تلاش‌ها برای بازیابی تصویر بوده است. هدف از گام بهره‌برداری، بازگشت به کاربران با حداکثر تعداد تصاویر مربوطه در یک منطقه از فضای ویژگی است، درحالی‌که گام اکتشاف باهدف جستجو در مناطق مختلف از فضای ویژگی است. این نوع سیستم‌ها به‌صورت پویا در هر تکرار، مجموعه‌ای از تصاویر را کنترل و نمایش می‌دهند. برای مثال، در [۱۶] یک الگوریتم masszoom ارائه‌شده است که مجموعه‌ای از تصاویر را نمایش داده است و بر روی تصاویر به‌عنوان مناسب‌ترین واسط کاربر متمرکز کرده است.

در [۱۳]، یک روش در نظر گرفته‌شده است که از اطلاعات به‌دست‌آمده توسط روش نزدیک‌ترین همسایه در مرحله بهره‌برداری، برای محاسبه استفاده می‌کند. به‌طور مشابه، یادگیری تقویتی یک روش امیدبخش است که می‌تواند به سیستم اجازه دهد از بازخورد جستجو ربط استفاده کند، و در همان زمان، از سوییچ بین بهره‌برداری و اکتشاف استفاده کند. یادگیری تقویتی در زمینه‌های مختلف از بازیابی اطلاعات اعمال می‌شود [۲، ۶]. بااین‌حال، بسیاری از سیستم بازیابی تصویر از استراتژی اکتشاف - بهره‌برداری [۱۳، ۱۶، ۱۹] استفاده می‌کنند.

در این مقاله، ما به طراحی یک سیستم جستجو در تصویر پرداخته‌ایم که این طراحی با ترکیب تکنیک‌های یادگیری تقویتی و الگوریتم ژنتیک انجام‌شده است. این ترکیب اجازه می‌دهد تا بهینه‌سازی در سیستم اکتشاف و بهره‌برداری به دست آید.

www.kahrobaonline.ir

هدف اصلی در این سیستم، کمک به کاربر در پیدا کردن تصاویر در یک پایگاه داده از تصاویر unannotated با سرعت‌بالا و بدون تایپ کردن هر سؤال است.

۲- روش

در این مقاله ما با استفاده از ترکیب الگوریتم ژنتیک و الگوریتم یادگیری تقویتی سعی بر کمتر کردن جستجو در بازیابی تصاویر داشته‌ایم. شکل ۱ فلوچارت روش پیشنهادی را نشان می‌دهد که در ادامه بر شرح کدام از مراحل خواهیم پرداخت.



شکل ۱. فلوچارت روش پیشنهادی

شکل ۱-فلوچارت روش پیشنهادی

۲-۱- پیش‌پردازش

در این مرحله تصاویر ورودی از سیستم گرفته‌شده و نویز زدایی بر روی تصاویر برای بالا رفتن کیفیت تصاویر با استفاده از فیلتر میانه انجام می‌شود. این فیلتر نویزهای موجود در سطح تصویر را تا حدودی از بین برده و تصویر را به یک تصویر باکیفیت نزدیک‌تر می‌کند.

۲-۲- یادگیری تقویتی

در این مرحله کاربر دستوری ارائه کرده است حال توسط الگوریتم تقویتی و با استفاده از بهره‌برداری و اکتشاف به جستجوی تصاویر پرداخته می‌شود. سپس تعدادی از تصاویر حذف و مجموعه‌ای دیگر از تصاویر در دسترس می‌باشد که به الگوریتم ژنتیک داده خواهد شد. الگوریتم تقویتی با سویچ بین اکتشاف و بهره‌برداری دایره جستجوی خود را کاهش می‌دهد. ولی این جستجو همچنان بهینه نیست و باید با الگوریتم ژنتیک ترکیب به وجود آید.

۲-۳- الگوریتم ژنتیک

در این مرحله تصاویر خارج‌شده از جستجوی الگوریتم تقویتی دوباره مورد جستجو با واژه کلیدی کاربر قرار خواهند گرفت. در اینجا با استفاده از عملگرهای برش و جهش و همچنین تابع هدف به محدود کردن فضای جستجو پرداخته می‌شود. هرکدام از این عملگرها در ادامه شرح داده می‌شوند. الگوریتم ژنتیک یکی از الگوریتم‌های بسیار عالی برای بهینه‌سازی می‌باشد.

۳-طراحی رابط

ایده اصلی ایجاد تعاملی در پشت رابط کاربر است که کاربر به‌جای تایپ کردن کلمات مربوط به تصویر موردنظر، سیستم از طریق محتویات قید "بستن" یا "مشابه" تصاویر را نمایش دهد. در تکرار بعدی، کاربر با یک مجموعه جدید از تصاویر ارائه‌شده است. رابط کاربری با یک مثال جستجو در شکل ۲ ارائه‌شده است.

جستجو با یک collage از تصاویر که به‌طور یکنواخت هستند به‌طور تصادفی از پایگاه داده نمونه شروع می‌شود. به‌منظور اینکه اطمینان حاصل شود که مجموعه اولیه از تصاویر یک نمایش خوب از کل فضای تصویر است، ما تمام تصاویر را با K خوشه‌بندی کردیم و k تعداد تصاویر نمایش داده‌شده است و سپس ما یک تصویر از هر خوشه را به‌عنوان نمونه برداشتیم. آزمایش‌های قبل نشان می‌دهد که این روش یک نقطه شروع خوب برای جستجو می‌باشد.

بازخورد نتایج در دامنه‌ای از ۰ تا ۱ می‌باشد، که در آن ۰ نشان‌دهنده یک تصویر داده‌شده متفاوت از هدف ایده آل است و ۱ نشان‌دهنده که یک تصویر داده‌شده بسیار شبیه به هدف ایده آل است. تصاویری که توسط کاربر امتیاز نیست، فرض می‌شود که نمره ۰ گرفته است. پس از هر تکرار، تصاویر جدید نمایش داده می‌شوند. هر تصویر را می‌توان به کاربر یک‌بار و در تمام طول جستجو نشان داد، یعنی یک تصویر نشان داده‌شده در تکرار بعدی نشان داده نمی‌شود. جست‌وجو ادامه دارد تا زمانی که کاربر به رضایت برسد.

در مثال ما، کاربر می‌خواهد یک تصویر برای نشان دادن "شهر در شب" پیدا کند. در آغاز جستجو (شکل a۲)، کاربر با یک collage از تصاویر یکنواخت از پایگاه داده تصویر انتخاب‌شده روبرو است، در تصاویر ارائه‌شده، تصویر پنجم در ردیف دوم و تصویر دوم در ردیف سوم به هدف ایده آل نزدیک‌ترند. کاربر با تکرار بعدی با فشار دادن دکمه "بعدی" در بالای صفحه حرکت می‌کند. در تکرار دوم (شکل B2)، تصاویر مربوط به "شب" یا "شب" ارائه‌شده است. بازهم، کاربران چهار عکس نزدیک به آنچه او به دنبالش است، انتخاب می‌کند. در تکرار ۳ و ۴ (شکل C۲ و D)، عکس‌های بیشتری مربوط به موضوع به کاربر ارائه‌شده و کاربر می‌تواند با انتخاب تصاویر خاص خود جستجو را محدود کند.


(a) Iteration 1 (b) Iteration 2

www.kahrobaonline.ir



شکل۲. چهار تکرار اول یک مثال جستجو برای "شهر در شب". در تکرار اول کاربر علائم دو عکس به‌عنوان مربوطه را می‌بیند. در جستجوهای بعدی کاربر بیشتر به هدف نزدیک می‌شود.

۴- الگوریتم ژنتیک

در اینجا از الگوریتم ژنتیک برای بهینه‌سازی جستجو استفاده‌شده است. به‌طوری‌که هر دوره از تصاویری که به کاربر نشان داده می‌شود به‌عنوان نسل اولیه در الگوریتم ژنتیک قرار داده می‌شود و پس‌ازآن با در نظر گرفتن تابع هدف و عملیات برش و جهش و محدوده جستجو کم می‌شود. در اینجا به‌وسیله تابع هدف (شب بودن و پیکسل‌های تیره در تصویر هدف ما هست) می‌توان محدود جستجو را کاهش داد. نرخ برش برابر با ۰/۳ و نرخ جهش برابر با ۰/۴ در نظر گرفته‌شده است[۲۱].

۵- مدل‌سازی کاربر

جستجو در سراسر پایگاه داده، سیستم یک مدل برای کاربران ایجاد می‌کند و بر اساس آن تصاویر K برای ارائه به کاربر در هر تکرار انتخاب می‌شود. ما فرض کنیم که کاربر به دنبال یک تصویر یا مجموعه‌ای از تصاویر مربوط به خود است. بازخورد کاربر توسط یک نمره مربوط $Y \epsilon [۰، ۱]$ ، که در آن ۱ نشان‌دهنده مربوط بودن بالای تصویر است. پروتکل تکرار به شرح زیر است:

- سیستم، k تصاویر انتخاب و به کاربر ارائه می‌کند.
- کاربر نمرات مربوطه را دربازه ۰ و۱ نشان می‌دهد. تصاویر که توسط کاربر امتیاز داده نمی‌شوند ۰ فرض می‌شود.
- تصاویری که ۱ انتخاب شدند به الگوریتم ژنتیک برای محدود شدن فضای جستجو داده می‌شوند.

۶- نتیجه‌گیری

برای حمایت از کاربران در جهت اکتشاف فضای جستجوی تصویر، ما یک سیستم تعاملی CBIR که شامل اکتشاف و بهره‌برداری با بازخورد کاربر طراحی کردیم. کاربران می‌توانند فضای تصویر را با بازخورد خود و کمک الگوریتم ژنتیک کم کنند. در این سیستم مجموعه داده با برچسب تصاویر نیاز نیست، کاربران نیاز به دانستن تصاویر ندارند. ابتدا به‌صورت تصادفی مجموعه‌ای از تصاویر انتخاب‌شده به کاربر نشان داده می‌شود و به‌وسیله بازخورد کاربر و الگوریتم ژنتیک این فضای جستجو محدود می‌شود تا نهایت کاربر به هدف خود می‌رسد.

مراجع

[1] Auer, P., Cesa-Bianchi, N., Fischer, P.: Finite-time analysis of the multiarmed bandit problem. Mach. Learn. 47, 235–256 (2002)

[2] Auer, P., Hussain, Z., Kaski, S., Klami, A., Kujala, J., Laaksonen, J., Leung, A.P., Pasupa, K., Shawe-Taylor, J.: Pinview: implicit feedback in content-based image retrieval. JMLR Workshop Conf. Proc. 11, 51–57 (2010)

[3] Cox, I., Miller, M., Minka, T., Papathomas, T., Yianilos, P.: The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments. Image Process. 9(1), 20–37 (2000)

[4] Datta, R., Li, J., Wang, J.: Content-based image retrieval: approaches and trends of the new age. In: Multimedia Information Retrieval, pp. 253–262. ACM (2005)

[5] G_lowacka, D., Hore, S.: Balancing exploration-exploitation in image retrieval. In: Proceedings of UMAP 2014 Posters, Demonstrations and Late-Breaking Results (2014)

[6] G_lowacka, D., Shawe-Taylor, J.: Content-based image retrieval with multinomial relevance feedback. In: Proceedings of ACML, pp. 111–125 (2010)

[7] Guiver, J., Snelson, E.: Learning to rank with softrank and Gaussian processes. In: Proceedings of SIGIR, pp. 259–266 (2008)

[8] Kato, T., Kurita, T., Otsu, N., Hirata, K.: A sketch retrieval method for full color image database-query by visual example. In: Pattern Recognition. Computer Vision and Applications, pp. 530–533 (1992)

[9] Kelly, D., Fu, X.: Elicitation of term relevance feedback: an investigation of term source and context. In: Proceedings of SIGIR (2006)

[10] Kosch, H., Maier, P.: Content-based image retrieval systems-reviewing and benchmarking. JDIM 8(1), 54–64 (2010)

[11] Laaksonen, J., Koskela, M., Laakso, S., Oja, E.: Picsom-content-based image retrieval with self-organizing maps. Pattern Recogn. Lett. 21(13), 1199–1207 (2000)

[12] Pham, T.-T., Maillot, N.E., Lim, J.-H., Chevallet, J.-P.: Latent semantic fusion model for image retrieval and annotation. In: Proceedings of CIKM (2007)

[13] Piras, L., Giacinto, G., Paredes, R.: Enhancing image retrieval by an exploration exploitation approach. In: Perner, P. (ed.) MLDM 2012. LNCS (LNAI), vol. 7376, pp. 355–365. Springer, Heidelberg (2012)

[14] Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press, Cambridge (2006)

www.kahrobaonline.ir

[15] Smeulders, A., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. Pattern Anal. Mach. Intell. 22(12), 1349– 1380 (2000)

[16] Suditu, N., Fleuret, F.: Iterative relevance feedback with adaptive exploration/exploitation trade-off. In: Proceedings of CIKM (2012)

[17] Veltkamp, R.C., Tanase, M.: Content-based image retrieval systems: a survey. Department of Computing Science, Utrecht University (2002)

[18] Villegas, M., Leiva, L.A., Paredes, R.: Interactive image retrieval based on relevance feedback. In: Sappa, A.D., Vitri`a, J., Multimodal Interaction in Image and Video Application, vol. 48, pp. 83–109. Springer, Heidelberg (2013)

[19] Yee, K.-P., Swearingen, K., Li, K., Hearst, M.: Faceted metadata for image search and browsing. In: Proceedings of CHI, pp. 401–408 (2003)

[20] Zhou, X., Huang, T.: Relevance feedback in image retrieval: a comprehensive review. Multimed. Syst. 8(6), 536–544 (2003)

[21] Miguel, A., Francesc, J., Salvador, M. : Improving distance based image retrieval using non dominated sorting genetic algorithm. Elsevier, 201