دانشگاه خوارزمی

دانشکده مهندسی کامپیوتر

بازیابی تصویر بر اساس محتوا در مقیاس بالا با استفاده از الگوی دودویی محلی و الگوریتم هشینگ چندی سازی تکرارشونده

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته مهندسی کامپیوتر
گرایش هوش مصنوعی

مونا شاکردنیوی

استاد راهنما

دکتر جمشید شنبه زاده

تابستان ۱۳۹۴

تأییدیه‌ی هیأت داوران جلسه‌ی دفاع از پایان‌نامه

نام دانشکده: دانشکده مهندسی کامپیوتر

نام دانشجو: مونا شاکردنیوی

عنوان پایان‌نامه: بازیابی تصویر بر اساس محتوا در مقیاس بالا با استفاده از الگوی دودویی محلی و الگوریتم هشینگ چندی سازی تکرارشونده

تاریخ دفاع: تابستان ۱۳۹۴

رشته: مهندسی کامپیوتر

گرایش: هوش مصنوعی

ردیف	سمت	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه	امضـــــــا
۱	استاد راهنما	دکتر جمشید شنبه زاده	دانشیار	دانشگاه خوارزمی	
۲	استاد مدعو خارجی	دکتر احسان اله کبیر	استاد	دانشگاه تربیت مدرس	

تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب مونا شاکردنیوی به شماره دانشجویی ۹۱۳۸۲۲۶۰۶ دانشجوی رشته مهندسی کامپیوتر مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه‌ی نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه‌برداری‌شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. درصورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ...) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض درخصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسؤولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذی‌صلاح (اعم از اداری و قضایی) به عهده‌ی اینجانب خواهد بود و دانشگاه هیچ‌گونه مسؤولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: مونا شاکردنیوی

تاریخ و امضا:

مجوز بهره‌برداری از پایان‌نامه

بهره‌برداری از این پایان‌نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما به شرح زیر تعیین می‌شود، بلامانع است:

☐ بهره‌برداری از این پایان‌نامه برای همگان بلامانع است.

☐ بهره‌برداری از این پایان‌نامه با اخذ مجوز از استاد راهنما، بلامانع است.

☐ بهره‌برداری از این پایان‌نامه تا تاریخ ..... ممنوع است.

استاد راهنما: دکتر جمشید شنبه زاده

تاریخ:

امضا:

تقدیم به:

خانواده عزیزم

خدایا...[1]

به من زیستنی عطا کن که در لحظه مرگ، بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مُردنی عطا کن که بر بیهودگیش، سوگوار نباشم. بگذار تا آن را، خود انتخاب کنم، اما آنچنان که تو دوست می‌داری.

تو می‌دانی و همه می‌دانند که شکنجه دیدن بخاطر تو، زندانی کشیدن بخاطر تو و رنج بردن به پای تو تنها لذت بزرگ زندگی من است، از شادی توست که من در دل می‌خندم، از امید رهایی توست که برق امید در چشمان خسته‌ام می‌درخشد و از خوشبختی توست که هوای پاک سعادت را در ریه‌هایم احساس می‌کنم. نمی‌توانم خوب حرف بزنم. نیروی شگفتی را که در زیر کلمات ساده و جمله‌های ضعیف و افتاده، پنهان کرده‌ام دریاب، دریاب.

تو می‌دانی و همه می‌دانند که زندگی از تحمیل لبخندی بر لبان من، از آوردن برق امیدی در نگاه من، از برانگیختن موج شعفی در دل من، عاجز است.

تو، چگونه زیستن را به من بیاموز، چگونه مردن را خود خواهم آموخت.

به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی‌همراه، جهاد بی‌سلاح، کار بی‌پاداش، فداکاری در سکوت، دین بی‌دنیا، مذهب بی‌عوام، عظمت بی‌نام، خدمت بی‌نان، ایمان بی‌ریا، خوبی بی‌نمود، گستاخی بی‌خامی، قناعت بی‌غرور، عشق بی‌هوس، تنهایی در انبوه جمعیت، و دوست داشتن بی‌آنکه دوست بداند، روزی کن.

اگر تنهاترین تنهاشوم، باز خدا هست

او جانشین همه نداشتن‌هاست...

---

[1] مناجاتی از دکتر علی شریعتی.

قدردانی

سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می دانم از زحمات بی دریغ استاد راهنمای خود، جناب آقای دکتر شنبه زاده، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی های ارزنده ایشان، این مجموعه به انجام نمی رسید.

در پایان، بوسه می زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می کنم وجود مقدس شان را و تشکر می کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امید بخش وجودشان، که بهترین پشتیبان من بودند.

مونا شاکر دنیوی
تابستان ۱۳۹۴

چکیده

تکنیک هشینگ، الگوریتمی کارا به منظور تخمین نزدیک ترین همسایه در پایگاه داده های بزرگ تصاویر می باشد.یادگیری کد باینری یکی از گام های کلیدی به منظور بهبود عملکرد این الگوریتم ها است و هنوز به عنوان چالش پیش رو در این حوزه مطرح می باشد.ورودی های این الگوریتم ها در عملکرد آن ها تاثیر بسزایی دارند. روش پیشنهادی در این رساله با به کارگیری ورودی های مناسب عملکرد الگوریتم های هشینگ را بهبود بخشیده است.در این روش به منظور استخراج ویژگی های تصویر از روش الگوی دودویی محلی استفاده شده است و سپس این بردار به عنوان ورودی به الگوریتم هشینگ چندی سازی تکرارشونده (ITQ) داده شده است که این امر منجر به تولید کدهای فشرده تر،حافظه کمتر ، هزینه محاسباتی پایین ترو عملکرد بهتر گردیده است.دلیل این امر ماهیت دودویی و عملکردمناسب الگوی دودویی محلی در استخراج ویژگی ها ی تصویر می باشد.در پایان به منظور مرتب سازی نتایج بر اساس بیشترین شباهت به تصویر درخواستی کاربر، تکنیک مرتب سازی مجدد با توجه به اهمیت هر بیت اعمال گردیده و به بیت هایی با اهمیت بیشتری وزن بالاتری تخصیص می بابد و وزن های کمتر به بیت های کم اهمیت نسبت داده می شود و با محاسبه فاصله همینگ وزن دار با توجه به اهمیت هر بیت نتایج بر اساس بیشترین شباهت مرتب می گردد. به منظور بررسی عملکرد روش فوق از دو پایگاه داده CIFAR-۱۰ و MNIST و منحنی precision-recall استفاده شده است.نتایج آزمایشات نشان می دهد روش فوق نسبت به روش های مشابه از نظرطول کد باینری،حافظه و هزینه محاسباتی کاراتر می باشد.

واژگان کلیدی: بازیابی تصاویر بر اساس محتوا،مقیاس بالا،الگوی دودویی محلی،الگوریتم هشینگ چندی سازی تکرار شونده

فهرست مطالب

فهرست تصاویر

فهرست جداول

فهرست الگوریتم‌ها

فصل ۱

۱ـ۱ مقدمه

در این فصل پیش زمینه ای از مفهوم بازیابی تصاویر و کاربردهای آن و هدف از انجام این کار ارائه خواهد شد.

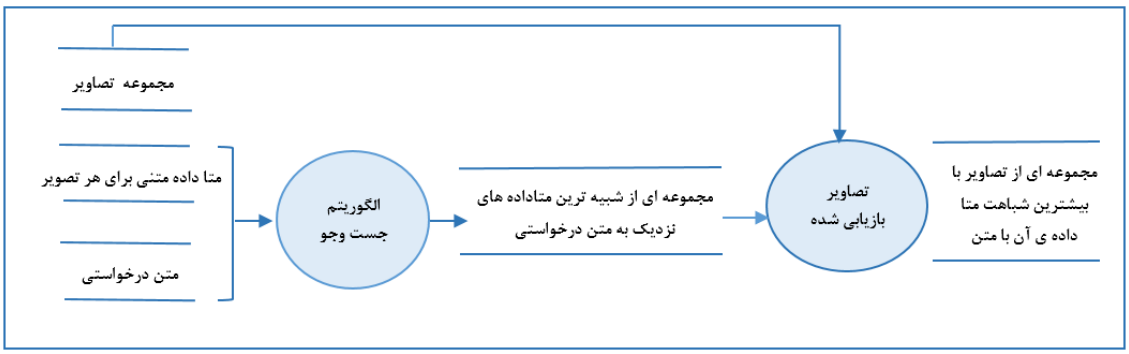
۱ـ۲ بازیابی تصویر

بازیابی تصویر ، فرآیند جست وجو و بازیابی تصاویر در یک پایگاه داده تصاویرمی باشد .بازیابی تصاویر به دو دسته ی کلی تقسیم می شود [۴۲] :

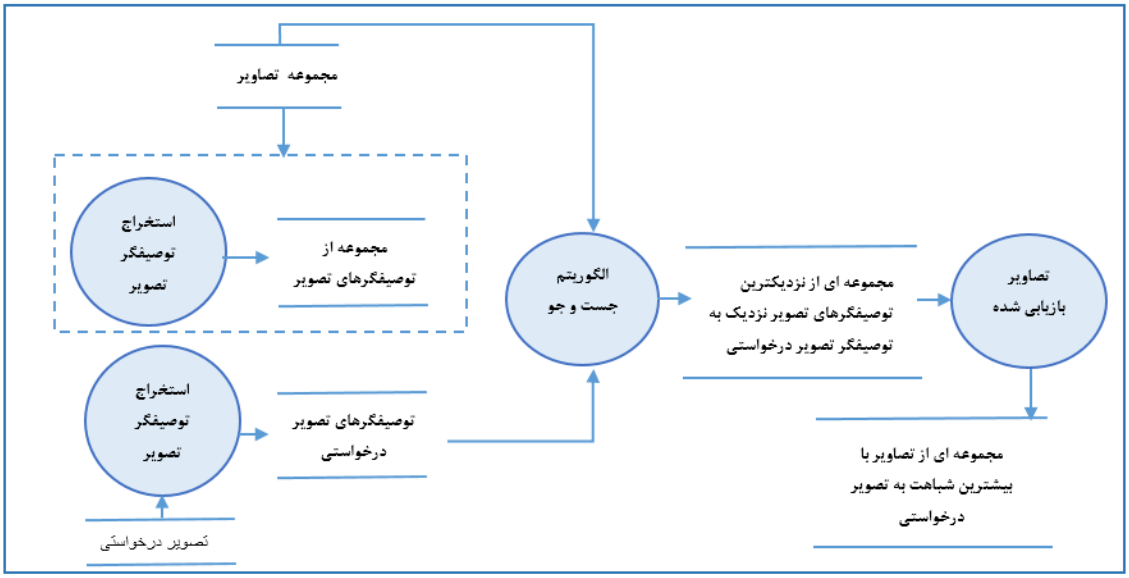
- بازیابی تصویر بر اساس متن[۱]
- بازیابی تصویر بر اساس محتوا[۲]

در بازیابی تصویر بر اساس متن(شکل ۱ـ۱)، یک کلمه کلیدی [۳] ،زیرنویس [۴] یا یک نوشته توصیفی [۵] به عنوان داده ماوراء به هر یک از تصاویر موجود در پایگاه داده اضافه می شود بنابراین جست و جو روی این داده ها صورت می گیرد.در این روش کلمه کلیدی یا عبارت به عنوان درخواست ارسال می گردد سپس تصاویری که داده ی ماورای آن مرتبط با درخواست کاربر است نمایش داده می شود .این داده ها اکثرا به صورت دستی

[۱] concept based image retrieval
[۲] Content based image retrieval
[۳] Key word
[۴] caption
[۵] Description text



شکل ۱ـ۱: سیستم بازیابی تصویر بر اساس متن [۲۳].



شکل ۱ـ۲: سیستم بازیابی تصویر بر اساس محتوا [۲۳].

به تصاویر الحاق می گردد که این یکی از محدودیت های این روش است. در بازیابی تصاویر بر اساس محتوا (شکل ۱ـ۲) ، از داده های متنی در فرآیند بازیابی تصاویر استفاده نمی شود در این روش،بازیابی براساس میزان شباهت مفاهیم تصاویر به تصویر درخواستی ارسال شده توسط کاربر صورت می گیرد.این مفاهیم شامل رنگ،بافت،شکل و یا هر اطلاعات دیگری که از تصویر استخراج می گردد می باشد بنابراین مفاهیم تصویر باید تجزیه گردد و با یک شیوه ی مناسب به منظور فرآیند جست وجو نمایش داده شود.

۱ـ۲ـ۱ محدودیت های روش بازیابی تصویر بر اساس متن

امروزه اکثر موتورهای جست وجوی تصاویر تحت وب ،از الگوریتم های جست وجو بر پایه فراداده استفاده می کنند.مشکل این روش آن است که تعداد زیادی از نتایج ،مطلوب ما نمی باشند دلایل این امر:

- زبان: ابهامات ؛ مانند چندمعنایی[۶] یا ترادف [۷]
- داده های متنی : اگرچه فرآیند جست وجو با تکنولوژی های موجود آسان است اما برای پایگاه داده های بزرگ ویا تصاویری که به صورت خودکار(تصاویر دوربین های مدار بسته) تولید می شوند این روش ها غیر عملی می باشد زیرا نیاز به افرادی است که تصاویر موجود در پایگاه داده را در قالب کلمات توصیف نمایند.
- برچسب گذاری انسان:علاوه بر پرهزینه بودن و ناکارا بودن این روش ها برای پایگاه داده های بزرگ می توان به خطا در برچسب گذاری نیز اشاره نمود.ازاین خطاها می توان به غلط املایی یا برچسب گذاری نادرست اشاره کرد.کاربران متفاوت برچسب های متفاوتی را به تصویر الحاق می کنند و ممکن است تمام مفاهیم موجود در تصویر را برچسب گذاری نکنند.

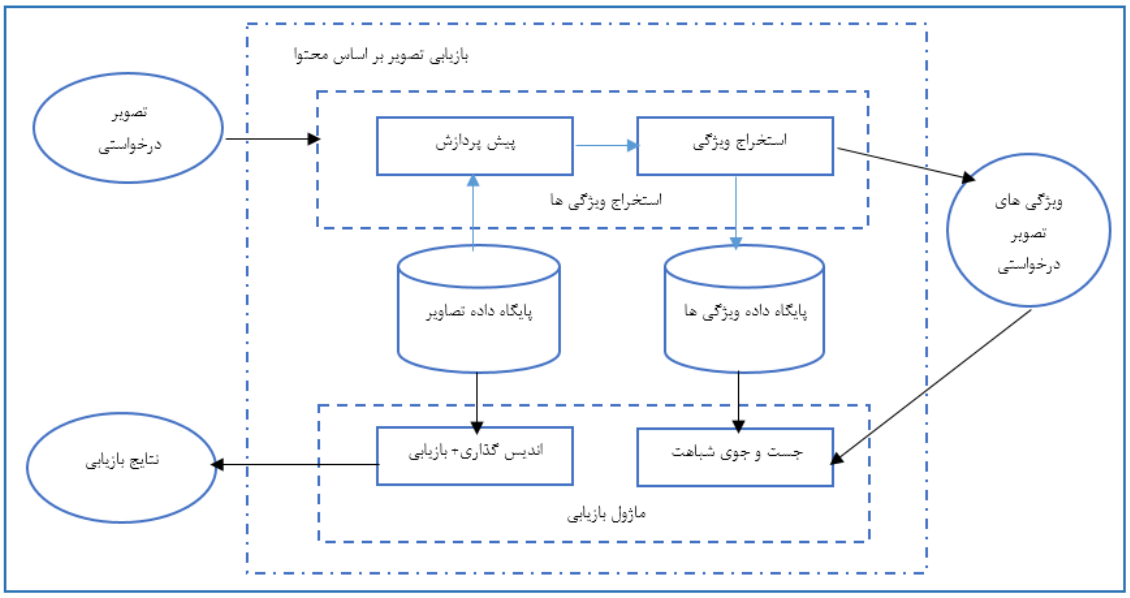
۱ـ۳ ساختار سیستم های بازیابی تصویر بر اساس محتوا

شکل ۱ـ۳ سیستم بازیابی تصویر بر اساس محتوا را نمایش می دهد فرآیند آموزش این سیستم به صورت آفلاین انجام می شود و با پیکان های آبی نمایش داده شده است و فرآیند درخواست کاربر به صورت آنلاین انجام می گردد و با پیکان های مشکی نمایش داده شده است.

۱ـ۳ـ۱ تصویر درخواستی

این بخش ، ورودی (تصویری است که به عنوان درخواست برای عملیات جست وجو گرفته شده است)را برای سیستم توصیف می کند.کاربر با انتخاب یک تصویر به دنبال یافتن تصاویر مشابه به آن در پایگاه داده می باشد.

[۶] polysemy
[۷] synonymy



شکل ۱ـ۳: ساختار سیستم بازیابی تصویر بر اساس محتوا.

۱ـ۳ـ۲ استخراج ویژگی ها

این بخش یکی از اجزای مهم این سیستم می باشد.در این بخش ویژگی های تصویر با هدف نمایش مفاهیم تصویر در قالب عددی استخراج می شود.

۱ـ۳ـ۳ پایگاه داده تصاویر

شامل مجموعه ای از تصاویر می باشد که در این سیستم نتایج درخواست کاربر را فراهم می کند.

۱ـ۳ـ۴ پایگاه داده ویژگی ها

مکانی است که کلیه اطلاعات استخراج شده از تصاویر آموزشی در آن ذخیره می شود .بردار ویژگی های کلیه تصاویر آموزشی استخراج و ذخیره می گردد.سیستم از این داده ها به منظور مقایسه ویژگی ها در فرآیند جست و جو استفاده می کند.ویژگی ها ممکن است در قالب یک لیست یا بردارهای ویژگی و یا ساختمان داده های پیچیده تری مانند درخت و... ذخیره گردد.

۱ـ۳ـ۵ جست و جو شباهت

اساس این جزء، مقایسه می باشد و نتایج را بر مبنای تکنیکی که به منظور مقایسه استفاده شده است برمی گرداند.سیستم میزان شباهت را برای هر یک از تصاویر موجود در پایگاه داده فراهم میکند.این فرآیند بر اساس معیارهای شباهت که به منظور محاسبه اختلاف میان توصیفگرهای تصویر درخواستی و پایگاه داده ویژگی ها می باشد انجام می گیرد.از دید سیستم ، شباهت دو تصویر وابسته به اختلاف در فضای ویژگی است.بنابراین زمانی که اختلاف میان توصیفگرهای تصویر کم باشد شباهت آن دو بیشتر می باشد.

۱ـ۳ـ۶ اندیس گذاری و بازیابی

در این بخش سیستم تعدادی از تصاویر را به عنوان نتیجه انتخاب و به کاربر نمایش می دهد.

۱ـ۴ بازیابی تصویر در مقیاس بالا

دیتاست های تصویر در مقیاس بالا شامل هزاران،میلیون ها یا بیشتر تصویر می باشند.نمونه هایی از این دیتاست ها صفحات وب [۸] ،تصاویر پزشکی [۹] ،فریم های ویدئو استخراج شده از دوربین های کاوشگرو... است.جست وجوی تصاویر در مقیاس بالا با استفاده از تکنیک های تطابق ناشیانه [۱۰] (تطابق تصویر درخواستی با تمامی تصاویر موجود در دیتاست) یک روش عملی [۱۱] نمی باشد.این روش در صورتی که پایگاه داده شامل چند صد تصویر باشد مناسب است دلایل رد روش فوق به شرح ذیل می باشد:

- تعداد بالای تصاویر موجود در دیتاست.
- ابعاد بالای توصیفگرهای تصویر که معمولا شامل ده ها یا صدها بعد برای هر یک از تصاویر می باشد.
- تکنیکی که به منظور تطابق جفت تصویرها مورد استفاده قرار می گیرد.
- پیچیدگی معیار فاصله ای که به منظور تطابق توصیفگرها مورد استفاده قرار می گیرد که یک عامل اصلی در افزایش سرعت و دقت بازیابی محسوب می گردد.

[۸] www
[۹] Medical image
[۱۰] Brute force matching
[۱۱] practical

۱ ـ ۵ کاربردهای سیستم های بازیابی تصویر بر اساس محتوا

بازیابی تصاویر بر اساس محتوا مولد بسیاری از برنامه های کاربردی در دنیای واقعی است.این برنامه ها شامل شناسایی اشیاء [۱۲] ،پیدا کردن کپی های غیر مجاز تصاویر در صفحات وب [۱۳] ،جست وجو کردن فریم های ویدئو [۱۴] ،دسته بندی کردن تصاویر [۱۵]،تصاویر پزشکی می باشد.کاربردهای دیگر آن شامل گالری های هنری و مدیریت موزه ها،مهندسی معماری،طراحی داخلی،مدیریت منابع زمین،سیستم های اطلاعات جغرافیایی،مدیریت پایگاه داده های علمی،پیش بینی آب وهوا،طراحی مد،مدیریت پایگاه داده های تجاری،اجرای قانون و پیشگیری جرم،آرشیو تصاویر و سیستم های اجتماعی و سرگرمی های خانگی می باشد. یکی از کاربردهای جذاب بازیابی تصاویر بر اساس محتوا،موتور جست وجوی تصاویر آنلاین می باشد به عنوان نمونه ای از این موتورهای جست وجو تصاویر می توان بهTinyEye google، Reverse images Image Search اشاره نمود.

۱ ـ ۶ اهداف پایان نامه

اهداف این پایان نامه به شرح ذیل می باشد:

- ارائه نمایی کلی از حوزه ی بازیابی تصویر
- مروری بر آخرین الگوریتم ها و متدهای ارائه شده به منظور رفع مشکلات مطرح شده در حوزه ی بازیابی تصاویر در مقیاس بالا
- ارائه روشی جدیدو کارا به منظور بازیابی تصاویر بر اساس محتوا در مقیاس بالا

۱ ـ ۷ ساختار پایان نامه

در فصل ۲ به مرور روش هایی که به منظور حل چالش های بازیابی تصویر بر اساس محتوا در مقیاس بالا پرداخته است اشاره دارد .این فصل به بیان روش هایی که به منظور نمایش ویژگی های تصویر و روش هایی که به منظور جست و جوی نزدیک ترین همسایه ارائه شده است پرداخته است.

[۱۲] Object recognition

[۱۳] Finding partial image duplicate on the web

[۱۴] Searching indivisual video frame

[۱۵] Image classification

فصل ۳ به تشریح الگوریتم های به کار رفته در این کار پرداخته است.

فصل ۴ به بیان روش پیشنهادی در این پایان نامه اشاره دارد.

فصل ۵ به تشریح آزمایشات انجام شده ،نتایج به دست آمده و مقایسه و ارزیابی روش پیشنهادی با روش های دیگر پرداخته است .در انتها به بیان نتیجه گیری و کارهای پیشنهادی برای آینده اشاره دارد.

فصل ۲

۲ـ۱ مقدمه

این فصل به مرورتحقیقات وکارهای اخیر انجام شده به منظور حل چالش های موجود در حوزه ی بازیابی تصویر بر اساس محتوا در مقیاس بالا می پردازد.از چالش های موجود در این حوزه می توان به مساله ابعاد بالای توصیفگر های تصویر و پیچیدگی و سرعت الگوریتم های جست وجو اشاره نمود که در پایگاه داده های تصاویر با ابعاد بالا منجر به افزایش حافظه ی مورد نیاز و کاهش سرعت بازیابی می گردند.به منظور رفع مشکلات فوق راهکارهایی ارائه گردیده است که منجر به کاهش هزینه ی محاسباتی،کاهش حافظه ی مورد استفاده و افزایش سرعت بازیابی گردد. در بخش اول به بیان الگوریتم های ارائه شده به منظور نمایش ویژگی های تصویر پرداخته شده است و در بخش دوم روش هایی که به منظور جست وجوی تصاویر ارائه شده بیان گردیده است.

۲ـ۲ نمایش تصویر

متدهای نمایش تصویر قدیمی [۱] از رنگ ، بافت یا شکل تصویر برای توصیف مفاهیم آن استفاده می کند.به عنوان نمونه در سیستم QBIC [۲] از هیستوگرام رنگ تصویر، برای نمایش تصاویر استفاده می شود.هیستوگرام رتگ تعداد وقوع هر رنگ در تصویر را نمایش می دهد(شکل ۲ـ۱)

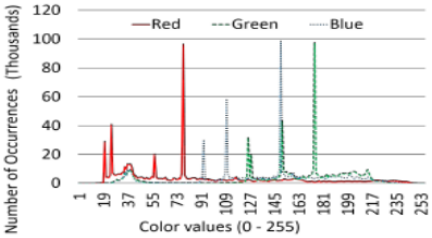
این موضوع بسیار بدیهی است که هیستوگرام رنگ،شامل اطلاعاتی درباره ی ساختار تصویر نمی باشد و تنها توزیع رنگ در تصویر را توصیف می کند.متدهای نمایش تصویر که اخیرا ارائه شده است از تکنیک

[۱] traditional

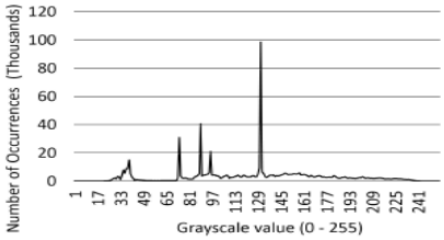
[۲] Query By Image Content

(آ) تصویر رنگی

(ب) تصویر در سطح خاکستری



(ج) هیستوگرام رنگی



(د) هیستوگرام سطح خاکستری

شکل ۲ـ۱ : نمونه ای از هیستوگرام تصویر

های پیشرفته ای مانند توصیفگرهای نقاط کلیدی تصویر [۳] و یا توصیفگرهای نقاط مورد علاقه [۴] استفاده می کنند.نقاط کلیدی تصویر به نقاطی اطلاق می گردد که در جایگاه های خوش تعریف [۵] در تصویر قرار دارد و نسبت به تغییرات سراسری و محلی در حوزه ی مقیاس [۶] ،چرخش [۷] ، نور و...ثابت هستند.تعدادی از نقاط مورد علاقه با قابلیت اطمینان و درجه ی بالایی از تکرارپذیری [۸] محاسبه می شوند این نقاط کلیدی به منظور تولید یک نمایش عددی [۹] برپایه ی اطلاعات پیکسل های موجود درناحیه ای از تصویر پیرامون نقطه کلیدی استفاده می شود.نتیجه ی این نمایش عددی بردارهای توصیفگر نقاط کلیدی [۱۰] یا بردار ویژگی [۱۱] نامیده می شود. تجزیه کردن یک تصویر به ویژگی ها [۱۲] یا نقاط کلیدی محلی مورد علاقه [۱۳] تکنیکی است که به

[۳] Image key point descriptor
[۴] Interest point descriptors
[۵] Well defined position
[۶] scaling
[۷] rotate
[۸] reproducibility
[۹] Numerical representation
[۱۰] Keypoint descriptor vecotr
[۱۱] Feature vector
[۱۲] features
[۱۳] Local keypoint of interest

صورت گسترده برای نمایش تصاویر دیجیتال مورد استفاده قرار می گیرد. بسیاری از کاربردهای کامپیوتر شامل آن هایی که مرتبط با بازیابی تصویر براساس محتوا می باشند وابسته به حضور پایدار ویژگی های نماینده یک تصویر می باشند.این کار هدف اصلی شناساگرهای نقاط کلیدی تصویر و توصیفگرهای آن است.هدف شناساگرنقاط کلیدی ایده آل ، شناسایی نواحی از تصویر است که در صورت اعمال تغییرات در نور،زاویه دید یا مقیاس،دوباره همان نقاط به عنوان نقاط کلیدی شناسایی گردد.به طور کلی این نواحی تصویر نسبت به تمام جابه جایی های [۱۴] ممکن یک تصویرمقاوم هستند.

۲ـ۲ـ۱ ویژگی های محلی

ویژگی های محلی محتویات بصری نواحی مورد علاقه را توصیف می کند [۲۸].یک توصیفگر محلی خوب باید جداساز و مقاوم در برابر تغییرات نمایش که منجر به تغییرات شناساگر می گردد باشد ضعف این روش این است که یک تصویر با مجموعه زیادی از توصیفگرها ،توصیف می شود و در جستجو برای بازیابی تصاویر مشابه باید حجم زیادی از مقایسات صورت گیرد زیرا توصیفگرهای تصویر درخواستی با توصیفگرهای تصاویر موجود در دیتاست مقایسه می گردد.

۲ـ۲ـ۲ انبان کلمات بصری

هدف این روش جایگزینی هر یک از توصیفگرهای محلی یک تصویر با یک کلمه بصری که از لغات [۱۵] از پیش تعریف شده است . هدف این روش به کارگیری تکنیک های بازیابی متن در بازیابی بر اساس محتوا می باشد.

مراحل اصلی مدل

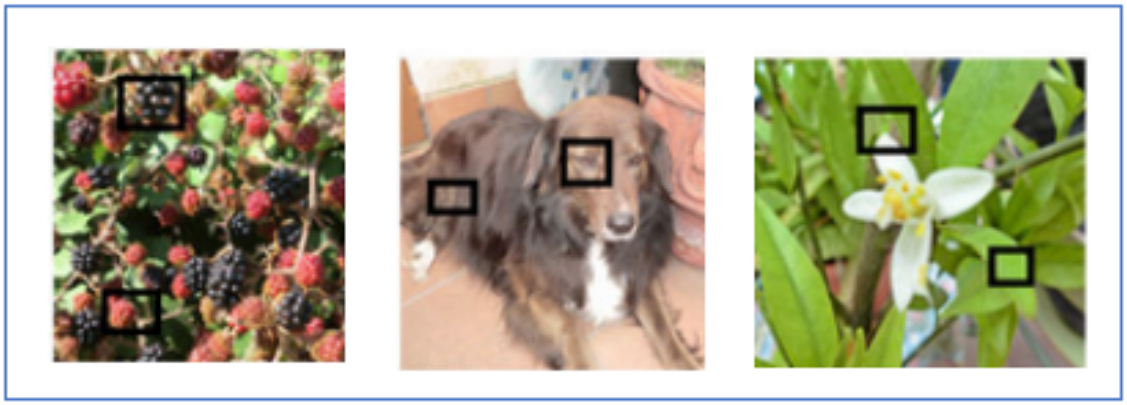
مراحل اصلی در مدل انبان کلمات بصری به صورت زیر است:

استخراج ویژگی

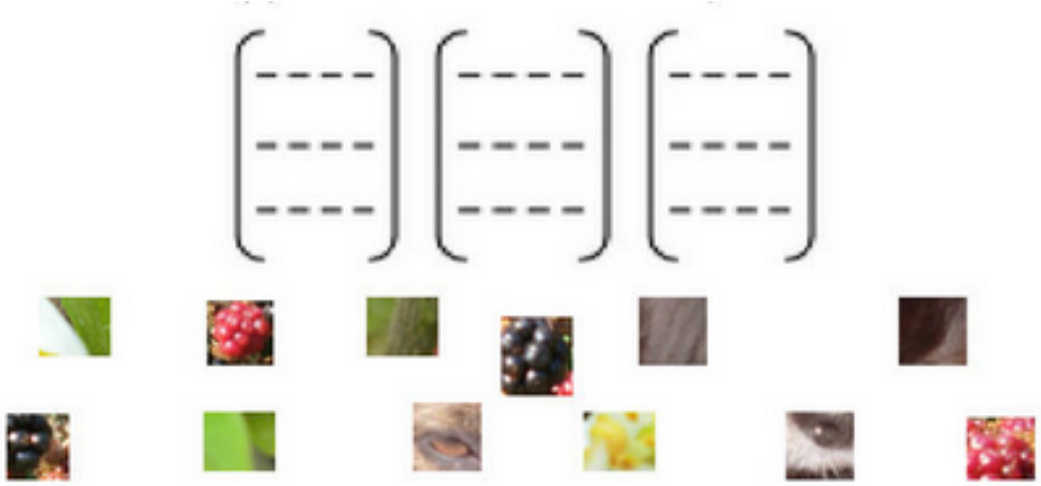
در مرحله استخراج ویژگی باید از قسمت هایی از تصویر به عنوان کلمات بصری نمونه گیری کنیم(شکل ۲ـ۲).

[۱۴] transformation

[۱۵] vocabulary



شکل ۲ـ۲: استخراج ویژگی های تصویر.

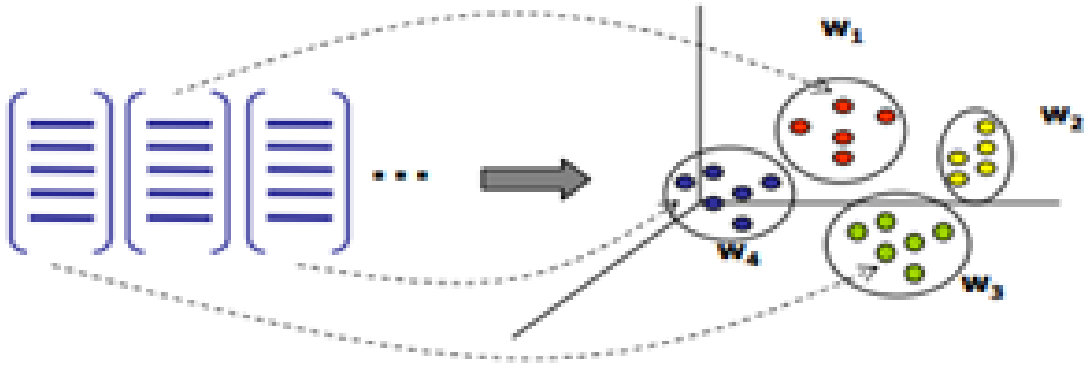


شکل ۲ـ۳: یادگیری واژه نامه بصری.

یادگیری واژه نامه بصری

انواع روش های خوشه بندی راهکارهایی هستند که در یادگیری واژه نامه [۱۶] از آن استفاده می شود و معمولا نوعی یادگیری بدون ناظر هستند یکی از معمول ترین روش ها برای یادگیری واژه نامه الگوریتم k-means است بعد از اجرای الگوریتم k-means مرکز هر خوشه [۱۷] به عنوان کلمه کلیدی [۱۸] در نظر گرفته می شود(شکل ۲ـ۳).

[۱۶] codebook
[۱۷] centroid
[۱۸] codewords



شکل ۲ـ۴: کوانتیزه کردن بردار ویژگی با استفاده از واژه نامه بصری.

شکل ۲ـ۵: بازنمایی تصاویر با استفاده از تعداد تکرار کلمات بصری.

کوانتیزه کردن بردارهای ویژگی با استفاده از واژه نامه بصری

در این مرحله از واژه نامه برای چندی سازی بردارهای ویژگی استفاده می شود چندی ساز برداری یک بردار ویژگی را به عنوان ورودی گرفته و آن را به نزدیک ترین بردار واژه [۱۹] درون واژه نامه انتساب می دهد معیار فاصله می تواند اقلیدسی یا هر متریک دیگری باشد(شکل ۲ـ۴).

بازنمایی تصاویر با استفاده از تعداد تکرار کلمات بصری

در این مرحله ویژگی ها را بر اساس آنچه در مرحله قبل به آن انتساب داده ایم شمارش می کنیم و یک بردار ویژگی جدید یا یک هیستوگرام از کلمات بصری تشکیل می دهیم (شکل ۲ـ۵).

[۱۹] codevector

۲ـ۲ـ۳ بردار فیشر

کرنل های فیشر چگونگی انحراف مجموعه ای از توصیفگر ها را از توزیع متوسط توصیف می کنند[۱۹]. با استفاده از یک مدل سراسری پارامتریک مدل می شود.کرنل های فیشر به منظور طبقه بندی تصاویر [۳۳] و جست وجوی تصاویر در مقیاس بالا مورد استفاده قرار می گیرند .هم چنین اثبات شده است که بردار فیشر بر گرفته از روش انبان کلمات بصری است .به طوری که روش انبان کلمات بصری تعداد توصیفگرهایی که به هر ناحیه نسبت داده شده است می شمارد بردار فیشر موقعیت تقریبی توصیفگرها در هر ناحیه را کد می کند.نتایج حاصل از [۲۱] نشان می دهد بهر ه گیری از بردار فیشر و تکنیک هشینگ دارای عملکرد بهتری نسبت به مدل انبان کلمات بصری است.

۲ـ۲ـ۴ بردار ویژگی های تجمع شده محلی

این روش برگرفته از نام $VectorofLocallyAggregatedDescriptors$ می باشد مانند روش انبان کلمات بصری در ابتدا یک واژه نامه $\{\mu_1, \mu_2, ..., \mu_k\}$ با استفاده از یک الگوریتم خوشه بندی مثل k-means ایجاد می شود سپس هر یک از توصیفگرهای محلی x_t به نزدیکترین کلمه بصری موجود در واژه نامه نسبت داده می شوند سپس از رابطه زیر:

$$v_i = \sum_{x_t : NN(x_t)=i} x_t - \mu_t \qquad (۱ـ۲)$$

بردار $V = [v_1^T, v_2^T, ..., v_k^T]$ به دست می آید و برای بهبود کارآیی این روش نرمال سازی انجام می شود بعد از انجام این پروسه دو توصیفگر سراسری V_1 و V_2 که مربوط به دو تصویر می باشند می توانند با استفاده از ضرب داخلی با هم مقایسه شوند.

۲ـ۳ جستجوي نزديكترين همسايه

مجموعه اي از نقاط به نام P در فضاي بعد بالا داريم.می خواهیم ساختمان داده اي بسازيم تا با گرفتن هر نقطه ي ورودي q ، نزديكترين نقاط به q در P را بدهد.اين مسأله جستجوي نزديكترين همسايه [۲۰] نام دارد و اهميت زيادي در بخش هاي مختلف علوم كامپيوتر (شامل شناسايي الگو، جستجو در داده هاي مالتی مديا، آمار محاسباتی، و داده كاوي) دارد. بسياري از اين كاربردها با مجموعه هاي بزرگ (مثلاً يک پايگاه داده

[۲۰] nearest neighbor search

شامل صفحات وب ميتواند بيش از يک بيليون متن داشته باشد) سروکار دارند. علاوه بر اين، بعد نقطه ها نيز می تواند بسيار بزرگ باشد (مثلاً چند صد يا حتی چند هزار بعد). بنابراين، طراحی الگوريتم هايی که بتوانند با اندازه ي پايگاه داده و تعداد ابعاد داده به خوبی مقياس پيدا کنند، بسيار پراهميت است. مسأله ي نزديکترين همسايه يک نمونه از دسته ي بزرگتري از مسائل به نام مسائل مجاورت [۲۱] می باشد. اين ها مسائلی هستند که تعريفشان وابسته به فاصله ي بين نقاط ورودي می باشد. ابتدا مسأله ي جستجوي نزديکترين همسايه ي دقيق [۲۲] را تعريف می کنيم.

۲ـ۳ـ۱ جستجوي نزديکترين همسايه ي دقيق

با داشتن مجموعه ي P از نقاط در فضاي R^d به دنبال ساختمان داده ای می باشيم که با گرفتن هر نقطه ي ورودي q ، نقطه اي درP که نزديکترين فاصله را تا q دارد پيدا کند. مسأله کامل تعريف نشده است مگر اينکه فاصله ي بين هر دو نقطه ي p و q تعريف شود. معمولاً فرض می شود که فاصله توسط نرم L_s ايجاد می شود.يعنی فاصله ی p و q به صورت $\|p-q\|_s$ است که در آن $\|X\| = (\sum_{i=1}^{d} \lceil x_i \rceil^s)^{1/s}$ است.تعريف ديگر و کلی تر از فاصله نيز امکان پذير است. يک راه حل ساده و پيش پا افتاده براي اين مسأله به شرح زير است: با گرفتن نقطه ي ورودي q ،فاصله ي q تا همه ي نقاط در P را حساب کن و نقطه اي که کمترين فاصله را داشت اعلام کن. اين روش که بررسی خطی [۲۳] نام دارد،پيچيدگی زمانی $\Theta(dn)$ دارد. اين براي مجموعه داده هاي کوچک قابل قبول می باشد ولی به شدت براي مجموعه داده هاي بزرگتر غير کاراست. هدف تحقيقات در اين حوزه اين است که الگوريتمی طراحی شود که اين مسأله را در زمان زيرخطی (يا حتی لگاريتمی) حل کند. مسائل مجاورت جزء اولين مسائلی بودند که در هندسه ي محاسباتی مورد بررسی قرار گرفتند. نتيجه ي اين تحقيقات اين بوده که تعداد زيادي راه حل هاي کارا براي حالتی که نقاط در فضايی با بعد ثابت قرار داشته باشند ارائه شده است. به عنوان مثال، مسأله ي نزديکترين همسايه براي هر نقطه ي ورودي با$o(\log n)$ زمان و $o(n)$ حافظه قابل حل است. نتايج مشابه براي مسائل ديگر نيز به دست آمده است. متأسفانه با رشد بعد، اين الگوريتم ها کارايی خود را از دست می دهند. به طور دقيق تر، پيچيدگی زمانی و حافظه اي آنها به صورت نمايی نسبت به بعد افزايش می يابد. به خصوص، مسأله ي نزديکترين همسايه راه حلی به زمان $o(d\log n)$ برای هر نقطه ورودی ولی با $n^{o(d)}$ حافظه دارد. اگر بخواهيم پيچيدگی حافظه را خطی يا نزديک خطی کنيم، بهترين کران بالاي به دست آمده براي پيچيدگی زمانی براي ورودي تصادفی $min(۲^{o(d)}, dn)$ است،که حتی برای d هايی که خيلی بزرگ نباشند تقريبا نسبت به n خطی است. علاوه بر

[۲۱] proximity problems

[۲۲] exact nearest neighbor search

[۲۳] linear scan

اين، وابستگی نمايی حافظه و/يا زمان به بعد (كه مصيبت بعد [۲۴] نام دارد) از ديدگاه كاربردي نيز بررسی شده است. به خصوص، دانسته شده است كه بسياري از ساختمان داده هاي پركاربرد (كه ازحافظه هاي خطی يا نزديک خطی استفاده می كنند) زمانی كه بعد از يک مقدار آستانه اي بيشتر می شود(معمولاً بين ۱۰ الی ۱۲ كه به تعداد نقاط داده بستگی دارد)، از خود پيچيدگی زمانی خطی نسبت به نشان می دهند. عدم موفقيت در از بين بردن وابستگی نمايی به بعد باعث شده كه بسياري از محققان ادعا كنند كه هيچ پاسخ كارايی براي اين دسته از مسائل زمانی كه بعد به اندازه ي كافی بزرگ باشد وجود ندارد. در همين راستا، سؤالی مطرح می شود: در صورتی كه اجازه بدهيم پاسخ ها تقريبی باشند، آيا ممكن است كه وابستگی نمايی از بين برود؟ مفهوم تقريب زدن براي جستجوي نزديكترين همسايه به خوبی تعريف ميشود: به جاي اينكه خروجی را نقطه ي p كه از همه به q نزديک تر است اعلام كند، الگوريتم اجازه دارد كه هر نقطه اي كه در فاصله ی $(1+\varepsilon)$ برابر فاصله ی p تا q قرار دارد را اعلام كند. تعريف هاي مشابهی قابل اعمال به مسائل مجاورت ديگر نيز می باشد. توجه كنيد كه اين رويكرد مشابه طراحی الگوريتم هاي تقريبی كارا براي مسائل ان پی هارد می باشد. در سالهاي اخير، تعدادي از محققان نشان دادند كه در واقع در بسياري از حالات، تقريب باعث كاهش وابستگی به بعد از نمايی به چندجمله اي می شود.

۲ـ۳ـ۲ جستجوي نزديكترين همسايه ي تقريبی

تقريباً همه ي الگوريتم هاي مربوط به مسائل مجاورت در فضاهايي با بعد بالا ابتدا مسئله را به مسئله ي يافتن نزديک ترين همسايه هاي تقريبی كاهش می دهند. بنابراين ما نيز با تعريف اين مسئله شروع می كنيم.

جستجوی نزدیک ترین همسایه ی تقریبی یا NN-(r،c)

با داشتن مجموعه ی P از n نقطه در فضای متری M=(X،D)، به دنبال طراحی ساختمان داده ای می باشیم كه بتواند عمل زير را انجام دهد: برای هر نقطه ی ورودی $q \in X$، اگر $p \in P$ وجود داشته باشد به قسمی كه $D(p,q) \leq r$ باشد،نقطه ی $\acute{p} \in P$ را پيدا كنيد طوری كه $D(\acute{p},q) \leqslant cr$ باشد. معيار متريک همينگ: يک معيار متری (Σ, D) كه در آن Σ مجموعه ای از نمادهاست و برای هر $(p,q) \in \Sigma$، $D(\acute{p},q)$ برابر با تعداد $i \in 1,...,d$ است به طوری كه $p_i \neq q_i$ باشد.

[۲۴] curse of dimensionality

رویکرد تصویر کردن تصادفی

اولین الگوریتم ها برای (r،c)-NN در ابعاد بالا به کمک تکنیک تصویر کردن تصادفی به دست آمده اند.این تکنیک زمانی کاربرد دارد که معیار متری D توسط نرم L_pایجاد شده باشد،زمانی که $p \in [۰, ۲]$ باشد.ابتدا به حالتی می پردازیم که تمام نقاط و نقطه ی ورودی بردارهای دودویی از $\{۰, ۱\}^d$ باشند و D فاصله ی همینگ باشد(یعنی معیار متری ای که توسط نرم $L_۱$ ایجاد شود).

کاهش بعد

یک روش تصادفی شده می باشد که بعد فضای همینگ را از d به $k = O(logn/\varepsilon^۲)$ کاهش دهد به طوری که با $(۱ + \varepsilon)$ ضریب فاصله های بین نقطه های داده و نقطه ی ورودی را حفظ می کند.این روش به ما اجازه می دهد تا مساله (r،c)-NN را در فضای d– بعدی به مساله ی (r،c)-NN در فضای k– بعدی کاهش دهیم.رویکرد مشابهی را می توان برای حل مساله نزدیک ترین همسایه در فضای اقلیدسی به کار برد.به خصوص حل مسئله ی $(r, c + \varepsilon) - NN$ در $L_۲^d$ به کمک $n(\frac{۱}{d})^{o(d)}$ حافظه قابل انجام است.

۲ـ۳ـ۳ هشینگ حساس به همسایگی(ال اس اچ)

ال اس اچ یک الگوریتم تصادفی شده است. از خاصیت تصادفی در ساخت ساختمان داده استفاده می شود. علاوه بر این، الگوريتم براي حل مسئله ي همسايه ي نزديک [۲۵] به کار می رود نه مسئله ي نزديكترين همسايه. مسئله ي همسايه ي نزديک نسخه ي تصميم مسئله ي نزديكترين همسايه است.تعریف مسئله به صورت زیر می باشد.دراین تعاریف نقطه ی p ، همسایه ی R– نزدیک q می باشد.اگر فاصله از p به q حداکثر R باشد.

۲ـ۳ـ۴ نزدیکترین همسایه ي c– تقریبی تصادفی شده

با داشتن مجموعه ی P از نقاط در فضای R^d و پارامترهای $\delta > ۰$R>۰،ساختمان داده ای بساز که با گرفتن هر نقطه ی ورودی q ،هر همسایه ی R– نزدیک q در p را با احتمال $۱ - \delta$ اعلام کند. توجه کنید که در تعریف دوم،احتمال های اعلام همسایه های R– نزدیک مختلف ممکن است از هم مستقل نباشد.علاوه بر این،توجه کنید که در هر دو مسئله مقدار R در زمان پیش پردازش دانسته شده است.بنابراین،با مقیاس کردن مختصات همه ی نقاط می توانیم فرض کنیم که ۱=R است. ایده ي اصلی الگوريتم ال اس اچ این است که

[۲۵] decision version

نقاط داده را با استفاده از چند تابع هش طوري هش كند كه براي هر تابع، احتمال برخورد [۲۶] براي نقاطی كه به هم نزديک ترند بسيار بالاتر از نقاطی باشد كه از هم دورترند. در اين صورت می توان همسايه هاي نزديک را با هش كردن نقطه ي ورودي q و بازيابی نقاط موجود در جعبه ي شامل q به دست آورد.

۲ـ۳ـ۵ رويكرد تقسيم و تخسير

تكنيک هاي كاهش بعد و ال اس اچ محدوديت هايی دارند. به خصوص، نمی توان براي حل مسئله ي نزديک ترين همسايه ي تحت نرم از آن ها استفاده كرد. خوشبختانه، اين نرم خاصيت هاي جالب ديگري دارد كه طراحی ساختمان داده براي نزديک ترين همسايه ي تقريبی را ميسر می سازد. تنها الگوريتم شناخته شده برای حل (r،c)-NN تحت نرم L_∞^d دارای پارامترهای زير برای هر $\rho > ۰$ می باشد:

- ضريب تقريب:($c = O(۴\lceil log_(۱+\rho) log ۴d \rceil)$ اگر $\rho = logd$ آن گاه c=۳
- حافظه:$dn^{۱+\rho}$
- زمان:O(dlogn)

ايده ي اصلی الگوريتم استفاده از رويكرد تقسيم و تخسير است. به خصوص، ابرصفحه هاي H كه شامل همه ي نقاطی كه يک مختصشان (مثلاً i امين) با هم برابر باشد را در نظر بگيريد. الگوريتم سعی می كند ابرصفحه ي اي پيدا كند كه داراي اين خاصيت باشد كه مجموعه ي نقاط $P_L \subset P$ كه در سمت چپ آن و در فاصله ی r≥ از آن قرار دارد خيلی كوچک تر از مجموعه ی P_M از نقاطی باشد كه در فاصله ی r از آن قرار دارد .شرايط مشابهی بايد برای مجموعه ی P_Rاز نقطه ها در طرف راست H نيز برقرار باشد.اگر چنين H ای وجود داشته باشد،P را به دو مجموعه ی $P_۱ = P_L \cup P_M$و $P_۲ = P/P_L$ تقسيم می كنيم و ساختمان داده را به صورت بازگشتی روی $p_۱$ و $P_۲$ می سازيم.به راحتی می توان ديد كه هنگام پردازش نقطه ی ورودی q ،وابسته به اين كه q در كدام سمت H قرار داشته باشد ،الگوريتم را به صورت بازگشتی روی $P_۱$ و $P_۲$ ادامه می دهيم.علاوه بر اين،می توان اثبات كرد كه افزايش در استفاده از حافظه كه ناشی از تكرار P_M می باشد خيلی زياد نيست.

۲ـ۳ـ۶ درخت های كی دی

درخت كی دي ساختمان داده اي است كه توسط جان بنتلی در سال ۱۹۷۵ اختراع شد. با آن كه اين ساختمان داده بسيار قديمی است ولی هم چنان جزء ساختمان داده هاي پركاربرد براي جستجو در فضاهاي چندبعدي

[۲۶] collision

[۲۷] به خصوص جستجو در حافظه ي اصلی به شمار می رود.
با داشتن مجموعه اي ازn نقطه در فضاي d بعدي، درخت کی دي به صورت بازگشتی به شرح زير ساخته می شود. ابتدا، ميانه ي مقادير مختص i ام نقاط را پيدا ميکنيم (از i=۱ شروع می کنيم). يعنی مقدار M اي محاسبه ميشود که ۵۰ درصد نقاط داراي مختص i ام بزرگتر يا مساوي M بوده و ۵۰ درصد ديگر نقطه ها داراي مختص i ام کوچک تر از M باشند. مقدارx ذخيره می شود و مجموعه ی P به P_L و P_R افراز می شود به قسمی که P_L فقط شامل نقاطی می شود که مختص i امشان کوچک تر يا مساوی M باشد و $|P_R| = |P_L| \pm ۱$.اين پروسه به صورت بازگشتی روی P_L و P_R تکرار می شود با اين تفاوت که i با i=۱ جايگزين می شود (يا ۱ اگر i = d باشد).زمانی که مجموعه ی نقاط يک نود اندازه ی يک داشته باشد،الگوريتم بازگشتی خاتمه می يابد. ساختمان داده حاصل يک درخت دودويی است با n برگ و عمق $\lceil logn \rceil$.به خصوص برای d=۱ ،يک درخت جستجو دودويی متوازن به دست می آوريم.با توجه به اين که ميانه ی n نقطه در زمان O(n) به دست می آيد،کل ساختمان داده در زمان $O(nlogn)$ قابل ساخت است.
چند راه حل برای يافتن نزديک ترين همسايه های نقطه ی ورودی q در مجموعه نقاط P وجود دارد.اولين راه حل در همان مقاله ی سال ۱۹۷۵ ارائه شده است.يک راه حل را ما در اينجا توضيح می دهيم.جستجو از ريشه ی درخت آغاز می شود و بازگشتی است.در هر زمان ،الگوريتم فاصله ی R به نزديک ترين نقطه به q را نگه می دارد،در ابتدا R=∞ است.در نود برگ (که مثلا شامل نقطه ی $\acute{P}$است) الگوريتم چک می کند که $\|q - \acute{p}\| < R$ باشد.اگر اين طور بود مقدار R ، $\|q - \acute{p}\|$ قرار داده می شود و$\acute{p}$ به عنوان کانديد نزديک ترين همسايه ذخيره می شود.برای يک نود داخلی (غير برگ) الگوريتم به ترتيب زير عمل می کند فرض کنيد M مقدار ميانه ذخيره شده در نود باشد که برای مختص های i ام حساب شده است.الگوريتم چک می کند که مختص i ام کوچک تر يا مساوی M است اگر کوچک تر بود، الگوريتم روی نود چپ ادامه می دهد،در غير اين صورت روی نود راست ادامه می دهد.بعد از اتمام الگوريتم بازگشتی،الگوريتم تستی را انجام می دهد يعنی چک می کند که اگر کره ای با شعاع R شامل نقطه ای در R^d می باشد که مختص i ام آن در طرف مخالف M نسبت به q باشد.اگر اين طور بود،الگوريتم بازگشتی روی فرزند (سرنزده ی) نود فعلی ادامه پيدا می کند.در غير اين صورت الگوريتم پايان می يابد.در پايان،الگوريتم آخرين کانديد نزديک ترين همسايه را به عنوان جواب اعلام می کند.
نشان داده شده است اگر داده و نقطه ی ورودی به صورت مستقل و تصادفی از يک توزيع تصادفی انتخاب شوند،آن گاه الگوريتم در زمان $G(d)logn$برای يک تابع G انجام می شود.ولی G(d) نسبت به d نمايی است (برای d هايی که نسبت به $logn$ کوچک هستند) ولی در بدترين حالت پيچيدگی زمانی O(dn) است.زمانی که بعد فضای جستجو از چند بعد بيشتر شود مجبور می شويم که تقريبا همه ی نودهای مجموعه داده را

[۲۷] multidimensional spaces

بررسی کنیم.

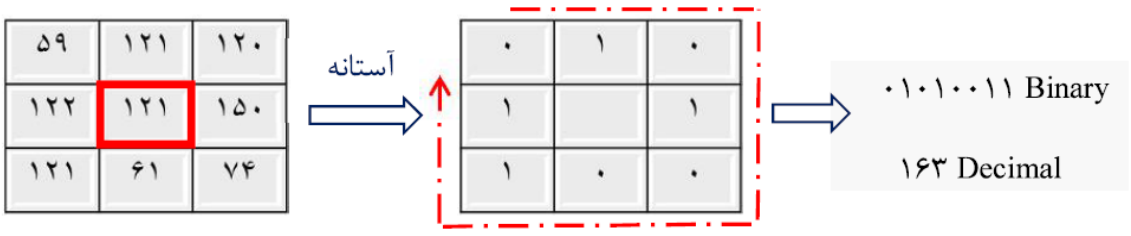
فصل ۳

۳ـ۱ مقدمه

در این فصل به بیان پیشینه ی الگوریتم های به کار رفته در این پایان نامه اشاره می گردد و عملکرد آن ها به تفصیل بیان می گردد.

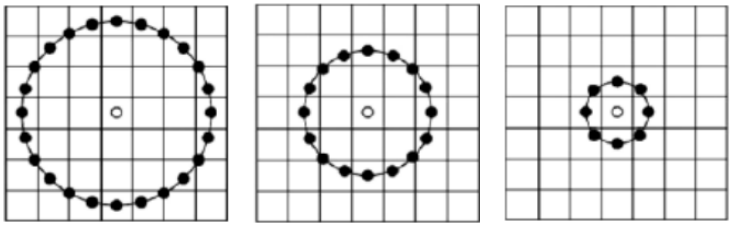
۳ـ۲ الگوی دودویی محلی

عملگر LBP اصلی به عنوان توصیفگری قدرتمند برای بافت تصویر معرفی شده است [۲۹].این عملگر برای هر پیکسل با توجه به برچسب پیکسل های همسایگی ۳×۳ یک عدد دوویی تولید می کند.برچسب ها با آستانه سازی مقدار پیکسل های همسایه با مقدار پیکسل مرکزی به دست می آیند. به این صورت که برای پیکسل های با مقدار بزرگتر یا مساوی مقدار پیکسل مرکزی برچسب ۱ و برای پیکسل های با مقادیر کوچکتر از مقدار پیکسل مرکزی برچسب ۰ قرار می گیرد. سپس این برچسب ها به صورت چرخشی در کنار هم قرار گرفته و یک عدد ۸ بیتی تشکیل می دهند.نحوه ی کار این عملگر در شکل ۳ـ۱ آمده است. محدودیت عملگر LBP



شکل ۳ـ۱: عملگر الگوهای دودویی محلی [۲۹].

پایه همسایگی کوچک (۳×۳) آن است که باعث می شود بر تصاویر با مقیاس بزرگ تسلط نداشته باشد.بدین منظور بعدا این عملگر با توسعه ای بر اندازه همسایگی مطرح شد که به صورت دایره ای با شعاع R پیکسل بر روی P پیکسل نشان داده می شود.این عملگر به صورت $LBP_{P,R}$ نشان داده می شود و حداکثر می تواند ۲^P مقدار مختلف، مطابق با ۲^P الگوی دودویی تولید شده توسطP پیکسل موجود بر روی شعاع همسایگی R تولید کند.برای مثال می توانید به شکل ۳ـ۲ مراجعه کنید که نحوه ی انتخاب پیکسل های همسایه در این نوع الگوی دودویی محلی را به ازای سه شعاع مختلف را نشان می دهد. بعد از برچسب گذاری یک تصویر



(آ) $(P=۸,R=۱)$ (ب) $(P=۱۶,R=۲)$ (ج) $(P=۲۴,R=۳)$

شکل ۳ـ۲: عملگر LBP با شعاع ها و تعداد همسایگی های مختلف [۱۴]

توسط عملگر LBP ،هیستوگرامی از برچسب ها به صورت زیر تعریف می شود:

$$H_i = \sum_{xy} I(f_i(x,y)=i), i = ۰,...,n-۱ \qquad (۱ـ۳)$$

که n تعداد برچسب های تولید شده توسط عملگر LBP و تابع I به صورت رابطه زیر تعریف شده است:

$$I_(A) = \begin{cases} ۱ & AisTrue \\ ۰ & Aisfalse \end{cases} \qquad (۲ـ۳)$$

اپراتور الگوی دودویی محلی پایه برای تغییر شکل های مقیاس خاکستری یکنواخت ، تغییر نمی کند که می تواند رده تمرکز پیکسل را در مجاورت محلی حفظ نماید و تصویری از برچسب های الگوی دودویی محلی محاسبه شده در یک منطقه به عنوان توصیفگر بافت نشان دهد.

این اپراتور می تواند مقادیر خروجی متفاوتی مربوط به الگوهای محلی متفاوت را توسط P پیکسل مجاور نشان دهد.در صورتی که تصویر بچرخد این پیکسل های اطراف ،در هر کدام از قسمت های مجاور به همین ترتیب در مجاورت دایره حرکت می کنند و به این ترتیب نتیجه آن یک مقدار الگوی دودویی متفاوت می باشد

و تنها استثنای الگوها یک و صفر خواهد بود.به منظور حذف چرخش تصویر ،یک الگوی دودویی محلی ثابت چرخشی در [۳۰] پیشنهاد شده است. که آن را در فرمول زیر مشاهده می کنید:

$$LBP^{ri}_{P,R} min\{ROR(LBP_{P,R}, i)|i = ۰, ۱, ..., p - ۱\} \qquad (۳ـ۳)$$

که در آن ROR(x،i) به عنوان یک متغیر سمت راست دایره ای کوچک در تعداد P بیت به صورت (x،i) اعلان می گردد.اپراتور $LBP^{ri}_{P,R}$ مقادیر آماری الگوهای انتشاری متغیر غیر چرخشی گوناگون را نشان می دهد که به ریزویژگی های مشخص شده درون تصویر بستگی دارد و در حقیقت می تواند الگوهایی را نشان دهد که به عنوان آشکار ساز ویژگی [۱] لحاظ شده است.با این وجود نشان داده شده است که چنین اپراتور الگوی محلی چرخشی ثابت ضرورتا نمی تواند اطلاعات متمایزی را ارائه دهد زیرا تناوبات الگوهای انتشاری،جداگانه تلفیق شده و همان طور که در $LBP^{ri}_{P,R}$ نشان داده شده تفاوت های عمده ای با یکدیگر داشته و فقط قابلیت تعیین مقادیر خام فضای زاویه دار در فواصل ۴۵ درجه امکان پذیر می گردد.

در [۳۰] تعدادی از الگوهای بنیادی به نام الگوهای یکنواخت [۲] نام گذاری شده اند که عبارتند از الگوهایی که در آن ها با چرخش بیت ها حداکثر ۲ تغییر ۰ به ۱ یا برعکس داشته باشیم.به عنوان مثال ارقام ۰۰۰۰۰۰۰۰، ۰۰۱۱۱۰۰۰ و ۱۱۱۰۰۰۰۱ یکنواخت هستند .ذخیره ی الگوهای یکنواخت عملگر LBP جدیدی به نام $LBP^{u۲}_{P,R}$ تولید می کند که برای همسایگی ۸ پیکسل ۵۹ الگو خواهد داشت ، به طوری که برای ۵۸ بین [۳] اول هیستوگرام تعریف شده مربوط به الگوهای یکنواخت و بین ۵۹ ام برای ذخیره ی مجموع دیگر الگوهای غیر یکنواخت است.این عدد برای LBP استاندارد ۲۵۶ متناظر با تعداد حالت های مختلف عدد دودویی ۸ بیتی است.۵۸ الگوی یکنواخت تولید شده توسط عملگر LBP پایه در شکل ۳ـ۳ نشان داده شده اند.

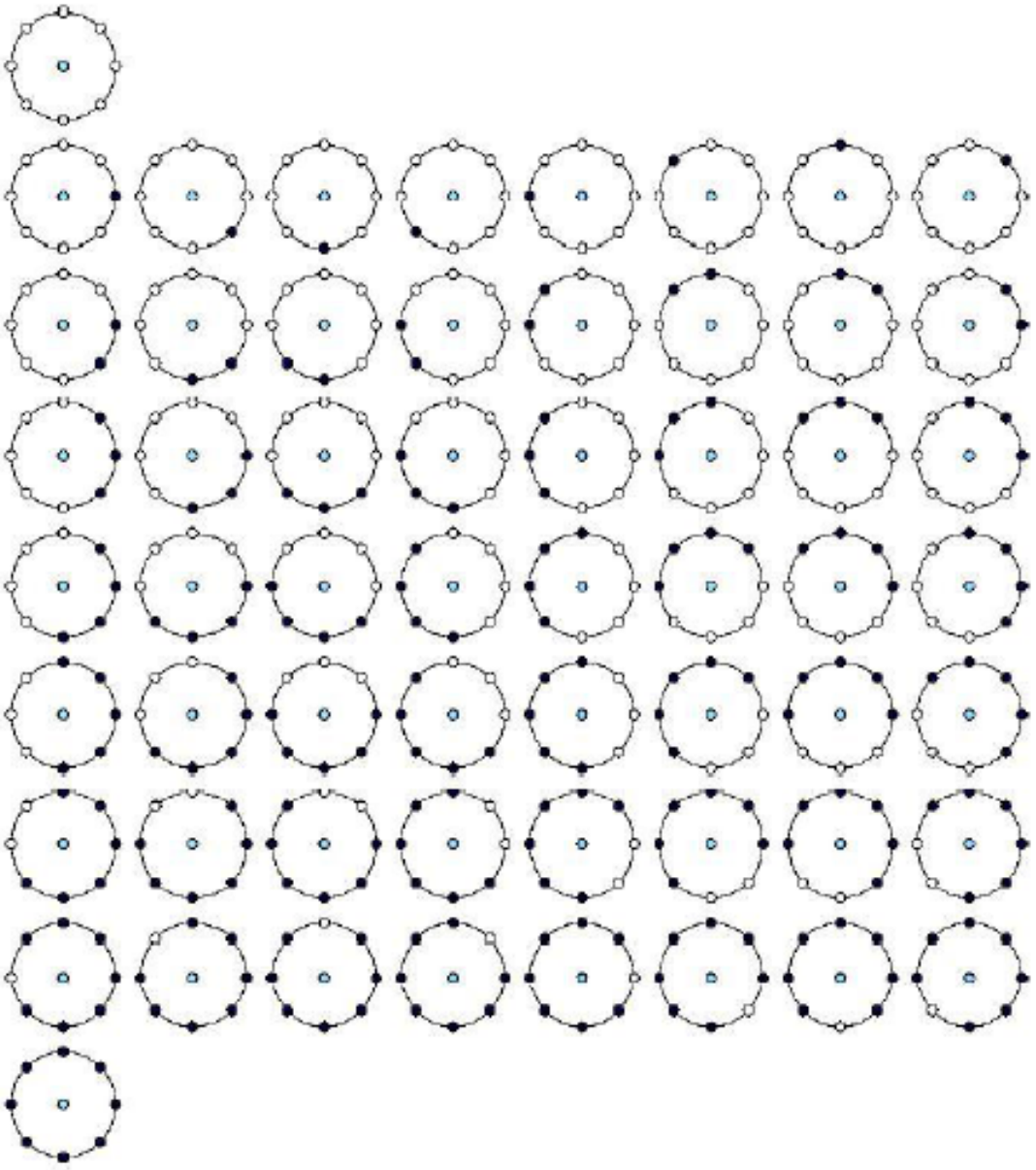
آزمایشات بر روی تصاویر نشان می دهد که ۹۰ % الگوهای (۱و۸) و ۷۰ % الگوهای (۲و۱۶) یکنواخت هستند[۳۰].آزمایشات مشابه بر روی تصاویر چهره نشان می دهد که ۹۰,۶ % الگوهای (۱و۸)و۸۵,۲ % الگوهای (۲و۸) یکنواخت هستند .

هیستوگرام $LBP^{u۲}$ به دست آمده از تصویر شامل اطلاعاتی درباره ی ریزالگوهای محلی ازقبیل لبه، نقطه، نواحی صاف و ... روی تصویر است.بدین ترتیب بافت تصویر که همان چین و چروک های آن می باشد توصیف خواهد شد.شکل ۴ـ۳ نمونه هایی از ریزالگوهای توصیف شده در هیستوگرام $LBP^{u۲}$ را به خوبی نشان می دهد،در این شکل دایره های مشکی نشان دهنده عدد ۰ و دایره های سفید نشان دهنده ی عدد ۱ هستند.

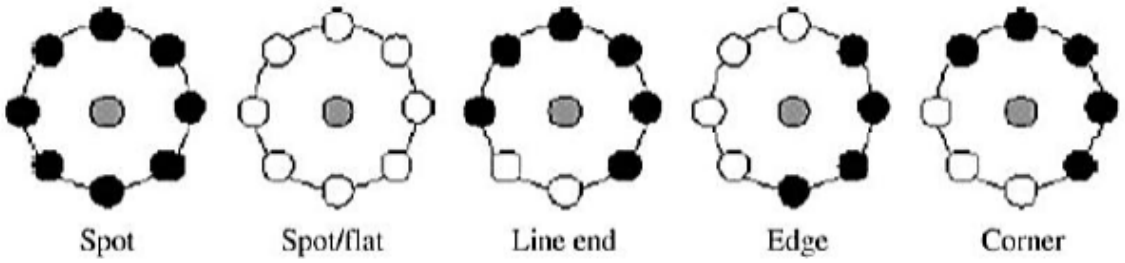
[۱] Feature Detector

[۲] uniform

[۳] bin



شکل ۳ـ۳: الگوی یکنواخت در عملگر LBP با شعاع ۱ و تعداد همسایگی ۸ [۳۰].



شکل ۳ـ۴: مثال هایی از مفهوم ریزالگوها در LBP

جدول ۳ـ۱: لیستی از آخرین تغییرات الگوی دودویی محلی

گزینش	تغییرات	ویژگی ها	منبع
بهبود قدرت جداسازی	LBP Improved	در نظر گرفتن تاثیرات پیکسل های مرکزی،ارائه الگوهای ساختاری کامل	[۵]،[۳۵]،[۲۲]
	LBP Hamming	ترکیب الگوهای غیریکنواخت در الگوهای یکنواخت	[۴۷]
	LBP Extended	جداسازی الگوهای محلی یکسان،افزایش ابعاد	[۱۷]،[۱۵]
	LBP Completed	در نظر گرفتن علامت و مقدار نواحی محلی	[۱۱]
افزایش تحمل در برابر تغییرات	Patterns Ternary Local	مقاومت نسبت به تغییرات سطح خاکستری	[۳۹]
	LBP Soft	مقاومت نسبت به تغییرات یکنواخت سطح خاکستری، افزایش پیچیدگی محاسباتی	[۳]
انتخاب همسایه	LBP Elongated	استخراج اطلاعات ناهمگن ،مقاومت نسبت به چرخش	[۲۶]
	LBP Multi-Block	گرفتن اطاعات ساختار کوچک و کلان	[۴۸]،[۲۵]
	LBP Patch Three/Four	کدگذاری نوع پچ اطلاعات متنی	[۴۵]
گسترش به سه بعدی	LBP ۳D	گسترش الگوی دودویی محلی به داده های سه بعدی	[۶]،[۳۲]
	(LBP-TOP) LBP Volume	توصیف بافت پویا،افزایش ابعاد توصیفگر	[۵۱]،[۵۰]
ترکیب با دیگر روش ها	Wavelet Gabor and LBP	ترکیب فواید گابور و الگوی دودویی محلی،افزایش زمان محاسباتی و ابعاد توصیفگر	[۴۹]،[۱۲]،[۳۶]،[۴۶]،[۴۰]،[۳۷]،[۲۴]
	SIFT and LBP	به کارگیری فواید توصیفگر SIFT ،کاهش طول ابعاد توصیفگر	[۱۳]،[۱۶]
	Fourier Histogram LBP	مقاومت نسبت به چرخش	[۲]

۳ـ۲ـ۱ تغییرات اخیر الگوی دودویی محلی

متد LBP در سال های اخیر به منظور بهبود عملکرد در کاربردهای مختلف دستخوش تغییراتی گردیده است.این تغییرات بر جنبه های مختلف الگوی دودویی محلی اصلی تمرکز کرده است .

- بهبود قدرت جداسازی آن
- افزایش تحمل در برابر تغییرات [۴]
- نحوه ی انتخاب همسایه
- گسترش به داده های سه بعدی
- ترکیب آن با دیگر روش ها

در جدول ۳ـ۱ لیستی از آخرین تغییرات الگوی دودویی محلی آورده شده است.

[۴] robust

۳ـ۳ ال اس اچ

قبل از این که مفهوم ال اس اچ (هش حساس به همسایگی) را توضیح بدهیم ابتدا به مفهوم هش می پردازیم:

۳ـ۳ـ۱ هش

با ساختن یک جدول هش، (یعنی ساختمان داده اي که به ما اجازه می دهد به سرعت از یک نماد به یک مقدار برسیم)، زمانی که یک ورودي می گیریم، می توانیم یک تابع شبه تصادفی از آن نماد را استفاده کنیم که به ما یک عدد صحیح بدهد که به عنوان اندیس جدول هش استفاده کنیم. بدین صورت نمادي با تعداد زیادي حروف و بیت به یک اندیس نسبتاً کوچک از جدول هش نگاشت می شود. برخورد زمانی رخ می دهد که دو نقطه به یک مقدار هش شوند. یک جدول هش اي که خوب طراحی شده باشد داراي زمان جستجوي $O(۱)$و حافظه ي $O(N)$ خواهد بود که N تعداد اقلام جدول است. علاوه بر این، یک تابع هش مناسب دو نماد که به هم نزدیک هستند را به دو جعبه ي مختلف هش می کند. این کار باعث می شود که جدول هش براي یافتن «تطبیقهاي دقیق» مناسب باشد. براي یافتن تطبیق هاي نزدیک (یا تقریبی) به صورت کارا، می توان از ال اس اچ استفاده کرد که در ادامه توضیح می دهیم.

۳ـ۳ـ۲ هش حساس به همسایگی

خانواده ای از تابع های هش $h : \{۰,۱\}^d \to U$ را $(r_۱, r_۲, P_۱, P_۲)$ حساس گوییم[۵] که برای $(r_۱ < r_۲, P_۱ < P_۲)$ اگر برای هر $q, p \in \{۰,۱\}^d$

- اگر $D(p,q) \leq r_۱$ آنگاه $pr[h(q) = h(p)] \geq P_۱$
- اگر $D(p,q) > r_۲$ آنگاه $pr[h(q) = h(p)] \leq P_۲$

مفهوم ال اس اچ می تواند برای هر فضای متری D تعریف شود.ولی ،برای فضای همینگ،خانواده ی ال اس اچ بسیار ساده است،کافی است که همه ی تابع های h_i،$i = ۱, ..., d$ را که $h_i(p) = p_i$ ، $p \in \{۰,۱\}^d$ باشد برداریم.از آن جا که $Pr[h(q) = h(p)] = ۱ - D(p,q)/d$ می توان نتیجه گرفت که این خانواده از توابع هش حساس است.

یک هشینگ حساس به همسایگی استاندارد ،نگاشت ها را به صورت تصادفی انجام می دهد.هر یک از توابع

[۵] Hash functions

هش به صورت مستقل با انتخاب یک بردار تصادفی w_k از یک توزیع گاوسی با میانگین صفر و کوواریانس مشخص تولید می شود.سپس تابع هش به صورت زیر در نظر گرفته می شود:

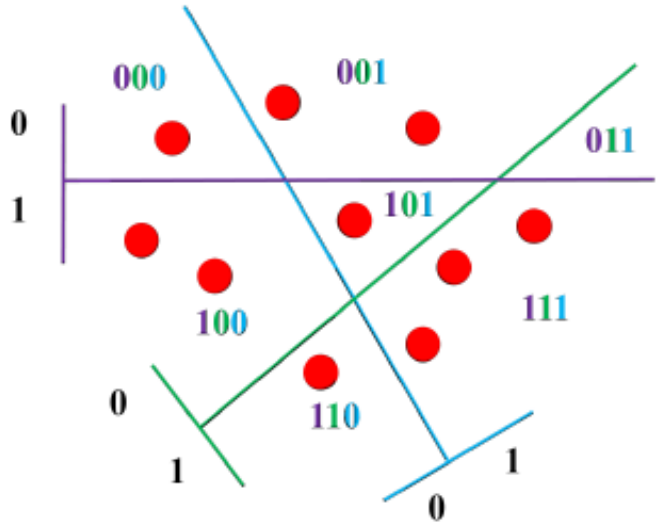
$$h_k(x) = sign(xw_k) \qquad (۴ـ۳)$$

$$Y = sign(XW) \qquad (۵ـ۳)$$

اگر هر بیت در کد با استفاده از نگاشت خطی تصادفی و آستانه تصادفی محاسبه گردد،آن گاه به دنبال افزایش تعداد بیت ها ،فاصله ی همینگ میان کدها به سمت کسینوس زاویه ی میان آنها همگرا می شود.این بدان معناست که فاصله ی همینگ میان کدها به فاصله ی اقلیدسی میان آیتم ها نزدیک می شود.بنابراین این متد توپولوژی داده های ورودی را به دنبال افزایش تعداد بیت ها حفظ می کند. نگاشت تصادفی در LSH و دیگر متدها استفاده می شود.این متدها مستقل از دیتاست می باشند و هیچ فرض پیشین درباره ی توزیع داده ها در نظر نمی گیرندولی برای دستیابی به دقت بازیابی خوب نیاز به تعداد کدهای بیشتری دارند به طور کلی یادگیری توابع دودویی هش مزایای زیر را دارد:

- کدها بیشتر فشرده می شوند.
- کلاس های عمومی تر از معیارهای شباهت می تواند حفظ شود.

در شکل ۳ـ۵ نمونه ای از نگاشت دودویی دو بعدی برای LSH نمایش داده شده است:
اگر خانواده ی ال اس اچ با فاصله ی بزرگی بین $P_۱$ و $P_۲$ داشته باشیم،مسئله ی $(r_۲/r_۱, r_۱) - NN$ را می توان به صورت زیر حل کرد.هنگام پیش پردازش ،تمام نقاط داده مانند P به جعبه ی h(p) هش می شوند .برای پاسخ گویی به نقطه ی ورودی q ،الگوریتم تمام نقاطی که در جعبه ی h(q) قرار دارند را بازیابی می کند و چک می کند که کدام به q نزدیک تر است.اگر فاصله ی بین $P_۱$ و $P_۲$ به اندازه ی کافی بزرگ باشد،می توان نشان داد که این روش زمان زیر خطی برای پاسخ گویی به نقطه ی ورودی نیاز دارد.متاسفانه فاصله ای که توسط خانواده ی ال اس اچ به وجود می آید به اندازه ی کافی بزرگ نیست،ولی با به هم چسباندن چند تابع هش می توان این فاصله را بزرگ کرد.
برای k و L ای که بعدا مشخص می شود،L تابع را $g_j(q) = (h_{۱,j}(q), ..., h_{k,j}(q))$ ،که در آن $h_{t,j}(۱ \leq t \leq ۱, ۱ \leq j \leq L)$ به صورت مستقل و یکنواخت از H انتخاب شده اند را در نظر می گیریم.در زمان پیش پردازش، هر $p \in P$ (نقاط مجموعه داده) را در جعبه ی $g_j(p)$ برای $j = ۱, ..., L$ قرار می دهیم.



شکل ۳ـ۵: نگاشت دودویی دوبعدی .LSH

برای پاسخ گویی به نقطه ی ورودی در،q جعبه های $g_1(q), ..., g_L(q)$ جستجو می کنیم.دو راه حل زیر وجود دارد:

- جست وجو را پس از یافتن اولین L' نقطه متوقف کنیم.
- جست وجو را تا زمان بازیابی همه ی نقاط از جعبه ها ادامه دهیم.

نشان داده شده است که در اولین راه حل،اگر $L' = ۳L$ باشد می تواند مساله همسایه ی نزدیک c– تقریبی تصادفی را با پارامترهای R و $\delta < ۱$ حل کند.برای این کار باید مقدار L را $\Phi(n^\rho)$ قرار دهیم که در آن $\rho = \frac{ln۱/p_۱}{ln۱/p_۲}$ ،توجه کنید که این باعث می شود الگوریتم در زمان متناسب با n^ρاجرا شود که زمانی که $P_۱ > P_۲$ باشد نسبت به n زیر خطی است.

از طرف دیگر، در عمل، راه حل دوم پرکاربردتر به نظر می رسد. مخصوصاً به خاطر این که در آن نيازي به مشخص کردن پارامترهاي اضافی نیست. علاوه بر این، هیچ دلیلی براي مشخص کردن نیست، زیرا که الگوریتم قادر است مسأله ي همسايه ي نزديک دقيق را حل کند، که در ادامه توضیح می دهیم. می توان نشان داد که راه حل دوم می تواند مسأله ي همسايه ي نزديک دقيق تصادفی را براي پارامتر R حل کند. مقدار δ وابسته به انتخاب پارامترهاي k و L می باشد، براي هر δ می توان پارامترهاي k و L را طوري تعيين کرد که احتمال خطا کوچکتر از δ باشد. پيچيدگی زمانی نيز وابسته به k و L است: می توان در بدترين حالت $\phi(n)$ باشد، ولی براي بسياري از مجموعه هاي داده و انتخاب هاي پارامترها ميتواند در زمان زيرخطی اجرا شود. کیفیت تابع ال اس اچ توسط پارامتر ρ تعیین می شود.نشان داده شده است که اگر فاصله بر اساس نرم $L_۱$

اندازه گیری شود،آن گاه خانواده ی توابع (R, cR, P_1, P_2)ـ حساس وجود دارد به طوری که $\rho = 1/c$.

۳ـ۳ـ۳ حل مسأله ي نزديک ترين همسايه ي تقريبی

الگوریتم بالا را می توان با تغییرات کمی براي حل مسئله ي نزديک ترين همسايه ي تقريبی به کار برد،بدون این که پیچیدگی الگوریتم تغییر پیدا کند. مساله ی نزدیک ترین همسایه ی تقریبی تحت نرم L_p ،برای $p \in [1, 2]$،می تواند به همین مسئله در فضای همینگ کاهش پیدا کند(تبدیل شود).این کار به راحتی برای L_d^1 انجام می شود.اگر فرض کنیم که مختصات نقاط p دارای مقادیر $\{1, ..., M\}$ باشند،اگر $U(p) = (U(p_1), ..., U(p_d))$ را تعریف کنیم که در آن $U(x)$ یک رشته از x ۱ وM-x ۰ دنبال می شود باشد،آن گاه خواهیم داشت $\|p - q\|_1 = D(U(p), U(q))$.در حالت کلی،M می تواند بسیار بزرگ باشد،ولی می تواند مسئله ی نزدیک ترین همسایه ی تقریبی به $d^{o(1)}$ کاهش یابد.بنابراین می توانیم $(r, c) - NN$ را تحت L_1 به $(r, c) - NN$ در فضای همینگ کاهش دهیم. برای به دست آوردن الگوریتم برای نرم L_p زمانی که $p \in [1, 2]$، باشد از این موضوع استفاده می کنیم که $L_1^{o(d)}$ با تغییرات محدودی [۶] جاسازی [۷] شود.از طرف دیگر برای p=۲ ،می توان مساله را مستقیم در فضای اقلیدسی حل کرد.

۳ـ۳ـ۴ نقش نگاشت هاي تصادفی در ال اس اچ

ال اس اچ مبتنی بر این ایده ي ساده است که اگر دو نقطه به هم نزدیک هستند، آن گاه پس از نگاشته شدن [۸] ، این نقاط هم چنان به هم نزدیک می مانند. براي شرح بیشتر مطلب از توضیح نگاشت تصادفی آغاز می کنیم که یک نقطه را از یک فضاي با بعد بالا به یک زیرفضاي با بعد پایین نگاشت می کند.

نگاشت هاي تصادفی: ضرب داخلی

در قلب ال اس اچ، نگاشت اسکالر(یا همان ضرب داخلی) $h(V) = V^T.X$ می باشد که V نقطه ای ورودی در فضای با بعد بالا بوده و X برداری است که مولفه های آن به صورت تصادفی از یک توزیع گاوسی مثلا $N(0, 1)$ انتخاب شده است. این نگاشت اسکالر سپس به چند جعبه ي هش گسسته سازي می شود، با این هدف که اقلام نزدیک در فضاي اصلی به یک جعبه هش شوند. تابع هش اي که نتیجه می شود عبارت است

[۶] bounded distortion

[۷] embedding

[۸] projection

از:

$$h^{X+b}(V) = \lceil \frac{X^T.V + b}{w} \rceil \qquad (۶ـ۳)$$

که در آن w عرض هر جعبه ي نتيجه ي گسسته سازي است و يک متغير تصادفی است که بين ۰ و w توزيع يکنواخت دارد و کمک می کند تا خطاي گسسته سازي بدون اضافه کردن به هزينه ي عملکرد، راحت ترتحليل شود. براي آن که عملگر نگاشت براي کار ما مناسب باشد، می بايست نقاط نزديک را به مکان هاي نزديک به هم تصوير کند. به عبارت ديگر:

- برای هر دو نقطه p و q در R^d که نزديک به هم هستند،احتمال بالای $P_۱$ای وجود داشته باشد که در جعبه ی يکسان هش شوند:

$$P_H[h(p) = h(q)] \geq P_۱ for \|p - q\| \leq R_۱ \qquad (۷ـ۳)$$

- برای هر دو نقطه p و q در R^d که از هم دور هستند،احتمال کم $P_۲ < P_۱$ ای وجود داشته باشد که در جعبه ی يکسان هش شوند:

$$P_H[h(p) = h(q)] \leq P_۲ for \|p - q\| \geq cR_۱ = R_۲ \qquad (۸ـ۳)$$

در اين تعاريف $\|.\|$ نرم $L_۲$ است و $R_۲ > R_۱$ است.توجه شود که به واسطه ی خطی بودن ضرب داخلی،اختلاف بين دو نقطه ی هش شده $\|h(p) - h(q)\|$ اندازه ای دارد که توزيع آن با $\|p - q\|$ متناسب است و بنابراين $P_۱ > P_۲$.

هاي تصادفی: k ضرب داخلی

براي بزرگ کردن اختلاف بين $P_۱$ و $P_۲$ ، ضرب داخلی را به صورت موازي انجام می دهيم. اين کار نسبت احتمالی را که نقاط با فاصله هاي متفاوت به يک جعبه هش شوند را افزايش می دهد زيرا $(P_۱/P_۲)^k >$ $P_۱/P_۲$. نگاشت نهايی با انجام k ضرب داخلی مستقل براي تبديل نقطه ي ورودي v به k عدد حقيقی به دست می آيد. همان طور که براي نگاشت اسکالر انجام شده، k ضرب داخلی را با هدف اين که نقاط مشابه به جعبه ي يکسان هش شوند، گسسته سازي می کنيم.

افزايش عرض گسسته سازيw باعث افزايش تعداد نقاطی می شود که به هر جعبه هش می شوند. براي رسيدن به نتيجه ي نهايی همسايه ي نزديک می بايست يک جستجوي خطی روي همه ي نقاطی که در جعبه ي يکسان با نقطه ي ورودي قرار می گيرند انجام دهيم. بنابراين تغييرw يک موازنه بين جدول بزرگ تر با جستجوي خطی کوچک تر، يا يک جدول کوچک تر با تعداد بيشتري نقاط در جستجوي نهايی.
براي هر مجموعه از k ضرب داخلی، اگر نقطه ورودي و نزديک ترين همسايه در يک جعبه قرار بگيرد، ما موفق شده ايم. اين با احتمال $P_۱^k$ اتفاق می افتد که با استفاده از ضرب داخلی هاي بيشتر کاهش پيدا می کند. براي کاهش تأثير گسسته سازي نادرست در هر نگاشت، L نگاشت مستقل تشکيل می دهيم و همسايه ها را با کمک همه ي آنها به دست می آوريم. اين بدين دليل است که احتمال اينکه نزديک ترين همسايه واقعی براي همه ي نگاشت ها در جعبه ي نادرست قرار گيرد بسيار کم است. با افزايش L می توانيم نزديک ترين همسايه ي واقعی را با هر احتمالی به اندازه ي دلخواه بالا، به دست آوريم.

۳ـ۳ـ۵ پياده سازي هش

عمل نگاشت و گسسته سازي هر نقطه را در جعبه اي قرار می دهد که با k انديس صحيح مشخص می شود. با توجه به اين که اين فضاي k— بعدي تنک [۹] است، می توانيم از روش هاي معمول (و دقيق) هشينگ استفاده کنيم تا به صورت کارا بفهميم که چه زمانی نقاط در يک جعبه افتاده اند. ما در اينجا روش $E^۲LSH$ را شرح می دهيم ولی روش هاي پيچيده تر نيز وجود دارد.
جستجوي ساده براي يافتن نقاطی از مجموعه داده که در جعبه ي يکسان با نقطه ورودي قرار دارند ميتواند $O(logN)$ زمان صرف کند ولی ما اين را به $O(۱)$ کاهش می دهيم چون از توابع هش معمولی استفاده می کنيم. ابتدا از يک تابع هش معمولی استفاده می کنيم تا نگاشت k— بعدي گسسته سازي شده را به يک انديس خطی تبديل کنيم

$$T_۱ = (\Sigma_i H_i k_i) mod P_۱ \qquad (۹ـ۳)$$

که در آن H_i وزن هاي صحيح هستند و $P_۱$ سايز جدول هش است (يک عدد اول بزرگ). هدف اين جدول هش اين است که هر نقطه ی جداگانه را در فضای k— بعدی در يک خانه ی جداگانه از $P_۱$ خانه جدول قرار دهد،تا اين که برخوردی بين نقاط متفاوت رخ ندهد(مقدار يکسان هش به آنها نسبت داده نشود)، با آن که جدول هش خوب اقلام را به خوبی به طور يکنواخت توزيع می کند، ولی با کوچک شدن جدول (براي آن که

[۹] sparse

بتوان آن را در حافظه ي اصلی جاي داد) ریسک برخورد نقطه هاي نامرتبط افزایش می یابد.
برای حل این مشکل هش دومی به نام $T_۲$ نقاط k– بعدی را به کار می گیریم تا بررسی کنیم که نقاطی که از جدول هش بازیابی شده اند واقعا با نقاط ورودی تطبیق دارند یا خیر. $T_۲$ همان شکل $T_۱$ را دارد ولی وزن ها وسایزهای مختلفی استفاده می کند.مقادیر دومین هش را در جعبه هایی که توسط اولین هش انتخاب شده اند ذخیره می کنیم.در زمان بازیابی،مقدار ذخیره شده را با اندیس جعبه مقایسه می کنیم تا تطبیق هاي واقعی را در صورت وجود بیابیم. با توجه به اینکه مقادیر هش دوم کوتاه هستند (مثلاً مقادیر ۱۶ ـ بیتی) مقایسه شان خیلی سریع تر از مقایسه ي k – بعدي نقاط اولیه می باشد. علاوه بر این، احتمال برخورد پی در پی تحت هر دوی $T_۱$ و $T_۲$ به شدت کاهش می یابد،حتی زمانی که جدول هش کوچک است.

۳ـ۳ـ۶ بررسی عملکرد

دقت

دقت ال اس اچ توسط احتمال این که نزدیک ترین همسایه ی واقعی را پیدا می کند تعیین می شود.برای تحلیل این موضوع،مفهوم توزیع s ـ مقاوم را معرفی می کنیم.توزیع s D ـ مقاوم است اگر برای هر مجموعه از متغیرهای تصادفی،$X_۱, ..., X_n$ مستقل با توزیع یکسان D و هر مجموعه از اعداد حقیقی $v_۱, ..., v_n$ متغیر تصادفی $\Sigma_i V_i X_i$ توزیع یکسان با توزیع متغیر تصادفی $(\Sigma_i |V_i|^s)^{۱/sX}$ داشته باشد که در آن X از D انتخاب شده است.برای s=۲ (یعنی همان نرم $L_۲$)،توزیع گاوسی s ـ مقاوم است.
برای تحلیل دقت نگاشت هایمان،توجه کنید که نگاشت های دو نقطه ی نزدیک p و q که با فاصله ی $u = \|p - q\|$ از هم جدا شده اند، همیشه به هم نزدیک خواهد بود.ولی به دلیل گسسته سازی ممکن است در دو جعبه ی متفاوت قرار گیرند.احتمال این که این دو نقطه با گسسته سازی به یک جعبه نسبت داده شوند عبارت است از:

$$p(u) = Pr_{a,b}[h_{a,b}(p) = h_{a,b}(q)] = \int_۰^w \frac{۱}{u} f_s(\frac{t}{u})(۱ - \frac{t}{w})dt \qquad (۱۰ـ۳)$$

که در آن f_s تابع چگالی احتمال H است.برای هر جعبه با عرض w، این احتمال با افزایش فاصله ی u کاهش می یابد. احتمال P_H را درفضای $L_۲$ با $R_۱ = w$ به صورت زیر حساب می کنیم:

$$P = ۱ - ۲F(-w/c) - \frac{۲}{\sqrt{۲\pi} w/c} \times (۱ - e^{-(w^۲/۲c^۲)}) \qquad (۱۱ـ۳)$$

در این جا$F(.)$ تابع چگالی احتمال تجمعی یک متغیر تصادفی گاسی است و c نسبت فاصله هاست اگر c=۱ قرار دهیم،$P_۱$ را به دست می آوریم.
احتمال این که نقطه ای خاص با نقطه ی ورودی در یک جعبه قرار گیرد عبارت است از $(P_۱)^k$،از طرف دیگر احتمال این که همه ی نگاشت ها نتوانند برخوردی بین نقطه ی ورودی و نزدیک ترین همسایه ی واقعی تولید کنند عبارت است از $(۱ - P_۱^k)^L$.اگر بخواهیم که احتمال این که ال اس اچ نتواند نزدیک ترین همسایه ی واقعی را پیدا کند بیش تر از δ نباشد،با داشتن یک مقدار برای k ،بایستی L حداقل مقدار زیر باشد:

$$L = \frac{[log\delta]}{log(۱ - P_۱k)} \qquad (۱۲\text{ـ}۳)$$

الگوریتم $E^۲LSH$ بهترین مقدار را برای k با آزمودن هزینه محاسبات برای نمونه های در مجموعه داده به دست می آورد.

سرعت

زمان لازم براي يافتن نزديک ترين همسايه، زمان T_g برای محاسبه ی نگاشت ها و هش کردن،به اضافه ی زمان T_c مورد نیاز برای جست وجو در جعبه ها برای برخوردها می باشد.با توجه به این که KL نگاشت وجود دارد،$O(nkL)T_g$ است،که در آن n بعد فضای داده ی اصلی است.از طرف دیگر T_c به صورت خطی با امید ریاضی تعداد برخوردها (یعنی $T_c = O(dLN_c)$ که در آن d میانگین تعداد نقاط در هر جعبه می باشد)افزایش می یابد. N_c،که امید تعداد برخوردهای یک نگاشت می باشد،توسط عبارت زیر داده شده است:

$$N_c = \sum_{q' \in D} p^k(\|q - q'\|) \qquad (۱۳\text{ـ}۳)$$

که در آن $p(.)$ احتمال کمک هر نقطه به ایجاد برخورد می باشدو D کل نقاط در پایگاه را نشان می دهد.به راحتی می توان دید که T_g متناسب با k افزایش می یابد،در حالی که T_c کاهش می یابد چون $P_k < P$ برای $P < ۱$ و $k > ۱$.

۳ـ۴ کرنل های ثابت به جابه جایی

کدگذاری دودویی حساس به همسایگی با استفاده از کرنل های ثابت به جابه جایی ویژگی هایی را استخراج می کند که تضمین شده است نگاشت دودویی داده اصلی نسبت به جابه جایی مقاوم هستند.ویژگی های تصادفی فوریه [۱۰] به صورت زیر تعریف می شوند:

$$\phi_{w,b}(x) = \sqrt{۲}\cos(w^T x + b) \qquad (۱۴ـ۳)$$

که در آن $w \sim P_k$ و $b \sim Unif[۰, ۲\pi]$.به عنوان نمونه برای کرنل $K(S) = e^{-\gamma\|s\|^۲/۲}$ و $w \sim Normal(۰, \gamma I)$.می توان نشان داد $E_{w,b}[\phi_{w,b}(x)\phi_{w,b}(y)] = K(x,y)$.کد باینری به صورت زیر محاسبه می شود:

$$sign(\phi_{w,b}(X) + t) \qquad (۱۵ـ۳)$$

که در آن t آستانه تصادفی و $t \sim Unif[-۱, ۱]$ می باشد.

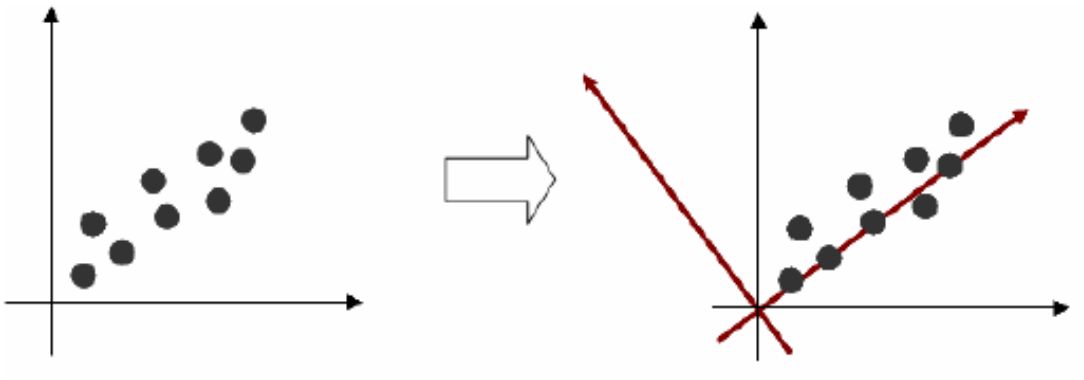
۳ـ۵ چندی سازی تکرار شونده

برای تشریح الگوریتم هشینگ چندی سازی تکرار شونده [۹]،[۱۰] نیاز به درک مفاهیم تکنیک آنالیز اجزای اصلی ، تحلیل همبستگی کانونیک و تجزیه مقادیر تکین می باشد بدین منظور در ابتدا به بیان این مفاهیم می پردازیم.

۳ـ۵ـ۱ تکنیک آنالیز اجزای اصلی

در این روش محورهای مختصات جدیدی برای داده ها تعریف شده و داده ها بر اساس این محورهای مختصات جدید بیان می شوند.اولین محور باید در جهتی قرار گیرد که واریانس داده ها ماکسیمم شود(یعنی در جهتی که پراکندگی داده ها بیشتر است).دومین محور باید عمود بر محور اول به گونه ای قرار گیرد که واریانس داده ها ماکسیمم شود.به همین ترتیب محورهای بعدی عمود بر تمامی محورهای قبلی به گونه ای قرار می گیرند که داده ها در آن جهت دارای بیشترین پراکندگی باشند در شکل ۳ـ۶ این مطلب برای داده های دو بعدی

[۱۰] Random Fourier features



شکل ۳ـ۶: انتخاب محورهای جدید برای داده های دو بعدی.

نشان داده شده است. روش PCA به نام های دیگری نیز معروف است مانند:

- (KLT) Transform Loeve Karhunen
- Transform Hotelling
- (EOF) Function Orthogonal Empirical

الگوریتم PCA

مرحله ۱ ـ انتخاب داده

مرحله ۲ ـ کم کردن میانگین از داده ها در این مرحله میانگین هر بعد از مقادیر آن بعد کم می شود تا میانگین داده ها در هر بعد صفر گردد.

مرحله ۳ ـ محاسبه ماتریس کواریانس

مرحله ۴ ـ محاسبه بردارهای ویژه و مقادیر ویژه در این مرحله بردارهای ویژه و مقادیر ویژه ماتریس کواریانس محاسبه می شود.طبق قضایای جبر خطی،یک ماتریس متقارن n×n دارای n بردار ویژه ی مستقل و n مقدار ویژه می باشد.مقادیر ویژه میزان پراکندگی داده ها در راستای بردار ویژه مربوطه را نشان می دهد.می توان گفت بردار ویژه ای که دارای بزرگترین مقدار ویژه است مولفه اصلی داده های موجود می باشد.

مرحله ۵ـ انتخاب مولفه ها و ساختن Vector Feature در این مرحله مفهوم کاهش ابعاد داده وارد می شود.بردار ویژه ای که در مرحله ی قبل به دست آمده است بر اساس مقادیر ویژه ی آنها از بزرگ به کوچک مرتب می شود(مقادیر ویژه ی ماتریس کواریانس همگی بزرگتر یا مساوی صفر هستند).بدین ترتیب مولفه ها ی داده ها از پراهمیت به کم اهمیت مرتب می شوند.در این مرحله به منظور کاهش داده ها ،مولفه های کم

اهمیت حذف می شود.البته این کار با از دست دادن مقدار کمی اطلاعات همراه است.کاری که در این مرحله باید انجام شود ایجاد یک Vector Feature است که در واقع ماتریسی از بردارها می باشد.این ماتریس شامل آن بردارهای ویژگی ای است که ما میخواهیم آنها را نگه داریم.

$$FeatureVector = (eig_1 eig_2 eig_3 ... eig_n) \qquad (۳ـ۱۶)$$

مرحله ۶ـ به دست آوردن داده های جدید در آخرین مرحله از PCA فقط باید ترانهاده ی ماتریس Feature Vector که در مرحله ی قبل به دست آمده را در ترانهاده ی داده های نرمالسازی شده ضرب شود.

$$FinalData = RowFeatureVector \times RowDataAdjust \qquad (۳ـ۱۷)$$

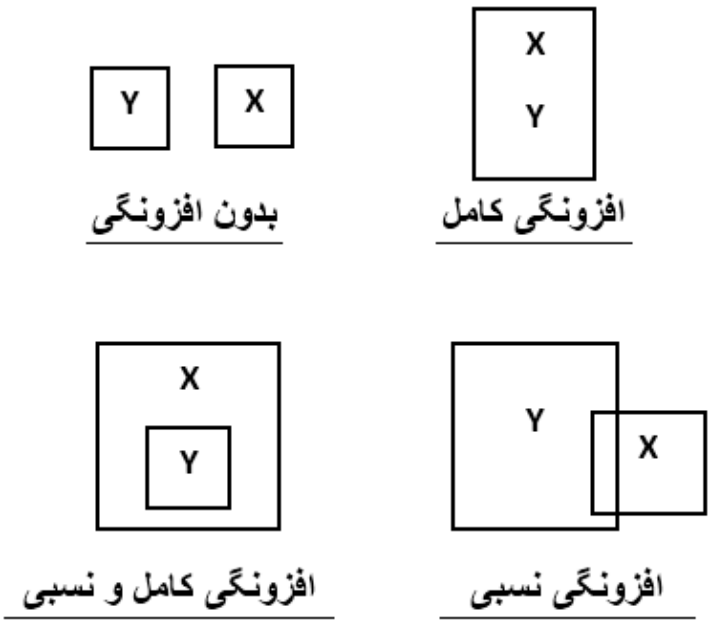
که RowFeatureVector ماتریسی است که بردارهای ویژه در سطرهای آن به ترتیب مقادیر ویژه از بالا به پایین قرار گرفته اند و RowDataAdjust ماتریسی است که شامل داده هایی است که میانگین هر بعد از آن بعد کم شده است.در این ماتریس ،داده ها در ستون های آن ذخیره شده و هر سطر آن مربوط به یک بعد است.

۳ـ۵ـ۲ تحلیل همبستگی کانونیک

تحلیل همبستگی کانونیک یکی از متداول ترین روشهای تحلیل چند متغیره بوده و هدف آن تعیین ارتباط خطی بین متغیرهای چند بعدی است.در این روش در نظر گرفتن یک مجموعه از متغیرها به عنوان متغیرهای مستقل و مجموعه ی دیگر به عنوان نتغیرهای وابسته می تواند بسیار مفید باشد.از جمله امتیازات روش تحلیل همبستگی کانونیک در مقایسه با تحلیل همبستگی معمولی [۱۱] این است که روش OCA به سیستم مختصاتی که در آن تعریف شده است ،وابسته می باشد.این بدان معنی است که حتی اگر رابطه ای قوی بین دو متغیر چند بعدی وجود داشته باشد،این ارتباط ممکن است با انتخاب سیستم مختصات استفاده شده در روش OCA دیده نشود.در حالی که در یک سیستم مختصات [۱۲] دیگر،این رابطه خطی میزان همبستگی بالایی را ارائه دهد.روش CCA سیستم مختصاتی را می یابد که در آن میزان همبستگی دارای مقدار بهینه است.در بعضی از دسته داده های چند متغیره،متغیرها به طور طبیعی به دو گروه مجزا مثل $X_1, X_2, X_3, ..., X_p$ و $Y_1, Y_2, Y_3, ..., Y_p$ تقسیم می شوند.هدف اصلی CCA ،ساختن دو گروه جدید از مولفه ها به صورت U=aX

[۱۱] Ordinary Correlation Analysis (OCA)

[۱۲] Coordinating System



شکل ۳ـ۷: تصویری از ضرایب افزونگی حاصل از تحلیل همبستگی کانونی .

و V=bY است که ترکیب خطی از متغیرهای اولیه هستند. همبستگی کانونی با دو مجموعه از داده ها آغاز می شود که شامل بردارهایی از مشاهدات انجام شده بر کلیه متغیرها می باشد. ، هدف همبستگی کانونی، با ایجاد X بعنوان یک بردار m بعدی از متغیرهای پیش بین و Y بعنوان یک بردار P بعدی از متغیرهای ملاک، دستیابی به یک ترکیب خطی از متغیرهای پیش بین است که دارای حداکثر همبستگی با یک ترکیب خطی از متغیرهای ملاک می باشد. تحلیل همبستگی کانونی اندازه رابطه بین دو مجموعه از متغیرها را با ضرایب افزونگی [۱۳] تعیین می کند. ضرایب افزونگی، درجه همپوشانی بین دو مجموعه متغیر را نشان می دهد؛ به بیان دقیق تر ضرایب افزونگی، نمایه میانگین درجه انحراف در یک مجموعه متغیر می باشد که با متغیرهای کانونی در سایر مجموعه ها سهیم یا قابل پیش بینی از آنها می باشد.

استفاده از یک مجموعه متغیر برای پیش بینی مجموعه دوم از متغیرها بیانگر این است که مجموعه دوم با وجود مجموعه اول زائد [۱۴] است. یک شاخص نامتقارن به این معنا است که ضریب افزونگی یک مجموعه از متغیرها عموما با ضریب افزونگی مجموعه دیگری از متغیرها برابر نمی باشد. ضرایب افزونگی یا بطور مجزا یا بصورت ادغام شده در توابع کانونی آزمون می شوند. شکل ۳ـ۷ برای درک بهتر، تصویری از ضرایب افزونگی حاصل از تحلیل همبستگی کانونی را نشان می دهد. بنابراین همانطور که پیشتر ذکر شد تحلیل

[۱۳] Redundancy Coefficients

[۱۴] Redundant

همبستگی کانونی اندازه رابطه بین دو مجموعه از متغیرها را با ضرایب افزونگی (Rd) تعیین می کند. ضرایب افزونگی، درجه همپوشانی بین دو مجموعه متغیر را نشان می دهد؛ به بیان دقیق تر ضرایب افزونگی، نمایه میانگین درجه انحراف در یک مجموعه متغیر می باشد که با متغیرهای کانونی در سایر مجموعه ها سهیم یا قابل پیش بینی بوسیله آنها می باشد.

۳ـ۵ـ۳ تجزیه مقادیر تکین

یکی از مهم ترین ابزارهای تفکیک و تجزیه در تسهیل حل دستگاه های بزرگ خطی،روش تجزیه مقادیر منفرد است.این روش،خود تعمیمی از روش تجزیه مقادیر ویژه ی اپراتورهای مربعی است.با این تفاوت که در این روش هر ماتریس با هر بعد را می توان به حاصل ضرب سه ماتریس که یکی از آن ها قطری و دو ماتریس دیگر نیز ماتریس هایی اورتوگونال و یا معکوس پذیر می باشد،تبدیل نمود.به عبارتی دیگر در این روش برای هر ماتریس با هر بعد دو ماتریس یافت می گردد،به طوری که با اعمال دو ماتریس یکی از چپ و دیگری از راست،یک ماتریس قطری حاصل می شود.از این روش در زمره ی روشهای قطری سازی [۱۵] نیز می باشد. اگر معادله گسسته شده مساله به صورت زیر باشد:

$$Ax = bA \in R^{m\times n}, x \in R^n, b \in R^m \qquad (۱۸ـ۳)$$

آن گاه یکی از ابزارهای کار با این مساله روش تجزیه مقادیر تکین ماتریس A است.برای آسانی کار فرض می نمائیم که $m \geq n$ باشد.در این حالت تجزیه مقادیر منفرد ماتریس A به صورت ذیل خواهد بود:

$$A_{mn} = U_{mm}\Sigma_{mm}V^T_{nn} = \sum_{i=۱}^{n}(u_i\delta_i v_i^T) \qquad (۱۹ـ۳)$$

در معادله ۳ـ۱۹ $U_{mm} = u_۱, u_۲, .., u_m$ و $V_{mm} = v_۱, v_۲, ..., v_m$ ماتریس هایی با ستون های ارتونرمال است،

$$V^TV = I_{nn}, U^TU = I_{mm} \qquad (۲۰ـ۳)$$

[۱۵] Bi-diagonalization

ستون های ماتریس های U و V را به ترتیب بردارهای منفرد چپ و راست ماتریس A می نامند.عناصر قطر اصلی ماتریس $\Sigma_{mn} = [\sigma_i \delta ij]_{.\leq i \leq m, \cdot \leq j \leq n}$ غیر منفی بوده و غالبا به ترتیب ذیل قرار می گیرند:

$$\sigma_۱ \geq \sigma_۲ \geq ... \geq \sigma_n \geq ۰ \qquad (۲۱\_۳)$$

مقادیر قطر اصلی Σ ،مقادیر منفرد ماتریس A و نسبت $\sigma_۱/\sigma_n$ عدد شرط ماتریس نامیده می شود،از روابط $A^TA = V^T\Sigma^T\Sigma V$ و $AA^T = U\Sigma^T\Sigma U^T$ دیده می شود که ماتریس A به شدت به تجزیه مقادیر ویژه ماتریس های A^TA و AA^T وابسته است.این وابستگی یکتایی تجزیه مقادیر منفرد را نشان می دهد که از یکتایی تجزیه مقادیر ویژه ماتریس های متقارن نتیجه می گردد. در ارتباط بامسائل بدوضع گسسته دومشخصه اصلي براي تجزيه مقادير منفرد اغلب يافت مي گردد: مقادير منفرد به تدريج بدون گسستگي به سمت صفر میل میکند و افزایش ابعاد A ، تعداد مقادیر منفرد کوچک را افزایش خواهد داد. در حالی که $_i$ با افزایش اندیس کاهش می یابد بردارهای منفرد u_i, v_i نیز با افزایش i مرتبا تغییر علامت می دهند. هم چنین با توجه به رابطه ی اصلی تجزیه مقادیر منفرد در ابتدای بخش روابط زیر نیز حاصل می گردد:

$$\begin{cases} Av_i = \delta_i u_i, & \|Av_i\|_۲ = \sigma_i. \\ A^Tu_i = \delta_i v_i, & \|Au_i\|_۲ = \sigma_i. \end{cases} \quad i = ۱, ۲, ..., n \qquad (۲۲\_۳)$$

۳ـ۵ـ۴ الگوریتم

این روش یکی دیگر از طرح های ساده و کارآمد برای یافتن مناسبترین چرخش برای داده هایی به مرکز صفر به منظور به حداقل رساندن خطای کواتیزاسیون به هنگام نگاشت داده ها به رئوس ابرمکعب دودویی می باشد. به منظور یافتن کدهایی با بیشترین واریانس و ناهمبستگی میان آن ها ،بردارهای ورودی با یک روش بدون ناظر مانند تکنیک آنالیز اجزای اصلی [۱۶] نگاشت می شوند اگرچه می توان از تکنیک های با ناظر مانند CCA [۱۷] نیز استفاده نمود.

اگر $W \in R^{d \times q}$ ماتریس ضرایب باشد که با روش PCA به دست آمده است آن گاه هر بیت $k = ۱, ..., q$ می تواند با رابطه ی زیر محاسبه نمود:

$$h_k(x) = sgn(xw_k) = sgn(v) \qquad (۲۳\_۳)$$

[۱۶] PCA

[۱۷] Canonical correlation analysis

کلیه ی فرآیند رمزنگاری [۱۸] به صورت:

$$Y = sgn(XW) = sgn(V) \qquad (۲۴ـ۳)$$

اگر W را ه حل بهینه باشد آن گاه WR ، که در آن R یک ماتریس متعامد q×q است نیز بهینه می باشد.بنابراین داده های نگاشت شده V=XR نیز به صورت متعامد تبدیل می شوند.ITQ به صورت متعامد داده ها را تبدیل می کند تا خطای چندی سازی را به حداقل برساند.

$$Y = sgn(XWR) = sgn(VR) \qquad (۲۵ـ۳)$$

فرض کنید $v \in R^d$ یک بردار در فضای نگاشت باشد.بسیار ساده است که نشان داده شود $sgn(v)$ یکی از رئوس ابر مکعب $\{-۱, ۱\}^q$ می باشد که در فضای اقلیدسی به v نزدیک است.
فقدان چندی سازی اختلاف میان بردار v و نگاشت آن به کد دودویی در ابرمکعب $\{-۱, ۱\}^q$ می باشد:

$$\|sgn(v) - v\|_۲ \qquad (۲۶ـ۳)$$

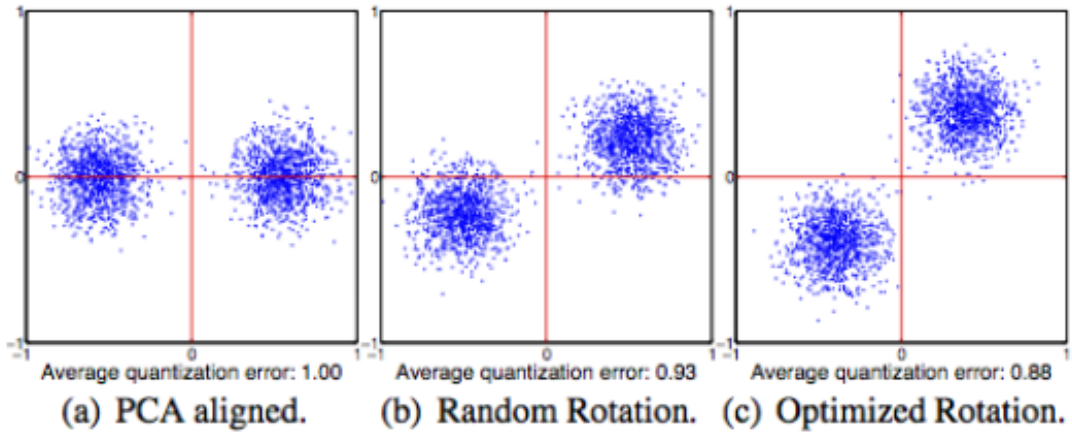
در صورتی که مقدار این فقدان کمتر باشد کدهای باینری بهتری تولید می گردد بنابراین در این روش به دنبال یافتن چرخش متعامدی می باشیم به گونه ای که نقاط نگاشت شده به کدهای باینری نظیرشان نزدیک باشند.

$$min(Q(Y, R)) \qquad (۲۷ـ۳)$$

$$Q(Y, R) = \|Y - VR\|_F \qquad (۲۸ـ۳)$$

شکل ۳ـ۸ نمایش از عملکرد الگوریتم ITQ را نشان می دهد ایده ی اصلی این روش کوانتیزه کردن هر یک از نقاط داده به نزدیک ترین رئوس ابرمکعب $(\pm ۱, \pm ۱)$ می باشد.در شکل اول بردارهای x و y مطابق با جهت PCA داده ها می باشند.کوانتیزاسیون نقاط داده در یک خوشه را به رئوس متفاوت نسبت می دهد.شکل دوم چرخش تصادفی داده ها در این روش واریانس داده ها متعادل شده است و خطای چندی سازی کمتر شده

[۱۸] encoding



شکل ۳ـ۸: تصویری از عملکرد الگوریتم کواتیزاسیون تکرارشونده [۱۰].

است در شکل سوم چرخش بهینه با استفاده از ITQ به دست آمده و خطای چندی سازی کمتر شده است . در این جا کار را با اعمال الگوریتم PCA روی نقاط داده آغاز می کنیم و مساله یادگیری یک کد باینری خوب و مناسب را به صورت کمینه کردن خطای چندی سازی داده های تصویر شده به هر یک از رئوس ابرمکعب فرمول می کنیم. در این بخش به مساله یادگیری کدهای باینری بدون هیچ اطلاعات با نظارت به برچسب کلاس ها اشاره می کنیم در ابتدا یک کاهش ابعاد خطی به داده ها اعمال می گردد سپس چندی سازی دودویی را در فضای نتیجه اعمال می کنیم.

بیشینه کردن واریانس

ما به دنبال کدهای کارآیی هستیم که در آن واریانس هر بیت بیشینه است و بیت ها دو به دو ناهمبسته هستند ما می توانیم به این هدف با بیشینه کردن تابع هدف زیر دست یابیم:

$$\tau(w) = \sum_k var(h_k(x)) = \sum_k var(sgn(xw_k)) \qquad (۲۹ـ۳)$$

$$\frac{۱}{n} B^T B \qquad (۳۰ـ۳)$$

زمانی واریانس بیشینه می گردد که توابع کدگذاری دقیقا بیت های متوازن تولید کنند یعنی زمانی که دقیقا برای نیمی از داده ها $h_k(x) = ۱$ و برای نیمی دیگر از داده ها $h_k(x) = -۱$

شرط بیشینه کردن واریانس

یک تابع هش با بیشینه آنتروپی $H(h_k(x))$ باید واریانس مقادیر هش را بیشینه کند و بر عکس:

$$H(h_k(x)) \leftrightarrow maxvar(h_k(x)) \qquad (۳۱\_۳)$$

فرض می کنیم h_k احتمال انتساب مقدار $h_k(x) = ۱$ به نقطه داده برابر p باشد و احتمال انتساب $h_k(x) =$ ۱− برابر p-۱ مقدار آنتروپی $H(h_k(x))$ از رابطه ی

$$H(h_k(x)) = -p\log(p) - (۱-p)\log(۱-p) \qquad (۳۲\_۳)$$

خیلی ساده است که نشان دهیم زمانی مقدار آنتروپی بیشینه است که بخش بندی به صورت متعادل انجام شود $p = \frac{۱}{۲}$در ادامه نشان می دهیم بخش بندی متعادل منجر به بیشینه کردن واریانس بیت ها می شود.

$$E(x) = M = ۲P - ۱ \qquad (۳۳\_۳)$$

$$var[h_k(x)] = (h_k(x) - M)^۲ = ۴P(۱-P) \qquad (۳۴\_۳)$$

بیشینه مقدار واریانس در $p = \frac{۱}{۲}$ اتفاق می افتد که این مقدار به مفهوم بخش بندی متعادل است.

۳ـ۶ مرتب سازی نتایج بر اساس اهمیت هر بیت

برای بسیاری از متدهای هشینگ مانند LSH ، SH،... ItQH توابع هش مستقل می باشند .هر یک از توابع هش یک بیت از کد باینری را تولید می کند بنابراین k کد باینری نیز از هم مستقل می باشند.برای این k بیت ما نمی توانیم تصمیم گیری انجام دهیم که کدام یک از این کدها نسبت به دیگر کدها دارای اهمیت بیشتری است بنابراین هنگام محاسبه فاصله ی همینگ به هر یک از بیت هاوزن یکسان تعلق می گیرد[۸]. در واقع می توان هر یک از توابع هشینگ را مانند یک کلاسبند دو کلاسه در نظر گرفت در فرآیند کلاس بندی ،بخشی از تصاویر که به دسته ی +۱ تعلق می گیرد با کد باینری ۱ و دسته ی دیگر به ۱ ــ تعلق می گیرد با کد باینری ۰ نمایش داده می شود.

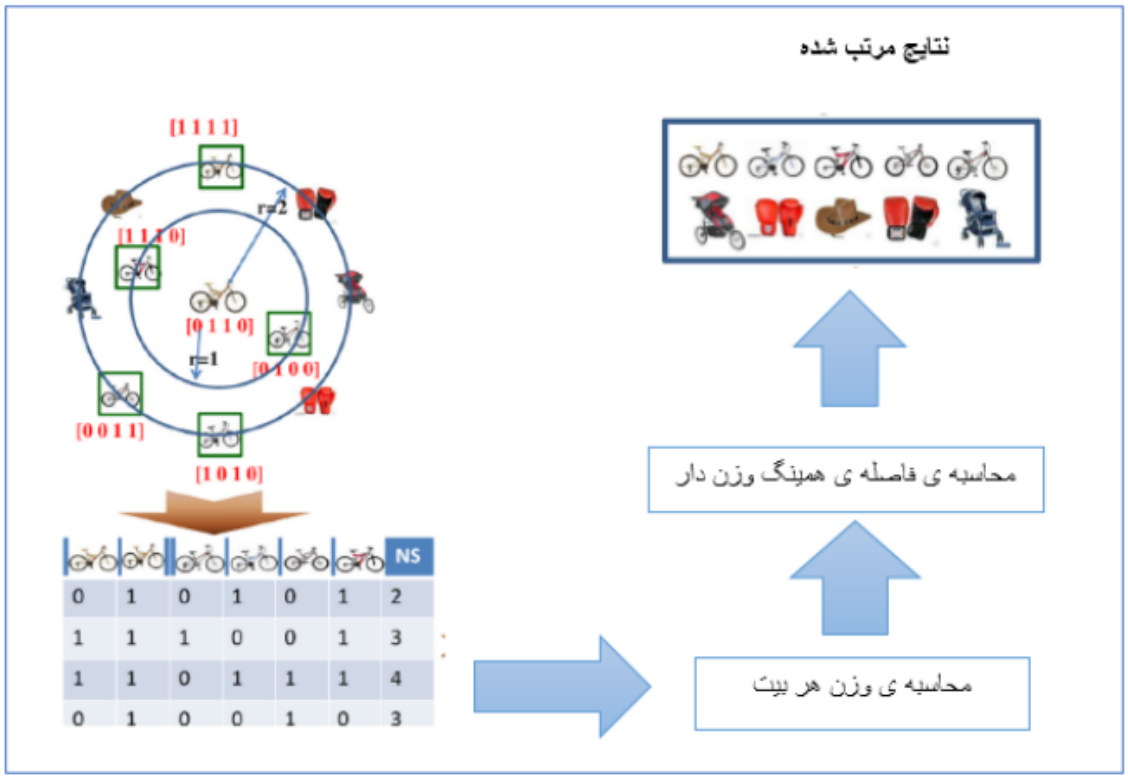
جدول ۳ـ۲: مثالی از k بیت کد باینری تصویر درخواستی با پنج تصویر

w	NS	$I_۵$	$I_۴$	$I_۳$	$I_۲$	$I_۱$	q	bit
$w_۱$	۱	۱	۰	۱	۱	۱	۰	$b_۱$
$w_۲$	۲	۱	۰	۰	۱	۰	۱	$b_۲$
$w_۳$	۳	۰	۱	۱	۰	۱	۱	$b_۳$
$w_۴$	۴	۰	۱	۰	۰	۰	۰	$b_۴$

اگر برای یک تابع هش تعداد زیادی از تصاویر مشابه دارای کد یکسان با آن در تصویر درخواستی باشند بنابراین این تابع هش از اهمیت بیشتری برخوردار است بنابراین باید وزن بیشتری به کد باینری متعلق به این تابع هش تخصیص داده شوددر ادامه نحوه ی محاسبه اهمیت هر بیت توضیح داده می شود. در ابتدا کد باینری تصاویر مشابه به دست آمده با کد باینری تصویر درخواستی به منظور مشخص نمودن اهمیت هر بیت مقایسه می گردد. جدول ۳ـ۲ مثالی از k بیت کد باینری تصویر درخواستی با پنج تصویر $\{I_i\}_{i=۱}^{۵}$ که فاصله ی همینگ آن ها با تصویر درخواستی برابر ۲ است را نمایش می دهد. NS تعداد تصاویری است که دارای کد باینری یکسان با تصویر درخواستی کاربر در آن بیت می باشد و $w = \{w_۱, ..., w_k\}$ وزن آن بیت می باشد.

با مشاهده جدول فوق می توان دریافت که در برخی از بیت ها مانند بیت چهارم تصاویر بیشتری کد هش یکسانی را به اشتراک گذاشته اند و در برخی از بیت ها مانند بیت اول تصاویر کمتری کد هش یکسان را به اشتراک قرار داده اند.برای $b_۴$ مقدار $NS = ۴$ ، برای $b_۳$ مقدار $NS = ۳$، برای $b_۲$ مقدار $NS = ۲$ و برای $b_۱$ مقدار $NS = ۱$ بنابراین ما به این نتیجه می رسیم که اهمیت بیت چهارم بیشتر از بیت سوم،دوم و اول است بنابراین $w_۴ > w_۳ > w_۲ > w_۱$ مرتب سازی جدد نتایج براساس اهمیت هر بیت در شکل زیر نمایش داده شده است .برای تصویر درخواستی تکنیک بازخور [۱۹] به منظور یافتن تصاویر مشابه از تصاویر بازگردانده شده اعمال می شود.در ابتدا پایگاه داده تصاویر بر پایه فاصله ی همینگ جست و جو می شود و تصاویری با کمترین فاصله ی همینگ نسبت به تصویر درخواستی استخراج می شود سپس تصاویری که بیشترین شباهت را با تصویر درخواستی کاربر دارند انتخاب می شوند نحوه ی تخصیص وزن به هر یک از هش بیت ها در بخش بعد توضیح داده می شود.

[۱۹] Feedback technique



شکل ۳ـ۹: مرتب سازی مجدد نتایج بر اساس اهمیت هر بیت [۸].

۳ـ۶ـ۱ محاسبه وزن

فرض کنیم تصویر درخواستی کاربر q و m تعداد تصاویری که بیشترین شباهت را به تصویر درخواستی کاربر دارد و با استفاده از تکنیک هشینگ محاسبه شده است.این m تصویر می توانند بر پایه ی فاصله ی همینگ به صورت یک شکل متحدالمرکز نمایش داده شوندو مرکز آن تصویر درخواستی و تصاویر دیگری است که دارای فاصله ی همینگ ۰ می باشند.کد باینری m تصویر به صورت $H = \{H_1, ..., H_m\}$ که در آن $H_i = \{H_i^{(1)}, H_i^{(2)}, ..., H_i^{(k)}\}$ می باشد کد باینری تصویر درخواستی به صورت $H_q = \{H_q^{(1)}, H_q^{(2)}, ..., H_q^{(k)}\}$ می باشد وزن K بیت هش کد تصویر درخواستی به صورت $w = \{w_1, ..., w_k\}$ می باشد که مقدار اولیه $w_k = 1$ است .

در این بخش مقادیر H_iوH_q بیت به بیت با هم مقایسه می شوند.برای کد باینری K ام اگر $H_i^{(k)} = H_q^{(k)}$ آن گاه وزن w_k افزایش می یابد در غیر این صورت این وزن کاهش می باشد.وزن w به صورت بازگشتی که تعداد تکرار آن برابر با m می باشد به روزرسانی می گردد.برای j امین تکرار w_k با رابطه ی ۳ـ۳۵ محاسبه می شود:

$$w_k = \begin{cases} \varepsilon w_k & if H_q^k \neq H_i^k \\ (1+\varepsilon) w_k & otherwise \end{cases} \quad 0 < \varepsilon < 1 \qquad (۳ـ۳۵)$$

در این رابطه ε یک پارامتر است که درباره ی مقدار آن در بخش مشاهدات بحث می شود.بعد از m تکرار می توانیم مقدار w را به دست آوریم و از آن برای محاسبه ی فاصله ی همینگ وزن دار استفاده کنیم و با مرتب سازی نتایج فاصله ی همینگ وزن دار تصاویری که بیشترین شباهت را به تصویر درخواستی کاربر دارند نمایش داده شود.نحوه ی محاسبه ی این وزن ها در الگوریتم زیر ارائه شده است:

الگوریتم ۳ـ۱ الگوریتم محاسبه وزن هر بیت در مرتب سازی مجدد بر اساس اهمیت هر بیت.

ورودی: کد باینری تصویر درخواستی کاربر $H_q = \{h_q^{(۱)}, h_q^{(۲)}, ..., h_q^{(k)}\}$؛ m تصویر که بیشترین شباهت را به تصویر درخواستی کاربر بر اساس فاصله ی همینگ دارند .کد باینری i امین تصویر بازگردانده شده $H_i = \{H_i^{(۱)}, H_i^{(۲)}, ..., H_i^{(k)}\}$؛ پارامتر ε .

خروجی: وزن $w = \{w_۱, ..., w_k\}$ برای H_q

۱: مقدار اولیه وزن $w_k = ۱, , ۱$

۲: برای $i = ۱ : m$.

۳: برای $k = ۱ : K$.

۴: آیا k امین بیت H_q و H_iبا هم برابر است؟

اگر بله،$w_k = (۱ + \varepsilon)w_k$

در غیر این صورت $w_k = \varepsilon w_k$

۵: پایان for

۶: پایان for

فصل ۴

۴ـ۱ مقدمه

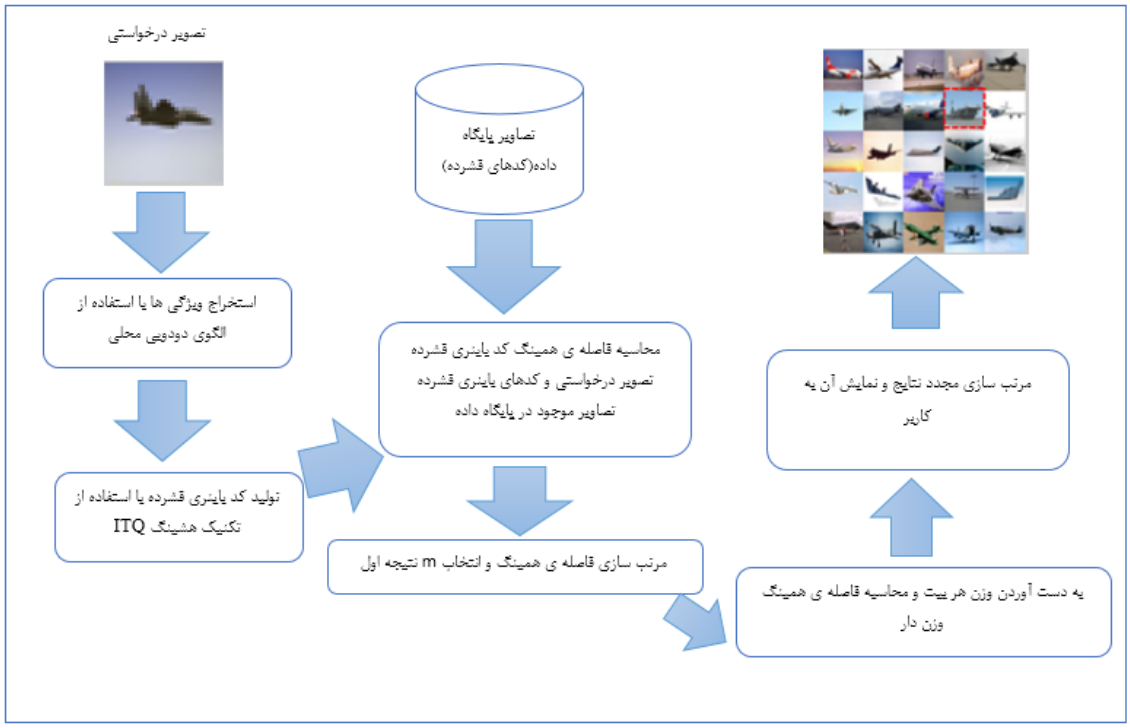
در فصل گذشته به تشریح الگوریتم های به کار رفته در این پایان نامه پرداخته شد.در این فصل برآنیم تا به بیان الگوریتم پیشنهادی و بررسی دقیق تر تکنیک های به کار رفته در این کار بپردازیم. دیاگرام روش پیشنهادی در شکل ۴ـ۱ نمایش داده شده است .به طور کلی الگوریتم پیشنهادی به چهار مرحله ی کلی استخراج ویژگی های تصاویر،ایجاد کد باینری فشرده ،به دست آوردن نزدیک ترین همسایه و مرتب سازی مجدد نتایج تقسیم بندی می شود در این فصل به تشریح کامل هر یک از این مراحل پرداخته می شود.

۴ـ۲ روش پیشنهادی

۴ـ۲ـ۱ مرحله اول،استخراج ویژگی های تصویر

به منظور توصیف یک تصویر باید ویژگی های آن استخراج گردد .ویژگی های محلی یک تصویر انتخاب مناسب تری نسبت به ویژگی های سراسری به منظور توصیف یک تصویر می باشد،زیرا ویژگی های محلی قدرت جداسازی بیشتری نسبت به ویژگی های سراسری دارند اما هزینه ی محاسباتی بیشتری نسبت به ویژگی های سراسری دارند. به دلیل حجم بالای تصاویر موجود در پایگاه داده ، روشی که به منظور استخراج ویژگی های محلی استفاده می گردد باید دارای ویژگی های ذیل باشد:

- قدرت جداسازی بالا
- هزینه محاسباتی پایین



شکل ۴ـ۱ : دیاگرام روش پیشنهادی.

در میان روش های استخراج ویژگی های محلی،الگوی دودویی محلی به دلیل سرعت تولید ، ماهیت دودویی،هزینه محاسباتی پایین و قدرت جداسازی بالا انتخاب مناسبی می باشد.الگوی دودویی محلی به عنوان یک توصیفگر غیر پارامتری مطرح است .از آن جایی که الگوی دودویی محلی هم از مشخصه های آماری و هم ساختاری بافت استفاده می کند یک ابزار قدرتمند به منظور توصیف یک تصویر می باشد.هم چنین این توصیفگر دارای مزایایی چون تحمل پذیری نسبت به تغییرات یکنواخت و سادگی محاسباتی می باشد.

با توجه به دلایل ذکر شده، در این مرحله به منظور استخراج ویژگی های تصاویر موجود در پایگاه داده از روش الگوی دودویی محلی استفاده و بردار ویژگی تصاویر تولید گردید.

۴ـ۲ـ۲ مرحله دوم، تولید کد باینری فشرده

به دلیل حجم بالای تصاویر موجود در پایگاه داده استفاده از روش تطابق ناشیانه،که در آن یک تصویر با تمام تصاویر موجود در پایگاه داده مقایسه می شود یک روش عملی نمی باشد زیرا این روش نیاز به حافظه ی زیادی دارد و هزینه ی محاسباتی آن بالا است.تکنیک هشینگ روشی مناسب به منظور تخمین زدن نزدیک ترین همسایه می باشد که در آن به ازای هر تصویر یک کد باینری فشرده تولید می شود که در نهایت عملیات مقایسه

تصاویر با استفاده از معیار فاصله ی همینگ انجام می گردد که این امر موجب افزایش سرعت محاسباتی و کاهش حافظه ی مورد نیاز می شود.در این مرحله ما از الگوریتم هشینگ ITQ به منظور تولید کد باینری فشرده استفاده نمودیم این الگوریتم یکی از کارآمدترین الگوریتم ها به منظور کاهش خطای چندی سازی می باشد.بردار ویژگی های تولید شده در مرحله ی قبل به عنوان ورودی این الگوریتم در نظر گرفته می شود و پس از کاهش ابعاد این بردار با استفاده از تکنیک آنالیز اجزای اصلی در داده های بدون برچسب و روش همبستگی کانونی در داده های برچسب گذاری شده ،مناسب ترین چرخش برای داده ها با استفاده از روش تجزیه مقادیر منفرد به دست می آید سپس هر یک از داده ها به نزدیک ترین رئوس ابر مکعب باینری تصویر می شوند و کد باینری فشرده برای آن ها به دست می آید.

۴ـ۲ـ۳ مرحله سوم، به دست آوردن نزدیک ترین همسایه

به دلیل ماهیت دودویی کدهای تولید شده در مرحله ی قبل ،می توان از فاصله ی همینگ به عنوان معیار فاصله به منظور اندازه گیری فاصله ی میان تصویر درخواستی و تصاویر موجود در پایگاه داده استفاده نمود.از آن جایی که این معیار یک عملگر بیتی محسوب می شود از سرعت بالای برخوردار می باشد.این امر در پایگاه داده هایی با مقیاس بالا از اهمیت ویژه ای برخوردار است. در این مرحله فاصله ی کد باینری فشرده تصویر درخواستی با کد باینری فشرده تصاویر موجود در پایگاه داده با استفاده از فاصله ی همینگ به دست می آید و نتایج مرتب سازی می گردد.

۴ـ۲ـ۴ مرحله ی چهارم، مرتب سازی مجدد بر اساس اهمیت هر بیت

همان طورکه در بخش قبل اشاره گردید،فاصله ی همینگ به منظور اندازه گیری شباهت میان دو کد دودویی مورد استفاده قرار می گیرد.برای k بیت کد دودویی ،فاصله ی همینگ یک مقدار صحیح می باشد و حداکثر مقدار آن برابر k می باشد.بنابراین تعداد زیادی از تصاویر فاصله ی همینگ یکسانی با تصویر درخواستی دارند.در معیار اندازه گیری همینگ تمامی بیت ها ارزش یکسانی داشته و با آن ها برخورد یکسانی انجام می شود ولی در عمل برخی از بیت ها از اهمیت بیشتری برخوردار می باشند بدین منظور اهمیت هر بیت در این مرحله با استفاده از الگوریتم مرتب سازی مجدد بر اساس اهمیت هر بیت محاسبه می گردد.بدین منظور ابتدا m تصویر اول به دست آمده از مرحله ی قبل انتخاب می گردد و بیت ها ی کد باینری فشرده آن با بیت های کد باینری تصویر درخواستی مقایسه می گردد و در صورت یکسان بودن به وزن آن بیت افزوده شده و در غیر این صورت از وزن آن کاسته می شود با اجرای این الگوریتم به ازای تصاویر انتخابی ،وزن بیت های آن به دست می آید حال با استفاده از معیار فاصله همینگ وزن دار ، محاسبات انجام شده و نتایج مرتب می گردد

و n تصویر اول به کاربر نمایش داده می شود.

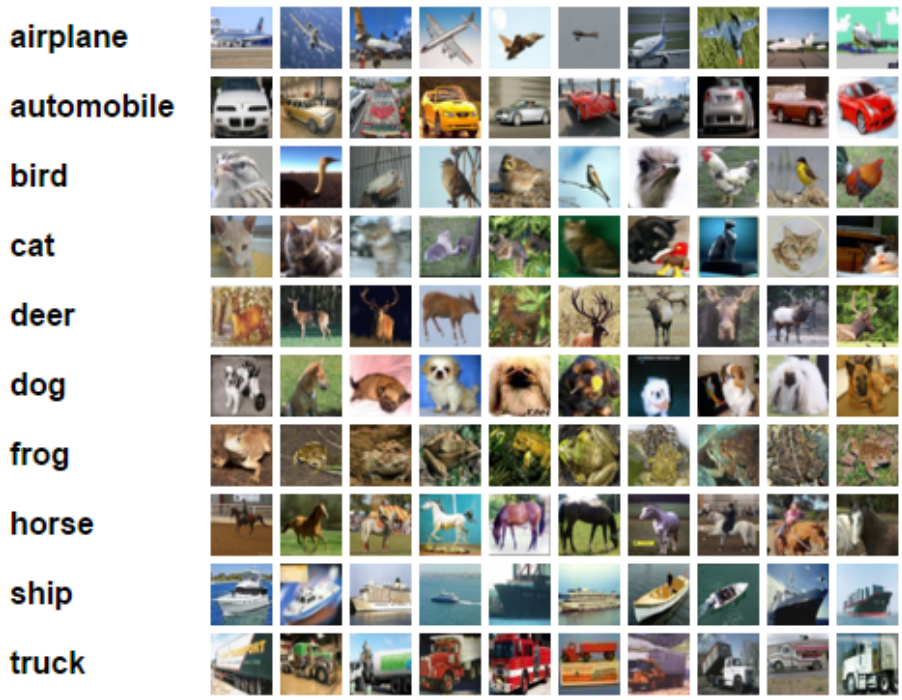
فصل ۵

۵ـ۱ مقدمه

در این فصل به بررسی مسائلی از قبیل معرفی دادگان مورد استفاده،معیارهای ارزیابی ،شرح انجام آزمایشات و ارزیابی و مقایسه روش پیشنهادی با سایر روش های ارائه شده می پردازیم.

۵ـ۲ مجموعه دادگان

در این پایان نامه به منظور مشاهده و ارزیابی نتایج روش پیشنهادی از دو مجموعه دادگان CIFAR۱۰ و MNIST استفاده شده است.مجموعه ی CIFAR۱۰ شامل ۶۰۰۰۰ تصویر رنگی در ابعاد ۳۲×۳۲ می باشد.این تصاویر در ۱۰ کلاس مختلف تقسیم بندی گردیده است .در هر کلاس از این مجموعه ۶۰۰۰ تصویر وجود دارد و این مجموعه شامل ۵۰۰۰۰ داده آموزش و ۱۰۰۰۰ داده تست می باشد.در شکل ۵ـ۱ نمونه ای از تصاویر موجود در این پایگاه داده نمایش داده شده است. مجموعه دادگان MNIST شامل ارقام ۰ تا ۹ دست نویس با ابعاد ۲۰×۲۰ می باشد این مجموعه شامل ۶۰۰۰۰ داده آموزش و ۱۰۰۰۰ داده تست می باشد در شکل ۵ـ۲ نمونه ای از تصاویر موجود در پایگاه داده MNIST نمایش داده شده است.



شکل ۵ـ۱: نمونه ای از تصاویر موجود در پایگاه داده CIFAR-۱۰.

شکل ۵ـ۲: نمونه ای از تصاویر موجود در پایگاه داده MNIST.

۵ـ۳ معیارهای ارزیابی

۵ـ۳ـ۱ دقت

نسبت تعداد تصاویر مرتبط بازیابی شده به تعداد کل تصاویر بازیابی شده می باشد.به عنوان نمونه اگر ۸ تصویر بازیابی شده باشد و تنها ۴ تصویر متعلق به دسته مورد جست وجو ما باشد دقت آن برابر ۵۰% می باشد.

۵ـ۳ـ۲ فراخوانی

نسبت تعداد تصاویر مرتبط بازیابی شده به تعداد کل تصاویر موجود در پایگاه داده.به عنوان نمونه اگر ۸ تصویر مرتبط از پایگاه داده ای از تصاویر شامل ۸۰ تصویر بازیابی شود فراخوانی آن برابر ۱۰% می باشد.

۵ـ۳ـ۳ منحنی دقت در مقابل فراخوانی

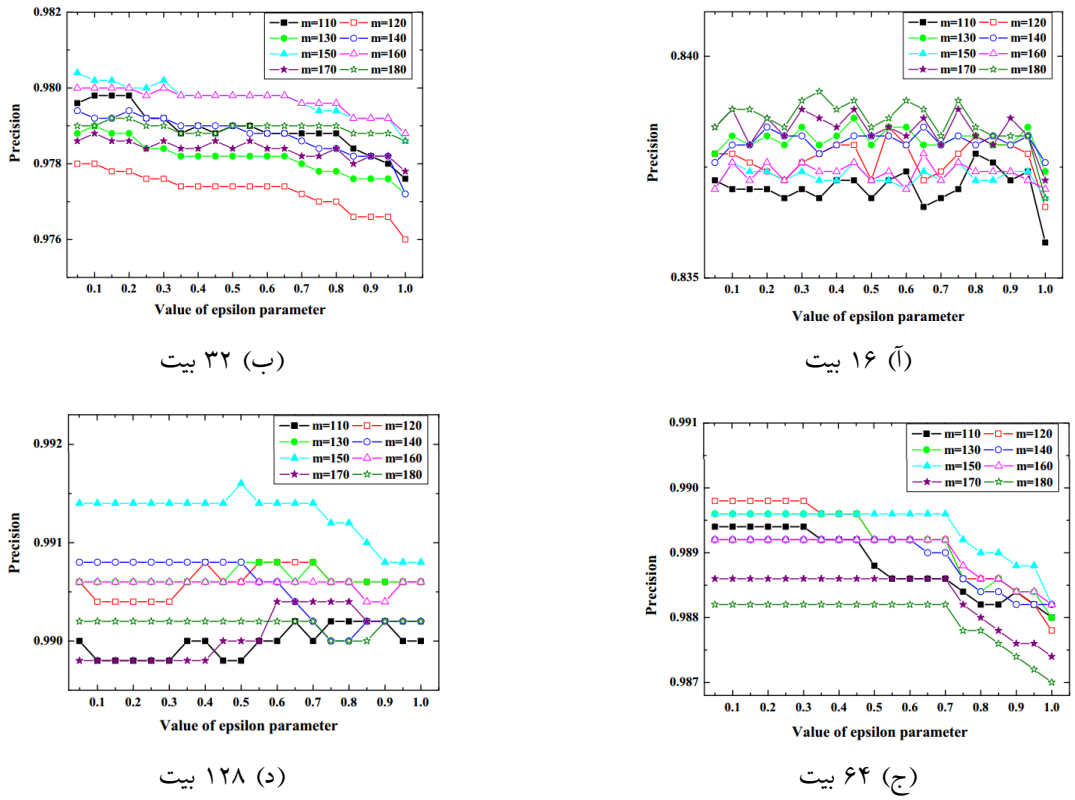
دقت یک امر مهم در فرآیند بازیابی و فراخوانی امری مهم در شناسایی می باشد.این منحنی به منظورنمایش و مقایسه عملکرد الگوریتم ها مورد استفاده قرار می گیرد.این منحنی چگونگی کاهش دقت ، زمانی که کسر بیشتری از تصاویر موجود در پایگاه داده بازیابی می شود را نشان می دهد.منحنی دقت در مقابل فراخوانی ایده آل ،دارای مقدار دقت ۱۰۰ % به ازای تمامی مقادیر فراخوانی می باشدکه به معنای بازیابی تمام تصویر مرتبط می باشد.

۵ـ۴ شرح آزمایشات انجام شده

۵ـ۴ـ۱ تعیین پارامترهای مساله

تعیین پارامترهای الگوریتم مرتب سازی مجدد بر اساس اهمیت هر بیت

این الگوریتم دارای دو پارامتر m و ε می باشد مقدار مناسب برای این دو پارامتر در هر یک از مجموعه دادگان با اجرای آزمایشات به دست آمد . شکل ۵ـ۳ تغییرات دقت بازیابی را بر اساس مقادیر مختلف m و ε برای مجموعه دادگان MNIST نشان می دهد.محدوده ی مقادیر ε بین ۰,۱ تا ۱ می باشد و مقادیر m بین ۱۱۰ تا ۱۸۰ در نظر گرفته شده است.در شکل (الف) این نمودار برای تعداد ۱۶ بیت، در (ب) برای ۳۲ بیت ، در



(آ) ۱۶ بیت

(ب) ۳۲ بیت

(ج) ۶۴ بیت

(د) ۱۲۸ بیت

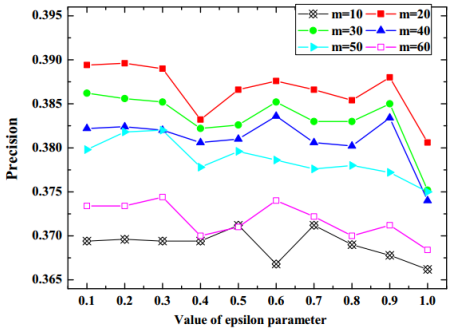
شکل ۵ـ۳: نمودار دقت بازیابی بر اساس مقادیر مختلف m و ε برای مجموعه دادگان MNIST

(ج) برای ۶۴ بیت و در (د) برای ۱۲۸ بیت کد باینری به دست آمده است. با توجه به نتایج به دست آمده برای مجموعه دادگان MNIST مقدار m=۱۵۰ و ε = ۰/۱ در نظر گرفته شد. شکل ۵ـ۴ تغییرات دقت بازیابی را بر اساس مقادیر مختلف m و ε برای مجموعه دادگان CIFAR-۱۰ نشان می دهد.محدوده ی مقادیر ε بین ۰,۱ تا ۱ می باشد و مقادیر m بین ۱۰ تا ۶۰ در نظر گرفته شده است.در شکل (الف) این نمودار برای تعداد ۱۶ بیت، در (ب) برای ۳۲ بیت ، در (ج) برای ۶۴ بیت و در (د) برای ۱۲۸ بیت کد باینری به دست آمده است. با توجه به نتایج به دست آمده برای مجموعه دادگان CIFAR-۱۰ مقدار m=۳۰ و ε = ۰/۳ در نظر گرفته شد.

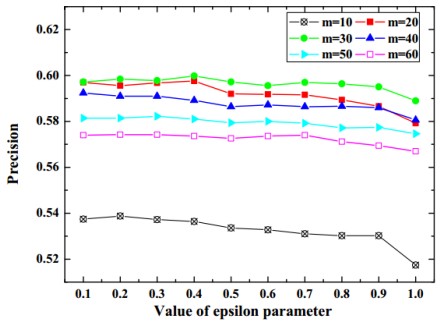
کلیه آزمایشات با پارامترهای به دست آمده در این قسمت انجام شده است.

تعیین پارامتر الگوریتم هشینگ ITQ

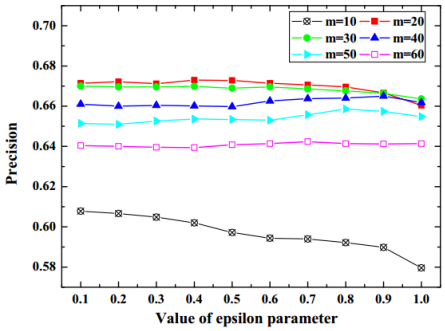
به منظور به دست آوردن تعداد تکرار مناسب برای الگوریتم هشینگ ITQ ،به ازای تعداد تکرار ۱، ۵۰ و ۱۵۰ این الگوریتم را تکرار نمودیم و منحنی دقت در مقابل فراخوانی را به دست آوردیم .طبق نتایج به دست آمده



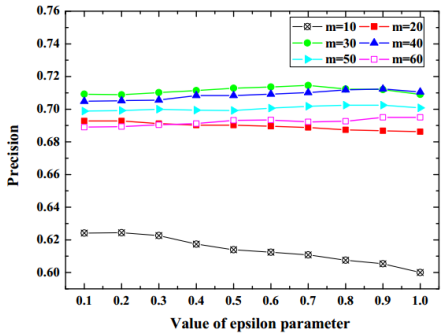
(آ) ۱۶ بیت



(ب) ۳۲ بیت



(ج) ۶۴ بیت



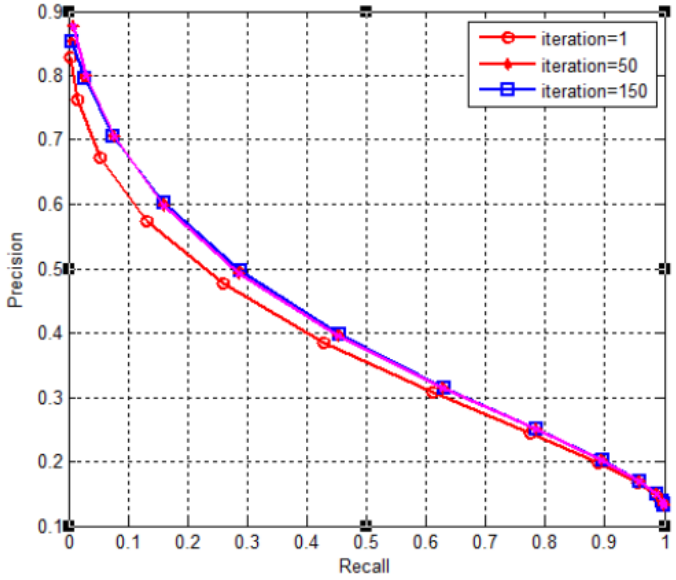
(د) ۱۲۸ بیت

شکل ۵ـ۴: نمودار دقت بازیابی بر اساس مقادیر مختلف m و ε برای مجموعه دادگان CIFAR-۱۰

در شکل ۵ـ۵ بهترین تکرار برای این الگوریتم ۵۰ می باشد و آزمایشات در این بخش با تعداد تکرار ۵۰ به انجام رسیده است.

۵ـ۴ـ۲ تأثیر ویژگی الگوی دودویی محلی در بهبود عملکرد الگوریتم

بردار ویژگی های تصویر به عنوان ورودی الگوریتم هشینگ ITQ ، نقشی کلیدی در عملکرد این الگوریتم دارد.در شکل ۵ـ۶ نموداری ارائه گردیده است که مقایسه ای را بین عملکرد الگوی دودویی محلی و توصیفگر سراسری GIST به عنوان ورودی الگوریتم هشینگ ITQ انجام داده است. همان طور که در نمودار ۵ـ۶ مشاهده می شود عملکرد این الگوریتم با به کارگیری الگوی دودویی محلی به عنوان ورودی بهبود یافته است .این الگو به دلیل ماهیت دودویی خود ،از یک سو هزینه محاسبات را کاهش داده است و از سوی دیگر به دلیل کارآیی این روش در استخراج ویژگی های تصویر ،موجب افزایش دقت بازیابی گردیده است.با بهره گیری از این روش با تعداد هش بیت کمتری سیستم به دقت بالاتری دست یافته است که این امر در سیستم های بازیابی تصویر بر اساس محتوا در مقیاس بالا بسیار پراهمیت است.



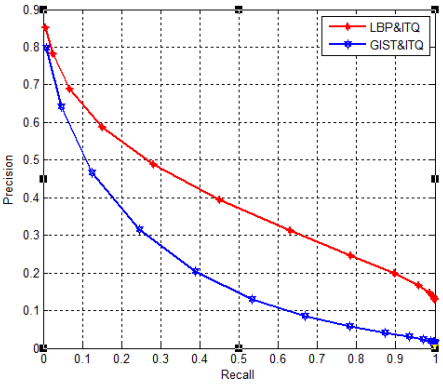
شکل ۵ـ۵: منحنی دقت در مقابل فراخوانی به ازای تعداد تکرار مختلف الگوریتم هشینگ .ITQ

۵ـ۴ـ۳ مقایسه ی روش پیشنهادی با روش های دیگر

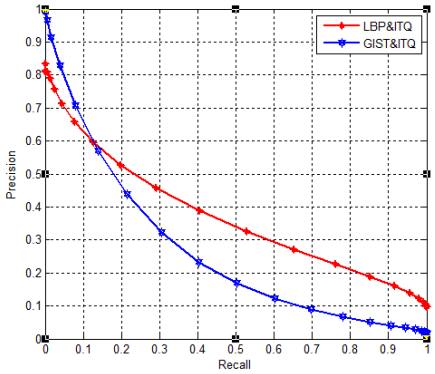
در این بخش الگوریتم پیشنهادی با سه الگوریتم هشینگ پایه ای دیگر که روند هشینگ $B = sgn(X\tilde{W})$ را دنبال می کند مقایسه شده است که در آن ماتریس تصویر $\tilde{W}$ با روش های مختلف به دست می آید:

- هشینگ حساس به همسایگی:$\tilde{W}$ یک ماتریس تصادفی گاوسی است .
- چرخش تصادفی: در این روش $\tilde{W} = WR$ می باشد که در آن W ماتریس جهت های PCA است و R یک ماتریس متعامد تصادفی است که ایده ی روش ITQ برگرفته از این روش می باشد.
- کرنل های ثابت به جابه جایی: این روش بر پایه ی ویژگی های فوریه ی تصادفی برای تخمین زدن کرنل های گاوسی می باشد.

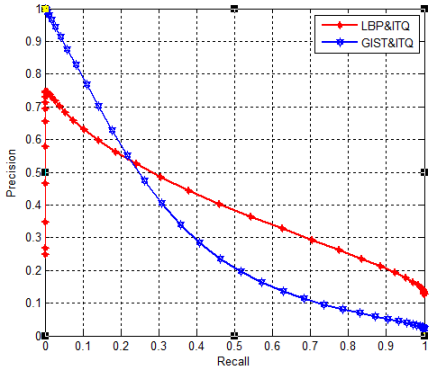
شکل ۵ـ۷ به مقایسه عملکرد الگوریتم پیشنهادی در مجموعه دادگان CIFAR-۱۰ و شکل ۵ـ۸ به مقایسه عملکرد الگوریتم پیشنهادی در مجموعه دادگان MNIST پرداخته است و نمودار دقت در مقابل فراخوانی رابرای روش های ذکر شده و روش پیشنهادی ترسیم نموده است.همان طور که در این نمودار مشاهده می شود دقت الگوریتم پیشنهادی از روش های ارائه شده بالاتر است .در این روش با تعداد هش بیت کمتر،سیستم به دقت بالاتری دست یافته است .



(آ) منحنی دقت براساس فراخوانی @۱۶ بیت



(ب) منحنی دقت براساس فراخوانی @۳۲ بیت



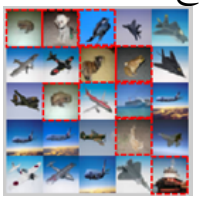
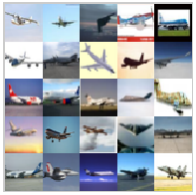
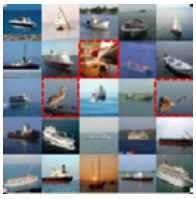
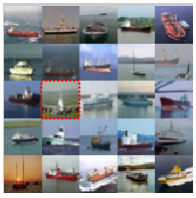
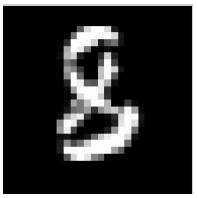
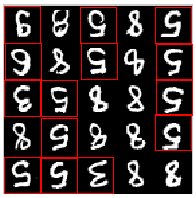
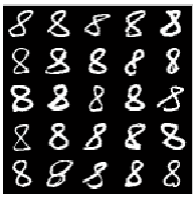
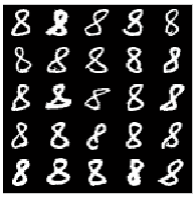
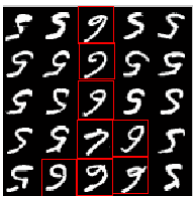
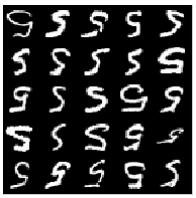
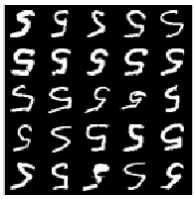
(ج) منحنی دقت براساس فراخوانی @۶۴ بیت

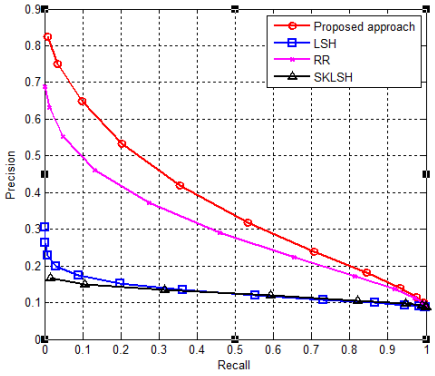
شکل ۵ـ۶: تاثیر الگوی دودویی محلی به عنوان ورودی به الگوریتم هشینگ ITQ

۵ـ۴ـ۴ نمایش خروجی به کاربر

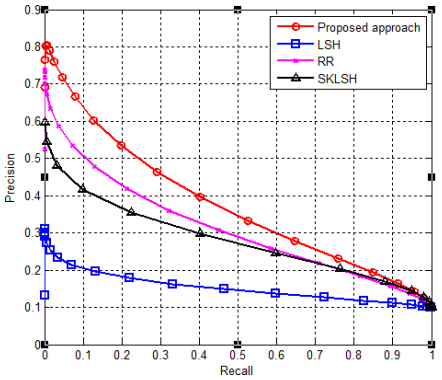
در این بخش با دریافت یک تصویر به عنوان تصویر درخواستی کاربر، با اجرای الگوریتم پیشنهادی ،۲۵ تصویر نخست بازیابی شده را به کاربر نمایش داده ایم.در جدول ۵ـ۱ نتایج به دست آمده با استفاده از تکنیک آنالیز اجزای اصلی و تحلیل همبستگی کانونی و دقت محاسباتی هر یک نمایش داده شده است .همان طور که مشاهده می شود دقت الگوریتم پس از اعمال مرتب سازی مجدد بر اساس اهمیت هر بیت افزایش یافته است .هم چنین نتایج بر اساس بیشترین شباهت به تصویر درخواستی نمایش داده شده است.

جدول ۵ـ۱: ۲۵ تصویر نخست بازیابی شده

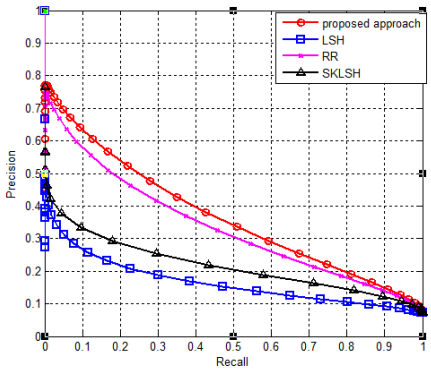
تصویر درخواستی	نتایج با استفاده از PCA	نتایج با استفاده از CCA	نتایج پس از مرتب سازی مجدد
			
دقت	۶۴٪	۹۶٪	۱۰۰٪
			
دقت	۶۰٪	۸۸٪	۹۶٪
			
دقت	۴۴٪	۱۰۰٪	۱۰۰٪
			
دقت	۷۲٪	۱۰۰٪	۱۰۰٪



(آ) منحنی دقت براساس فراخوانی @۱۶ بیت



(ب) منحنی دقت براساس فراخوانی @۳۲ بیت

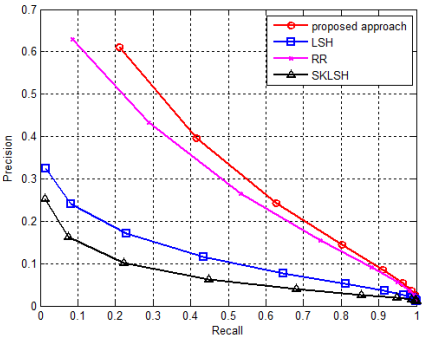


(ج) منحنی دقت براساس فراخوانی @۶۴ بیت

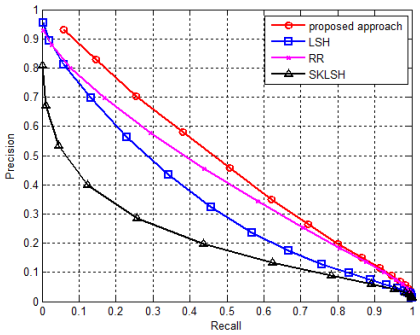
شکل ۵ـ۷: مقایسه عملکرد الگوریتم پیشنهادی با روش های دیگر در مجموعه دادگان CIFAR-۱۰

۵ـ۵ نتیجه گیری

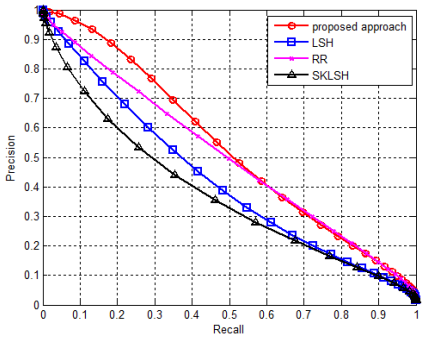
این پایان نامه به ارائه ی روشی کارا به منظور بازیابی تصاویر بر اساس محتوا در مقیاس بالا پرداخته است.این روش بر پایه ی الگوی دودویی محلی و الگوریتم هشینگ ITQ می باشد.در این کار به بررسی نقش کلیدی بردار ویژگی تصویر به عنوان ورودی الگوریتم هشینگ پرداخته است.همان طور که در نتایج آزمایشات در بخش ۵ـ۴ـ۲ نشان داده شد به کارگیری بردار ویژگی مناسب موجب بهبود عملکرد الگوریتم هشینگ گردیده است.همان طور که پیش تر مطرح گردید به منظور استخراج ویژگی های تصویر در روش پیشنهادی از الگوی دودویی محلی استفاده شده است .این الگو یکی از قدرتمندترین توصیفگرها برای ارائه ساختارهای محلی به شمار می آید .به دلیل مزایایی چون تحمل پذیری نسبت به تغییرات یکنواخت و سادگی محاسباتی ،به صورت موفقیت آمیزی در بسیاری از حوزه های آنالیز تصویر به کار برده می شود .دستیابی به مجموعه ی کوچک ویژگی ها،از متمایزترین خصوصیات الگوی دودویی محلی است که باعث عملکرد بهتر و کاهش ابعاد آن می



(آ) منحنی دقت براساس فراخوانی @۱۶ بیت



(ب) منحنی دقت براساس فراخوانی @۳۲ بیت



(ج) منحنی دقت براساس فراخوانی @۶۴ بیت

شکل ۵ـ۸: مقایسه عملکرد الگوریتم پیشنهادی با روش های دیگر در مجموعه دادگان MNIST

گردد.

به دلیل بالا بودن حجم تصاویر موجود در پایگاه داده انتخاب روشی کارا برای استخراج ویژگی های یک تصویر بسیار حائز اهمیت می باشد.زیرا روش انتخابی در عین کارایی در استخراج ویژگی ها باید هزینه ی محاسباتی پایینی نیز داشته باشد.الگوی دودویی محلی به دلیل آن که یک روش غیر پارامتری برای استخراج ویژگی ها می باشد از سادگی محاسباتی نیز برخوردار است .بدین ترتیب با بهره گیری از این روش می توانیم به ویژگی های مناسب با هزینه ی پایین دست یابیم.

در ادامه به منظور تخمین نزدیک ترین همسایه از تکنیک ITQ استفاده گردیده است .این روش نیز یکی از روش های ساده و کارا به منظور تخمین نزدیک ترین همسایه می باشد.نتایج آزمایشات در بخش ۵ـ۴ـ۳ نشان داده است که روش پیشنهادی نسبت به دیگر روش های ارائه شده از کارایی بالاتری برخوردار است.این بدان مفهوم می باشد که ما به استفاده از تعداد کد هش کمتر توانسته ایم به دقت بالاتری دست یابیم .هش کردن یک تصویر با تعداد کد کمتر منجر به استفاده از حافظه ی کمتر برای نمایش یک تصویر می شود و از سوی دیگر هزینه ی محاسباتی به منظور مقایسه ی دو تصویر را کاهش می دهد.تحقق این امر در پایگاه داده

هایی با ابعاد بالا از اهمیت زیادی برخوردار است زیرا نیاز به تکنیک هایی که حافظه ی مصرفی و هزینه ی محاسباتی را کاهش دهد احساس می گردد.
از آن جا که اهمیت هر بیت در کد باینری تولید شده یکسان نمی باشد با بهره گیری از تکنیک مرتب سازی مجدد نتایج بر اساس اهمیت هر بیت، وزن هر یک از بیت ها به دست آمده و با بهره گیری از معیار فاصله ی همینگ وزن دار نتایج به دست آمده مرتب سازی می گردد.این گام در الگوریتم پیشنهادی به موجب نمایش نتایج بر اساس بیشترین شباهت به تصویر درخواستی در نظر گرفته شده است.
با توجه به نتایج به دست آمده در بخش آزمایشات دقت روش پیشنهادی نسبت به روش های موجود بالاتر است.این روش با به کارگیری تعداد هش کد کمتر توانسته به این دقت دست یابد که این امر نتیجه ای مطلوب به منظور بازیابی تصاویر بر اساس محتوا در مقیاس بالا می باشد.

۵ـ۶ کارهای آینده

به دنبال مطالعه در یک حوزه،بررسی مشکلات موجود در آن و ارائه راه حل برای آن سپس تست و آزمایش راه حل پیشنهادی ،با مشکلات موجود در این مسیر بیشتر آشنا شده و راه حل هایی به ذهن محقق خطور می کند که به دلیل محدودیت زمانی ،امکان بررسی آن برای فرد وجود ندارد .هدف از این بخش ارائه مشکلات این مسیر و راه حل های محقق می باشد به گونه ای که انگیزه ی ادامه این راه را برای خود نویسنده یا خوانندگان فراهم آورد.در این بخش به ارائه ی چندین نمونه که می تواند در تکمیل این کار مؤثر واقع شود می پردازم.
اغلب روش های بافت محور به منظور استخراج ویژگی های یک تصویر حساسیت زیادی نسبت به تغییرات نورپردازی شدید داشته و هم چنین نسبت به تصاویر مبهم و مختل شده حساسیت نشان می دهند به منظور رفع مشکلات ارائه شده می توانیم قبل از مرحله ی استخراج ویژگی ها یک مرحله پیش پردازش روی تصویر انجام داده و سپس خروجی این مرحله را به عنوان ورودی الگوی دودویی محلی در نظر بگیریم.البته در پایگاه دادهایی که در این کار مورد استفاده قرار گرفته است مشکلات فوق وجود ندارد ولی به منظور عمومی سازی این الگوریتم می توان این مرحله را به مراحل موجود اضافه نمود.
به منظور بهبود عملکرد الگوی دودویی محلی در سالهای اخیر تغییرات زیادی به این الگوریتم اعمال شده است .این تغییرات بر جنبه های مختلف الگوی دودویی محلی اصلی تمرکز کرده است که به منظور بهبود قدرت جداسازی ،افزایش تحمل در برابر تغییرات،نحوه ی انتخاب همسایه،گسترش به داده های سه بعدی و ترکیب با دیگر روش ها است که لیستی از این تغییرات در جدول ۳ـ۱ آورده شده است.به عنوان یکی از کارهای پیشنهادی برای آینده می توان به کارگیری این روش ها در الگوریتم پیشنهادی در مرحله ی نخست اشاره کرد و به بررسی و مقایسه نتایج به دست آمده با این روش ها پرداخت.

به کارگیری پایگاه داده هایی با حجم تصاویر بیشتر که شامل میلیون ها تصویر باشد پیشنهاد دیگرارائه شده در این بخش است.می توان عملکرد روش پیشنهادی را در این پایگاه داده ها بررسی و مقایسه نمود.

مراجع

[1] Ahonen, T., Hadid, A., and Pietikainen, M. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 28*, 12 (2006), 2037–2041.

[2] Ahonen, T., Matas, J., He, C., and Pietikäinen, M. Rotation invariant image description with local binary pattern histogram fourier features. in *Image Analysis*. Springer, 2009, pp. 61–70.

[3] Ahonen, T., and Pietikäinen, M. Soft histograms for local binary patterns. in *Proceedings of the Finnish signal processing symposium, FINSIG* (2007), volume 5, p. 1.

[4] Andoni, A., and Indyk, P. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. in *Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on* (2006), IEEE, pp. 459–468.

[5] Bai, G., Zhu, Y., and Ding, Z. A hierarchical face recognition method based on local binary pattern. in *2008 congress on image and signal processing* (2008), IEEE, pp. 610–614.

[6] Fehr, J. Rotational invariant uniform local binary patterns for full 3d volume texture analysis. in *Finnish signal processing symposium (FINSIG)* (2007).

[7] Fergus, R., Weiss, Y., and Torralba, A. Semi-supervised learning in gigantic image collections. in *Advances in neural information processing systems* (2009), pp. 522–530.

[8] Fu, H., Kong, X., and Wang, Z. Binary code reranking method with weighted hamming distance. *Multimedia Tools and Applications* (2014), 1–18.

[9] Gong, Y., and Lazebnik, S. Iterative quantization: A procrustean approach to learning binary codes. in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (2011), IEEE, pp. 817–824.

[10] Gong, Y., Lazebnik, S., Gordo, A., and Perronnin, F. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 35*, 12 (2013), 2916–2929.

[11] Guo, Z., Zhang, L., and Zhang, D. A completed modeling of local binary pattern operator for texture classification. *Image Processing, IEEE Transactions on 19*, 6 (2010), 1657–1663.

[12] He, L., Zou, C., Zhao, L., and Hu, D. An enhanced lbp feature based on facial expression recognition. in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the* (2006), IEEE, pp. 3300–3303.

[13] Heikkilä, M., Pietikäinen, M., and Schmid, C. Description of interest regions with local binary patterns. *Pattern recognition 42*, 3 (2009), 425–436.

[14] Huang, D., Shan, C., Ardabilian, M., Wang, Y., and Chen, L. Local binary patterns and its application to facial image analysis: a survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 41*, 6 (2011), 765–781.

[15] Huang, D., Wang, Y., and Wang, Y. A robust method for near infrared face recognition based on extended local binary pattern. in *Advances in Visual Computing*. Springer, 2007, pp. 437–446.

[16] Huang, D., Zhang, G., Ardabilian, M., Wang, Y., and Chen, L. 3d face recognition using distinctiveness enhanced facial representations and local feature hybrid matching. in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on* (2010), IEEE, pp. 1–7.

[17] Huang, Y., Wang, Y., and Tan, T. Combining statistics of geometrical and correlative features for 3d face recognition. in *BMVC* (2006), Citeseer, pp. 879–888.

[18] Indyk, P., and Motwani, R. Approximate nearest neighbors: towards removing the curse of dimensionality. in *Proceedings of the thirtieth annual ACM symposium on Theory of computing* (1998), ACM, pp. 604–613.

[19] Jaakkola, T., Haussler, D., et al. Exploiting generative models in discriminative classifiers. *Advances in neural information processing systems* (1999), 487–493.

[20] Jégou, H., Douze, M., Schmid, C., and Pérez, P. Aggregating local descriptors into a compact image representation. in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (2010), IEEE, pp. 3304–3311.

[21] Jégou, H., Perronnin, F., Douze, M., Sanchez, J., Perez, P., and Schmid, C. Aggregating local image descriptors into compact codes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 34*, 9 (2012), 1704–1716.

[22] Jin, H., Liu, Q., Lu, H., and Tong, X. Face detection using improved lbp under bayesian framework. in *Multi-Agent Security and Survivability, 2004 IEEE First Symposium on* (2004), IEEE, pp. 306–309.

[23] Kamel, A., Mahdi, Y. B., and Hussain, K. F. Multi-bin search: improved large-scale content-based image retrieval. in *Image Processing (ICIP), 2013 20th IEEE International Conference on* (2013), IEEE, pp. 2597–2601.

[24] Lei, Z., Liao, S., He, R., Pietikainen, M., and Li, S. Z. Gabor volume based local binary pattern for face representation and recognition. in *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on* (2008), IEEE, pp. 1–6.

[25] Liao, S., Zhu, X., Lei, Z., Zhang, L., and Li, S. Z. Learning multi-scale block local binary patterns for face recognition. in *Advances in Biometrics*. Springer, 2007, pp. 828–837.

[26] Liao, S., and Chung, A. C. Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude. in *Computer Vision–ACCV 2007*. Springer, 2007, pp. 672–679.

[27] Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision 60*, 2 (2004), 91–110.

[28] Mikolajczyk, K., and Schmid, C. A performance evaluation of local descriptors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 27*, 10 (2005), 1615–1630.

[29] Ojala, T., Pietikäinen, M., and Harwood, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition 29*, 1 (1996), 51–59.

[30] Ojala, T., Pietikäinen, M., and Mäenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 24*, 7 (2002), 971–987.

[31] Oliva, A., and Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision 42*, 3 (2001), 145–175.

[32] Paulhac, L., Makris, P., and Ramel, J.-Y. Comparison between 2d and 3d local binary pattern methods for characterisation of three-dimensional textures. in *Image Analysis and Recognition*. Springer, 2008, pp. 670–679.

[33] Perronnin, F., and Dance, C. Fisher kernels on visual vocabularies for image categorization. in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (2007), IEEE, pp. 1–8.

[34] Raginsky, M., and Lazebnik, S. Locality-sensitive binary codes from shift-invariant kernels. in *Advances in neural information processing systems* (2009), pp. 1509–1517.

[35] Ruiz-del Solar, J., and Quinteros, J. Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches. *Pattern Recognition Letters 29*, 14 (2008), 1966–1979.

[36] Shan, S., Zhang, W., Su, Y., Chen, X., and Gao, W. Ensemble of piecewise fda based on spatial histograms of local (gabor) binary patterns for face recognition. in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on* (2006), volume 4, IEEE, pp. 606–609.

[37] Singh, R., Vatsa, M., and Noore, A. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition. *Pattern Recognition 41*, 3 (2008), 880–893.

[38] Sivic, J., and Zisserman, A. Video google: A text retrieval approach to object matching in videos. in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on* (2003), IEEE, pp. 1470–1477.

[39] Tan, X., and Triggs, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. in *Analysis and Modeling of Faces and Gestures*. Springer, 2007, pp. 168–182.

[40] Tan, X., and Triggs, B. Fusing gabor and lbp feature sets for kernel-based face recognition. in *Analysis and Modeling of Faces and Gestures*. Springer, 2007, pp. 235–249.

[41] Tsaneva, M., Krezhova, D., and Yanev, T. Development and testing of a statistical texture model for land cover classification of the black sea region with modis imagery. *Advances in Space Research 46*, 7 (2010), 872–878.

[42] Vasconcelos, N. From pixels to semantic spaces: Advances in content-based image retrieval. *Computer*, 7 (2007), 20–26.

[43] Wang, J., Shen, H. T., Song, J., and Ji, J. Hashing for similarity search: A survey. *arXiv preprint arXiv:1408.2927* (2014).

[44] Weiss, Y., Torralba, A., and Fergus, R. Spectral hashing. in *Advances in neural information processing systems* (2009), pp. 1753–1760.

[45] Wolf, L., Hassner, T., and Taigman, Y. Descriptor based methods in the wild. in *Workshop on Faces in'Real-Life'Images: Detection, Alignment, and Recognition* (2008).

[46] Yan, S., Wang, H., Tang, X., and Huang, T. Exploring feature descritors for face recognition. in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (2007), volume 1, IEEE, pp. I–629.

[47] Yang, H., and Wang, Y. A lbp-based face recognition method with hamming distance constraint. in *Image and Graphics, 2007. ICIG 2007. Fourth International Conference on* (2007), IEEE, pp. 645–649.

[48] Zhang, L., Chu, R., Xiang, S., Liao, S., and Li, S. Z. Face detection based on multi-block lbp representation. in *Advances in biometrics*. Springer, 2007, pp. 11–18.

[49] Zhang, W., Shan, S., Gao, W., Chen, X., and Zhang, H. Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* (2005), volume 1, IEEE, pp. 786–791.

[50] Zhao, G., and Pietikainen, M. Dynamic texture recognition using local binary patterns with an application to facial expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 29*, 6 (2007), 915–928.

[51] Zhao, G., and Pietikäinen, M. Experiments with facial expression recognition using spatiotemporal local binary patterns. in *Multimedia and Expo, 2007 IEEE International Conference on* (2007), IEEE, pp. 1091–1094.

[52] Zhao, H., Wang, Z., and Liu, P. The ordinal relation preserving binary codes. *Pattern Recognition* (2015).

واژه‌نامه انگلیسی به فارسی

بازیابی تصویر بر اساس متن ..... Concept based image retrieval
بازیابی تصویر بر اساس محتوا ..... Content based image retrieval
کلمه کلیدی ..... Key word
زیرنویس ..... Caption
نوشته توصیفی ..... Description text
تجزیه ..... Analyzed
چندمعنایی ..... Polysemy
ترادف ..... Synonymy
تصویر درخواستی ..... Query image
استخراج ویژگی ها ..... Feature extractor
پایگاه داده ویژگی ها ..... Features vector
جست و جو شباهت ..... Similarity search
اندیس گذاری و بازیابی ..... Indexing and retrieval
صفحات وب ..... www
تصاویر پزشکی ..... Medical image
تطابق ناشیانه ..... Brute force matching
عملی ..... Practical
شناسایی اشیاء ..... Object recognition
کپی های غیر مجاز تصاویر در صفحات وب ..... Finding partial image duplicate on the web
جست وجو کردن فریم های ویدئو ..... Searching indivisual video frame
دسته بندی کردن تصاویر ..... Image classification
نمایش تصویر ..... Image Representation
قدیمی ..... Traditional
نقاط کلیدی تصویر ..... Image key point descriptor
توصیفگرهای نقاط مورد علاقه ..... Interest point descriptors
خوش تعریف ..... Well defined position

واژه‌نامه انگلیسی به فارسی

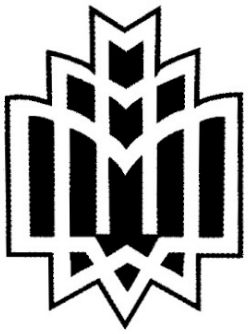
مقیاس ..... Scaling
چرخش ..... Rotate
تکرارپذیری ..... Reproducibility
نمایش عددی ..... Numerical representation
توصیفگر نقاط کلیدی ..... Keypoint descriptor vecotr
بردار ویژگی ..... Feature vector
ویژگی ها ..... Features
نقاط کلیدی محلی مورد علاقه ..... Local keypoint of interest
جابه جایی ها ..... Transformation
نواحی مورد علاقه ..... Local interest region
جداساز ..... Distinctive
جستجوي نزديكترين همسايه ..... Nearest neighbor search
مسائل مجاورت ..... Proximity problems
نزدیکترین همسایه ي دقیق ..... Exact nearest neighbor search
بررسی خطی ..... Linear scan
مصیبت بعد ..... Curse of dimensionality
تقریب زدن ..... Approximation
ان پی هارد ..... NP-hard
برخورد ..... Collision
رویکرد تقسیم و تخسیر ..... Divide and conquer approach
درخت های کی دی ..... Kd-trees
فضاهاي چندبعدي ..... Multidimensional spaces
درخت جستجو دودویی متوازن ..... Balanced binary search tree
عملگر الگوی دودویی محلی اصلی ..... Original LBP operator
آستانه سازی ..... Threshold
پایه ..... Basic
آشکار ساز ویژگی ..... Feature Detector
ریزالگوها ..... Micro pattern
تحمل ..... Robust
تابع های هش ..... Hash functions
تغییرات محدود ..... Bounded distortion
جاسازی ..... Embedding
نگاشت هاي تصادفی ..... Random projections
جعبه ي هش ..... Hash bins

تنک Sparse
کرنل های ثابت به جابه جایی Shift invariant kernels
چندی سازی تکرار شونده Iterative quantization
ویژگی های تصادفی فوریه Random Fourier features
تحلیل همبستگی معمولی Ordinary Correlation Analysis (OCA)
سیستم مختصاتی Coordinating System
ضرایب افزونگی Redundancy Coefficients
زائد Redundant
قطری سازی Bi-diagonalization
تکنیک آنالیز اجزای اصلی Principal component analysis
تحلیل فاصله ی کانونی Canonical correlation analysis
رمزنگاری encoding
فقدان چندی سازی Quantization loss
مرتب سازی نتایج بر اساس اهمیت هر بیت Reranking based on bit importance weight
تکنیک بازخور Feedback technique

Abstract:

Hashing technique is an efficient algorithm for approximating nearest neighbor in large-scale image dataset. Learning binary code is a key step to improve its performance and still an ongoing challenge. The inputs of hashing affects its performance. Proposed method in this thesis improve the efficiency of learning binary code by using suitableness input.in this method, image features extracts by local binary pattern then, they use as input vector of iterative quantization (ITQ) that leads to compact codes, less memory and computational requirement. The reasons behind these achievements are the binary nature and high efficiency in feature generation of local binary pattern. Finally in order to rerank results based on most similarity to query image, bit importance reranking method is used and then big weights are assigned to important bits and small weights are assigned to minor bits and weighted hamming distance is used to compute similarity between query image and top results. To evaluate performance of proposed method we use CIFAR-10 and MNIST as dataset and precision vs recall as evaluation criteria. The simulations compare the new algorithm with three state of the art and along the line algorithms from three points of view; the hashing code size, memory space and computational cost, and the results demonstrate the effectiveness of the new approach.

Keywords: large-scale image retrieval, Local Binary Pattern, iterative quantization, image hashing algorithm



kharazmi University

Computer Engineering Department

Large-scale content based image retrieval using local binary patterns and itereative quantization

A Thesis Submitted in Partial Fulfillment of the Requirement for the Degree of Master of Science in Computer Engineering

By:

Mona shakerdonyavi

Supervisor:

Dr.Jamshid Shanbehzadeh

August 2015