مروری بر نمایه‌سازی متنی، محتوایی و فازی تصاویر رقمی: نظام بازیابی تصاویر رقمی

محمدباقر دستغیب[1]

چکیده

شبکه جهانی اینترنت، مجموعه‌ای ناهمگون از اطلاعات متنی، صوتی و تصویری‌ست. مطالعات اخیر در زمینه بازیابی اطلاعات، بیشتر در حوزه بازیابی اطلاعات متنی بوده است. بازیابی تصاویر در حوزه‌هایی نظیر هنر، معماری و پزشکی کاربرد بسیار دارد.

در حال حاضر حدود سه ترابایت تصویر قابل دسترس عموم در شبکه اینترنت وجود دارد و اندکی باز این مجموعه بزرگ نمایه شده است. نمایه‌سازی تصاویر برپایه متن همراه با تصویر، برپایه محتوای تصویر و یا با درنظر گرفتن نسبی بودن مقایسه‌های هوشمند، به‌صورت فازی انجام می‌شود مدلسازی عدم قطعیت ذاتی نظام‌های بازیابی اطلاعات، از کاربردهای منطق فازی‌ست.

کلیدواژه‌ها

بازیابی تصویر رقمی، نمایه‌سازی فازی، نمایه‌سازی متنی، نمایه‌سازی محتوایی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقدمه

شبکه جهانی اینترنت، مجموعه‌ای گسترده و ناهمگون از اطلاعات متنی و غیرمتنی از قبیل صوت، تصویر و چندرسانه‌ای[2] هاست. در سال‌های اخیر تحقیقات گسترده‌ای در زمینه بازیابی اطلاعات متنی انجام شده که کاوشگرهای متنی را به ارمغان آورده است ولی فعالیت برروی نظام‌های بازیابی اطلاعات تصویری[3]، به‌صورت پراکنده انجام شده است. دلیل اصلی مشکلات جستجوی تصاویر در شبکه جهانی اینترنت، نسبت به جستجوی تصاویر در پایگاه داده‌ها[4]، گسترده بودن اینترنت، اضافه شدن روزانه اطلاعات و گستردگی نوع تصاویر موجود در شبکه جهانی‌ست (۳).

طبق آمار سال ۱۹۹۹، ۱۸۰ میلیون تصویر در دسترس عموم[5] برای نمایه‌سازی[6] وجود دارد. حجم این داده‌های تصویری حدود ۳ ترابایت[7] تخمین زده می‌شود و روزانه در

۱. کارشناس ارشد مهندسی کامپیوتر کتابخانه منطقه‌ای علوم و تکنولوژی شیراز
dstghaib@srlst.com

۷. هر ترابایت 10^{12} بایت است.

2.Multimedia
3.Image Retrieval System
4.Database
5.Public access
6.Indexing

حدود یک میلیون تصویر، تولید می‌شود (۳). با وجود این، طبق آمار ارائه شده، جستجوگر آلتاویستا، تا سال ۲۰۰۲ فقط ۲۹ میلیون تصویر را نمایه کرده است. با توجه به افزایش روزافزون تصاویر چندرسانه‌ای با قاب‌های گوناگون، تخمین زده می‌شود تا پایان ۲۰۰۵ بیش از یک تریلیون تصویر قابل نمایه وجود داشته باشد. تا پایان ۲۰۰۳، جستجوگر گوگل، ۴۵ میلیون تصویر را نمایه کرده است و از تعداد تصاویر نمایه شده تا پایان سال ۲۰۰۵ نیز هنوز آمار دقیقی ارائه نشده است، ولی با توجه به میزان افزایش آمارهای ارائه شده نسبت به سال‌های قبل، تعداد تصاویر نمایه شده نسبت به تصاویر تولید شده، مقدار ناچیزی‌ست (۴).

موضوع مهم دیگری که در بحث جستجوی اطلاعات وجود دارد، دقت جواب‌هایی‌ست که موتور جستجو[8] به کاربر عرضه می‌کند. وقتی جستجو برای یافتن تصاویر باشد، ممکن است کاربر خصوصیات تصویر مانند رنگ، بافت[9]، نوع تصویر و ابعاد تصویر را مشخص کند. بنابراین جستجو و نمایه‌سازی تصاویر، با اطلاعات متنی متفاوت است (۷).

تحقیقات اخیر در زمینه جستجو در اینترنت، سه عمل را لازمه جستجوی موفق دانسته است. اولین مورد نمایه‌سازی‌ست. در نمایه‌سازی تصاویر، خصوصیات تصویر و نشانی محل ذخیره تصویر در وب[10]، ذخیره می‌گردد. دومین مورد، دسته‌بندی[11] یا جمع‌آوری تصاویر مشابه در یک مجموعه است. دسته‌بندی به نمایه‌سازی کمک می‌کند، و امکان سازماندهی پاسخ‌ها را فراهم می‌آورد. بازخورد ربط[12]، مورد سوم است. بازخورد ربط، موجب قرارگرفتن جواب‌های صحیح در ابتدای صفحه می‌شود، بنابراین کاربر نیازی به حرکت دادن صفحه و مراجعه به صفحات بعد برای یافتن اطلاعات مرتبط ندارد (۲).

هر تصویر را به روش‌های گوناگون می‌توان نمایه‌سازی کرد. وجود روش‌های مختلف، موجب ابهام نمایه‌سازی تصاویر می‌شود. تعیین معیار مشابهت تصاویر، در زمان دسته‌بندی آسان نیست، و به نیاز کاربران وابسته است. برای حفظ سرعت بازیابی، دسته‌بندی و نمایه‌سازی تصاویر باید قبل از مرحله بازیابی انجام شود (۶).

هدف این مقاله مرور روش‌های نوین نمایه‌سازی تصاویر است. امروزه در موتورهای کاوش از این روش‌ها یا ترکیب آنها استفاده می‌شود. در ادامه برخی روش‌های مرسوم نمایه‌سازی تصاویر و سپس کاربرد منطق فازی[13]، در بازیابی تصاویر مرور می‌گردد.

۲. نمایه‌سازی اطلاعات تصویری

۲-۱. نمایه‌سازی متنی تصاویر

بیشتر نظام‌های بازیابی اطلاعات متنی هستند، ولی اغلب تصاویر، اطلاعات متنی مختصری در کنار تصویر دارند. توانایی نمایه‌سازی بر پایه متن، معرفی عمومی یا خاص شیء، در سطوح مختلف پیچیدگی‌ست. قبل از رقمی‌شدن تصاویر، جستجو در مجموعه تصاویر، با مراجعه به کتابداران و جستجو در خصوصیات متنی و کدهای طبقه‌بندی موضوعی انجام می‌شده است. امروزه، در برخی موارد، از ترکیبی از سازوکارهای دستی و خودکار برای نمایه‌سازی استفاده می‌شود. برای مثال در یک گنجینه هنر و معماری[14]، بالغ بر ۱۲۰ هزار تصویر با نمایه متنی سلسله مراتبی تا هفت مرحله (مراحل: مفاهیم وابسته، خصوصیات فیزیکی، دوره و سبک، نماینده دسته، کاربری، جنس، و اشیاء) وجود داشته است (۳).

مشکل بازیابی دقیق به کمک توصیف تصاویر با یک مجموعه واژگان کنترل شده، و یا دسته‌بندی الگوها، حل نمی‌شود. توصیف متنی تصویر مشکل‌ساز است، زیرا رساندن محتوای تصویر و چگونگی آن با کلمات، آسان نیست. بازیابی با ابرداده[15]، روشی در دسته‌بندی تصاویر است که به‌صورت تحقیقاتی به‌کار رفته است. ابرداده شامل خصوصیات تصویر مانند نام سازنده، نوع قالب‌بندی، تاریخ ایجاد، و عنوان صفحه است. ائتلاف جهانی وب[16]، استانداردی برای ابرداده تعیین کرده است. این استاندارد که "قالب توصیف منبع"[17] نام دارد، علاوه بر اطلاعات متنی، اطلاعات دیگری مانند هیستوگرام[18] رنگ را نیز معین می‌کند، ولی استفاده از ابرداده هنوز گسترش کافی نیافته است (۳).

۸.Search Engine
۹.Texture
۱۱.Clustering
۱۲.Relevancy Feedback
۱۳.Fuzzy Logic
۱۴.Getty's art and architecture thesaurus
۱۶.W_3C
۱۷.RDF = Resource Description Framework

۱۰. Web (شبکه جهانی)
۱۵. Metadata: داده‌ای که درباره داده دیگری توضیح می‌دهد.
۱۸. نموداری برای نمایش تعداد تکرار داده در محیط آماری

۲-۲. نمایه‌سازی محتوایی تصاویر

مشکلات دسترسی به تصاویر با نمایه‌سازی متنی، پژوهشگران را برانگیخته است، تا راه حل‌های بهتری برای نمایه‌سازی تصاویر بیابند. یکی از روش‌هایی که اخیراً به‌کار برده می‌شود، نمایه‌سازی محتوایی تصاویر است. نمایه‌سازی محتوایی تصاویر، براساس خصوصیات بدوی تصاویر مانند رنگ، شکل، بافت و هر خصوصیتی که مستقیماً تصویر را توصیف کند، کار می‌کند. اولین نظام‌های تجاری بازیابی تصاویر بر پایه محتوا، در سال ۱۹۹۵ میلادی به کار گرفته شدند (۳).

پرسش‌هایی که از نظام بازیابی تصاویر بر اساس محتوا می‌شود، مانند خصوصیات تصویر (رنگ و نوع قالب‌بندی و ...) است. برای مثال، کاربران ممکن است درخواست برای یک طرح را با مشخص نمودن پالت رنگ و یا شکل مورد نظر خود، به کاوشگر ارسال دارند. نظام در میان مجموعه تصاویر نمایه‌سازی شده جستجو می‌کند و تصاویری که دارای بالاترین معیار شباهت هستند، با یک آستانه[19] فهرست می‌کند (۲).

خصوصیات تصاویر

رنگ

بازیابی تصاویر، براساس شباهت طیف رنگی، با محاسبه هیستوگرام رنگ برای هر تصویر، انجام می‌شود. در هیستوگرام، سهم نقاط[20] هر طیف رنگی در تصویر مشخص می‌شود. تحقیقات جدید در این زمینه، به روی تطبیق منطقه‌ای رنگ تأکید می‌کند. هیستوگرام رنگی یک تابع چهار بعدی‌ست، که شدت رنگ‌های اصلی را برای هر نقطه توصیف می‌کند. وقتی تفاضل دو هیستوگرام رنگ، به اندازه کافی کوچک باشد، دو هیستوگرام شبیه هستند (۷).

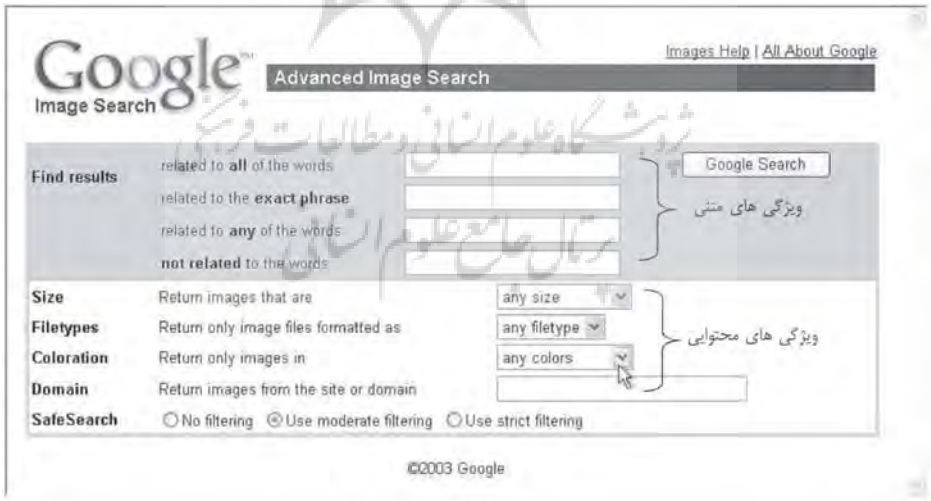
بافت

توصیف بافت تصویر مشکل است. توصیف بافت یک تصویر، با تغییر طیف خاکستری مشخص می‌شود. میزان روشنایی دو نطقه مجاور و تمایز آن، بافت را توصیف می‌کند. مشکل عمده، شناسایی و دسته‌بندی الگوی بافت (مثلاً نرم و یا خشن) است. برای رفع این مشکل در سال ۱۹۹۸، مجموعه‌ای از الگوهای بافت، تشکیل شده است (۲).

نوع و شکل تصویر

درخواست[21] برای یافتن اشکال، معمولاً با مثال انجام می‌شود. اصول بازیابی اشکال، براساس خصوصیات ویژه مانند خطوط، مرز شکل، نسبت وضعیت[22] و تشخیص لبه‌های تصویر است. در این قسمت علاوه بر رنگ و بافت، ویژگی‌های بصری تصویر و نوع آن مشخص می‌شود (۲).

مثال ۲-۱ موتور جستجوی تصاویر را درسایت گوگل نشان می‌دهد. این صفحه قابلیت جستجوی محتوایی و متنی به



مثال ۲-۱. موتور کاوشگر گوگل و بازیابی محتوایی تصاویر

۲۲. Aspect ratio : نسبت طول و عرض تصویر نسبت به صفحه نمایش را مشخص می‌کند.

19.Threshold
20.Pixels
21.Query

صورت ترکیبی را دارد. در این کاوشگر تصویر، می‌توان ابعاد تصویر، نوع و قالب فایل و رنگ آن را مشخص نمود.

مثال ۲-۲ موتور جستجوی تصاویر آلتاویستا را نشان می‌دهد. در این موتور جستجو نیز قابلیت جستجوی متنی و محتوایی به‌صورت ترکیبی وجود دارد.

۲-۳. نمایه‌سازی هوشمند برپایه منطق فازی

نمایه‌سازی نیازمند ایجاد رابطه میان داده‌ها، برای جستجوی کارآمد است. الگوریتم‌های زیادی برای نمایه‌سازی اطلاعات متنی وجود دارد. مثال ۲-۱ و ۲-۲ نشان می‌دهد که کاوشگر گوگل و آلتاویستا برای نمایه‌سازی تصاویر، از ترکیبی از روش متنی و محتوایی استفاده می‌کنند. کاربر، برای جستجو در وب به دنبال یک تصویر، برپایه پرسش متنی‌ست. آمار داده‌های چندرسانه‌ای در اینترنت نشان می‌دهد که میزان ظهور تصاویر ۷ برابر صوت و ۳/۵ برابر تصاویر متحرک است. وقتی تصاویر براساس هیستوگرام رنگ نمایه‌سازی شوند، نتیجه برای یک تصویر، قطعی[۲۳] و غیر قابل انعطاف است (۶).

منطق فازی

منطق فازی (منطق غیردقیق)، در مقابل منطق قطعی مطرح می‌شود. در منطق فازی، عدم قطعیت دنیای طبیعی شبیه‌سازی می‌شود. این شبیه‌سازی به کمک مجموعه‌های ریاضی که دارای هم‌پوشانی هستند، الگو می‌شود. برای مثال برای مدلسازی فازی متغیر سن، از سه مجموعه پیر، میانسال و جوان استفاده می‌شود. این سه مجموعه با هم همپوشانی دارند. برای مثال از ۱۰ الی ۲۰ سال جوان، ۱۸ الی ۴۰ میانسال و از ۳۵ الی ۶۰ سال پیر در نظر گرفته می‌شود. در این صورت فرد ۳۰ ساله دارای درجه عضویت در هرسه مجموعه است. یا به عبارت دیگر در هر سه مجموعه پیر، جوان و میانسال عضواست. بنابراین نسبیت در این مدل به‌خوبی پیاده‌سازی می‌شود. شکل ۳-۱ نمودار مجموعه‌های فازی برای مقایسه سن را نشان می‌دهد. طبق شکل ۳-۱ خواهیم داشت:

درجه عضویت فرد ۳۰ ساله در مجموعه جوان = ۰/۲

درجه عضویت فرد ۳۰ ساله در مجموعه میانسال = ۰/۸

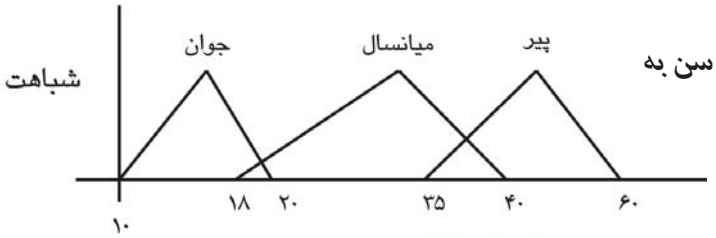
درجه عضویت فرد ۳۰ ساله در مجموعه پیر= ۰/۳

بنابراین غیردقیق و نسبی بودن متغیر سن به‌خوبی در محاسبات دیده می‌شود.

مثال ۲-۲. موتور جستجوی تصاویر آلتاویستا

23.Crisp
24.Discrete

برای مقایسه تصاویر نیاز به معیاری برای مشخص‌کردن تفاوت‌های تصاویر است. این تابع معمولاً یک تابع گسسته[۲۴] ریاضی‌ست. در نظر گرفتن معیار شباهت به‌صورت قطعی، مناسب نیست. امکانات موجود در منطق فازی، امکان مقایسه غیرقطعی را فراهم می‌کند. هر تصویر دارای یک درجه عضویت یا درجه شباهت در مجموعه جواب‌های یک پرسش است. این معیار برای ارزش‌گذاری تصاویر بازیابی شده استفاده می‌شود. به‌طور کلی باید توجه داشت، شباهت و عدم شباهت،



شکل ۳-۱. نمودار متغیر سن به صورت فازی

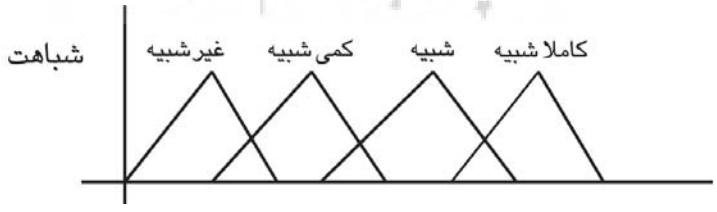
به‌ویژه برای تصاویر که با در نظر گرفتن خصوصیات بصری نمایه می‌شود، مفهوم نسبی‌ست. بنابراین محاسبه آن به‌صورت قطعی با واقعیت منطبق نیست. منطق فازی می‌تواند محاسبه تطبیق رنگی تصاویر را با عدم‌قطعیت محاسبه کند و مقایسه فازی درصد شباهت را در گروه‌هایی که دارای همپوشانی هستند محاسبه می‌نماید (۵).

شکل ۳-۲ نتیجه مقایسه فازی را نشان می‌دهد. این مجموعه‌ها نشان می‌دهند که نتیجه هر مقایسه یک مفهوم نسبی‌ست. برای بیان نسبت شباهت، دامنه شباهت در این مثال به مجموعه‌های غیرمرتبط، نسبتاً شبیه، شبیه و کاملاً شبیه تقسیم شده است. در هر مقایسه جواب به‌صورت یک محدوده، نسبیت را نشان می‌دهد. در حقیقت مزیت استفاده از مقایسه‌ها به‌صورت فازی، امکان محاسبه غیردقیق بودن محاسبه شباهت است. محاسبه غیردقیق بودن محاسبه‌ها به‌صورت مجموعه‌های فازی دارای همپوشانی، موجب می‌شود نتیجه بازیابی با واقعیت بهتر تطبیق داشته باشد، بنابراین بیشتر رضایت کاربر را جلب خواهد کرد.

مجموعه ویژگی‌های تصویر، بردار[۲۵] مشخصه تصویر را می‌سازد. این بردار نمایه تصویر در فهرست فایل نمایه است. در مقایسه فازی می‌توان بردارها را مقایسه کرد. منطق فازی، جواب مقایسه بردارها را به‌صورت غیرقطعی الگو می‌کند.

استخراج مشخصه‌های تصویر، نیازمند محاسبات خاص است. معمولاً در ابتدا، با آشکارسازی لبه‌های تصویر[۲۶]، اندازه، رنگ، شکل و بافت تصویر مشخص می‌شود، سپس بردار مشخصه در فایل نمایه ذخیره می‌شود. برای مدل سازی اشیاء و رابطه میان آنها، اطلاعات دسته‌بندی می‌شوند. دسته‌بندی اطلاعات، بر سرعت و کیفیت بازیابی، تأثیرگذار است. اگر داده‌ها دسته‌بندی شده باشد، جستجو برای تصاویر مرتبط در دسته‌های شبیه انجام می‌شود. جستجوی سلسله‌مراتبی سرعت و کیفیت بهتری نسبت به روش‌های کلاسیک دارد (۵).

در دسته‌بندی فازی، هر عنصر، با یک درصد عضویت،



شکل ۳-۲.نمایش فازی شباهت

25.Vector
26.Edge detection

عضو چند مجموعه است. دسته‌بندی فازی، در مجموعه‌های پیچیده و ناهمگون مانند اینترنت، بهتر از دسته‌بندی قطعی، اطلاعات را دسته‌بندی می‌کند. وجود درجه عضویت عناصر، ارزش‌گذاری نتایج جستجو را تسهیل می‌کند. به‌طور کلی دسته‌بندی فازی، از یک حد آستانه فازی و نه قطعی برای دسته‌بندی اشیاء، استفاده می‌کند. در این دسته‌بندی از مقیاس‌های نسبی که دارای همپوشانی مشابه شکل ۳-۲ است، استفاده می‌شود.

بازخورد ربط

بازخورد ربط براساس میزان مرتبط بودن جواب برای کاربر است. اگر پرسش ترکیبی از روش‌های مختلف نمایه‌سازی تصاویر (متنی، محتوایی و فازی) باشد، میانگین محاسبه می‌شود.

بازخورد، روش اصلی ارتباط نظام بازیابی اطلاعات با کاربر است. بنابراین استفاده از بازخورد برای انتقال دیدگاه‌های کاربران به نظام بازیابی اطلاعات، بسیار مفید است. بسیاری از نظام‌های بازیابی اطلاعات، در چند سطح از بازخورد ربط استفاده می‌کنند. اگر بازخورد علاوه بر تصاویر مرتبط، تفاوت تصاویر بازیابی شده از تصویر دلخواه را نیز بیان کند، دقت[۲۷] بازیابی مراحل بعد بیشتر می‌شود. برای مثال ساختن پرسش، با ترکیب تصاویر دلخواه، کارآیی نظام بازیابی را اضافه می‌کند. برای اعمال بازخورد به نظام موجود، بازیابی فازی، نمایه فازی و دسته‌بندی فازی، می‌تواند با تصحیح فاکتورهای عدم قطعیت، نمایه را بازسازی نماید. نظام بازیابی اطلاعات پس از دریافت بازخورد، بردارهای نمایه تصاویر هم دسته با بازخورد را طوری اصلاح می‌کند که نتایج به‌دست آمده، با نتایج برگرفته از بازخورد یکسان باشد. در این مرحله، کلیه مقایسه‌ها، به‌صورت فازی‌ست (۱؛ ۶).

نتیجه گیری

با توجه به افزایش روزافزون اطلاعات چندرسانه‌ای در وب و اهمیت بازیابی تصاویر در علوم مختلف، کارآیی و دقت نظام‌های بازیابی تصاویر مهم است. در میان روش‌های مختلف نمایه‌سازی و بازیابی اطلاعات، نظام‌های فازی به دلیل تطبیق با عدم قطعیت ذاتی بازیابی اطلاعات، مقایسه، محاسبه شباهت و ربط را بسیار بهتر از نظام‌های قطعی انجام می‌دهند. نظام‌های غیر فازی، مانند نمایه محتوایی یا متنی معمولاً دارای جواب‌های غیرمرتبط زیادی هستند. این مشکل زاده ناهمگون بودن و فضای گسترده اطلاعات موجود در وب است. در میان نظام‌های غیرفازی، نظام‌های ترکیبی بهترین کارآیی را دارند. امروزه حرکت موتورهای جستجوی تجاری به سمت نمایه‌های فازی شروع شده است.

منابع

1. Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier. ***Modern Information Retrieval***. ACM press, 1999.

2. Eakins, John P.; Graham, Margaret E. ***Content-based Image Retrieval A report to the JISC Technology Applications***. Newcastle: Institute for Image Data Research, University of North Umbria, 1999. [on-line]. Available: http// :www.citeseer.psu.edu/context/1537566.

3. Goodrum, Abby A. "Image Information Retrieval: An Overview of Current Research". ***Information science***, Vol. 3, No. 2 (2000).

4. Search engine report number 101, number 104, number 107. [on-line]. Available: http://www.searchenginewatch.com/sereport/article.php/3553646.

5. Tosukhowong, Panrit … [et al]. "Flexible Image Search Engine". [on-line]. Available: http://www.kom.e-technik.tu-darmstadt.de/acmmm99/ ep/andres/#title.

6. Walker, Ellen L. "Image Retrieval on the Internet - How can Fuzzy Help?". [on-line]. Available: http://cs.hiram.edu/~walkerel/papers/nafips 2002.pdf.

7. Ze Wang, James … [et al]. "Content-based image indexing and searching using Daubechies wavelets". ***International Journal of Digital Libraries***, Vol. 1, No. 4 (1997).

27.Precision