مجله علمی مطالعات پژوهشی در علوم کامپیوتر آینده، ۱۴۰۳، دوره ۱، شماره ۱، ص ص ۱-9

مجله علمی مطالعات پژوهشی

در علوم کامپیوتر آینده

www.journalhi.com/com/

ارائه یک روش جدید برای بازیابی تصاویر با استفاده از دسته‌بندی معنایی آن‌ها

نوید آریانا [۱]

۱- کارشناسی ارشد رشتهٔ مهندسی کامپیوتر- گرایش نرم‌افزار، مؤسسه آموزش عالی آفرینش علم گستر بروجرد.

چکیده

با گسترش فناوری‌های دیجیتال در سال‌های اخیر، تعداد تصاویر دیجیتال بطور چشم‌گیری افزایش پیدا کرده است. سهولت استفاده از این فناوری‌ها و در دسترس بودن ابزارهای مؤثر برای ذخیره‌سازی و انتقال تصاویر باعث شده که حجم عظیمی از تصاویر در دسترس عموم قرار گیرد. اما افزایش تعداد تصاویر، مشکلاتی را نیز به همراه داشته است که از مهم‌ترین آن‌ها می‌توان به جستجوی یک تصویر خاص در بین مجموعه‌هایی از تصاویر اشاره کرد. یافتن این تصویر خاص بصورت دستی و با نگاه کردن به تک تک تصاویر، کاری زمانبر، حتی در برخی اوقات غیرممکن است. بنابر این نیاز به یک سیستم بازیابی تصویر که بتواند تصویر مورد نظر را از میان تصاویر آن مجموعه استخراج کند، بیش از پیش احساس می‌شود. مهم‌ترین چالش موجود در این راه، دست‌یابی به روشی کارا برای بررسی میزان شباهت تصاویر به یکدیگر و بازیابی آن‌ها است. امروزه به دلیل افزایش روز افزون اهمیت بازیابی تصاویر دیجیتالی نیاز به سامانه‌ای که روند بازیابی را سریع و دقیق انجام دهد بیشتر احساس می‌شود. لذا وجود سیستم هوشمندی که بتواند این امر را تحقق بخشد ضرورت دارد. در این مقاله روشی جدید برای بازیابی تصاویر با استفاده از دسته‌بندی معنایی آن‌ها ارائه شده است. در روش ارائه شده ویژگی‌های تصاویر با استفاده از توصیفگرهای SIFT[۱] (ویژگی‌های مقاوم در برابر تغییر مقیاس) و LBP[۲] (الگوی دودویی محلی) در قالب روش بستهٔ واژگان استخراج شد. در نهایت بررسی‌های مختلف منجر به ارائهٔ روشی جدید با ترکیب توصیفگرهای SIFT و LBP در قالب بستهٔ واژگان تصویر برای بازیابی تصاویر شد. کارایی روش پیشنهادی با استفاده از دسته‌بند SVM[3] (ماشین بردار پشتیبان) غیرخطی با کرنر گوسی روی پایگاه دادهٔ معروف OT که شامل ۲۶۸۸ تصویر در ۸ دستهٔ مختلف می‌باشد، اندازه‌گیری شد. نتایج نشان داد که روش ارائه شده کارایی بالایی در بازیابی تصاویر با استفاده از دسته‌بندی معنایی آن‌ها داشته است.

کلمات کلیدی: بازیابی تصاویر، دسته‌بندی معنایی، LBP، SIFT، BOW و دسته‌بند SVM.

Presenting a new method for retrieving images using their semantic classification

Navid Ariana [1]

1- Department of Cmputer Engineering،Faculty of Computer Engineering ،INstitute of Higher Education Afarinesh Elmgostar Borojerd.

Abstract

With the growth and proliferation of digital technologies in recent years, the number of digital images has dramatically increased. ease of use of these technologies and the availability of effective tools for storing and transferring images has caused huge volumes of images to be made available to the general public. but the increasing number of images has caused problems, including the most important of which can be cited as a special image among the sums of images. Finding this particular image manual and looking at individual images is time - consuming, boring, and even impossible in some cases. Therefore, the need for an image recovery system that can extract desired image from the images is more than ever. the most important challenge in this approach is achieving an efficient way to investigate the similarity of images to each other and retrieve them. today, due to increasing importance of retrieving digital images, the need for a system that will perform the recovery process faster and more accurately. so the existence of an intelligent system that can achieve this will be necessary. in this thesis, a new method for retrieving images using their semantic classification has been presented. in the presented method, the image features were features against scale change)and LBP (local binary pattern)in terms of word packet method. finally, (Resistant extracted using Descriptors different investigations led to the creation of a novel method by combining Descriptors and LBP in a closed form of image vocabulary for image retrieval. the performance of the proposed method was measured using the svm classifier (support vector machine)with gaussian (support vector machine)on the famous OT database which consists of 2688 images in 8 different categories. the results showed that the proposed method has a high performance in retrieval of images using their semantic classification.

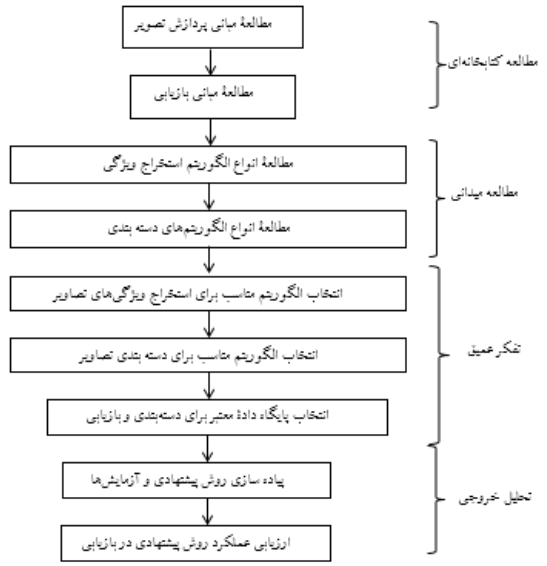
Keywords: SVM classification, BOW, SIFT, LBP, OT database

[1] Scale Invariant Feature Transform (SIFT)
[2] Local Binary Pattern (LBP)
[3] Support Vector Machine (SVM)

۱- مقدمه

امروزه مجموعه‌های عظیمی از تصاویر مختلف به‌صورت منابع گوناگون چون صفحات وب، آرشیوهای خانوادگی و... در دسترس است که چالش‌های فنّی بسیاری را در سامانه‌های رایانه‌ای برای انجام پردازش‌های گوناگون بوجود آورده است. جست‌وجوی تصاویر دیجیتال در پایگاه داده‌های بسیار بزرگ یکی از مهم‌ترین این چالش‌هاست که در دهه‌های اخیر توجه بسیاری از دانشمندان علوم کامپیوتر را به خود معطوف کرده است. در روش‌های قدیمی برای جست‌وجوی تصاویر از برچسب‌گذاری [1] استفاده می‌شد بدین ترتیب که تصاویر در یک پایگاه داده توسط انسان برچسب‌گذاری و سپس جست‌وجو براساس این کلمات کلیدی صورت می‌گرفت (Chan and Hsu 1992). این روش بسیار وقت‌گیر و ناکارآمد بود. بنابراین پژوهشگران به فکر پیداکردن راه‌های مؤثرتری برای یافتن تصاویر مشابه در یک پایگاه داده افتادند. آنان تلاش می‌کنند جست‌وجویی مؤثر و مبتنی بر محتوای دلخواه کاربر انجام دهند. پس سامانه‌ای‌که ورودی آن یک تصویر و خروجی‌اش تصویر یا تصاویری مشابه با این ورودی باشد باید این قابلیت را داشته باشد که بتواند میزان تشابه بین تصویر ورودی و تصاویری را که قرار است به‌عنوان خروجی تولید شوند به‌طور مؤثر اندازه‌گیری کند. اگر بتوان میزان شباهت بین اشیا را اندازه‌گیری کرد، می‌توان یک شیء را از شیء دیگر متمایز ساخت، می‌توان اشیا را بر اساس میزان تشابه (یا عدم تشابه) دسته‌بندی کرد، سپس می‌توان رفتار این گروه‌ها را تجزیه و تحلیل کرده، به نتایج مفیدی رسید طوری که بتوان رفتار این گروه‌ها را بر اساس الگویی خاص توضیح داد. می‌توان اطلاعات را بسیار کاراتر ذخیره، جست‌وجو و بازیابی کرد طوری که بتوان شیء جدید را به راحتی در گروه مربوط به خود قرار داد و حتی بتوان رفتار اشیای جدید را پیش‌بینی کرد. در نهایت می‌توان با توجه به صفات، ویژگی‌ها و رفتار این گروه‌ها پردازش‌های مورد نظر را مؤثرتر انجام داد.

شکل ۱- فرایند اجرای تحقیق



۲- تعاریف

بازیابی تصویر: بازیابی تصویر به معنای کشف و استخراج تصاویر مشابه با تصویر وارد شده می‌باشد.

LBP: به‌عنوان روشی جدید برای استخراج ویژگی‌های الگوی تصویر به منظور توصیف بافت معرفی شد. با انجام عملیاتی روی یک پیکسل و همسایگی آن ویژگی‌هایی به‌صورت هیستوگرام به‌دست می‌دهد، و یک رشتهٔ دودویی تولید می‌کند.

SIFT: روشی برای توصیف تصویر با استفاده از نقاط کلیدی و ویژگی‌های استخراج شده از این نقاط است. با استفاده از این روش می‌توان اشیای مشابه را بازشناسی کرد. از مزایای این روش می‌توان به نامتغیربودن نسبت به چرخش، تغییر مقیاس، تغییرات روشنایی تصویر و اعوجاج اشاره کرد. به این ترتیب که اگر شیء در تصویری وجود داشته باشد با کمک روش SIFT می‌توان همان شیء را اگر چرخیده، تغییر مقیاس داده شده، جابه‌جا شده و یا کج و معوج شده باشد، بازشناسی‌کرد. ویژگی‌های استخراج شده با استفاده از SIFT بسیار از یکدیگر متمایز هستند که در یافتن تطابق‌ها در تصاویر بسیار حائز اهمیت است.

BOW: بدون تردید امروزه این روش یکی از کاراترین روش‌ها برای بازیابی معنایی تصاویر است. ابتدا تصویر به قطعاتی تقسیم و سپس ویژگی‌های مختلف (SIFT، LBP و...) از هر یک از این قطعات استخراج می‌شود. هر یک از این ویژگی‌ها یک «واژه» (لغت) نام دارد. در گام بعدی برای کاهش تعداد واژه‌ها یک روش خوشه‌بندی روی همهٔ واژگان استخراج شده از همهٔ تصاویر اعمال می‌شود. مراکز خوشه‌ها جداسازی شده و «دایرهٔ لغت/ لغت‌نامه/ دایرهٔ واژگان» را تشکیل می‌دهد. سپس برای توصیف هر تصویر «هیستوگرامی» روی دایرهٔ واژگان برای هر تصویر محاسبه و به‌عنوان بردار ویژگی تصویر معرفی می‌شود.

SVM: از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی پرسپترون نشان داده است که در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب شود که حاشیه اطمینان بیشتری داشته باشد. مبنای کاری دسته‌بند SVM دسته‌بندی خطی داده‌ها است. بردارهای پشتیبان به زبان ساده، مجموعه‌ای از نقاط در فضای n بعدی داده‌ها هستند که مرز دسته‌ها را مشخص می‌کنند و مرزبندی و دسته‌بندی داده‌ها براساس آن‌ها انجام می‌شود و با جابه‌جایی یکی از آن‌ها، خروجی دسته‌بندی ممکن است تغییر کند. SVM، یک دسته‌بند یا مرزی است که با معیار قرار دادن بردارهای پشتیبان، بهترین دسته‌بندی و تفکیک بین داده‌ها را برای ما مشخص می‌کند.

پردازش تصویر دیجیتال به مطالعه نظریه‌ها، مدل‌ها و الگوریتم‌ها برای دست‌کاری تصاویر (معمولاً از طریق کامپیوتر) می‌پردازد. ورودی یک تصویر عکس یا صحنه‌ای از یک فیلم است. خروجی پردازشگر تصویر می‌تواند تصویر یا یک مجموعه از نشانه‌های ویژه یا متغیرهای مربوط به تصویر باشد. پردازش تصاویر امروزه بیشتر به موضوع پردازش تصویر دیجیتال گفته می‌شود که شاخه‌ای از دانش رایانه است که با پردازش سیگنال دیجیتال که نماینده تصاویر برداشته شده با دوربین دیجیتال یا پویش شده توسط پویشگر هستند سر و کار دارد. پردازش تصویر شامل سه مرحله اصلی است که شامل: ۱- گرفتن تصویر با اسکنرهای نوری یا با دوربین‌ها و حسگرهای دیجیتال. ۲- تجزیه و تحلیل تصویر که شامل فشرده‌سازی اطلاعات، بهبود تصویر، بازسازی، پردازش مورفولوژیک، استخراج یک بخش مورد نظر از تصویر،

[1] labeling

تشخیص الگوها و استخراج ویژگی و.... ۳- آخرین مرحله خروجی است که می‌تواند تصویر یا گزارش باشد که از نتیجه تجزیه و تحلیل تصویر حاصل شده است. تصاویر آنالوگ دریافتی از ورودی به تصاویر دیجیتال تبدیل می-شوند. تصاویر ورودی دو نوع خاکستری و رنگی می‌باشد. برای تبدیل تصویر آن‌ها به‌صورت ماتریس دوبعدی در نظر گرفته می‌شوند. اگر تصویر خاکستری باشد هر پیکسل آن یک درایه از ماتریس دو بعدی می‌شود و مقدار هر درایه را با f (x,y) نشان می‌دهند. اگر تصویر رنگی باشد هر پیکسل آن یک درایه از سه ماتریس دوبعدی است که یک ماتریس برای رنگ قرمز، یکی برای رنگ سبز و آخر برای رنگ آبی در نظر گرفته می‌شود. ترکیب درایه‌های متناظر این سه ماتریس تشکیل‌دهنده پیکسل‌های نهایی هستند. تصاویری که توسط نویزها آسیب دیده، یا تخریب شده بازسازی می‌شوند. در این مرحله تصاویری که به‌طور ناخواسته روی تصویر اصلی افتاده‌اند و جزء تصویر نمی-باشند را حذف و انحنا و پیچیدگی‌های هندسی ناخواسته اصلاح می‌شوند. در این مرحله حرکت تصویر که ناشی از حرکت دوربین هنگام تصویربرداری است نیز اصلاح می‌شود. فشرده‌سازی تصویر رنگی، تشخیص لبه رنگ و نیز تبدیل تصاویر رنگی به خاکستری از جمله کارهایی است که روی تصاویر رنگی انجام می‌شود. کاربردهای پردازش تصویر در تشخیص الگو[1]، تشخیص حرکت[2]، کاربردهای نظامی، سیستم‌های امنیتی و تشخیص چهره، کاربردهای صنعتی، سیستم‌های سنجش از دور، انیمیشن کامپیوتری و کاربرد پزشکی است.

استخراج ویژگی[3] فرایندی است که در آن با انجام عملیاتی بر روی تصاویر، ویژگی‌های بارز و تعیین‌کنندهٔ آن مشخص می‌شود. هدف آن حذف نویز، جداسازی اجزای مستقل تصویر و بازیابی تصاویر است. در روش اول با استفاده از تشخیص‌گرهای نقطهٔ مورد علاقه[4] نقاط مهم تصویر مانند گوشه‌ها، کیمنه‌های سراسری و محلی و... تشخیص داده می‌شوند. نقاط تشخیص داده شده، با استفاده از توصیفگرهای[5] مختلف، توصیف می‌شوند یعنی ویژگی‌هایشان استخراج می‌شود. در روش دوم برخلاف روش اول، نقاط مهم تصویر استخراج نمی‌شوند و از همان ابتدا از توصیفگرهای تصویر برای استخراج ویژگی‌ها استفاده می‌شود. توصیفگرهای تصویر برای استخراج ویژگی‌های تصویر و بیان عددی ویژگی‌ها به‌کار می‌روند. برای توصیف یک تصویر روش‌های خاصی وجود دارد. برخی از این روش‌ها تصویر را مجموعه‌ای متشکل از تعدادی نواحی مهم، برخی تصویر را به‌صورت مجموعه‌ای متشکل از اشیای مهم و برخی نیز تصویر را به‌صورت مجموعه‌ای از نقاط مهم (نقاط کلیدی) توصیف می‌کنند. از توصیفگرهای SIFT، LBP و BOW در انجام آزمایش‌های مختلف استفاده شده است. . برای استخراج ویژگی‌ها با استفاده از روش SIFT چهارگام اصلی اعمال می‌شود:

الف. **تشخیص نقاط اکسترمم فضای مقیاس**: در این مرحله تصویر با توابع گوسی مختلف فیلتر می‌شود در نتیجه نقاط کلیدی تصویر که نسبت به تغییر مقیاس نامتغیر هستند، استخراج می‌شوند.

(۱)

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y)$$
$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$
$$D(x,y,\sigma) = \big(G(x,y,k\sigma) - G(x,y,\sigma)\big) * I(x,y)$$
$$= L(x,y,k\sigma) - L(x,y,\sigma)$$

ب. **تعیین نقاط کلیدی**: پس از انتخاب تعدادی نقطه به‌عنوان نقاط نامزد که نقاط کلیدی و با ارزش تصویر محسوب می‌شوند، نقطه‌ای مانند X از این مجموعه نقاط نامزد شده، در صورتی انتخاب می‌شود که این نقطه بین ۲۶ همسایه‌اش بیشینه یا کمینه باشد.

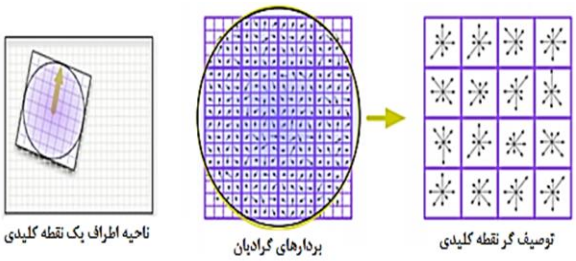
ج. **تخصیص جهت**: اکنون تعدادی نقاط مناسب در اختیار است. در این مرحله هدف اختصاص جهت یا جهت‌هایی به نقاط انتخاب شده با استفاده از بردار گرادیان محلی تصویر است. از نتایج این مرحله توصیف نقاط به قسمی است که تحت تغییر مقیاس و چرخش، بدون تغییر باشند.

(۲)

$$m(x,y) = \sqrt{\big(L(x+1,y) - L(x-1,y)\big)^2 + \big(L(x,y+1) - L(x,y-1)\big)^2}$$
$$\theta(x,y) = \tan^{-1}\big(L(x,y+1) - L(x,y-1) / L(x+1,y) - L(x-1,y)\big)$$

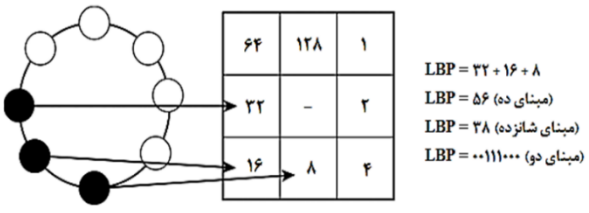
د. **توصیف نقاط کلیدی**: به یک توصیفگر محلی در اطراف نقاط کلیدی برای توصیف آن نقطه نیاز است. به همین منظور هیستوگرام گرادیان روی پنجره‌ای به ابعاد ۴×۴ و در ۸ جهت محاسبه می‌شود. بنابراین بردار ویژگی با ۴×۴×۸=۱۲۸ بُعد برای هر نقطه ساخته می‌شود (شکل ۲).

شکل ۲- تشکیل توصیفگر نقطه کلیدی



در نسخهٔ اصلی LBP از یک همسایگی ۳×۳ در اطراف پیکسل موردنظر استفاده می‌شود. روش کار به این صورت است‌که پیکسل مرکزی با همسایه‌هایش مقایسه می‌شود اگر از همسایه‌ای بزرگتر بود، به‌جای آن عدد ۱ و در غیراین‌صورت به‌جای همسایه عدد ۰ قرار داده می‌شود. به این ترتیب برای نمایش یک همسایگی، یک الگوی دودویی تولید می‌شود. از آنجا که در یک همسایگی ۳×۳، پیکسل مرکزی با هشت پیکسل دیگر همسایه است، ۲۵۶= 2^8 الگوی دودویی مختلف با توجه به مقادیر پیکسل‌ها متصوراست (گیوکی ۱۳۹۵). برای محاسبهٔ عدد دودویی متناظر با الگوی دودویی تولید شده باید وزن هر یک از همسایه‌ها در مبنای ۲ مشخص گردد (شکل ۳).

شکل ۳- تعیین وزن‌های همسایه‌ها در محاسبهٔ عدد دسیمال متناظر با الگوی تولید شده



[1] Pattern Recognition

[2] Motion Tracking
[3] Feature extraction
[4] Interst point detectors
[5] Descriptor

(۳)

$$LBP_{R,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(g_p - g_c)2^p$$

برای نمایش یک تصویر با استفاده از روش بسته واژگان (BOW) به‌طور معمول سه مرحلهٔ زیر در نظر گرفته می‌شود (Fei-fei and Perona 2005).

۱- تشخیص ویژگی

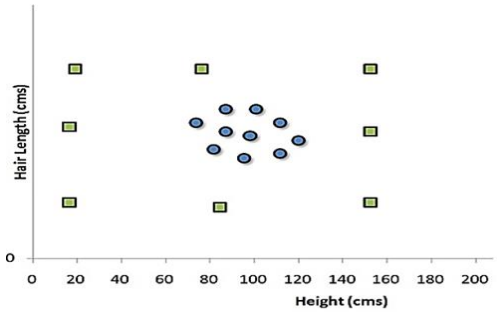
۲- توصیف ویژگی

۳- تولید دایرهٔ واژگان

یکی از مؤثرترین روش‌ها برای تولید دایرهٔ واژگان، خوشه‌بندی واژگان به روش K-means است. پس از انجام خوشه‌بندی می‌توان کدواژه‌ها را به‌عنوان مراکز خوشه‌ها در نظر گرفت.

در SVM فقط داده‌های قرار گرفته در بردارهای پشتیبان مبنای یادگیری ماشین و ساخت مدل قرار می‌گیرند و این الگوریتم به سایر نقاط داده حساس نیست و هدف آن هم یافتن بهترین مرز در بین داده‌هاست به گونه‌ای که بیشترین فاصله ممکن را از تمام دسته‌ها (بردارهای پشتیبان آن‌ها) داشته باشد. اگر داده‌ها به‌صورت خطی قابل تفکیک باشند، الگوریتم بردار پشتیبان می‌تواند بهترین ماشین را برای تفکیک داده‌ها و تعیین دسته یک رکورد داده، ایجاد کند اما اگر داده‌ها به‌صورت خطی توزیع شده باشند همانند (شکل ۴) آن‌ها را تفکیک می‌کند (Wu et al. 2009).

شکل ۴- تفکیک داده‌ها به‌صورت غیرخطی با استفاده از بردارپشتیبان



نقاط ضعف SVM

۱-این نوع الگوریتم‌ها، محدودیت‌های ذاتی دارند مثلاً هنوز مشخص نشده است که به ازای یک تابع نگاشت، پارامترها را چگونه باید تعیین کرد. ۲- ماشین‌های مبتنی بر بردار پشتیبان به محاسبات پیچیده و زمانبر نیاز دارند و به دلیل پیچیدگی محاسباتی، حافظه زیادی نیز مصرف می‌کنند. ۳- داده‌های گسسته و غیرعددی هم با این روش سازگار نیستند و باید تبدیل شوند.

رنگ یکی از مهم‌ترین و پراستفاده‌ترین ویژگی‌های تصویر در سیستم‌های بازیابی تصویراست. تحقیقات در این زمینه را می‌توان در یکی از سه گروه زیر جای داد: ۱- استفاده از فضای رنگی مناسب بر اساس هدفی که از قبل برای کاربردی خاص تعیین شده‌است. ۲- معرفی الگوریتم‌های مناسب برای استخراج ویژگی رنگ. ۳- مطالعه ارزیابی تشابه‌سنج‌های مختلف مبتنی بر رنگ. می‌توان اطلاعات رنگ را در فضاهای رنگی سه بعدی (مانند YIQ, RGB, HSV و...) پردازش کرد (Bimbo 1999). ممان‌های رنگ اطلاعات آماری مفید و مهمی در ارتباط با چگونگی توزیع رنگ‌ها در تصویر به‌دست می‌دهند.

در روش ممان‌های رنگ معمولا میانگین (دادهٔ آماری مرتبهٔ اول μ_i)، واریانس (دادهٔ آماری مرتبهٔ دوم σ_i) و چولگی[1] (دادهٔ آماری مرتبهٔ سوم S_i) برای استخراج بردار ویژگی مورد استفاده قرار می‌گیرند. این ممان‌ها به ترتیب به‌صورت زیر تعریف می‌شوند (Stricker and Orengo 1995):

(۴)

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij}$$

$$\sigma_i = \frac{1}{N}\sum_{j=1}^{N}((f_{ij} - \mu_i)^2)^{\frac{1}{2}}$$

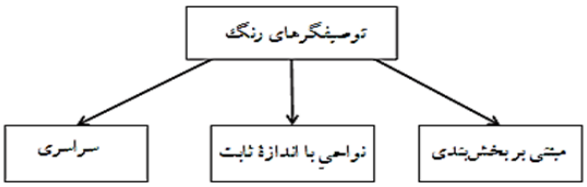
$$s_i = \frac{1}{N}\sum_{j=1}^{N}((f_{ij} - \mu_i)^3)^{\frac{1}{3}}$$

هیستوگرام گسستهٔ رنگ پر استفاده‌ترین توصیف‌گر رنگ در بازیابی تصاویر است. هیستوگرام یعنی شمارش تعداد پیکسل‌ها بر اساس تعداد رنگ‌ها در هر مؤلفهٔ رنگی. در روش بردار انسجام رنگ ابتدا هیستوگرام رنگ ساخته می‌شود و سپس هر بازهٔ هیستوگرام (بین) به دو قسمتِ «منسجم[2]» و «نامنسجم[3]» افراز می‌شود. قسمت منسجم قسمتی است که پیکسل‌های آن مربوط به ناحیهٔ نسبتا وسیع و یکدست/همگن از تصویر با رنگ‌های یکنواخت است، در غیر این صورت آن قسمت نامنسجم است. کرلوگرام رنگ روشی است که علاوه بر مشخص‌کردن توزیع رنگ‌ها قابلیت مشخص‌کردن همبستگی مکانی آن‌ها را نیز دارد.

(۵)

$$\gamma_{i,j}^{(k)} = \Pr[p_2 \in I_{c(j)} \mid |p_1 - p_2| = k]$$
$$p_1 \in I_{c(i)}, p_2 \in I_{c(j)}$$

شکل ۵- دسته‌بندی روش‌های مبتنی بر رنگ

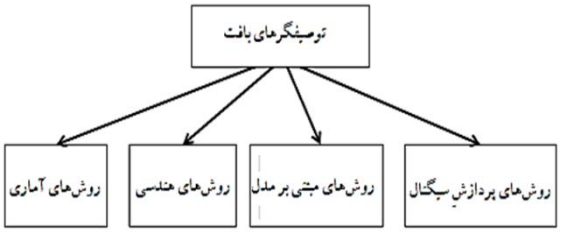


تامورا و همکارانش با توجه به مطالعات روانشناسی‌خود، برای یک بافت شش ویژگیِ «زبری[4]»، «تباین/تضاد[5]»، «جهت[6]»، «شبیه خط بودن[7]»، «منظمی[8]» و «سختی[9]» را مطرح کردند (Tamura, Mori and Yamawaki 1978). سه ویژگی اول در بازیابی تصاویر بسیار مؤثر هستند. سختی مربوط به فاصلهٔ تغییرات قابل ملاحظه در سطح خاکستریِ تصویر است. به بیان بهتر سختی اندازهٔ عناصر تشکیل دهندهٔ بافت (تکسل‌ها)[10] را معیّن می‌کند. برای محاسبهٔ زبری متوسط سطح خاکستری تصویر با استفاده از شش پنجرهٔ

[1] skewness
[2] coherent
[3] Incoherent
[4] coarseness
[5] contrast
[6] Directionality
[7] Linelikeness
[8] Regularity
[9] Roughness
[10] Texels

متحرک با اندازه‌های مختلف‌که در سرتاسر تصویر در حال حرکت هستند و با هم همپوشانی ندارند، محاسبه می‌شود.

شکل ۶- دسته بندیِ روش‌های مبتنی بر بافت



پیشینه تحقیق

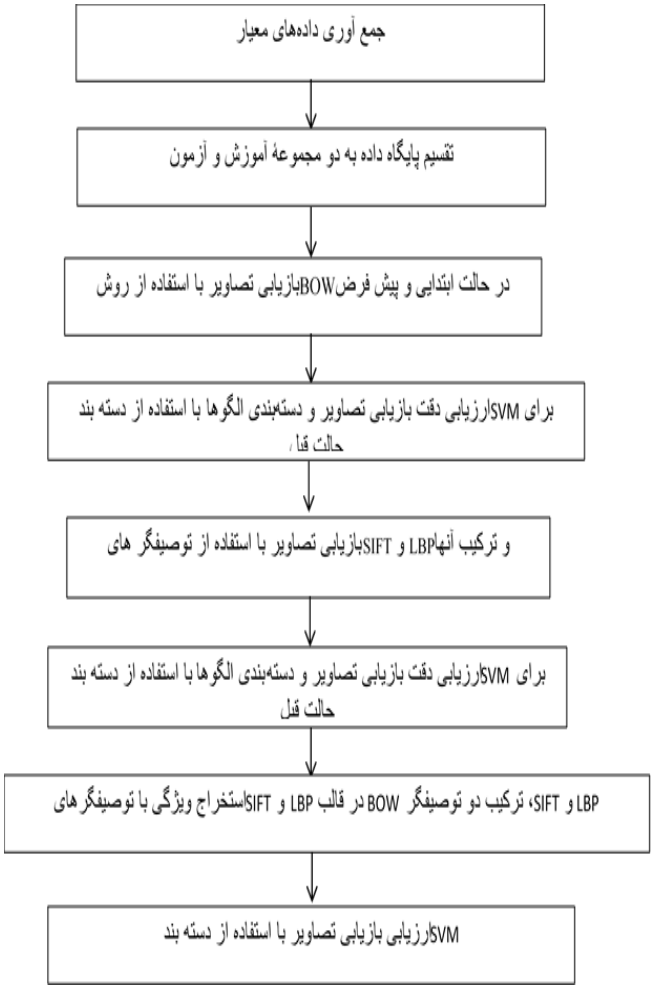
دیگری کریشنان و همکارانش (Krishnan, Banu and Christiyana 2007) با استفاده از مفهومِ رنگ چیره توانستند میزان بازیابی تصاویر با استفاده از روش هیستوگرام رنگ را افزایش دهند. رنگ چیره رنگی است که مربوط به یک ناحیهٔ یک رنگ و یکدست از تصویر است. مشکل استفاده از رنگ چیره این است که اگر تصاویری از اشیایی مشابه اما با اندازه‌های متفاوت تشکیل شده باشند، این روش این تصاویر را مشابه با یکدیگر در نظر نمی‌گیرد. برای رفع این مشکل کریشنان و همکارانش مفهوم رنگ چیره برای اشیای «پیش‌زمینه» را معرفی می‌کنند. ایدهٔ آن‌ها این است که پس‌زمینهٔ تصویر در مقایسه با پیش‌زمینهٔ آن از اهمیت کمتری برخوردار است. در این روش ابتدا تصویر رنگی به تصویر خاکستری و سپس تصویر خاکستری به تصویر دودویی تبدیل می‌شود طوری که در آن سفید بیانگر شیء اصلی موجود در تصویر و سیاه رنگ پس‌زمینه است. پس از تشخیص شیء پیش‌زمینه، تنها رنگ این شیء از سفید به رنگی بازیابی می‌شود.

بوش و همکارانش در (Buch, Jain and Pradhan 2011) برای بازیابی تصاویر ماهواره‌ای روش جدیدی را مبتنی بر ویژگی‌های رنگ و بافت ارایه کردند. آن‌ها از ماتریس هم‌رخدادی و فضای رنگی HSV برای ساختن بردارهای ویژگی استفاده کردند. روش کار به این صورت است که ابتدا تصویر رنگی RGB به تصویر خاکستری تبدیل می‌شود سپس برای هر تصویر خاکستری ویژگی‌های بافت توسط ماتریس هم‌رخدادی استخراج می‌شود. سپس تصویر RGB را به HSV تبدیل و ویژگی‌های رنگ از آن استخراج می‌شود. استفاده از ترکیب دو ویژگی رنگ و بافت برای تصاویر ماهواره‌ای به‌طور محسوسی باعث افزایش کارایی سیستم بازیابی تصاویر می‌شود. محدودیت این روش فضای جست‌وجوی تصاویر است.

فولکرس و سامت (Folkers and Samet 2002) روشی مؤثر را برای بازیابی تصاویر لگو با استفاده از توصیفگرهای فوریه معرفی کردند. در این سیستم ابتدا کاربر تصویر یا طرحی را که مشابه با لگوی مورد نظرش است ترسیم می‌کند سپس سیستم تصاویر مشابه با این شکل را بازیابی می‌کند. تصاویر ترسیم شده توسط کاربر به شکل‌هایی چون دایره- بیضی- مستطیل و... محدود می‌شود. اما مزیت این روش استفاده از توصیفگرهای فوریه است که نسبت به انتقال- چرخش- تغییر مقیاس و نقطهٔ شروع نامتغییر هستند.

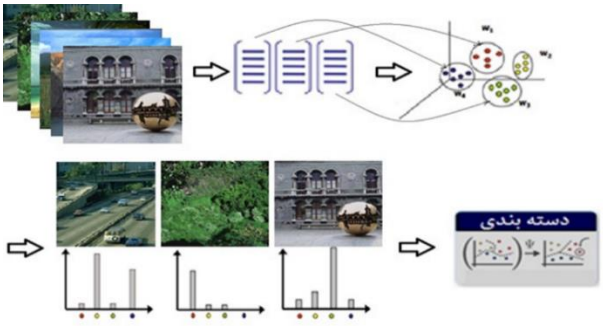
۳- روش انجام کار

شکل ۷- روند نمای انجام کار



ویژگی‌های هر تصویر با استفاده از توصیفگرهای SIFT و LBP استخراج می‌گردد. ویژگی‌های استخراج شده در قالب BOW برای ساخته شدن بردار ویژگی نهایی که همان هیستوگرام واژگان است، استفاده می‌شود. در نهایت بردارهای ویژگی استخراج شده با استفاده از SVM غیرخطی[1] با کرنر گوسی دسته‌بندی می‌شوند (شکل ۸).

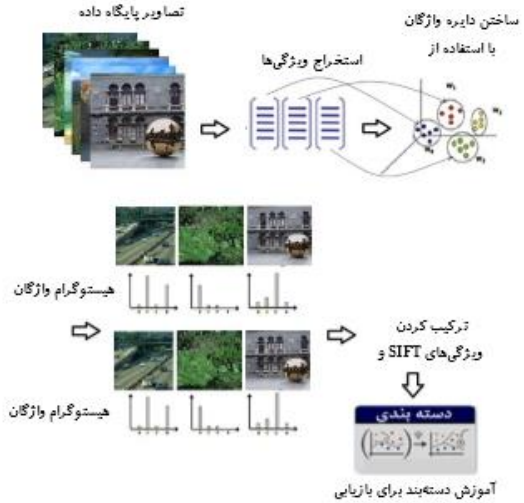
شکل ۸- مراحل بازیابی تصاویر با استفاده از توصیفگرهای SIFT و LBP در قالب روش BOW



[1] Non-Liner Support Vector Machine

مراحل انجام روش پیشنهادی به این صورت است که ابتدا ویژگی‌های هر تصویر با استفاده از توصیفگرهای SIFT و LBP استخراج می‌گردد. ویژگی‌های استخراج شده در قالب BOW برای ساخته شدن بردار ویژگی نهایی که همان هیستوگرام واژگان است، استفاده می‌شود. پس از استخراج هر یک از ویژگی‌های SIFT و LBP، بردار ویژگی نهایی با ترکیب[1] این بردارهای ویژگی ساخته می‌شود. در نهایت بردارهای ویژگی استخراج شده با استفاده از SVM غیرخطی با کرنر گوسی دسته‌بندی می‌شوند (شکل ۹).

شکل ۹- مراحل بازیابی تصاویر با استفاده از روش پیشنهادی

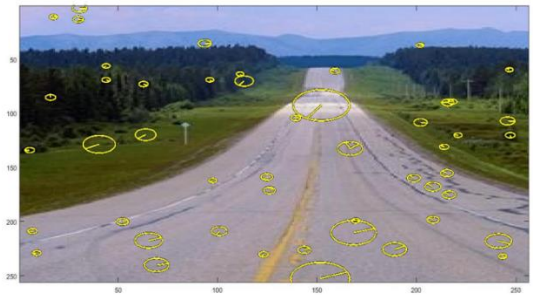


در مرحله استخراج ویژگی، ویژگی‌های تصاویر پیش پردازش شده با استفاده از توصیفگر SIFT و LBP استخراج می‌شود.

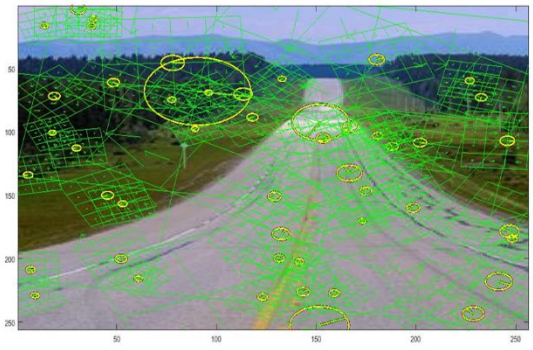
شکل ۱۰- شکل اصلی

در مرحلهٔ بعد نقاط کلیدی تصاویر با استفاده از تشخیص نقاط کمینه و بیشینه و هیستوگرام تصویر کشف و استخراج می‌شود. در نهایت نقاط کلیدی که یا به اشتباه تشخیص داده شده‌اند یا از درجهٔ ارزش پایینی برخوردار هستند حذف می‌شوند.

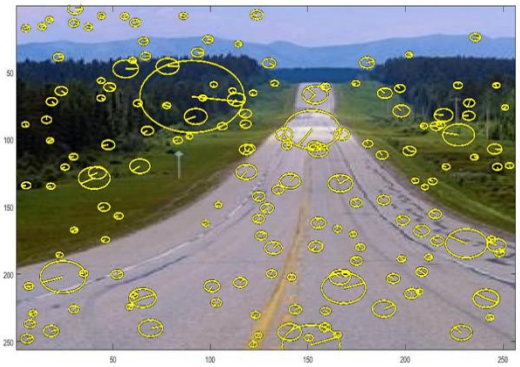
شکل ۱۰- ۵۰ نقطهٔ کلیدی استخراج شده توسط توصیفگر SIFT از تصویر پیش پردازش شده



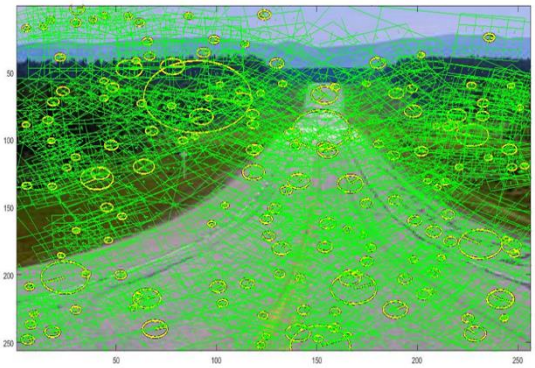
شکل ۱۱- تصویر پیش پردازش شدهٔ پایگاه داده پس از اعمال الگوریتم SIFT با ۵۰ نقطهٔ کلیدی



شکل ۱۲- تمامی نقاط کلیدی استخراج شده توسط توصیفگر SIFT از تصویر پیش پردازش شده



شکل ۱۳- تصویر پیش پردازش شدهٔ پایگاه داده پس از اعمال الگوریتم SIFT با تمامی نقاط کلیدی



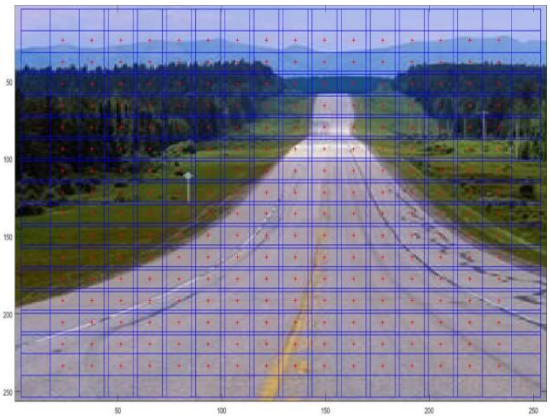
که از شکل ۱۲ برداشت تمامی نقاط کلیدی در تصویرکشف و استخراج ویژگی از تمامی این نقاط در شکل ۱۳ انجام شده است.

در مرحلهٔ بعد الگوریتم LBP را روی تصاویر اصلی اعمال می‌کنید.

[1] Concatenation

شکل ۱۴- تصویر اصلی

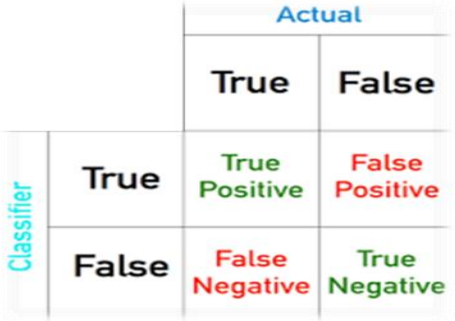
شکل ۱۵- تصویر پایگاه داده پس از اعمال الگوریتم LBP



معیارهای آماری که در همهٔ دنیا در ارزیابی عملکرد سیستم‌های بازشناسی و دسته‌بندی مورد استفاده قرار می‌گیرد نمودار متوسط میانگین دقت [1] و ماتریس درهم ریختگی [2] هستند.

$$ACC. = \frac{TP+TN}{TP+TN+FP+FN} \quad (۶)$$

شکل ۱۶- نمودار درهم ریختگی



فرض کنید P (مثبت) نشان دهندهٔ تایید یک دسته و N (منفی) نشان دهندهٔ رد یک دسته است. در این صورت TP [3] نشان دهندهٔ تعداد نمونه‌هایی است که به درستی تایید شده‌اند. TN [4] نشان دهندهٔ تعداد نمونه‌هایی است که به درستی رد شده‌اند. FP [5] نشان دهندهٔ نمونه‌هایی است که به اشتباه تأیید شده‌اند و FN [6] نشان دهندهٔ نمونه‌هایی است که به اشتباه رد شده‌اند. این مفاهیم برای دسته‌بندی‌های دو دسته‌ای تعریف می‌شوند. برای دسته‌بندی‌های بیشتر از دو دسته یک دسته به‌عنوان مثبت در نظرگرفته شده و سایر دسته‌ها همگی به‌عنوان منفی در نظرگرفته می‌شوند و معیار ماتریس درهم ریختگی و متوسط میانگین دقت برای هر دسته به‌طور جداگانه محاسبه می‌شود.

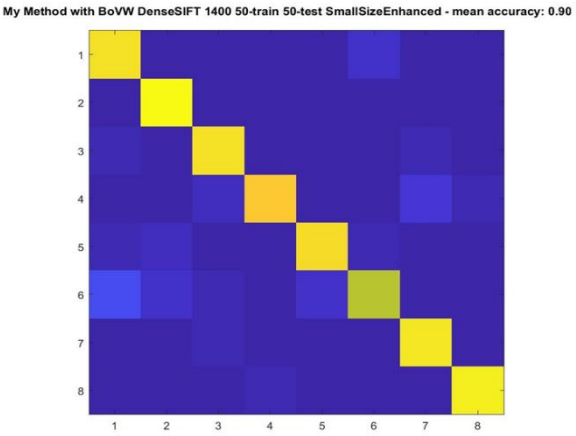
جدول ۱.- پارامترها و نتایج دسته بند SVM

توصیفگر	تعداد تصاویر آموزش	تعداد تصاویر آزمون	تعداد واژگان	درصد دقت آزمون
SIFT	٪۵۰	٪۵۰	۱۰۰۰	٪۸۹
LBP	۵۰ ٪	٪۵۰	۱۰۰۰	٪ ۸۶
SIFT	۵۰ ٪	٪۵۰	۱۱۰۰	٪ ۸۴
LBP	۵۰ ٪	٪ ۵۰	۱۱۰۰	٪ ۸۶
SIFT	۵۰ ٪	٪ ۵۰	۱۲۰۰	٪ ۹۰
LBP	۵۰ ٪	٪ ۵۰	۱۲۰۰	٪ ۸۷
SIFT	۵۰ ٪	٪ ۵۰	۱۳۰۰	٪ ۹۰
LBP	۵۰ ٪	٪ ۵۰	۱۳۰۰	٪ ۸۶
SIFT	۵۰ ٪	٪ ۵۰	۱۴۰۰	٪ ۹۰
LBP	۵۰ ٪	٪ ۵۰	۱۴۰۰	٪ ۸۷
SIFT	۵۰ ٪	٪ ۵۰	۱۵۰۰	٪ ۸۹
LBP	۵۰ ٪	٪ ۵۰	۱۵۰۰	٪ ۸۷

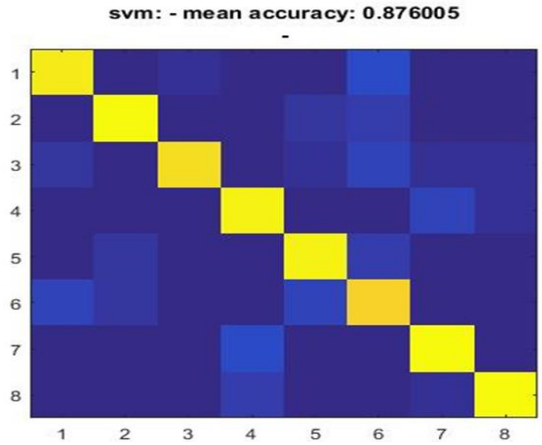
با توجه به جدول ۱ بهترین حالت به‌دست آمده با استفاده از توصیفگر SIFT با ۵۰ درصد تصاویر برای آموزش، ۵۰ درصد تصاویر برای آزمون و طول بردار ویژگی ۱۴۰۰ بوده است که به دقت ۹۰ درصد روی داده‌های آزمون رسیده است. ماتریس درهم ریختگی روش مربوطه در شکل ۱۷ نشان داده شده است. همچنین بهترین دقت به‌دست آمده با استفاده از توصیفگر LBP با ۵۰ دصد تصاویر برای آموزش، ۵۰ درصد تصاویر برای آزمون و طول بردار ویژگی ۱۴۰۰ بوده است که به دقت ۸۷ درصد روی داده‌های آزمون رسیده است. ماتریس درهم ریختگی روش مربوطه در شکل ۱۸ نشان داده شده است.

[1] mean Average Precision (mAP)
[2] Confusion matrix
[3] True Positive
[4] True Negative
[5] False Positive
[6] False Negative

شکل ۱۷-نمودار در هم ریختگی در حالت اولیه روش BOW با توصیفگر SIFT

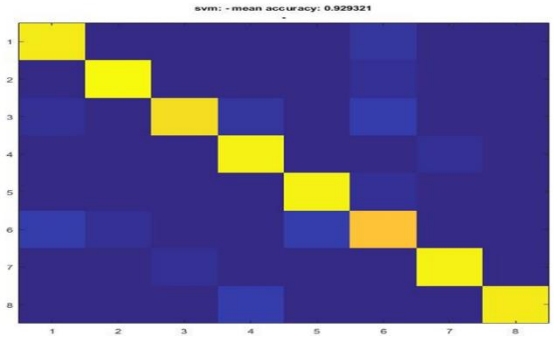


شکل ۱۸- نمودار در هم ریختگی در حالت اولیه روش BOW با توصیفگر LBP

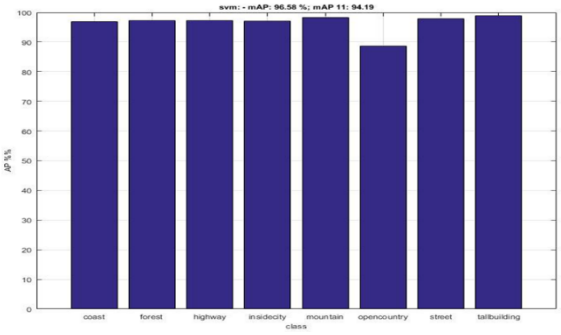


همان‌طور که در شکل‌های ۱۷ و ۱۸ دیده می‌شود قطر اصلی که با مربع‌های کوچک نشان داده شده در واقع شمارهٔ حقیقی دسته‌هایی است که داده‌های آزمون به آن‌ها تعلق دارند. مربع‌هایی که روی قطر اصلی قرار نگرفته‌اند بیانگر داده‌های آزمونی هستند که به نادرستی در دسته‌های دیگر دسته‌بندی شده‌اند. شکل ۱۷ تشان می‌دهد که دسته‌های ۹، ۲۰ و ۳۰ در قطر اصلی قرار نگرفته‌اند و این به معنای آن است که دسته‌بند به‌طور کامل نتوانسته است با استفاده از ویژگی SIFT دسته‌های فوق را به درستی تشخیص دهد. در برخی موارد روش بستهٔ واژگان تصویری به نادرستی عمل می‌کند. برای بهبود عملکرد این روش استفاده از روش پیشنهادی توصیه می‌شود. با استفاده از روش پیشنهادی با ترکیب دو توصیفگر SIFT و LBP دسته‌بندی داده‌ها با دقت بیشتری انجام خواهد شد. در روش پیشنهادی از ترکیب توصیفگرهای SIFT و LBP برای استخراج ویژگی در قالب روش BOW و از SVM با کرنر گوسی برای دسته‌بندی استفاده شده است. همچنین از 50% تصاویر در هر دسته برای آموزش و 50% تصاویر از هر دسته برای آزمون دسته‌بند استفاده شده است. روش ارائه شده با رسیدن به دقت 92% و نرخ خطای 0.8% در دسته‌بندی و بازشناسی تصاویر پایگاه دادهٔ OT، عملکرد بهتری نسبت به حالت اولیهٔ روش بستهٔ واژگان تصویری داشته است. روش ارائه شده طی آزمایش‌های متعدد با تغییر پارامترهای مختلف به نتیجه بهتری با دقت دسته‌بندی 92% و میانگین دقت دسته‌بندی 94% رسیده است. ماتریس در هم ریختگی و نمودار متوسط میانگین دقت روش ارائه شده در شکل ۱۹ و ۲۰ به نشان داده شده است. همچنین جدول ۲ دقت‌های دسته‌بندی به‌دست آمده در حالات مختلف را در قالب روش ارائه شده نشان می‌دهد.

شکل ۱۹- ماتریس درهم ریختگی برای داده‌های آزمون روش ارائه شده



شکل ۲۰- نمودار متوسط میانگین برای داده‌های آزمون روش ارائه شده



جدول ۲- نتایج به‌دست آمده از ترکیب توصیفگرهای SIFT و LBP در قالب روش BOW

توصیفگر	دسته‌بند	تعداد تصاویر آموزش	تعداد تصاویر آزمون	تعداد واژگان برای توصیفگر SIFT	تعداد واژگان برای توصیفگرLBP	درصد دقت آزمون
SIFT + LBP	SVM	٪50	٪50	1000	1000	٪87
SIFT + LBP	SVM	٪50	٪50	1000	1100	٪84
SIFT + LBP	SVM	٪50	٪50	1000	1200	٪88
SIFT + LBP	SVM	٪50	٪50	1000	1300	٪88
SIFT + LBP	SVM	٪50	٪50	1000	1400	٪92
SIFT + LBP	SVM	٪50	٪50	1000	1500	٪90

۴- نتیجه گیری

با گسترش فناوری‌های دیجیتال در سال‌های اخیر، تعداد تصاویر دیجیتال به‌طور چشم‌گیری افزایش پیدا کرده است. سهولت استفاده از این فناوری‌ها و در دسترس بودن ابزارهای مؤثر برای ذخیره‌سازی و انتقال تصاویر باعث شده است که حجم عظیمی از تصاویر در دسترس عموم مردم قرار گیرد. اما افزایش تعداد تصاویر، مشکلاتی را نیز به همراه داشته است که از جمله مهم‌ترین آن‌ها می‌توان به جستجوی یک تصویر خاص در بین مجموع‌های از تصاویر اشاره کرد. یافتن این تصویر خاص به‌صورت دستی و با نگاه کردن به تک تک تصاویر، کاری زمانبر، خسته‌کننده و دربرخی اوقات غیرممکن است. بنابراین نیاز به یک سیستم بازیابی تصویر که بتواند تصویر مورد نظر را از میان تصاویر آن مجموعه استخراج کند، بیش از پیش احساس می‌شود. مهم‌ترین چالش موجود در این راه دست‌یابی به روشی کارا برای بررسی میزان شباهت تصاویر به یکدیگر و بازیابی آن‌ها است تشخیص هویت هوشمند عملی است که امروزه معمولا با

استفاده از پردازش تصویر، شبکه‌های عصبی و دسته‌بندها انجام می‌شود. پردازش تصویر پرکاربردترین ابزار در زمینهٔ بازیابی تصاویر می‌باشد.

نتایج به‌دست آمده به شرح زیر است:

- ✓ از تصاویر پایگاه دادهٔ OT برای دسته‌بندی استفاده شده است و از این جهت یک روش دسته‌بندی بین المللی و کاربردی طراحی شده است.
- ✓ با استفاده از نتایج به‌دست آمده دریافتیم روش BOW قابلیت بالایی در بازیابی تصاویر دارد و در این زمینه بهتر از روش‌های دیگر عمل می‌کند.
- ✓ با استفاده از تحقیقات انجام شده، یک روش BOW جدیدی با استفاده از توصیفگرهای SIFT و LBP و SVM تولید شدکه توانایی بالاتری در دسته‌بندی الگوها نسبت به حالت اولیهٔ آن دارد.
- ✓ نتایج نشان داد که ترکیب ویژگی‌های مهم و با ارزش می‌تواند تأثیر خوبی در بازیابی تصاویر داشته باشد.

پیشنهادات

- ✓ به منظور بهره‌گیری بیشتر و بهتر، نتایج حاصل از روش ارائه شده به‌صورت نرم افزار یا وب گاهی طراحی شود که با دریافت تصاویر از کاربر، نتیجه را به او بگوید.

پیشنهاد می‌شود از روش ارائه شده برای پایگاه داده‌های پیچیده‌تر استفاده شود.

مراجع

- گیوکی، داور، رسالهٔ دکتری، "ارائه الگوریتم هوشمند برای بازیابی محتوا محور تصاویر قابل کاربرد در جویشگرهای تحت وب"، بهمن ۱۳۹۵.
- Bimbo. A. del. 1999. "Visual. Information Retrieval.". Morgan Kaufmann Publishers, San Francisco, CA, USA.
- Buch, A. Jain,S. and Pradhan. S. N. 2011. "Exploring a new direction in colour and texture based satellite image search and retrieval. system". Institute of Technology Nirma University, Ahmed Abad, pp: 1-5.
- Chan, Shi-Kuo. and Hsu. Arding. 1992. ''Image Information Systems: Where do we go from here?'' IEEE Trans. on Knowledge and Data Engineering. 4(5), pp: 431-442.
- Fei-Fei Li; Perona, P. (2005). "A Bayesian Hierarchical. Model for Learning Natural. Scene Categories". 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)2. p. 524.
- Folkers A. and Samet. H. 2002. "Content based image retrieval. using Fourier ddescripotors o a Logo database". Proc of the 16th Int. Conf. on Pattern Recognition, vol. 3, pp. 521-524.
- Krishnan, Dr. Sheerin Banu,M. and Cal.lins Christiyana. C. 2007. "Content Based Image Retrieval. Using Dominant Color Identification Based on Foreground Objects". International. Conference on Computational. Intelligence and Multimedia Applications, vol. 3, pp: 190-194.
- Stricker, M. A. and Orengo. M. 1995, "Similarity of Color Images". In Storage and Retrieval. for Image and Video Databases (SPIE), pp: 381–392.
- Tamura, H. Mori, S. and Yamawaki. T. 1978. "Textural. Features Corresponding to Visual. Perceptron". IEEE Transactions on Systems, Man and Cybernatics, 8(6), pp: 460–473.
- Wu, Y. and Wu. Y. 2009. "Shape based Image Retrieval. Using Combining Global. and Local. Shape Features". 2nd International. Congress on Image and Signal. Processing, pp. 1-5.