سیستم احراز هویت بیومتریک بر روی دستگاه های شخصی همراه

چکیده

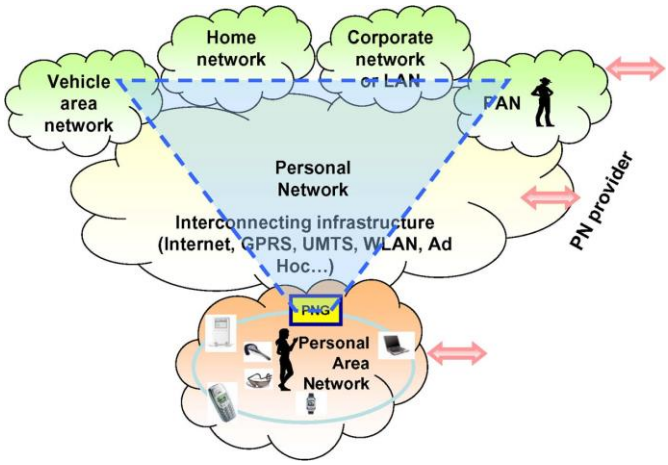
ما یک سیستم احراز هویت بیومتریک امن، قوی و کم هزینه برای دستگاه های همراه شخصی برای PN پیشنهاد کرده ایم. این سیستم از ۵ ماژول کلیدی زیر تشکیل شده است:

۱. تشخیص چهره
۲. ثبت چهره
۳. روشنایی و نرمال سازی
۴. تایید چهره
۵. ترکیب اطلاعات

در مورد وظیفه ی سنگین احراز هویت وسایل با منابع محدود، تاکید بیشتر بر روی اعتبار و قابلیت اجرای سیستم است. هر دو جنبه تئوری و کاربردی مورد توجه قرار گرفت است. سیستم نهایی توانایی رسیدن به نرخ خطای ۲٪ کمتر از پروتکل های سنجش نقادانه را دارد. هزینه ی کم نرم افزاری و سخت افزاری، باعث شد این سیستم، برای محدوده ی وسیعی از Applicationهای امنیتی مفید باشد.

۱- مقدمه

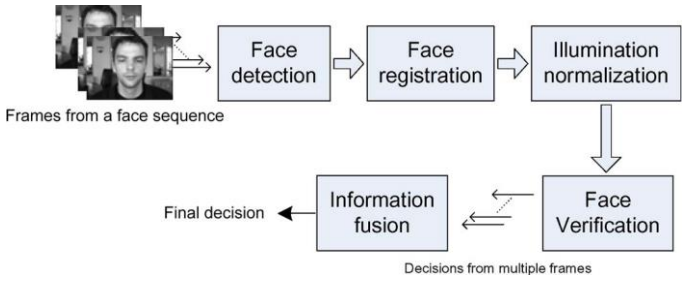
در دنیای مدرن، موقعیت های زیادی وجود دارند که هویت ما باید قابل اثبات باشد. اما هویت ما چیست؟ اغلب یک رمز عبور، یک پاسپورت یا شماره ی اجتماعی است.ارتباط بین این موارد و یک فرد ممکن است ضعیف باشد. چون ممکن است گم شوند، دزدیده شوند یا فراموش گردند. بیومتریک ها، خصوصیات رفتاری یا بیولوژیکی منحصر به فرد انسان، مثلا صورت، اثر انگشت، گفتار، عنبیه و ... یکی از مرسوم ترین و قابل قبول ترین جایگزین برای حل این مشکل است. بیومتریک ها مناسبند، زیرا مردم به طور طبیعی آنها را با خود دارند و قابل اعتمادند، زیرا تنها وسیله ی احراز هویت اند که حضور فیزیکی کاربر را تضمین می کنند. در این مقاله، موضوع احراز هویت بیومتریک روی دستگاه های همراه شخصی در PN[1] را مورد مطالعه قرار می دهیم. PN یک محیط ارتباطی برای ارتباط نامحدود خود کاربر و وسایل الکترونیک شخصی اش است. توصیف یک PN در تصویر زیر دیده می شود.



[1] Personal Network

سیستم احراز هویت بیومتریک به عنوان یک پیوند امن بین کاربر و PN در نظر گرفته می شود و دسترسی امن کاربر به PN را فراهم می کند. این نوع برنامه کاربردی ۳ نیاز سیستم احراز هویت بیومتریک را پیش روی ما قرار می دهد: امنیت، راحتی و پیچیدگی.

امنیت، اولین و اصلی ترین دلیل برای معرفی احراز هویت بیومتریک برای PN است. در سناریو های MPD[۲]دو گونه احراز هویت وجود دارد: احراز هویت در زمان ورود (log on) و احراز هویت در زمان اجرا (run time). علاوه بر احراز هویت زمان ورود، احراز هویت زمان اجرا نیز از اهمیت زیادی برخوردار است. زیرا این ویژگی می تواند از دسترسی کاربران غیر مجاز و استفاده ی آنها از MPD و نیز دسترسی به اطلاعات محرمانه ی کاربر از طریق PN جلوگیری کند. برای کمیت سنجی احراز هویت بیومتریک از دیدگاه امنیت، و به منظور به منظور تشخیص احتمال استفاده ی فرد فریبکار از دستگاه، از خاصیت FAR[۳] استفاده می شود. از نقطه نظر امنیتی، اندازه ی FAR باید low باشد. FRR[۴] احتمال reject شدن کاربر معتبر را تعیین می کند؛ این خاصیت نیز به راحتی کار کاربر مربوط است. FRR کاربر را مجبور می کند دوباره داده ی بیومتریک را وارد کند که باعث ایجاد ناراحتی قابل توجهی در کاربر است. این موضوع منجر به نیاز به یک FRR کم از سیستم احراز هویت بیومتریک می شود. به علاوه از نقطه نظر راحتی، اگر احراز هویت بیومتریک درست باشد، درجه ی بالایی از کاربر دوستی به دست می آید. این بدین معنی است که احراز هویت بدون اقدامات صریح کاربر قابل انجام شدن است. به طورکلی، یک وسیله ی همراه منابع محاسباتی محدودی دارد. MPD در PN کار می کند، اما امکان ذخیره ی الگوهای بیومتریک را در یک database مرکزی فراهم می کند و احراز هویت در شبکه انجام می شود. اگرچه محدودیت در پیچیدگی الگوریتمی از دقت کمتری برخوردار است، option ها ریسک امنیتی بالاتری را به همراه دارند. اولاً هنگامی که لازم باشد داده ی بیومتریک بر روی شبکه منتقل شود، نسبت به استراق سمع آسیب پذیر است. دوماً الگوهای بیومتریک باید در یک database ذخیره شوند و نسبت به جملات آسیب پذیرند. از نظر مفهوم، ترجیح می دهیم که احراز هویت MPD مستقل تر از دیگر بخش های PN باشد. بنابراین لازم است که احراز هویت بیومتریک به صورت محلی و در MPD انجام پذیرد. اختصاصاً، سخت افزار (مثلاً سنسور بیومتریک) باید ارزان، و نرم افزار (مثلاً الگوریتم) باید پیچیدگی محاسباتی کمی داشته باشد. در این مقاله، سیستم احراز هویت ارزان، راحت و امنی را برای MPD ها در PN توسعه داده ایم. بیومتریک انتخاب شده، تصویر دو بعدی چهره است که توسط دوربین MPD گرفته می شود. تصویر دوبعدی یکی از بهترین بیومتریک ها از نظر دقت، شفافیت و هزینه است. بنابراین احراز هویت کاربر، با آنالیز کردن تصویر چهره ی فردی که قصد ورود به PN یا استفاده از MPD را دارد، انجام می شود. تنها تجهیزات مورد نیاز برای استفاده از این سیستم، این است که کاربر تصویر چهره اش را مقابل دوربین قرار دهد. اگرچه که ریز مسئله های شناسایی چهره (مثل ثبت، روشنایی، کلاس بندی و ...) در چندین مقاله منتشر شده و آدرس دهی شده اند، اما سیستم کامل تشخیص چهره در مقاله های کمی توصیف شده اند. به ویژه برای دستگاه های همراه کوچک که منابع محاسباتی محدودی دارند. در اینجا ما یک سیستم احراز هویت بیومتریک که شامل ۵ ماژول کلیدی مهم است و در تصویر ۲ دیده می شود را شرح می دهیم.



[2] Mobile Personal Devices
[3] False Acceptance Rate
[4] False Rejection Rate

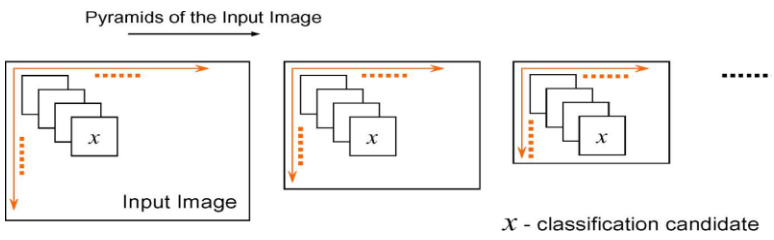
ماژول ها عبارتند از:

- تشخیص چهره
- ثبت چهره
- نرمال سازی روشنایی
- احراز هویت چهره
- ترکیب اطلاعات

بدین ترتیب این مقاله نیز به صورت زیر سازماندهی شده است: بخش های ۲ و ۳ شامل تشخیص چهره ی بلادرنگ و قدرتمند و الگوریتم های ثبت است. بخش ۴ متد روشنایی نرمال سازی، بخش ۵ متد تأیید چهره ی الگوهای چهره است و بخش ۶ ترکیب اطلاعات بین فریم های زمانی مختلف را شرح می دهد. بخش ۷ نتایج تجربی را نشان می دهد و بخش ۸ شامل نتیجه گیری است.

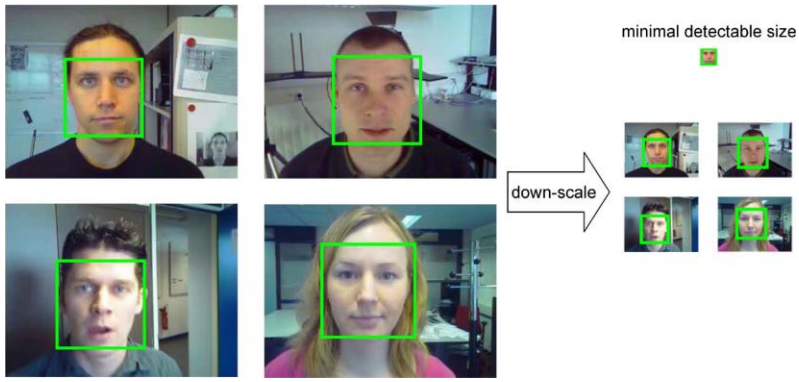
۲- شناسایی چهره

تشخیص چهره اولین گام برای احراز هویت چهره است. اگر چه که تشخیص چهره یک کار بصری راحت برای انسان است، اما با توجه به این واقعیت که چهره پویاست و تغییرات زیادی دارد، مسأله ای پیچیده برای بینایی کامپیوتر است. متدهای شناخته شده ی تشخیص چهره می توانند به دو گروه زیر تقسیم شوند: ۱- متدهای مبتنی بر اکتشافات و غیر مستدل ۲- متدهای مبتنی بر کلاس بندی. مثال هایی از دسته ی اول شامل متدهای رنگ پوست و متدهای هندسه ی صورت است. متدهای اکتشافی و تجربی معمولاً پیاده سازی آسانی دارند، اما قابل اعتماد نیستند. تجربه ها و اکتشافات، معمولاً نسبت به تغییرات بیرونی آسیب پذیرند. در مقایسه، متدهای مبتنی بر مقایسه، بیشتر قابلیت مواجه شدن با سناریو های پیچیده را دارند. چون آنها با تشخیص چهره، به عنوان یک مسأله ی کلاس بندی الگو رفتار می کنند. بنابراین به طور گسترده از منابع کلاس بندی الگوهای موجود بهره مند می شوند. هرچند بزرگ ترین نقطه ضعف متدهای مبتنی بر کلاس بندی، بارگیری بالای محاسباتی آنهاست. زیرا الگوها برای طبقه بندی باید مجموعه ای جامع از تکه های تصویر را در هر مقیاس و در هر محلی پوشش دهند.



آشکار ساز چهره ی Viola-Jones یکی از موفق ترین متدهای تشخیص چهره است. در این متد ۳ خصیصه وجود دارد: ۱- ویژگیهای haar-like که در تمام مقیاس ها به سرعت قابل محاسبه است. ۲- آموزش Adaboost برای انتخاب و اندازه گیری خصوصیات. ۳- ساختار کلاس بندی cascade برای سرعت بخشیدن به تشخیص.

آشکار ساز صورت فقط به یک بار آموزش نیاز دارد و پس از آن داده ها را برای تشخیص هر چهره ذخیره می کند. این خصیصه ها تشخیص قدرتمند چهره را امکان پذیر می سازند. برای استفاده از تشخیص چهره در MPD ها، ما استراتژی هایی را که باعث سرعت بخشیدن به تشخیص چهره می شود پیشنهاد می کنیم. دقت (وضوح) تصاویر چهره در MPD به توزیع اندازه های صورت در عکس های معمولی که توسط خود شخص از وسایل دستی گرفته شده، بستگی دارد. این اطلاعات محدودیت های مفیدی را در جست و جو فراهم می کند و پیاده سازی را به طور قابل توجهی سرعت می بخشد. در سمت چپ تصویر ۴ چند تصویر نمونه ی چهره که با PDA دستی گرفته شده، مشاهده می گردند.



فرض کنید چهره ی تشخیص داده شده با اندازه ی s در ناحیه ای بین s_{min} و s_{max} قرار گیرد. ما دو مرحله را برای کاهش مراحل تشخیص چهره پیشنهاد می کنیم. در ابتدا مقیاس تصویر اصلی را قبل از شناسایی کم کنید. فاکتور کوچک نمایی طبق s_{min}/s_{face} تنظیم می شود. $s_{template}$ اندازه ی نمونه ی آموزشی است، مثلاً حداقل اندازه ی قابل تشخیص. در آشکار سازهای آموزش دیده، $s_{template}=24$. در گام دوم؛ در تصویر کوچک شده، پنجره ی اسکن را محدود کنید. از حداقل سایز ۲۴ تا حداکثر سایز (s_{max} / s_{min}) ۲۴. با توجه به تصویر ۳، به راحتی دیده می شود که تعداد کاندیدها برای کلاس بندی، به صورت نمایی با سایز تصویر ورودی افزایش می یابد. بنابراین، اولین گام، تعداد دفعات طبقه بندی ممکن را کاهش می دهد. به علاوه، گام دوم، از جست و جوی غیر لازم برای تصاویر با اندازه های خیلی کوچک یا خیلی بزرگ، اجتناب می کند.این کار بعدها، تعداد unitهای کلاس بندی را تا اندازه ی زیادی کاهش می دهد. شکل ۴ نتایج شناسایی را در هر دو تصویر اصلی و تصویر کوچک شده نشان می دهد. مشاهده می کنیم که در تصویر دوم (آخر) نیز نتایج خوبی با محاسبات کمتر به دست آمده است. هر چند یکی از اشکالات down scaling [کم کردن مقیاس] این است که مقیاس چهره ی شناسایی شده درشت تر (نامطبوع تر) از تصویر اصلی است، زیرا scale های بسیار کمتری مورد بررسی قرار گرفته اند. این موضوع بر روی شناسایی نهایی چهره تأثیر نمی گذارد و در بخش بعد دیده می شود.

۳- ثبت چهره

به طور کلی، مکان چهره ی شناسایی شده برای آنالیز محتوای صورت، به اندازه ی کافی دقیق نیست. به تأکید، ثبت چهره گام الزامی بعد از شناسایی چهره است. دو راه معمول زیر برای ثبت چهره عبارتنداز: ۱- متدهای جامع (holistic) ۲- متدهای محلی. روش های ثبت جامع شامل هر دو نوع اطلاعات کلی بافت صورت و اطلاعات ترکیب صورت است. (مثل: محل چشم ها، بینی و دهان و ...). مثال هایی از این دسته عبارتند از:

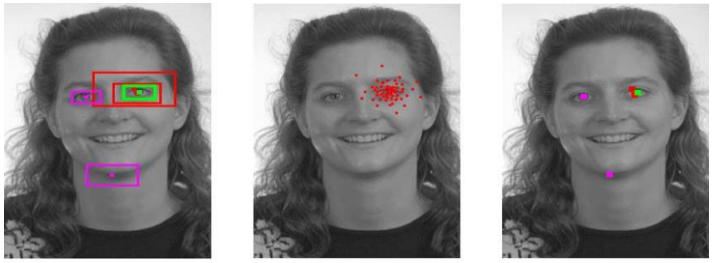
Active shape models - Active appearance models

Fit کردن این مدلها با تصاویر چهره ی ورودی معمولاً به صورت یک مسأله ی بهینه سازی تکراری[5] فرموله می شود. هرچند مطابق مسائل بهینه سازی پیچیده، ثبت نام holistic دو اشکال بالقوه دارد: امکان افتادن در دام مینیمم محلی و محاسبات بالا به ویژه برای MPD ها. در مقابل از طرف دیگر متد ثبت نام محلی، مستقیم تر و سریع تر است. زیرا برای محاسبه ی تغییر و تبدیلات[6]، از ویژگی های محلی چهره استفاده می کند. اما اشکال متد محلی این است که ویژگی های چهره با توجه به تنوع زیاد، به سختی قابل تشخیص است. در این مقاله، یک آشکار ساز بلادرنگ و قدرتمند چهره بر اساس متد Viola-Jones پیشنهاد می کنیم که با مدل novel error که دقیق و مختصر است ترکیب می شود. با توجه به آن که در واقعیت، ساخت یک آشکار ساز ویژگی های چهره قابل اعتماد با low FAR و low FRR غیر ممکن است، ما شناسایی ویژگی های چهره را در وهله ی اول با هزینه ی تعداد زیادی از شناسایی های اشتباه، ضمانت می کنیم. بنابراین، مسأله ی تشخیص ویژگی های چهره، به یک مسأله ی post

[5] iterative
[6] transformation

selection چندگانه ی تشخیص ویژگی های چهره تبدیل می شود. یک راه مرسوم برای انجام post selection ساخت مدل های آماری مثل ASM (active shape models) است. فرض اولیه ی این مدل ها این است که ویژگی های شناسایی شده، احتمالاً در اطراف محل های صحیح توزیع شده اند. هر چند این فرض در مورد شناسایی های چندگانه که از مدل Viola-Jones به دست می آید صحیح نیست. علاوه بر این، برای یک سیستم خاص کاربر، این مدل ارجح تر است. اما در عمل، این کار ممکن نیست. شکل ۵(a) یک مثال نمونه از تشخیص left-eye است.

ما یک error model جدید برای تشخیص های اشتباه (کاذب) ویژگی های چهره تولید کرده ایم. اساساً آشکار ساز Viola-Jones از ترکیب ساختار های محلی به عنوان الگو استفاده می کند، بنابراین تمام الگوهایی که ساختارهای کم و بیش مشابه داشته باشند، تقریباً قابل شناسایی اند. با توجه به این مکانیزم، دو نوع پذیرش false قابل تشخیص اند. False Acceptance نوع اول، پذیرش تکه های پس زمینه است که



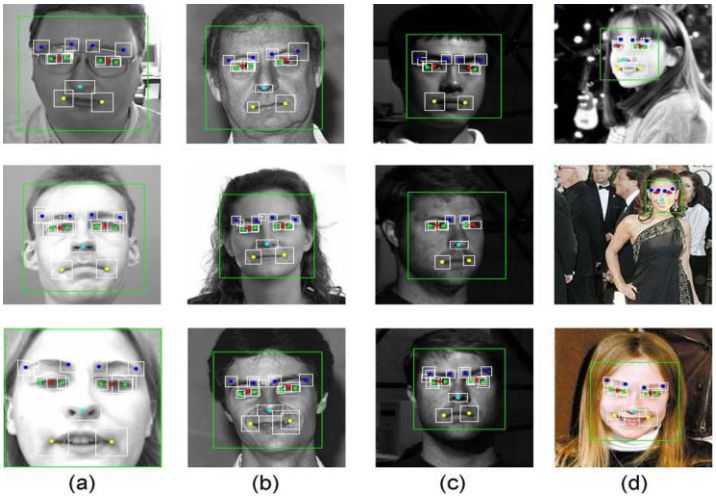
(a) (b) (c)

بطور اتفاقی ساختارهای محلی قابل مقایسه با ویژگی های چهره را دارد. سایه ی چانه که به غلط به عنوان چشم شناسایی شده، همان طور که در شکل ۵(a) دیده می شود، مثال خوبی برای خطای نوع اول است، چون سایه دارای الگوی روشن-تاریک-روشن[۷] است که شبیه بافت چشم به نظر می رسد. False Acceptance نوع دوم، پذیرش تکه های متمرکز تقریباً در همان مکان ویژگی های صورت و البته در اندازه های بزرگ تر است. خطای نوع دوم، با این واقعیت ایجاد می شود که جست و جو در مقیاس های مختلف صورت می گیرد و در مقیاس کمی درشت تر از True position؛ تکه های تصویر شناسایی شده معمولاً ساختارهای مشابه تکه های ساختار صورت را دارند. هردو نوع خطا در شکل ۵(a) قابل مشاهده اند. شکل ۵(c) مدل پیشنهادی را همراه با جزئیاتش در مقایسه با مدل احتمالی معمولی نشان داده شده در ۵(b) توصیف می کند. بدیهی است که مدل پیشنهادی، توزیع تشخیص های غلط را بهتر توصیف می کند.

قانونی برای حذف False Acceptance ها پیشنهاد شده است: تشخیص در مقیاس حداقل در داخل تشخیص در مقیاس حداکثر که به احتمال زیاد، محل صحیح را تشخیص می دهد. استدلال این اصل، مستقیماً به مکانیزم متد Viola-Jones مربوط است: اول اینکه تشخیص ها از بین تشخیص مقیاس حداکثر، شانس کمتری برای قرار گرفتن در خطای پذیرش نوع اول را دارد. زیرا آنها چندین بار با تشخیص های هم پوشانی شده، تأیید شده اند. دوم اینکه به نظر می رسد تشخیص مقیاس حداقل، درون تشخیص مقیاس حداکثر، همان تشخیص درست باشد. (به استثنای پذیرش غلط نوع دوم)

بنابراین پذیرش غلط نوع دوم به عنوان اطلاعات اضافی برای تأیید محل تصویر هستند؛ اما در انتها حذف می شوند. در مقابل، مدل آماری اضافی، می تواند تشخیص های غلط نوع اول را حذف کند، اما در اصل، نمی تواند با تشخیص های غلط نوع دوم مواجه شود، زیرا آنها به واقعیت بسیار نزدیکند. شکل ۶ تعدادی مثال از database ها و تصاویر بلادرنگ نشان می دهد.

[7] Bright-dark-bright



برای قابل قبول تر نشان دادن ثبت چهره در سیستم اتوماتیک، ۱۳ شاخص چهره[8] را از دیتابیس BioID آموزش داده ایم که در اولین عکس تصویر ۶ دیده می شود. در حالت تئوری، به منظور ثبت چهره، دو شاخص برای محاسبه ی تغییر و تحول کافی اند، اما landmark های بیشتر، در یک فضای کوچکتر مربعی، قدرتمند ترند. خطاهای گاه به گاه برای تشخیص شاخص های خاص، رخ می ده، اما در بیشتر موارد، تعداد نقاط شاخص شناسایی شده برای یک registration قابل اعتماد به اندازه ی کافی بزرگ است.

قدرتمندی و سرعت آشکار ساز ویژگی های چهره، از متد Viola-Jones به ارث برده شده اند و استراتژی post selection که در ادامه صحبت می کنیم، توانایی عملکرد خود را در MPD تقویت می کند. به طور خلاصه، آشکار سازهای پیشنهادی خصوصیات چهره، بسیار سریع و متکی به خود هستند و به هیچ مدل بافتی یا shape اضافی نیازی ندارند. همچنین نیازمند هیچ بهینه ساز تعاملی هم نیستند.

۴- روشنایی و نرمال سازی

تغییر پذیری تصاویر چهره که در اثر تغییرات نور به وجود می آید، یکی از بزرگ ترین موانع تشخیص چهره است. گفته شده که تغییرات بوجود آمده روی تصویر بر اثر تغییر نور، به راحتی از تغییرات حاصل از تغییر خود کاربر، پیشی گرفته است. [اون قدری که تغییرنور روی شناسایی تصویر موثر هست، تغییر چهره ی خود فرد موثر نیست] .

بنابراین، روشنایی و نرمال سازی، یک گام پیش پردازش بسیار مهم قبل از تشخیص چهره است. مطالعه ی فشرده ای درمورد این موضوع انجام شده و باعث به وجود آمدن دو متدلوژی مختلف شده است. گروه اول در مورد مسأله ی روشنایی مطالعه می کند که با مدل تصویر برداری فیزیکی و بازیابی مدل ۳ بعدی چهره به طور صریح یا ضمنی کار می کند. این دسته را متدهای ۳بعدی می نامیم که شامل متد شبه خطی، مخروط نور، هارمونیک کروی، خارج قسمت تصویر و ... است. هرچند دسته ی دوم بر recovering کامل اطلاعات ۳ بعدی تکیه نمی کنند و مستقیماً روی مقادیر پیکسل ۲بعدی کار می کنند. این دسته را متدهای ۲بعدی می نامیم. **متدهای ۳ بعدی، به بکار گیری اطلاعات 3D که برای روشنایی تصویر مفید هستند، کمک می کنند. هرچند تبدیل اشیاء ۳ بعدی به تصاویر ۲ بعدی، باعث از بین رفتن بخشی از اطلاعات می شود. روند معکوس این عمل، محدودیت ها و قواعدی را برای جبران آن اطلاعات از دست رفته، تولید می کند.**

با توجه به پیچیدگی الگوریتم، متدهای ۳ بعدی، در سیستم تشخیص چهره، محاسبات بسیار سنگینی دارند. **متدهای ۲بعدی ساده ترند. از سوی دیگر، به دلیل این صراحت و سادگی، برای متدهای ۲ بعدی امکان دستیابی به تغییر ناپذیری نور چنانکه قبلاً به صورت تئوری ثابت شده بود، وجود ندارد.** بنابراین ما در اینجا، **یک فیلتر غیر حساس نور ۲ بعدی** برای سیستم تشخیص هویت چهره، برای

[8] 13 facial landmarks

MPD به نام LBP[9] پیشنهاد کرده ایم. LBP در ابتدا برای حل مسأله ی texture [مثل پازل !!] پیشنهاد شد. ایده ی اصلی بدین صورت بیان می شود:

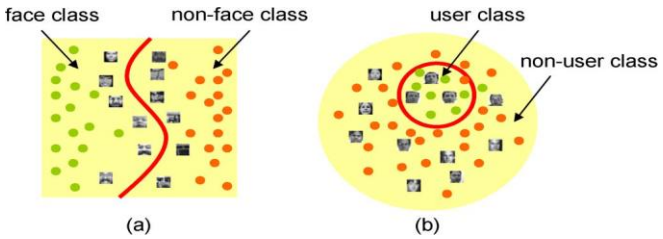
تصویر دارای بلوک های ۳*۳ در همسایگی هم توسط مقدار آستانه ی پیکسل مرکزی و به ۸ مقدار دودویی منتج می شود. پس از آن، این توالی دودویی ۸بیتی به یک محدوده ی دهدهی بین ۰ تا ۲۵۵ تبدیل می شود و بدین ترتیب نوع الگوی textureهای اطراف آن پیکسل مرکزی را نمایش می دهد. توزیع الگوهای LBP در سرتاسر تصویر به عنوان طرح کلی تصویر استفاده می شود. [تصویر را نمایش می دهند.]

هیستوگرام LBP به عنوان یک سنجش صحیح از بافت محلی به رسمیت شناخته شده و به نور و پارامترهای دوربین حساس است. این هیستوگرام، برای نمایش تصاویری که بافتهای یکنواخت کمتر یا بیشتری دارند مناسب است اما برای تصاویر چهره کافی نیست. توزیع، اتصال و ارتباط بین الگوها و مکان های وابسته در صورت را جدا می کند و تفاوتهای بین عناصر را کم می کند و آنها را در فضا مخلوط می کند. برای گنجاندن (اضافه کردن) اطلاعات مربوط به جایگیری LBP می تواند به عنوان یک متد پیش پردازش روی تصویر به کار رود. به واقع این جریان به عنوان یک فیلتر high-pass غیر خطی رفتار میکند. در نتیجه این موضوع باعث تشخیص واضح تر لبه های تصویر که تغییراتی محسوس پیکسلی دارند می شود. به همین ترتیب نیز باعث نمایش نویزهایی می شود که تغییرات کم و نامحسوسی در پیکسل ها بوجود می آورد. [باعث مشخص شدن نویزهایی می شود که تصویر را تغییر می دهند.] وزنهای مختلف LBP مثلاً نمایی ۲؛ در همسایگی ها دارای تغییرات زیادی است. مثلاً برای دو پیکسل در همسایگی هم، حداقل در دو حالت متفاوتند. اما در بدترین حالت، می توانند ۱۲۸ بار متفاوت باشند. از آنجا که با توجه به ۸ جهت، نویز در حالتهای تصادفی مختلفی رخ می دهد، تبدیل مقدار دودویی به مقدار دهدهی، فیلترینگ نویز LBP را حساس می کند. برای قدرتمند کردن فیلترینگ، پیشنهاد ما، ساده تر کردن مرحله ی weighting و نسبت دادن وزن های مساوی به هرکدام از ۸ همسایه است. بدین ترتیب نویز متوقف (suppress) می شود و روی کل اجزاء پخش می گردد. مشاهده کردیم که فیلترینگ simplified LBP؛ الگوهای پایدارتری را تحت نور معکوس تولید می کند. حتی در مجاورت نور خیلی زیاد. فواید فیلتر simplified LBP به شرح زیر است:

- LBP یک اندازه گیری محلی است. بنابراین، LBPها در ناحیه ی کوچک تحت تأثیر شرایط نوری نواحی دیگر قرار نمی گیرند.
- مقیاس با هیچ تبدیل یکنواخت مثل shifting و scaling پیکسل ها تغییر نمی کند.
- حساسیت LBP نسبت به نویز کم شده است.
- حتی در MPDهای با منابع محاسباتی محدود، فیلترینگ پیشنهادی بسیار پر سرعت است.

حائز اهمیت است که فیلترینگ simplified LBP ممکن است باعث فیلتر شدن بیش از حد اطلاعات تصویر چهره گردد. بنا براین هم اجزاء حساس به نور و هم برخی از اجزاء وابسته به چهره از بین می روند. این مشکل، با معرفی یک متد کلاسیفیکیشن که توانایی تشخیص بالایی دارد، حل می شود. بدین ترتیب، با توجه به این متد، اطلاعات از دست رفته توسط فیلترینگ simplified LBP ناچیز خواهد بود.

در بخش تأیید تصویر[10] با توجه به این موضوع، راجع به توانایی های کلاسیفایر در فضاهای با ابعاد بالاتر بحث می کنیم و فیلتر simplified LBP را به منظور نرمال سازی نور توجیه می کنیم. شایان ذکر است که با مسأله ی روشنایی (نور) می توان از دیدگاه سخت افزاری نیز برخورد کرد. به عنوان مثال، نور های نزدیک به infrared تصاویر یکپارچه تری در شرایط نورپردازی (نور رسانی) متفاوت در اختیار می گذارد و نیز احراز هویت چهره در



[9] Simplified Local Binary Pattern
[10] Face verification

شب را مقدور می سازد.

۵- تأیید چهره

در این بخش به حل مسأله ی تأیید چهره ی شناسایی شده، ثبت شده و نرمال سازی شده می پردازیم. این موضوع، یک مسأله ی کلاس بندی دو کلاسه است که دارای کلاس های user و nonuser است. اگرچه که مثل مسأله ی تشخیص چهره، هر دوی این مسائل دارای دو کلاس بندی هستند، مسأله ی تأیید چهره در توزیع کلاس ها، متفاوت عمل می کند. در تصویر ۸ دیده می شود که در حالت تأیید چهره، کلاس های user و nonuser به هم نزدیک ترند. به بیان دیگر، شانس اینکه تصویر یک nonuser به user شبیه باشد بیشتر از این است که یک بافت تصویری دیگر شبیه بافت چهره باشد. این موضوع بیانگر این است که متدهای کلاس بندی boundary-base که در مسأله ی تشخیص چهره به خوبی عمل می کردند (مثل متد Viola-Jones) برای مسأله ی تأیید چهره مناسب نیستند. راه حل بهتر به جای آنها، کلاس بندی در نواحی هم پوشانی شده از نقطه نظر آماری با خطای احتمالی کمتر است. پس ما استفاده از Likelihood Ratio را پیشنهاد می کنیم. Likelihood Ratio یک آماره ی بهینه ی نیومن-پیرسون است. در یک FAR داده شده Likelihood Ratio به FRR مینیمم منتهی شده، و یا در یک FRR داده شده؛ یک کلسیفایر به FAR مینیمم منتهی می شود. قانون کلاسیفیکیشن Likelihood Ratio به صورت زیر تعریف می شود:

$$L(x) = \frac{p(x|\omega)}{p(x|\bar{\omega})} > T$$

که x تصویر پیش پردازش شده است که در یک بردار قرار دارد (بصورت پشته) w کلاس user، w کلاس nonuser و T آستانه است. هنگامی که L(x)>T باشد، x به عنوان user اصلی accept می شود؛ در غیر این صورت reject می گردد. از آنجا که ما موضوع های نامتناهی بسیاری را در مجموعه های wUw متصور می شویم، کنار گذاشتن یک موضوع ساده (w) از آن عملاً توزیع x را تغییر نمی دهند. بنابراین رابطه ی زیر برقرار است:

$$p(x|\bar{\omega}) \backsimeq p(x)$$

که حتی مدل سازی را از دوکلاس به صورت ذهنی به شکل دو ابر هم پوشان در یک فضای چند بعدی آسان می سازد، همان طور که در شکل ۸ نشان داده شده است.

اکنون این دو کلاس عبارتند از: کلاس user-face و کلاس all-face (یا کلاس background). برای به دست آوردن Likelihood Ratio بردار ویژگی x ورودی، با توجه به دو کلاس w و w؛ توابع چگالی احتمال دو کلاس p(x|w) و p(x) ابتدا تخمین زده می شوند. معمولاً روی نمونه های بزرگ مجموعه داده، فرض گاوسین مورد قبول است. N نمونه بردار صورت به صورت x_i داده شده که μ و ∑ به صورت زیر به دست می آیند:

x_i ; i=1,2,...,N

$$\mu \models \frac{1}{N-1}\sum_{i=1}^{N} x_i \quad \Sigma = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)(x_i-\mu)^T. \quad (3)$$

برای اجتناب از تأثیر نمونه های خیلی دور که ممکن است در اثر نور خیلی زیاد بوجود آمده باشند، μ می تواند به صورت میانگین بردارهای ورودی در نظر گرفته شود.

$$\mu = \text{median}(x_1, \ldots, x_N).$$

کلاس های این مسأله عبارتند از کلاس user یا w_{user} و کلاس پس زمینه یا w_{bg}. فرمول زیر می تواند معادل فرمول ۱ نوشته شود:

$$\begin{aligned}\ln L(x) &= \ln p(x|\omega_{\text{user}}) - \ln p(x|\omega_{\text{bg}}) \\ &= \frac{1}{2}\left(\ln|\Sigma_{\text{bg}}| + (x-\mu_{\text{bg}})^T\Sigma_{\text{bg}}^{-1}(x-\mu_{\text{bg}})\right) \\ &\quad - \frac{1}{2}\left(\ln|\Sigma_{\text{user}}| + (x-\mu_{\text{user}})^T\Sigma_{\text{user}}^{-1}(x-\mu_{\text{user}})\right) \\ &= \frac{1}{2}\left(d_{\text{Maha}}(x|\omega_{\text{bg}}) - d_{\text{Maha}}(x|\omega_{\text{user}})\right) + c \qquad (4)\end{aligned}$$

که در آن μ_{user} ، μ_{bg} ،$\sum_{user}$ و $\sum_{bg}$ میانگین ها و کو واریانس های کلاس های user و bg هستند. جمله ی:

$C=1/2\ (\ln |\sum_{bg}| - \ln |\sum_{user}|)$

عبارت ثابتی است که می تواند به آستانه ی T در فرمول ۱ اضافه شود. همان گونه که فرمول ۴ نشان می دهد، لگاریتم؛ Likelihood Ratio را به تفاوت مقدار بین مربع فاصله های mahalanobis در کلاس های user و bg کم می کند. در اینجا به مطالعه ی توانایی های کلاسیفایر Likelihood Ratio می پردازیم: **توانایی تشخیص و توانایی تعمیم**. این دو، دو جنبه ی مهم در **تأیید** هستند. برای MPD App پیشنهادی، این دو ویژگی، به ترتیب راحتی و امنیت را توصیف می کنند. تا اینجا متوجه شدیم که تأیید چهره، در یک فضای با ابعاد بالاست. یک تصویر کوچک چهره، مثلاً سایز ۳۲*۳۲ دارای ۱۰۲۴ پیکسل است. یعنی بیانگر ۱۰۲۴ بردار ویژگی. فضای با ابعاد بالا، قدرت تشخیص قوی دارد؛ اما تعمیم آن دشوار است. این موضوع را با یک مثال ساده با استفاده از تصویر ۸(b) توضیح می دهیم. تصور کنید که هرکدام از کلاسهای user و bg یک ابر کره باشند.

and $r_{\text{bg}} = a \cdot r_{\text{user}}$, $a > 1$, in an N-dimensional space. For a single dimension, the ratio of volume between the two spaces is $V_{\text{bg}}/V_{\text{user}} = a$, which means that given an arbitrary point in the 1-D space, the chances that it belongs to the background class ω_{bg} is a times of the chance that it belongs to the user class ω_{user}. From all the N dimensions, however, the ratio becomes $V_{\text{bg}}/V_{\text{user}} = a^N$. When N is large, e.g., $N = 1000$, and a takes a moderate value, e.g., $a = 1.5$, $a^N = 1.5^{1000} \sim 10^{176}$ is almost infinite. This implies that for an arbitrary N-dimensional feature vector, the chance that it falls into the user class ω_{user} is almost none. In other words, the discrimination capability of

این موضوع بیانگر این است که برای یک بردار ویژگی N بعدی دلخواه، شانس اتفاق آن در کلاس w_{user} تقریباً صفر است. به بیان دیگر، توانایی تشخیص این نوع کلاسیفایر Likelihood Ratio در فضای با ابعاد بالا بسیار زیاد است. نظر به اینکه برای تعمیم بردار ویژگی تصویر user که در شرایط مختلفی گرفته می شود، باید توانایی ماندن در یک ناحیه ی بسیار کوچک را داشته باشد.کاهش زیاد اطلاعات تصویر با فیلتر simplified LBP باعث کوچکتر شدن هر دو کلاس پس از نرمال سازی نور می شود. کلاس bg در مقایسه با کلاس user، بطور قابل توجهی کاهش می یابد. چون این متد، اطلاعات خاصی را که برای تشخیص اشیاء مختلف مفید است، دور می ریزد. در نتیجه، حجم نسبی هر دو کلاس هم کم می شود. یا در واقع a کم شده است. وقتی a^N خیلی زیاد نباشد، تعمیم ساده تر می شود. مهم تر از آن اطلاعات دور ریز شده شامل محدوده ی بزرگی از اجزاء حساس به نور است که توانایی تعمیم را در بین روشنایی های مختلف افزایش می دهد. در ضمن قدرت کافی تشخیص به دلیل ابعاد بالای فضا محفوظ می ماند.

۶- ترکیب اطلاعات

ترکیب، یک مرحله عمومی برای افزایش اعتبار و قابلیت اطمینان تأیید بیومتریک است و در عملیات تشخیص چهره اعمال می شود. همان طور که در شکل ۲ دیده می شود، در سیستم احراز هویت پیشنهادی، ترکیب[11] بین فریم های مختلفی انجام می شود. این جریان تنها باعث بهبود کار سیستم نمی شود، بلکه احراز هویت مداوم را نیز متوجه می شود. از هر فریم، مقدار Likelihood Ratio اش را به دست می آوریم و برای تصمیم گیری آن را با آستانه مقایسه می کنیم. ما سه نوع مختلف از متدهای ترکیب زیر را مقایسه می کنیم:

- مجموع (sum) امتیازات
- AND تصمیم ها
- OR تصمیم ها

که برای قواعد min و max امتیازات، با هم معادل هستند. از جنبه ی تئوری مجموع لگاریتم Likelihood Ratio ها مشابه کلاسیفایر بیز عمل می کند و با وجود وابستگی های بین فریم های متوالی، می تواند تقریباً به اجرای بهینه نائل شود. هر چند در عمل، مشاهده کردیم که قانون OR تصمیم ها، منتهی به نتایج بهتری در سیستم تشخیص هویت پیشنهادی می شود. عملگر OR باعث می شود که کلاسیفایر تمایل بیشتری به accept داشته باشد، نه reject.

۷- تجربیات و نتایج

- جمع آوری داده

برای دانستن احتمال توابع چگالی کلاس user ($p(x|w_{user})$) و کلاس bg ($p(x|w_{bg})$) تعداد زیادی از نمونه ها مورد نیاز هستند. مجموعه نمونه های bg از دیتا بیس های عمومی چهره گرفته می شوند. از نظر تجربی، ۴ دیتابیس زیر مورد تأیید ماست:

BioID – FERET – YaleB – FRGC

تصاویر چهره با استفاده از متدهای ذکر شده در بخش های ۲ و ۳ detect و register می شوند. هر تصویر در سایز ۳۲*۳۲ register می شود و به یک بردار ویژگی ۱۰۲۴ بعدی تغییر حالت (تغییر سایز) می دهد. درکل، دیتابیس ها، بیشتر از ۱۰۰۰۰ نمونه به عنوان نمونه آموزشی در کلاس bg دارند. مجموعه نمونه های user از بین تصاویری است که توسط MPD گرفته شده است. ما از Eten M600 PDA به عنوان وسیله ی همراه استفاده کرده ایم. در عمل، جمع آوری مجموعه ی user به این صورت مناسب تر است:

در کل اطلاعات ۲۰ کاربر از بین داوطلبان جمع آوری شده است. از هر کاربر ۴ تصویر متفاوت که در زمانهای مختلف و در حالت های مختلف نوری (روشنایی) گرفته شده است. شکل ۹ مثال هایی از یک کاربر در ۴ حالت متفاوت را نشان می دهد. به علاوه، برای افزایش واریانس نمونه ها، و همچنین **tolerance to registration** مجموعه نمونه های آموزشی توسط شیفت دادن و چرخش و ویرایش مقیاس و اندازه ی تصویر گسترش پیدا کرده است.

- Test Protocol

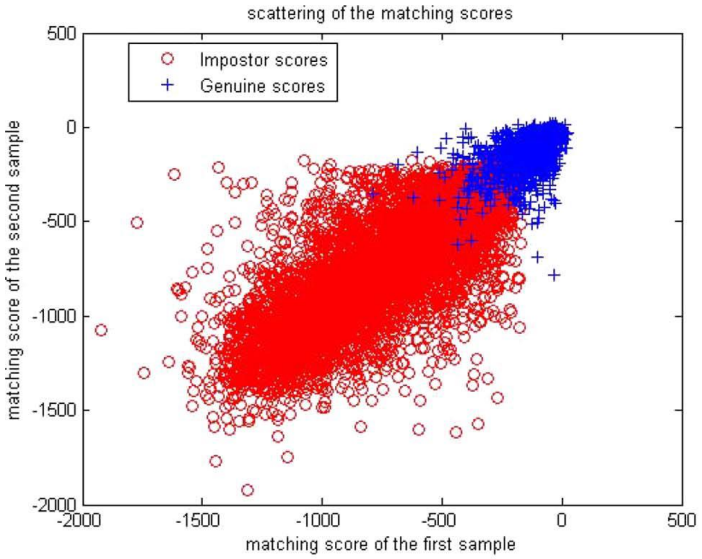
برای یک user خاص، تابع چگالی کلاس user را از یکی از آن ۴ تصویر، می دانیم و امتیازات آزمایشی واقعی را از جمله log likelihood ratio ها از ۳ تصویر دیگر محاسبه کرده ایم. تابع چگالی کلاس bg از دیتابیس عمومی مشخص است و امتیازات کاذب[12] از ۱۹ مورد جمع آوری شده ی دیگر محاسبه شده اند. به عنوان نتیجه، برای یک کاربر ۱۸۰۰ تصویر به عنوان نمونه ی آموزشی و ۵۴۰۰

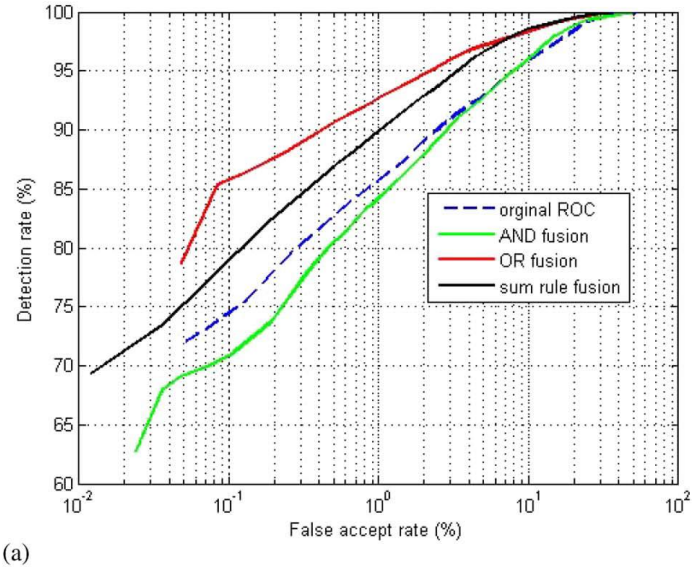
[11] fusion
[12] Impostor

تصویر برای اعتبار سنجی داریم. **On the imposter side** درحدود ۱۰۰۰۰ نمونه ی آموزشی از تصاویر عمومی چهره از دیتابیس و ۱۰۰۰۰۰ تصویر جمع آوری شده برای اعتبار سنجی داریم. توجه کنید که نمونه ی آموزشی و آزمایشی کلاس bg مستقل هستند. استفاده از دیتابیس های عمومی به عنوان نمونه های آموزشی برای پیاده سازی MPD مناسب است، چون پارامترهای bg می توانند فقط یک بار محاسبه شوند و برای تمام کاربر ها ذخیره گردند. از سوی دیگر به دست آوردن امتیازات کاذب از دیتابیس خودمان، **is of more interest** امتیازات حاصله از دیتابیس عمومی است. زیرا تصاویر چهره ی impostor از دیتابیس خودمان، تقریباً تحت شرایط یکسان جمع آوری شده اند، بنابراین برای عملیات verification معنی دار تر و حساس ترند.

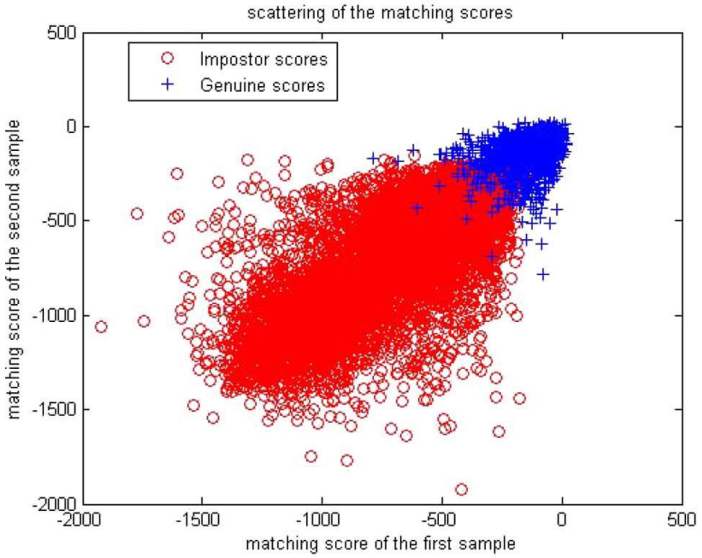
* Results on Mobile Data

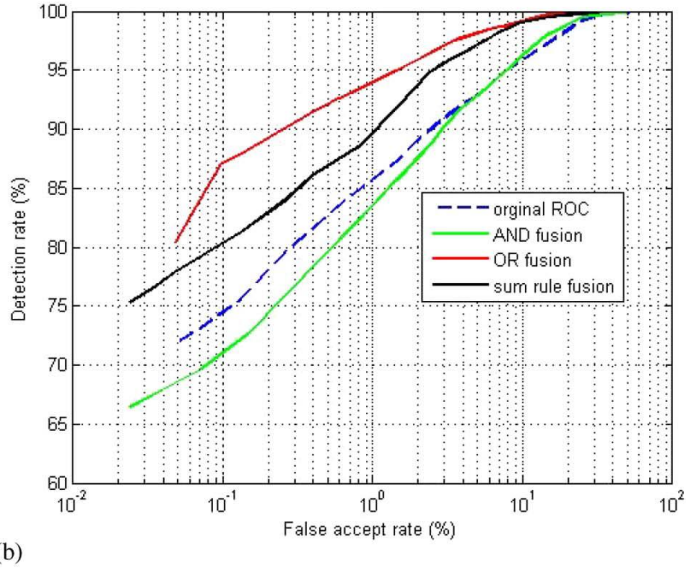
تا اینجا عملیات ترکیب دو فریم با فاصله زمانی مشخص t را نشان داده ایم. هرچقدر فاصله ی زمانی طولانی تر باشد، وابستگی کمتر است. فاصله زمانی های مختلف از ۰.۲ ، ۱ و ۳۰ ثانیه را تست کرده ایم. شکل ۱۰ نمودار scatter لگاریتم likelihood ratio های دو فریم را نمایش می دهد. مشاهده می شود که ترکیب تصمیمات با قانون AND پیشرفتی را نشان نمی دهند. در مقایسه با آن، ترکیب تصمیمات با OR به خوبی کار می کند. در مشاهدات بعدی دیده می شود که با افزایش t پیشرفت عملیات به وسیله ی ترکیب، **قطعی تر (برجسته تر – مشخص تر) می شود.**



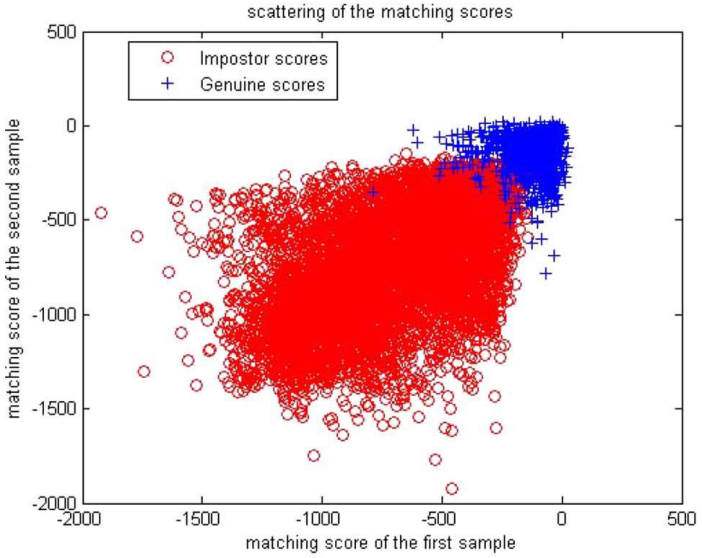


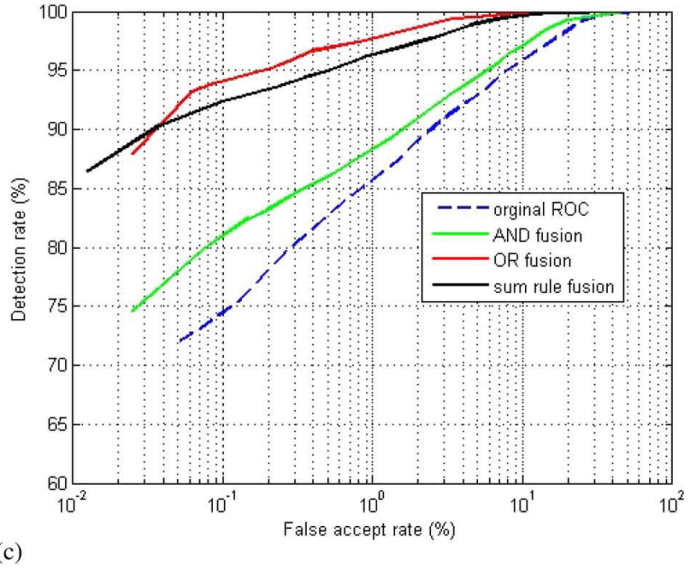
(a)





(b)



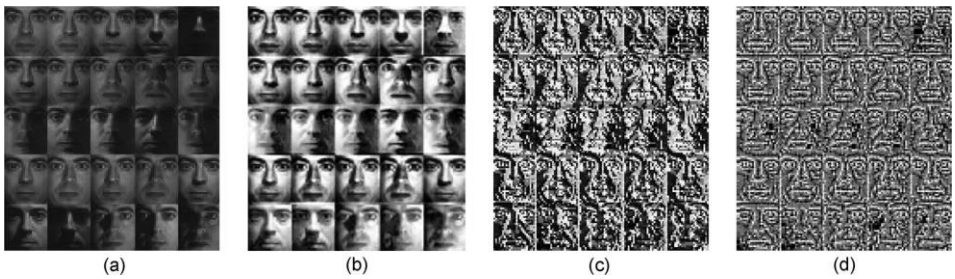


(c)

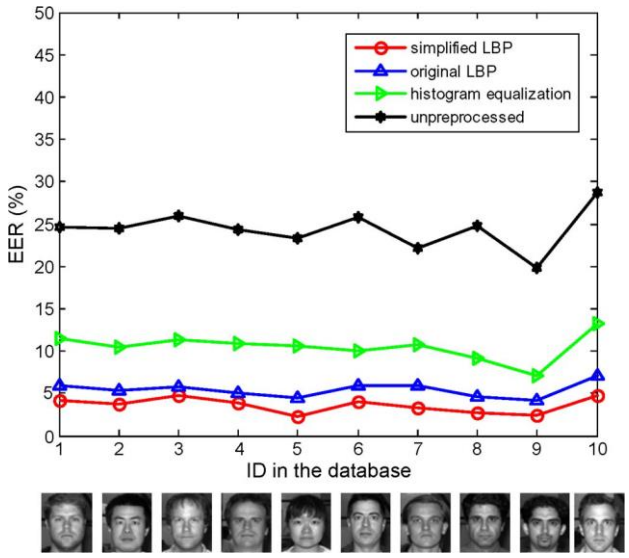
Results on YaleB Data *

الگوریتم همچنین روی دیتابیس YaleB هم آزمایش شد. این دیتابیس تصاویر ده subject را دارد که هرکدام تحت ۵۷۶ حالات مختلف دیده شده اند. (9 poses*64 illumination)

برای این دیتابیس که روی نور و روشنایی تأکید می کند، ۳ متد مختلف نرمال سازی نور را به نام های simplified LBP, original LBP, histogram equalization با هم مقایسه می کنیم. مثالهایی از این دیتابیس و تأثیرات نرمال سازی نور در شکل ۷ دیده می شوند. نتایج تصاویر پیش پردازش نشده نیز موجود هستند.



در آزمایش ما برای هر subject ۸۰ ٪ اطلاعات به صورت تصادفی برای آموزش و ۲۰٪ برای آزمایش هستند. داده های نه subject دیگر به عنوان impostor data استفاده می شوند. تأیید چهره نیز دقیقاً همانند موبایل است. برای شروع عملیات از EER[۱۳] به عنوان واحد سنجش استفاده می کنیم. عملیات متدهای مختلف روشنایی در شکل ۱۱ مقایسه شده اند.



در این شکل دیده می شود که برای تمام subject های این دیتا بیس، Simplified LBP به بهترین اجرا رسیده است. یعنی این متد قدرت بیشتری روی تغییرات وسیع نور دارد.

[13] Equal-Error Rate

* پیاده سازی

کارایی این سیستم پیشنهادی پیاده سازی واقعی آن را روی MPD امکان پذیر می سازد. ما pocket PC Eten M500 را برای اثبات انتخاب کرده ایم و الگوریتم های نوشته شده به زبان C را به platform windows mobile 5 تبدیل کرده ایم. برای تسهیل عملیات پیاده سازی، از کتابخانه ی Intel open CV استفاده کرده ایم. در تجربیات اولیه، کار روی PC است: MPD عکسهای متوالی کاربر را در مدت حدوداً ۲ دقیقه می گیرد و آنها را برای پردازش به PC انتقال می دهد. همچنین میانگین و کوواریانس user که برای محاسبه ی فاصله ی Mahalanobis در کلاس user استخراج شده اند، نیز به PC منتقل می شوند. میانگین و کوواریانس از قبل برای محاسبه ی فاصله ی Mahalanobis کلاس bg در دستگاه همراه ذخیره شده اند. سپس توالی های تصویر کاربر از دیاگرام شکل ۲ می گذرند تا تصمیم نهایی تأیید یا رد تصویر اتخاذ شود. حتی بدون بهینه سازی، سیستم ما به نرخ فریم ۱۰ فریم در ثانیه در laptop رسیده است و در موبایل ۸ فریم در ثانیه.

از این سیستم مشاهده می شودکه face detection و registration در مقایسه با illumination normalization و verification (که سریع ترند) هنوز هم قسمت های زمان بر هستند.