

بازیابی تصاویر چهره با استفاده از ترکیب هیستوگرام گرادیان و الگوی باینری محلی

محمد قاصری و حسین ابراهیم نژاد

چکیده

بازیابی چهره، یک موضوع تحقیقاتی مهم در پردازش تصویر است که هدف آن استخراج تصاویر چهره‌ای است که مشابه با یک تصویر جستار باشند. در این مقاله روشی برای بازیابی تصاویر چهره با استفاده از ترکیب هیستوگرام گرادیان و الگوی باینری محلی (LBP) پیشنهاد شده است. ترکیب این دو روش مقاومت در مقابل تغییرات موجود در تصاویر چهره را افزایش می‌دهد و در نتیجه عملکرد سیستم را در بازیابی تصاویر بهبود می‌بخشد. برای افزایش توانایی سیستم، یک طرح فیدبک ارتباطی مبتنی بر ماشین بردار پشتیبان (SVM) معرفی می‌کنیم. آزمایش‌ها بر روی پایگاه داده AR و در دو حالت بدون تصاویر با مانع و با تصاویر با مانع انجام شده است. نتایج آزمایش‌ها نشان می‌دهد که روش پیشنهادی ما به خوبی می‌تواند تصاویر چهره را بازیابی کند. در ادامه، روش پیشنهادی خود را با برخی از روش‌های موفق در توصیف چهره مقایسه کرده‌ایم. معیار دقت متوسط میانگین (MAP) برای روش پیشنهادی در حالت‌های اول و دوم آزمایش به ترتیب برابر است با ۹۴/۴۰٪ و ۶۸/۱۲٪. در حالی که بهترین نرخ برای روش‌های مقایسه‌شده برابر است با ۹۰/۳۷٪ و ۶۱/۹۹٪. این نتایج نشان می‌دهد روش پیشنهادی ما نسبت به این روش‌ها بهتر عمل می‌کند و یک روش خوب برای بازیابی تصاویر چهره است.

کلید واژه‌ها

الگوی باینری محلی، بازیابی چهره، فیدبک ارتباطی، ماشین بردار پشتیبان، هیستوگرام گرادیان.

تصویر که بتواند تصویر مورد نظر را از میان تصاویر آن مجموعه استخراج کند، بیش از پیش احساس می‌شود [۱، ۲].

بازیابی تصویر مبتنی بر محتوا (CBIR) که تصاویر را بر اساس محتوای ظاهری آن‌ها بازیابی می‌کند، در سال‌های اخیر به سرعت گسترش یافته است و به یک موضوع تحقیقاتی مهم تبدیل شده است. بازیابی تصویر مبتنی بر محتوا، تصاویر موجود در یک پایگاه داده را که با تصویر جستار^۲ مطابقت می‌کنند، بازیابی می‌کند. یکی از مهمترین شاخه‌های بازیابی تصویر، بازیابی چهره انسان است که هدف آن، کشف و استخراج تصاویر چهره در یک پایگاه داده می‌باشد که با یک تصویر مرجع یا یک سری از کلمات کلیدی مطابقت داشته باشند. بازیابی چهره یک موضوع تحقیقاتی فعال است که زمینه‌های دیگری مانند پردازش تصویر و ویدئو، بازشناسی الگو، بازشناسی چهره، بینایی کامپیوتر، یادگیری ماشین و شبکه‌های عصبی را دربرمی‌گیرد.

هدف ما در این مقاله، بازیابی تصاویر چهره با استفاده از

۱ مقدمه

با رشد و گسترش فناوری‌های دیجیتال در سال‌های اخیر، تعداد تصاویر دیجیتال بطور چشمگیری افزایش پیدا کرده است. سهولت استفاده از این فناوری‌ها و در دسترس بودن ابزارهای مؤثر برای ذخیره‌سازی و انتقال تصاویر باعث شده است که حجم عظیمی از تصاویر در دسترس عموم مردم قرار گیرد. اما افزایش تعداد تصاویر، مشکلاتی را نیز به همراه داشته است که از جمله مهم‌ترین آنها می‌توان به جستجوی یک تصویر خاص در بین مجموعه‌ای از تصاویر اشاره کرد. یافتن این تصویر خاص به صورت دستی و با نگاه کردن به تک تک تصاویر، کاری زمانبر، خسته‌کننده و حتی در برخی اوقات غیرممکن است. بنابراین نیاز به یک سیستم بازیابی

این مقاله در شهریورماه ۱۳۹۱ دریافت، در دی‌ماه ۱۳۹۱ بازنگری، و در بهمن‌ماه ۱۳۹۱ پذیرفته شد.

محمد قاصری، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق،

دانشگاه صنعتی سهند، sut.ac.ir@m_ghaseri

حسین ابراهیم نژاد، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق،

دانشگاه صنعتی سهند، sut.ac.ir@brahimnezhad

استفاده شده است. الگوی باینری محلی (LBP) [۷] یکی از این روش‌هاست که بطور وسیعی در تحقیقات مربوط به چهره و از جمله بازیابی چهره بکار رفته است. برای مثال در [۸] روشی برای جستجوی تصاویر چهره با استفاده از عملگر LBP ارائه شده است. در این کار، ویژگی‌های چهره با استفاده از LBP استخراج شده است و بازیابی تصاویر چهره با استفاده از این ویژگی‌ها صورت گرفته است. برای بهبود نتایج جستجو، دو روش فیدبک بانظارت و فیدبک بدون نظارت پیشنهاد شده است. در روش فیدبک بانظارت، مانند روش‌های معمولی فیدبک ارتباطی عمل می‌شود. کاربرد قضاوت خود را درباره نتایج بازیابی انجام می‌دهد و تصاویر مرتبط را انتخاب می‌کند. سیستم بر اساس این فیدبک به‌روز شده و نتایج جدیدی را به کاربر نمایش می‌دهد. در روش فیدبک بدون نظارت، کاربر هیچ دخالتی در عملکرد سیستم ندارد و سیستم بطور خودکار نتایج بازیابی در هر تکرار را ارزیابی می‌کند. برای این کار، یک مکانیزم امتیازدهی طراحی شده است. در شروع فرایند بازیابی امتیاز هر تصویر صفر است و بطور متوالی امتیازها به‌روز می‌شود. در نهایت تصاویر با بالاترین امتیاز بازگردانده می‌شوند.

در [۹] نیز بازیابی چهره در یک پایگاه داده ویدئویی با استفاده از روش LBP مورد مطالعه قرار گرفته است. در این کار ابتدا برای مجموعه چهره‌های یک شخص که در فریم‌های ویدئو ظاهر شده‌اند، یک چهره نماینده محاسبه می‌شود و سپس با استفاده از LBP، ویژگی‌های چهره نماینده استخراج می‌گردد. سپس با مقایسه ویژگی‌های متناظر با چهره‌های نماینده اشخاص مختلف، شباهت بین چهره‌ها محاسبه و تصاویر بازیابی می‌شوند.

از بانک فیلتر گابور^۹ نیز که به عنوان یک روش موفق در توصیف چهره شناخته می‌شود، در کارهای زیادی استفاده شده است. به عنوان مثال در [۱۰] ویژگی‌های چهره با استفاده از فیلتر گابور استخراج شده و سپس تطابق چهره توسط نقشه‌های خودسازمانده صورت می‌گیرد. در [۱۱] نیز با استفاده از ویژگی‌های گابور، مسئله بازیابی چهره با تغییرات سن بررسی شده است. در این کار، فیلترهای گابور بر روی نقاط کلیدی چهره اعمال شده و سپس نتایج بازیابی با استفاده از فیدبک ارتباطی بهبود داده شده‌اند. در [۱۲] Lee، و همکارانش از ایده ترکیب LBP و فیلتر گابور که در [۱۳] مطرح شده است، برای استخراج ویژگی استفاده کرده‌اند. آن‌ها ابتدا تصویر چهره را با استفاده از فیلترهای گابور در ۵ مقیاس و ۸ جهت فیلتر کرده و سپس عملگر LBP بر روی این تصاویر اعمال می‌شود.

در [۱] برای بازیابی چهره یک ویژگی جدید به نام ویژگی نقاط گوشه جهتی (DCP)^{۱۰} معرفی شده است. در این ویژگی پس از یافتن منحنی لبه، نقاط گوشه استخراج می‌شوند. بردار ویژگی برای

روشی مبتنی بر هیستوگرام گرادیان^۱ و الگوی باینری محلی (LBP)^۲ است. با استفاده از این روش، مسئله بازیابی تصاویر چهره که شامل تغییراتی مانند تغییرات روشنایی و وجود موانع هستند، را بررسی می‌کنیم. همچنین یک طرح فیدبک ارتباطی^۳ مبتنی بر ماشین بردار پشتیبان (SVM)^۴ برای بهبود عملکرد سیستم معرفی می‌کنیم. ادامه مقاله بصورت زیر نوشته شده است: در بخش ۲ مروری بر تحقیقات گذشته خواهیم داشت و در بخش ۳ روش پیشنهادی را توضیح خواهیم داد. بخش ۴ به نتایج آزمایش‌ها اختصاص دارد و در نهایت در بخش ۵ نتیجه‌گیری را ارائه می‌کنیم.

۲ مروری بر تحقیقات گذشته

تاکنون روش‌های مختلفی برای بازیابی چهره ارائه شده است. در روش‌های زیادی از ویژگی‌های مبتنی بر افکنش^۵ مانند تحلیل مؤلفه اصلی (PCA)^۶ و تحلیل جداساز خطی (LDA)^۷ استفاده شده است. برای مثال در [۳] این روش‌ها و نسخه‌های دو بعدی آنها مانند 2DLDA، 2DPCA و SVD مورد بررسی قرار گرفته‌اند و نویسندگان به این نتیجه رسیده‌اند که روش‌های دو بعدی به نرخ بازیابی بالاتری می‌رسند. در [۴] این نتیجه با استفاده از روشی بنام IMFDA به تأیید رسیده است. در [۵]، Kim و همکارانش برای کاهش تأثیرات ناشی از تغییرات روشنایی و وضعیت سر، از مدل مخلوط PCA مرتبه اول و دوم استفاده کرده‌اند و نشان داده‌اند که مدل مخلوط مرتبه اول در حالتی که تغییرات وضعیت و حالت چهره وجود داشته باشد، بهتر عمل می‌کند و مدل مخلوط مرتبه دوم برای بازیابی تصاویر با تغییرات روشنایی مناسب است.

در [۶] روشی به نام LDA زنجیره‌ای معرفی شده است که از دو مرحله LDA تشکیل می‌شود. در مرحله اول چند زیرتصویر از تصویر چهره انتخاب شده و هر زیرتصویر با استفاده از LDA نمایش داده می‌شود. برای کل تصویر نیز ضرایب LDA محاسبه می‌شود. سپس ضرایب بدست آمده از زیرتصاویر با ضرایب بدست آمده از کل تصویر ترکیب می‌شود و در مرحله دوم یک LDA دیگر به این مجموعه ویژگی‌ها اعمال می‌گردد تا یک توصیف فشرده از چهره بدست آید. در [۲] نیز سیستمی طراحی شده است که در آن تصاویر چهره با استفاده از PCA به یک فضای با بعد کمتر افکنش و سپس از نقشه‌های خودسازمانده (SOM)^۸ برای بازیابی چهره استفاده می‌شود.

در برخی کارها نیز از ویژگی‌های بافت برای استخراج ویژگی

1 Gradient Histogram

2 Local Binary Pattern

3 Relevance Feedback

4 Support Vector Machine

5 Projection

6 Principal Component Analysis

7 Linear Discriminant Analysis

8 Self-Organizing Maps

9 Gabor filters

10 Directional Corner Points

زمان نسبتاً زیادی برای استخراج ناحیه‌ی چهره صرف می‌کند و علاوه بر آن، این الگوریتم به ازای هر تصویر چهره، چند تصویر از ناحیه‌ی چهره در مقیاس‌های مختلف استخراج می‌کند. انتخاب بهترین تصویر به انتخاب مقیاس مناسب و تعامل کاربر بستگی دارد. علاوه بر این، استفاده از این الگوریتم در مواردی نیز با خطا در استخراج چهره و حذف ناقص پس‌زمینه همراه است. روش پیشنهادی ما برای بریدن تصاویر، علاوه بر این که ساده و سریع است، به هیچ دخالتی از سوی کاربر احتیاج ندارد و خطای آن نیز در بسیاری از موارد کمتر از خطای الگوریتم *viola-jones* است.

۲-۳ استخراج ویژگی

روش پیشنهادی ما برای استخراج ویژگی مبتنی بر ترکیب هیستوگرام گرادیان و الگوی باینری محلی (LBP) است. در هر پیکسل تصویر، یک همسایگی $w \times w$ در نظر می‌گیریم. این همسایگی‌ها را سلول می‌نامیم. در هر سلول یک هیستوگرام گرادیان با n بین محاسبه می‌شود. برای محاسبه هیستوگرام گرادیان، بصورت زیر عمل می‌کنیم.

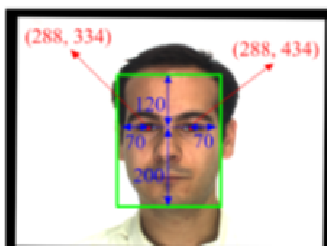
ابتدا تصویر با استفاده از کرنل‌های سوبل^۱ در جهت x و y فیلتر می‌شود تا گرادیان تصویر در راستای x و y بدست آید:

$$G_x = I * D_x \quad (1)$$

$$G_y = I * D_y \quad (2)$$



(الف)



(ب)



(پ)

شکل ۱: فرایند پیش‌پردازش (الف) تصویر اصلی (ب) تصویر تنظیم شده. اعداد روی شکل، مقادیر موردنیاز برای تنظیم و بریدن تصویر را نشان می‌دهد (پ) تصویر بریده شده.

که I تصویر اصلی، D_x و D_y کرنل‌های سوبل در راستای x و y و G_x

یک نقطه‌ی گوشه، از مختصات آن نقطه و زاویه‌ی تمایل آن نقطه به نقاط همسایه قبل و بعدش تشکیل می‌شود. در [۱۴] با استفاده از توصیفات معنایی مانند سن، جنسیت، نژاد، رنگ مو و غیره یک موتور جستجوی قدرتمند برای بازیابی چهره ساخته شده است. در [۱۵]، Wu و همکارانش نیز سیستمی طراحی کرده‌اند که در آن استخراج ویژگی با استفاده از ویژگی‌های محلی و سراسری صورت می‌گیرد. این ویژگی‌ها با استفاده از یک روش کوانتیزاسیون مبتنی بر هویت اشخاص کدگذاری می‌شوند و در نهایت برای بهبود نتایج یک طرح فیدبک ارتباطی که از چند تصویر مرجع برای بازیابی و مرتب کردن تصاویر استفاده می‌کند، بکار رفته است.

۳ روش پیشنهادی

۱-۳ پیش‌پردازش

مرحله پیش‌پردازش یک مرحله مهم و حیاتی برای افزایش دقت و کارایی روش پیشنهادی است. در این مرحله با تنظیم و نرمال‌سازی چهره‌ها، تصاویر را در یک محل خاص ثابت می‌کنیم. علت این کار کاهش تغییرات مقیاس، چرخش و جابجایی در تصاویر چهره است. برای تنظیم چهره‌ها، به طور دستی موقعیت مرکز چشم‌ها را در تصاویر علامت زده‌ایم. سپس تصاویر را با استفاده از دوران و انتقال طوری جابجا می‌کنیم تا مرکز چشم‌ها در همه تصاویر در موقعیت (۲۸۸، ۳۳۴) و (۲۸۸، ۴۳۴) پیکسل قرار گیرند. به این ترتیب فاصله بین مراکز چشم‌ها در همه تصاویر برابر با ۱۰۰ پیکسل می‌شود و در همه تصاویر مراکز چشم‌ها در یک موقعیت ثابت قرار می‌گیرند. این کار باعث می‌شود تغییرات مقیاس، چرخش و جابجایی در تصاویر به حداقل برسد.

پس از این که تصاویر تنظیم شدند، باید آن‌ها را برش دهیم تا فقط دربرگیرنده ناحیه چهره باشند و اثرات پس‌زمینه از بین برود. برای بریدن تصاویر نیز از روش ساده‌ای که مبتنی بر موقعیت جدید مرکز چشم‌هاست، استفاده می‌کنیم. برای این کار از مرکز چشم چپ ۷۰ پیکسل به سمت راست و از خط واصل مراکز چشم‌ها ۱۲۰ پیکسل به سمت بالا و ۲۰۰ پیکسل به سمت پایین را به عنوان ناحیه چهره در نظر می‌گیریم. این اعداد بصورت تجربی و با استفاده از چهره‌های مختلف بدست آمده است. در نتیجه تصاویر بریده شده با اندازه 320×240 پیکسل بدست می‌آیند. با توجه به اینکه تصاویر موجود در این پایگاه داده تمام‌رخ هستند، این روش برای استخراج ناحیه چهره مؤثر واقع می‌شود. در شکل ۱ مراحل نرمال‌سازی و بریدن تصاویر چهره نشان داده شده است.

در نهایت تصاویر را به اندازه 160×120 پیکسل تغییر اندازه می‌دهیم. با این کار اندازه تصاویر کوچکتر شده و در نتیجه فرایند استخراج ویژگی سریعتر صورت می‌گیرد. ما همچنین برای استخراج ناحیه چهره در تصاویر، الگوریتم *viola-jones* [۱۶] را که یکی از مشهورترین الگوریتم‌ها در تشخیص و استخراج چهره است، نیز آزمایش کرده‌ایم. ولی مشاهده کردیم که این الگوریتم

¹Sobel

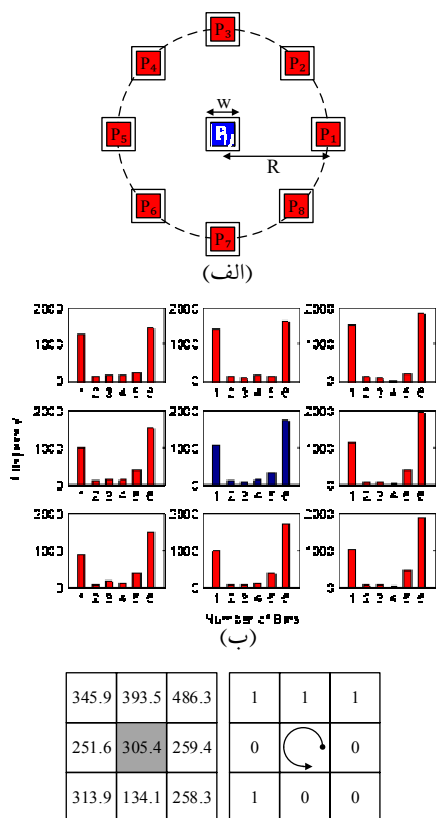
می‌آید. با تبدیل این رشته باینری به معادل دهدهی آن که مقداری بین صفر تا ۲۵۵ خواهد بود، مقدار کد برای آن موقعیت پیکسل بدست خواهد آمد. بطور خلاصه این روش کدگذاری را می‌توان با رابطه زیر بیان کرد:

$$GHBP_{w,R}(i,j) = \sum_{k=1}^8 f(d(h_k, h_{ij})) \cdot 2^{k-1} \quad (۶)$$

در این رابطه، h_k بردارهای همسایه موقعیت (i, j) و $d(\cdot, \cdot)$ فاصله اقلیدسی بین بردارها را نشان می‌دهد. تابع f نیز بصورت زیر تعریف می‌شود:

$$f(x) = \begin{cases} 1, & x \geq d_{av} \\ 0, & otherwise \end{cases} \quad (۷)$$

این روش کدگذاری در شکل ۲ نشان داده شده است. قسمت (الف) یک پیکسل و ۸ همسایه آن را در شعاع R ، قسمت (ب) هیستوگرام اختصاص یافته به هر پیکسل و قسمت (ج) نحوه کدگذاری این هیستوگرام‌ها را نشان می‌دهد. در قسمت (ج) در مربع سمت چپ، اعداد حاشیه‌ای نشان‌دهنده فاصله بین هیستوگرام متناظر با پیکسل مرکزی و ۸ همسایه آن و عدد مرکزی که با رنگی متفاوت مشخص شده است، مقدار میانگین این فاصله‌هاست. در مربع سمت راست نیز مقادیر آستانه گذاری شده، رشته باینری معادل و مقدار کد نشان داده شده‌اند.



شکل ۲: (الف) نمایش سلول متناظر با یک پیکسل و ۸ همسایه آن در شعاع R (ب) هیستوگرام متناظر با هر سلول (پ) کدگذاری هیستوگرام‌ها با روش LBP.

G_y گرادیان تصویر در راستای x و y و علامت * عمل کانولوشن را نشان می‌دهند. سپس اندازه و جهت گرادیان در هر پیکسل به صورت زیر بدست می‌آید:

$$|G(i,j)| = \sqrt{(G_x(i,j))^2 + (G_y(i,j))^2} \quad (۳)$$

$$\theta_G(i,j) = \tan^{-1} \left(\frac{G_y(i,j)}{G_x(i,j)} \right) \quad (۴)$$

که $|G|$ اندازه گرادیان، θ_G جهت گرادیان و i و j به ترتیب نشان‌دهنده شماره سطرها و ستون‌های تصویر هستند. برای تصاویر رنگی، گرادیان برای هر کانال رنگی به صورت جداگانه محاسبه می‌شود و بزرگترین مقدار برای هر پیکسل بعنوان بردار گرادیان آن پیکسل انتخاب می‌شود. برای محاسبه هیستوگرام گرادیان در هر سلول، ابتدا زاویه گرادیان را بصورت زیر به محدوده ۰-۱۸۰ درجه محدود می‌کنیم:

$$\theta'_G = \begin{cases} \theta_G & , 0 \leq \theta_G < 180 \\ \theta_G - 180 & , 180 \leq \theta_G < 360 \end{cases} \quad (۵)$$

سپس، فاصله بین ۰-۱۸۰ درجه به n فاصله مساوی تقسیم می‌شود که n تعداد جهت‌های گرادیان یا همان بین‌های هیستوگرام را نشان می‌دهد و هرکدام از این فاصله‌ها یک کانال هیستوگرام را تشکیل می‌دهند. دلیل این که بجای محدوده ۰-۳۶۰ درجه از محدوده ۰-۱۸۰ درجه استفاده می‌شود، این است که معمولاً در محدوده ۰-۳۶۰ درجه به تعداد بین‌های بیشتری برای استخراج اطلاعات نیاز است که این امر موجب صرف زمان بیشتری برای استخراج ویژگی می‌گردد. همچنین مشاهدات تجربی ما نشان داد که استفاده از محدوده ۰-۳۶۰ درجه نقش چندانی در بهبود نتایج نسبت به محدوده ۰-۱۸۰ درجه ندارد. این مسئله در بخش ۳-۴ نشان داده خواهد شد.

برای محاسبه هیستوگرام، هر پیکسل در داخل سلول بر مبنای زاویه گرادیان به یکی از کانال‌های هیستوگرام رأی می‌دهد. این رأی‌ها بر اساس اندازه گرادیان در آن پیکسل وزن دار می‌شوند. پس از محاسبه هیستوگرام گرادیان در هر سلول، این هیستوگرام به پیکسل مرکزی سلول اختصاص داده می‌شود. بنابراین برای هر موقعیت پیکسل، یک بردار n -بعدی که نشان‌دهنده هیستوگرام گرادیان همسایگی اطراف آن است، بدست می‌آوریم.

برای کد کردن این بردارها، از روش LBP استفاده می‌کنیم. برای هر موقعیت پیکسل، ۸ همسایه آن را در شعاع R در نظر می‌گیریم. برای مثال موقعیت (i, j) را در نظر بگیرید. اختصاص یافته به این موقعیت را h_{ij} می‌نامیم. سپس فاصله اقلیدسی^۱ بین بردار h_{ij} و هر یک از ۸ بردار متناظر با همسایه‌های آن را محاسبه می‌کنیم. علاوه بر آن، مقدار میانگین این ۸ فاصله را نیز بدست می‌آوریم و آن را d_{av} می‌نامیم. اگر فاصله بین هر یک از بردارهای همسایه و بردار h_{ij} ، از d_{av} بیشتر یا مساوی با آن باشد، به آن موقعیت، ۱ و در غیر این صورت صفر اختصاص می‌دهیم. بنابراین برای هر موقعیت پیکسل، یک رشته باینری ۸ رقمی بدست

۳-۴ فیدبک ارتباطی

یکی از مهم‌ترین مشکلات در سیستم‌های بازیابی تصویر، فاصله معنایی است که بصورت اختلاف بین توانایی بالای ادراک انسان و قدرت محدود ویژگی‌های تصویری تعریف می‌شود. با وجود فاصله معنایی، تصاویر بازیابی شده ممکن است تطابق زیادی با تصویر جستار نداشته باشند، به ویژه اگر تصاویر پایگاه داده از تنوع زیادی برخوردار باشند. یک راه‌حل مؤثر برای کاهش فاصله معنایی و برقراری ارتباط بین مفاهیم مدنظر کاربر و ویژگی‌های تصویری سطح پایین، استفاده از فیدبک ارتباطی است. فیدبک ارتباطی یک فرایند تعاملی و تکراری است که برای پالایش کردن نتایج جستجو بکار می‌رود. در فیدبک ارتباطی، در هر تکرار تصاویر مرتبط و یا تصاویر نامرتب انتخاب می‌شوند. سپس سیستم بر اساس این فیدبک، به روز شده و در تکرار بعدی تصاویر دیگری را به کاربر ارائه می‌کند. این فرایند تا جایی ادامه می‌یابد که تصاویر هدف در بین تصاویر بازگردانده شده ظاهر شود.

ما از یک روش مبتنی بر SVM برای بازخورد ارتباطی استفاده می‌کنیم. در این روش در هر تکرار، تصاویر مرتبط و نامرتب را مشخص کرده و سپس با استفاده از آنها یک SVM خطی را آموزش می‌دهیم. برای آموزش SVM، برچسب تصاویر مرتبط را +۱ و برچسب تصاویر نامرتب را -۱ در نظر می‌گیریم:

$$y_i = \begin{cases} +1 & , x_i \in S_{\text{Relevant}} \\ -1 & , x_i \in S_{\text{Irrelevant}} \end{cases} \quad (9)$$

که y_i برچسب تصاویر و S_{Relevant} و $S_{\text{Irrelevant}}$ مجموعه تصاویر مرتبط و نامرتب را نشان می‌دهد. پس از آموزش SVM، مقدار تصمیم‌گیری در الگوریتم SVM را برای تصاویر مدل محاسبه می‌کنیم. مقدار تصمیم‌گیری در الگوریتم SVM از رابطه زیر بدست می‌آید:

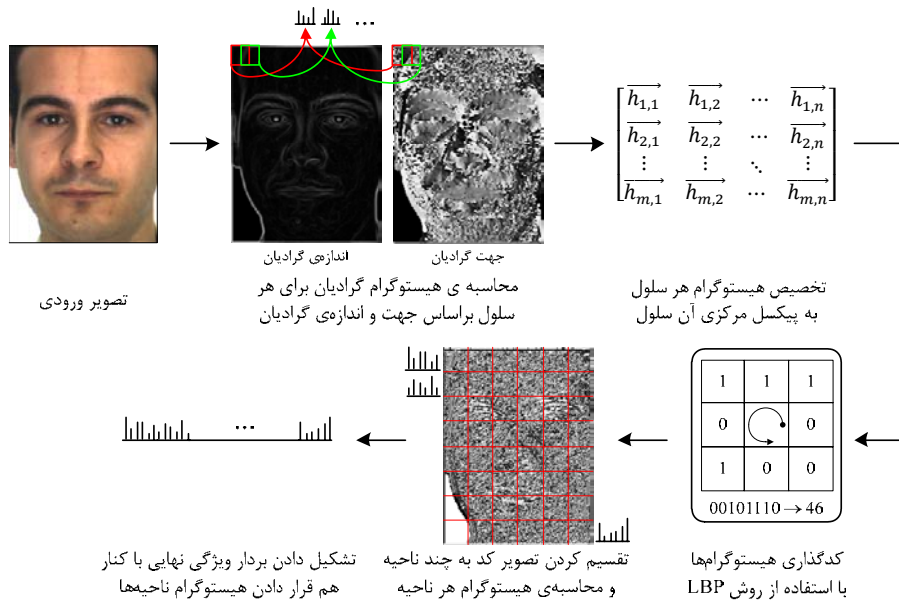
با اعمال این روش کدگذاری برای همه پیکسل‌های تصویر، تصویر کد بدست می‌آید. برای درنظر گرفتن اطلاعات مکانی تصویر چهره در بردار ویژگی، تصویر کد به چندین ناحیه تقسیم و هیستوگرام هر ناحیه با ۲۵۶ بین محاسبه می‌شود. بردار ویژگی نهایی با کنار هم قرار دادن هیستوگرام این نواحی تشکیل می‌شود. شکل ۳ فرایند استخراج ویژگی را بطور خلاصه نشان می‌دهد.

۳-۳ مرحله بازیابی

در مرحله بازیابی، تصاویر را با مقایسه بین بردارهای ویژگی متناظرشان بازیابی می‌کنیم. برای این کار بردار ویژگی برای تصویر جستار و تصاویر مدل محاسبه شده و با استفاده از یک اندازه شباهت، فاصله بین آنها بدست می‌آید. در نهایت M تصویر که کمترین فاصله را داشته باشند، به عنوان تصاویر مرتبط بازیابی می‌شوند. ما از فاصله اقلیدسی برای بدست آوردن فاصله بین بردارهای ویژگی استفاده می‌کنیم. اگر بردار ویژگی برای تصویر جستار و تصویر مدل را به ترتیب f_m و f_q بنامیم، فاصله اقلیدسی به صورت زیر تعریف می‌شود:

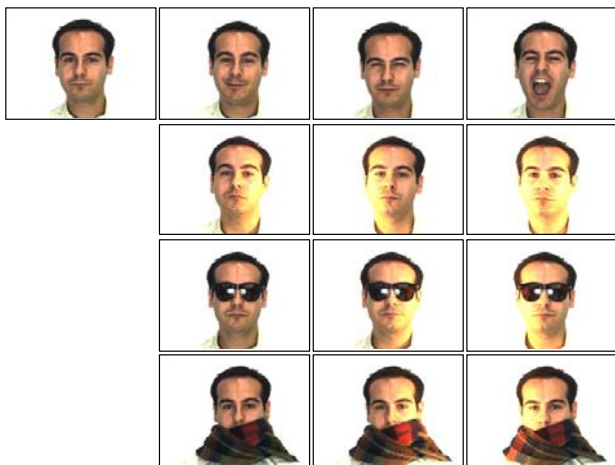
$$D_E(f_q, f_m) = \sqrt{\sum_{i=1}^L (f_{q_i} - f_{m_i})^2} \quad (8)$$

که زیرنویس i شماره ویژگی و L طول بردار ویژگی است. همچنین فاصله‌های دیگری از جمله فاصله cityblock، فاصله اشتراک هیستوگرام‌ها، فاصله خی مربع^۲، فاصله hellinger، فاصله earthmover و فاصله gower را تست کرده‌ایم، اما مشاهدات ما نشان می‌دهد که فاصله اقلیدسی علاوه بر این که بسیار ساده و سریع می‌باشد، نسبت به توابع فاصله دیگر عملکرد بهتری دارد. در بخش ۴-۳-۴ تأثیر توابع فاصله بررسی خواهد شد.



شکل ۳ فرایند استخراج ویژگی.

- 1 Histogram Intersection
- 2 chi2



شکل ۴ نمونه‌ای از تصاویر پایگاه داده AR

معیار دیگر، معیار فراخوانی^۳ است که بصورت نسبت تعداد تصاویر درست بازیابی شده، N_c ، به تعداد تصاویر مرتبط در پایگاه داده، N_{rel} ، تعریف می‌گردد:

$$Re = \frac{N_c}{N_{rel}} \quad (14)$$

مشابه بالا، فراخوانی متوسط با میانگین گرفتن از مقدار فراخوانی برای همه تصاویر جستار حاصل می‌شود. همانطور که از روابط فوق پیداست، وقتی که تصاویر بیشتری بازیابی شوند، دقت کاهش و فراخوانی افزایش می‌یابد. لذا برای مصالحه بین معیارهای دقت و فراخوانی به جای استفاده جداگانه از این معیارها، از یک منحنی بنام منحنی دقت-فراخوانی^۴ استفاده می‌شود. منحنی دقت-فراخوانی با رسم مقادیر دقت به ازای مقادیر فراخوانی متناظر بدست می‌آید. معمولاً سطوح فراخوانی به ۱۱ مقدار از صفر تا یک با گام ۰/۱ محدود می‌شود و مقادیر دقت به ازای این ۱۱ مقدار محاسبه می‌شود. در نهایت منحنی برای این زوج مقادیر رسم می‌گردد.

در یک سیستم بازیابی تصویر، رتبه تصاویر بازیابی شده نیز مهم است. اما هیچ کدام از معیارهای دقت و فراخوانی اطلاعاتی از رتبه تصاویر بازیابی شده را منعکس نمی‌کنند. بنابراین به معیار دیگری نیاز است تا بتواند رتبه تصاویر بازیابی شده را نیز در نظر بگیرد. یکی از مهم‌ترین معیارها، معیار دقت متوسط میانگین (MAP)^۵ است که بصورت زیر تعریف می‌شود:

$$MAP = \frac{1}{N_Q} \sum_{q=1}^{N_Q} \frac{1}{N_{rel}(q)} \sum_{k=1}^{N_{rel}(q)} Pr(k) \times rel(k) \quad (15)$$

که $N_{rel}(q)$ تعداد تصاویر مرتبط به تصویر جستار q ، $Pr(k)$ تعداد تصاویر بازیابی شده برای تصویر جستار q و $rel(k)$ دقت در رتبه k را نشان می‌دهد و $rel(k)$ تابعی است که برابر است با یک، اگر تصویر در رتبه k یک تصویر مرتبط باشد و در غیر اینصورت برابر است با صفر. [۱۹]

$$D(x) = w^T x + b \quad (10)$$

که w بردار ضرایب فوق صفحه جداکننده و b مقدار بایاس است که در فرایند آموزش SVM بدست آمده‌اند. در نهایت معیار شباهت در تکرار کنونی با اضافه کردن مقدار تصمیم‌گیری به آن در تکرار قبلی، به روز می‌شود:

$$S_{iter+1}(x) = S_{iter}(x) + D(x) \quad (11)$$

پس از به روز رسانی، این معیار شباهت بصورت نزولی مرتب شده و بر این اساس تصاویر جدیدی در هر تکرار بازگردانده می‌شوند. این روش تا جایی ادامه پیدا می‌کند که یا معیارهای ارزیابی در تکرار کنونی نسبت به تکرار قبل بهبود پیدا نکنند یا تعداد تکرارها از حد معینی بیشتر شود. حداکثر تعداد تکرارها را ۳ تکرار در نظر می‌گیریم. برای پیاده‌سازی SVM از جعبه‌ابزار LIBSVM [۱۷] استفاده می‌کنیم.

۴ نتایج آزمایش‌ها

۴-۱ پایگاه داده AR

پایگاه داده AR [۱۸] یک پایگاه داده با شرایط کنترل شده است. تصاویر این پایگاه داده در دو دوره مختلف با فاصله زمانی دو هفته گرفته شده است. در دوره اول ۱۳۵ شخص (شامل ۷۶ مرد و ۵۹ زن) و در دوره دوم، از ۱۳۵ شخص حاضر در دوره اول، ۱۲۰ شخص (شامل ۶۵ مرد و ۵۵ زن) حضور دارند. در هر دوره ۱۳ تصویر از هر شخص و تحت شرایط کنترل شده گرفته شده است. تصاویر بصورت تمام‌رخ و شامل تغییرات روشنایی، تغییرات حالت چهره و موانعی مانند عینک آفتابی و شال‌گردن هستند. تعداد کل تصاویر موجود در پایگاه داده، ۳۳۱۵ تصویر و اندازه تصاویر ۵۷۶×۷۶۸ پیکسل است. در شکل ۴ نمونه‌ای از تصاویر این پایگاه داده برای یک شخص خاص نشان داده شده است.

۴-۲ معیارهای ارزیابی

برای ارزیابی عملکرد یک سیستم بازیابی تصویر، معیارهای متفاوتی بکار می‌رود. یکی از رایج‌ترین معیارها، نرخ بازیابی^۱ یا دقت^۲ است که بصورت نسبت تعداد تصاویر درست بازیابی شده، N_c ، به تعداد تصاویر بازیابی شده، N_{ret} ، تعریف می‌شود:

$$Pr = \frac{N_c}{N_{ret}} \quad (12)$$

با میانگین گرفتن از نرخ بازیابی برای همه تصاویر جستار، نرخ بازیابی متوسط بدست می‌آید:

$$ARR = \frac{1}{N_Q} \sum_{q=1}^{N_Q} Pr(q) \quad (13)$$

که N_Q تعداد تصاویر جستار و $Pr(q)$ نرخ بازیابی برای تصویر جستار q را نشان می‌دهد.

3Recall

4Precision-recall Curve

5Mean Average Precision

1Retrieval Rate

2Precision

۳-۴ ارزیابی پارامترها

بصورت {۱، ۲، ۳، ۴} تغییر می‌دهیم. برای سایر پارامترها نیز از پارامترهایی که در آزمایش‌های قبل بدست آمده‌اند، استفاده می‌کنیم. در شکل ۵-ج نتایج آزمایش با تغییر شعاع LBP رسم شده است. می‌توان مشاهده کرد که بهترین نتیجه با شعاع ۲ بدست می‌آید. در این حالت معیار MAP برابر است با ۹۴/۴۰٪.

بنابراین بهترین پارامترها برای ادامه آزمایش‌ها عبارتند از: سلول‌های با اندازه ۹×۹ پیکسل، ۳ بین برای محاسبه هیستوگرام گرادیان در محدوده ۱۸۰-۰ درجه و شعاع LBP برابر با ۲.

۴-۳-۴ تأثیر توابع فاصله

در این قسمت نقش توابع فاصله اقلیدسی، cityblock، histogram intersection، chi2، hellinger، earthmover و gower را بررسی می‌کنیم. شکل ۵-د تأثیر این فاصله‌ها را نشان می‌دهد. همانطور که از این شکل پیداست، تابع فاصله اقلیدسی نسبت به سایر فاصله‌ها بهتر عمل می‌کند. هر چند فاصله‌های cityblock، histogram intersection، chi2، hellinger و earthmover نیز عملکرد تقریباً مشابهی با فاصله اقلیدسی دارند. فاصله‌های earthmover و gower نیز نسبت به فاصله‌های دیگر کمی ضعیفتر ظاهر می‌شوند. قابل ذکر است که فاصله اقلیدسی علاوه بر برتری در بازایی تصاویر، بسیار ساده و سریع است. بنابراین در ادامه از فاصله اقلیدسی استفاده می‌کنیم.

۴-۴ نتایج آزمایش بر روی پایگاه داده AR

در پایگاه داده AR یکی از مهم‌ترین مشکلات، وجود موانع است. برای بررسی این محدودیت، آزمایش‌ها را در دو حالت انجام می‌دهیم: الف) بدون در نظر گرفتن تصاویر با مانع که شامل ۲۵۵ تصویر جستار و ۱۵۳۰ تصویر مدل است و ب) با در نظر گرفتن تصاویر با مانع که شامل ۲۵۵ تصویر جستار و ۳۰۶۰ تصویر مدل می‌باشد.

نتایج آزمایش برای هر دو حالت در جدول ۱ آورده شده است. مشاهده می‌شود که در حالت اول آزمایش، نتایج خوبی بدست آمده است. در این حالت عمده تغییرات در تصاویر، تغییرات روشنایی است و روش پیشنهادی ما نسبت به تغییرات روشنایی مقاوم است. زیرا هیستوگرام گرادیان و LBP هر دو به تغییرات روشنایی مقاومند و بالطبع ترکیب آنها مقاومت در برابر تغییرات روشنایی را افزایش می‌دهد. در حالت دوم آزمایش که وجود موانع در تصاویر، عمده تغییرات را شامل می‌شود، عملکرد سیستم کاهش می‌یابد. زیرا در این حالت قسمت‌هایی از تصویر چهره که برای تمایز بین اشخاص حاوی اطلاعات مهمی هستند، پشت موانع قرار می‌گیرند و در نتیجه بخشی از اطلاعات از بین می‌رود.

جدول ۱: نتایج روش پیشنهادی برای هر دو حالت آزمایش.

بدون در نظر گرفتن تصاویر با مانع		با در نظر گرفتن تصاویر با مانع	
MAP (%)	ARR (%)	MAP (%)	ARR (%)
۹۴/۴۰	۶۶/۲۸	۶۸/۱۲	۹۱/۶۵

روش پیشنهادی ما دارای چند پارامتر قابل تنظیم است که تغییر آنها بر روی عملکرد سیستم تأثیر می‌گذارد. این پارامترها عبارتند از: اندازه سلول‌ها، تعداد بین‌های هیستوگرام، محدوده زاویه گرادیان (۱۸۰-۰ درجه یا ۳۶۰-۰ درجه) و شعاع LBP. برای ارزیابی پارامترها از تصاویر بدون مانع در پایگاه داده AR استفاده می‌کنیم. در پایگاه داده AR به ازای هر شخص، ۷ تصویر بدون مانع وجود دارد. از بین این تصاویر، برای هر شخص تصویر با حالت طبیعی را بعنوان تصویر جستار و مابقی تصاویر بدون مانع آن شخص را که شامل تصاویر با تغییرات حالات ظاهری و تغییرات روشنایی هستند، بعنوان تصاویر مدل در نظر می‌گیریم. بنابراین برای ارزیابی پارامترها، ۲۵۵ تصویر جستار و ۱۵۳۰ تصویر مدل در اختیار داریم.

۳-۳-۱ تأثیر اندازه سلول‌ها

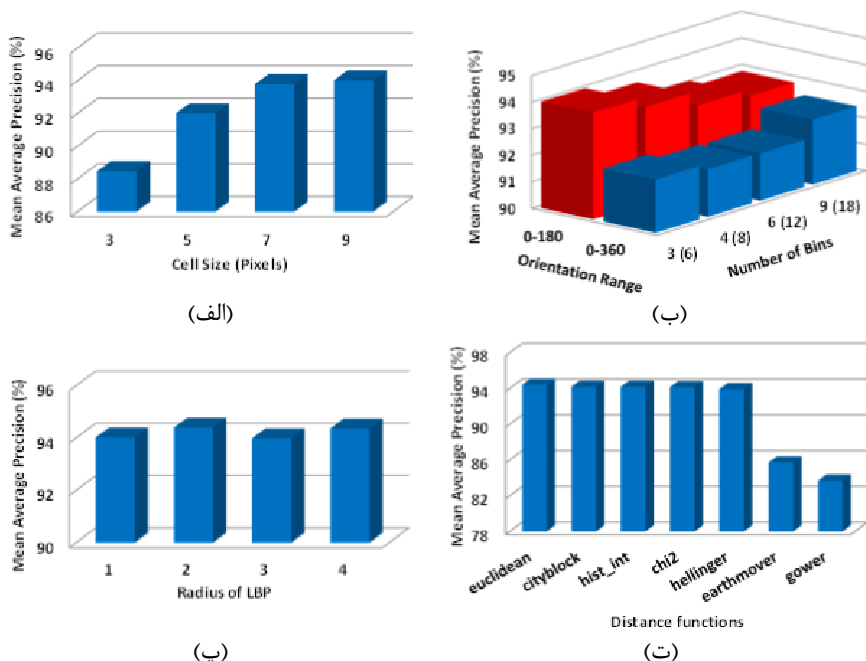
در اولین آزمایش، تأثیر اندازه سلول‌ها را بررسی خواهیم کرد. برای این کار ۴ سلول با اندازه‌های ۳×۳، ۵×۵، ۷×۷ و ۹×۹ انتخاب می‌کنیم. در این آزمایش، برای محاسبه هیستوگرام گرادیان از ۳ بین استفاده می‌کنیم و شعاع LBP را برابر با ۱ در نظر می‌گیریم. شکل ۵-الف نتایج آزمایش را با تغییر اندازه سلول‌ها نشان می‌دهد. از این شکل پیداست که بهترین عملکرد با سلول‌های ۹×۹ پیکسل حاصل می‌شود. بنابراین برای آزمایش‌های بعدی از سلول‌های ۹×۹ استفاده می‌کنیم.

۳-۳-۲ تأثیر محدوده زاویه گرادیان و تعداد بین‌های هیستوگرام

در این قسمت، تأثیر محدوده زاویه گرادیان و تعداد بین‌های هیستوگرام را بررسی می‌کنیم. بر اساس نتیجه آزمایش قبل، از سلول‌های ۹×۹ استفاده می‌کنیم. در اینجا نیز شعاع LBP را برابر با ۱ در نظر می‌گیریم. در حالتی که گرادیان در محدوده ۱۸۰-۰ درجه در نظر گرفته شده است، تعداد بین‌ها را بصورت {۳، ۴، ۶، ۹} تغییر می‌دهیم و در حالتی که گرادیان در محدوده ۳۶۰-۰ درجه در نظر گرفته شود، تعداد بین‌ها را دو برابر می‌کنیم. در شکل ۵-ب، معیار MAP با تغییر تعداد بین‌های هیستوگرام و محدوده زاویه گرادیان نشان داده شده است. در این شکل اعداد داخل پرانتز، تعداد بین‌ها را در حالتی که گرادیان در محدوده ۳۶۰-۰ درجه در نظر گرفته شده است، نشان می‌دهد. همانطور که پیداست، در حالتی که زاویه گرادیان در محدوده ۳۶۰-۰ درجه در نظر گرفته شود و همچنین با افزایش تعداد بین‌ها، عملکرد سیستم کاهش می‌یابد. بهترین عملکرد سیستم با ۳ بین و در حالتی که گرادیان در محدوده ۱۸۰-۰ درجه باشد، حاصل می‌شود. در نتیجه این پارامترها را برای ادامه آزمایش‌ها انتخاب می‌کنیم.

۳-۳-۴ تأثیر شعاع LBP

در این آزمایش، تأثیر شعاع LBP را بررسی می‌کنیم. شعاع LBP را



شکل ۵ تأثیر پارامترهای الگوریتم پیشنهادی بر روی نتایج آزمایش (الف) اندازه سلولها (ب) محدوده زاویه گرادیان و تعداد بین‌های هیستوگرام (پ) شعاع LBP (ت) توابع فاصله.

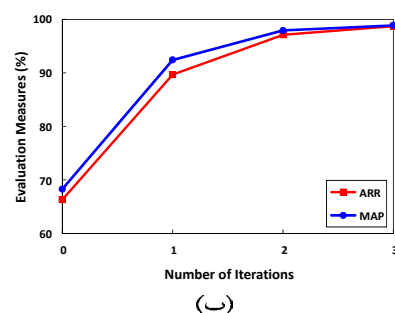
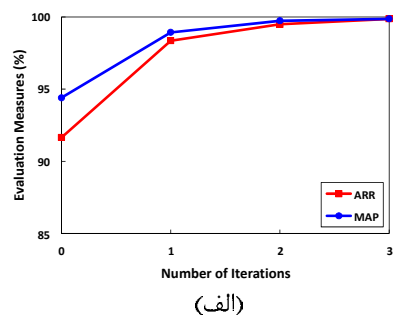
۴-۵ مقایسه با سایر روش‌ها

در این قسمت روش پیشنهادی خود را با برخی از روش‌های موفق و شناخته‌شده در توصیف چهره مقایسه می‌کنیم. توجه کنید که این مقایسه تنها از لحاظ قدرت بردار ویژگی استخراج شده صورت گرفته است و نقش فیدبک ارتباطی در این مقایسه لحاظ نشده است.

برای مقایسه از روش‌های eigenface [۲۰]، LBP [۷] و دو نسخه از آن بنام‌های TPLBP و FPLBP [۲۱]، ترکیب گابور و LBP [۱۳] و در نهایت HOG [۲۲] استفاده می‌کنیم. برای پیاده‌سازی روش eigenface، ۱۰۰ بردار ویژه را برای افکنش چهره‌ها انتخاب می‌کنیم. در روش LBP از الگوهای یکنواخت با شعاع ۲ که شامل ۸ نمونه می‌شوند، استفاده می‌کنیم. روش‌های TPLBP و FPLBP را نیز با پارامترهای ذکر شده در [۲۱] بکار می‌بریم. در روش Gabor+LBP ابتدا تصاویر را با استفاده از فیلترهای گابور در ۵ مقیاس و ۸ جهت فیلتر می‌کنیم و سپس عملگر LBP را با پارامترهایی که در بالا به آن اشاره شد، بر روی این تصاویر اعمال می‌کنیم. در روش HOG نیز از پارامترهایی که در [۲۲] بعنوان بهترین پارامترها گزارش شده است، استفاده می‌کنیم.

نتایج مقایسه در جدول ۲ نشان داده شده است. همان‌گونه از این جدول پیداست، در هر دو حالت آزمایش، روش پیشنهادی ما نسبت به سایر روش‌ها عملکرد بهتری از خود نشان می‌دهد. همچنین در شکل ۷ منحنی دقت-فراخوانی برای این روش‌ها رسم شده است.

برای بهبود عملکرد سیستم، از فیدبک ارتباطی استفاده می‌کنیم. شکل ۶ نتایج آزمایش را با استفاده از فیدبک ارتباطی نشان می‌دهد. از این شکل می‌توان مشاهده کرد که در حالت اول آزمایش، پس از ۳ تکرار معیارهای ARR و MAP هر دو به ۹۹/۸۷٪ می‌رسند که به ترتیب میزان ۸/۲۲٪ و ۵/۴۷٪ بهبود را نشان می‌دهد. در حالت دوم آزمایش، اثر فیدبک ارتباطی محسوس‌تر بوده و معیارهای ARR و MAP پس از ۳ تکرار به ترتیب ۳۲/۴۸٪ و ۳۰/۸۷٪ افزایش یافته و به ۹۸/۷۶٪ و ۹۸/۹۰٪ می‌رسند.



شکل ۶: اثر فیدبک ارتباطی در بهبود نتایج آزمایش (الف) بدون در نظر گرفتن تصاویر با مانع (ب) با در نظر گرفتن تصاویر با مانع.

جدول ۲: مقایسه روش پیشنهادی با سایر روش‌ها برای هر دو حالت آزمایش.

	با در نظر گرفتن تصاویر با مانع		بدون در نظر گرفتن تصاویر با مانع		
	MAP (%)	ARR (%)	MAP (%)	ARR (%)	
eigenface	۳۲/۸۳	۳۵/۸۳	۷۷/۴۲	۷۳/۸۱	
LBP	۶۰/۸۵	۵۹/۵۴	۸۵/۱۶	۸۱/۰۶	
TPLBP	۶۱/۵۳	۵۹/۹۱	۸۷/۲۵	۸۳/۷۰	
FPLBP	۶۶/۹۱	۶۴/۸۴	۹۰/۳۷	۸۶/۰۰	
Gabor+LBP	۶۰/۳۶	۵۹/۳۱	۸۵/۶۲	۸۰/۹۸	
HOG	۵۶/۱۹	۵۵/۴۸	۸۷/۷۱	۸۴/۳۹	
روش پیشنهادی	۶۸/۱۲	۶۶/۲۸	۹۴/۴۰	۹۱/۶۵	

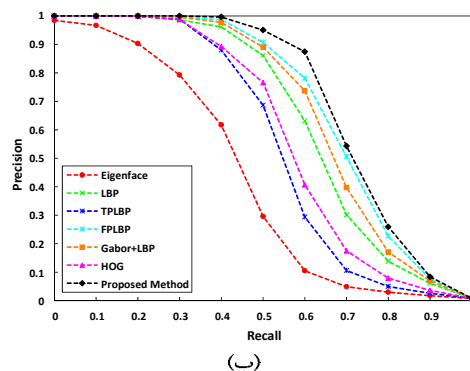
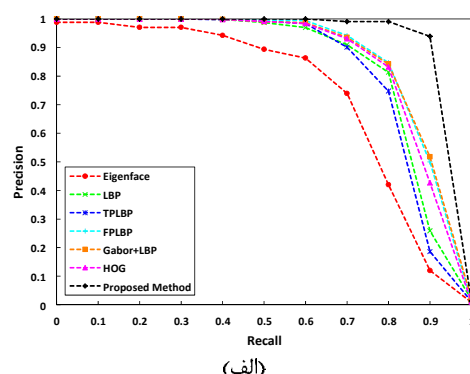
می‌کند و نسبت به تغییرات روشنایی نیز مقاوم‌تر از مقادیر سطح خاکستری تصویر است. در میان این روش‌ها، روش Gabor+LBP با وجود عملکرد نسبتاً خوب، زمان زیادی را برای استخراج ویژگی صرف می‌کند و دارای بردار ویژگی با طول بسیار زیادی است که نیاز به فضای زیادی برای ذخیره‌سازی دارد و همچنین منجر به کاهش سرعت در مرحله بازیابی می‌گردد. همچنین مزیت روش پیشنهادی ما نسبت به روش HOG که مبتنی بر هیستوگرام گرادینان است، ترکیب آن با روش LBP است که همان طور که اشاره کردیم ترکیب این دو روش نسبت به تغییرات روشنایی، مقاوم‌تر از حالتی است که فقط از یکی از آنها استفاده شود.

در جدول ۳ نیز مقایسه‌ای بین روش پیشنهادی ما و سایر روش‌ها، از لحاظ طول بردار ویژگی و زمان استخراج ویژگی صورت گرفته است. توجه کنید که چون استخراج ویژگی در روش eigenface به تعداد کل تصاویر بستگی دارد، در این روش زمان استخراج ویژگی در حالت‌های اول و دوم آزمایش تفاوت خواهد داشت و به همین دلیل در این روش، دو زمان برای استخراج ویژگی ذکر شده است که نشان‌دهنده زمان استخراج ویژگی برای حالت‌های اول و دوم آزمایش می‌باشد. این جدول نشان می‌دهد که بیشترین طول بردار ویژگی به روش Gabor+LBP تعلق دارد. طول بردار ویژگی در روش پیشنهادی ما برابر با روش TPLBP است و نسبت به سایر روش‌ها طول بیشتری دارد. از لحاظ زمان استخراج بردار ویژگی، روش LBP بسیار سریع عمل می‌کند، در حالی که روش Gabor+LBP زمان زیادی برای استخراج ویژگی صرف می‌کند. استخراج ویژگی در روش پیشنهادی ما نسبت به روش LBP و نسخه‌های آن زمان بیشتری به طول می‌انجامد، ولی نسبت به سایر روش‌ها زمان کمتری صرف می‌شود.

با توجه به اینکه روش پیشنهادی ما از ترکیب دو مرحله شامل استخراج گرادینان و کدگذاری آنها با استفاده از روش LBP تشکیل می‌شود، زمان صرف شده برای استخراج ویژگی، زمان نسبتاً مناسبی است و می‌توان نتیجه گرفت که روش پیشنهادی ما می‌تواند بصورت برخط نیز بکار رود. کلیه آزمایش‌ها بر روی یک کامپیوتر با فرکانس ۱/۷۳ گیگاهرتز و ۴ گیگابایت حافظه و با استفاده از نرم‌افزار MATLAB انجام شده است.

جدول ۳: مقایسه روش پیشنهادی با سایر روش‌ها از لحاظ زمان استخراج ویژگی و طول بردار ویژگی.

روش	زمان استخراج ویژگی (میلی ثانیه)	طول بردار ویژگی
eigenface	۱۵۰/۲۸-۳۳۰/۳۷	۱۰۰
LBP	۷/۰۶	۲۸۳۲
TPLBP	۱۴۶/۴۶	۱۲۲۸۸
FPLBP	۹۷/۱۶	۷۶۸
Gabor+LBP	۶۸۲/۴۳	۱۱۳۲۸۰
HOG	۲۵۸/۹۴	۹۵۷۶
روش پیشنهادی	۲۳۷/۳۲	۱۲۲۸۸



شکل ۷: منحنی دقت-فراخوانی برای روش‌های مختلف (الف) بدون در نظر گرفتن تصاویر با مانع (ب) با در نظر گرفتن تصاویر با مانع.

از میان روش‌های مقایسه شده، روش eigenface بدترین عملکرد را از خود نشان می‌دهد. این نتیجه مورد انتظار نیز بود، زیرا این روش در بازیابی تصاویر با تغییرات روشنایی و وجود موانع نمی‌تواند خوب عمل کند. همچنین روش ما نسبت به روش LBP و نسخه‌های آن و همچنین Gabor+LBP که بر اساس مقادیر سطح خاکستری پیکسل‌ها عمل می‌کنند، نتایج بهتری به همراه دارد. مزیت روش کدگذاری ما نسبت به این روش‌ها این است که بجای مقادیر سطح خاکستری پیکسل‌ها از هیستوگرام گرادینان همسایگی اطراف پیکسل‌ها استفاده شده است. هیستوگرام گرادینان، اطلاعات تصویر را در جهت‌های مختلف استخراج

۵ نتیجه گیری

بردارهای ویژگی عملکرد چندان خوبی حاصل نمی‌شود. در این حالت با استفاده از یک طرح فیدبک ارتباطی که بر اساس SVM کار می‌کند، عملکرد سیستم را بهبود دادیم و با حداکثر ۳ تکرار، توانستیم به نرخ بسیار بالایی در بازیابی تصاویر چهره دست پیدا کنیم. نتایج آزمایش‌ها ثابت کرد که در مواقعی که تصاویر چهره شامل تغییرات زیادی هستند و در نتیجه روش‌های معمولی که فقط با مقایسه ویژگی‌های استخراج شده از آن‌ها نمی‌توانند از پس این تغییرات برآیند، استفاده از این طرح فیدبک ارتباطی برای بهبود عملکرد سیستم و جبران نواقص ویژگی‌های استخراج شده، بسیار مؤثر و مفید است.

همچنین روش پیشنهادی خود را با برخی از روش‌های موفق در توصیف چهره مقایسه کردیم و مشاهده کردیم که روش پیشنهادی ما نسبت به این روش‌ها عملکرد بهتری در بازیابی تصاویر چهره از خود نشان می‌دهد. سرعت و دقت روش پیشنهادی ما باعث می‌شود تا این روش به عنوان یک روش خوب برای بازیابی تصاویر چهره معرفی شود.

در این مقاله یک توصیفگر چهره که مبتنی بر هیستوگرام گرادیان و الگوی باینری محلی (LBP) است، معرفی کردیم و با استفاده از آن به بازیابی تصاویر چهره پرداختیم. بخاطر ترکیب هیستوگرام گرادیان و LBP، روش پیشنهادی ما تا حد زیادی نسبت به تغییرات روشنایی مقاوم است. علاوه بر آن استفاده از هیستوگرام گرادیان، باعث می‌شود تا بتوانیم اطلاعات تصویر را در جهت‌های مختلف استخراج کنیم که این امر منجر به کدگذاری مؤثرتری می‌شود.

آزمایش‌ها بر روی پایگاه داده AR و در دو حالت انجام شده است: یک بار بدون در نظر گرفتن تصاویر با مانع و بار دیگر با در نظر گرفتن این تصاویر. در حالت اول آزمایش مشاهده کردیم که مقایسه ساده بین بردارهای ویژگی با استفاده از تابع فاصله اقلیدسی، نتایج خوبی به همراه دارد که این امر به دلیل عدم وجود تغییرات زیاد در تصاویر است. ولی در حالت دوم آزمایش دیدیم که تصاویر با مانع کمی کار را سخت‌تر می‌کنند و با مقایسه بین

مراجع

- [1] Y. Gao and Y. Qi, "Robust visual similarity retrieval in single model face databases", Pattern Recognition, vol. 38, pp. 1009-1020, 2005.
- [2] J. Ruiz-del-Solar and P. Navarrete, "FACERET: An interactive face retrieval system based on self-organizing maps", in Conference of Image and Video Retrieval, London, UK, pp. 45-64, 2002.
- [3] C. C. Chen, Y. S. Shieh, and H. T. Chu, "Face image retrieval by projection-based features", in International Workshop on Image Media Quality and Applications, Kyoto, Japan, pp. 138-144, 2008.
- [4] C. Zhang, H. Chen, M. Chen, and Z. Sun, "Image matrix fisher discriminant analysis (IMFDA)-2D matrix based face image retrieval algorithm", in International Conference on Advances in Web-Age Information Management, Hangzhou, China, pp. 894-899, 2005.
- [5] H. C. Kim, D. Kim, and S. Y. Bang, "Face retrieval using 1st-and 2nd-order PCA mixture model", in International Conference on Image Processing, Rochester, New York, pp. 605-608, 2002.
- [6] T. K. Kim, H. Kim, W. Hwang, and J. Kittler, "Component-based LDA face description for image retrieval and MPEG-7 standardisation", Image and Vision Computing, vol. 23, pp. 631-642, 2005.
- [7] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition", IEEE transactions on pattern analysis and machine intelligence, vol. 28, pp. 2037-2041, 2006.
- [8] A. Franco and D. Maio, "Similarity searches in face databases", in International Conference on Image Analysis and Processing, Vietri sul Mare, Italy, pp. 443-450, 2009.
- [9] T. N. Nguyen, T. D. Ngo, D. D. Le, S. Satoh, B. H. Le, and D. A. Duong, "An efficient method for face retrieval from large video datasets", in Conference of Image and Video Retrieval, Xi'an, China, pp. 382-389, 2010.
- [10] T. W. S. Chow and M. K. M. Rahman, "Face matching in large database by self-organizing maps", Neural processing letters, vol. 23, pp. 305-323, 2006.
- [11] N. Nitta, A. Usui, and N. Babaguchi, "Face image retrieval across age variation using relevance feedback", in Advances in Multimedia Modeling, Chongqing, China, pp. 152-162, 2010.
- [12] H. Lee, Y. Chung, J. Kim, and D. Park, "Face image retrieval using sparse representation classifier with gabor-LBP histogram", in International Workshop on Information Security Applications, Jeju Island, Korea, pp. 273-280, 2011.
- [13] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition", in International Conference on Computer Vision, Beijing, China, pp. 786-791, 2005.
- [14] N. Kumar, P. Belhumeur, and S. Nayar, "FaceTracer: A search engine for large collections of images with faces",

- [22] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", in Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, pp. 886-893, 2005.

محمد قاصری مدرک کارشناسی خود را در سال ۱۳۸۸ در رشته مهندسی برق-الکترونیک از دانشگاه یزد دریافت کرد و در سال ۱۳۹۱ تحصیلات مقطع کارشناسی ارشد خود را در رشته مهندسی برق-مخابرات در دانشگاه صنعتی سهند به پایان رساند. زمینه‌های تحقیقاتی مورد علاقه او پردازش تصویر و ویدئو، بینایی کامپیوتر و بازشناسی الگو است.



حسین ابراهیم‌نژاد مدرک کارشناسی و کارشناسی ارشد خود را به ترتیب در سال‌های ۱۳۷۲ و ۱۳۷۵ در رشته مهندسی برق-الکترونیک و برق مخابرات از دانشگاه تبریز و دانشگاه صنعتی خواجه نصیرالدین طوسی اخذ کرد. همچنین مدرک دکتری خود را در گرایش مخابرات سیستم در سال ۱۳۸۶ از دانشگاه تربیت مدرس دریافت کرد. زمینه‌های تحقیقاتی مورد علاقه ایشان بینایی کامپیوتر، پردازش مدل سه بعدی، پردازش تصویر، شناسایی الگو و محاسبات نرم بوده و در حال حاضر عضو هیأت علمی دانشگاه صنعتی سهند می باشد.



in European Conference on Computer Vision, Marseille, France, pp. 340-353, 2008.

- [15] Z. Wu, Q. Ke, J. Sun, and H. Y. Shum, "Scalable face image retrieval with identity-based quantization and multireference reranking", IEEE transactions on pattern analysis and machine intelligence, vol. 33, pp. 1991-2001, 2011.
- [16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", in Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, pp. I-511-I-518, 2001.
- [17] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines", ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, pp. 27, 2011.
- [18] A. M. Martinez and R. Benavente, "The AR face database", CVC Technical Report, vol. 24, 1998.
- [19] C. D. Manning, P. Raghavan, and H. Schutze, *Introduction to information retrieval*. Cambridge: Cambridge University Press, 2008.
- [20] M. Turk and A. Pentland, "Eigenfaces for recognition", Journal of cognitive neuroscience, vol. 3, pp. 71-86, 1991.
- [21] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, pp. 1978-1990, 2011.