





دانشکده مهندسی کامپیوتر و فناوری اطلاعات
پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه داده

نگارنده: محمد مهدی بخشی

استاد راهنما

دکتر حمید حسن پور

استاد مشاور

دکتر منصور فاتح

بهمن ۱۳۹۵



باسمه تعالی

شماره:

تاریخ:

ویرایش:

مدیریت تحصیلات

تکمیلی

دانشکده :

گروه :

پایان نامه کارشناسی ارشد آقای / خانم محمد مهدی بخشی به شماره دانشجویی: ۹۳۰۳۶۹۴

تحت عنوان:

استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه داده

در تاریخ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد

مورد ارزیابی و

با درجه مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور	امضاء	اساتید راهنما
	نام و نام خانوادگی : دکتر منصور فاتح		نام و نام خانوادگی : دکتر حمید حسن پور

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور
	نام و نام خانوادگی :		نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :

پاس‌گزاری

خدای رابی ساکرم که از روی کرم، پرومادی فداکار نصیم ساخته تا در سایه درخت پر بار وجودشان بیایم
و از ریشه آن ها شاخ و برگ گیرم و از سایه وجودشان در راه کسب علم و دانش تلاش نمایم. والدینی که
بودشان تلج افتخاری است بر سرم و نشان دلیلی است بر بودنم، چرا که این دو وجود، پس از پروردگار،
مایه هستی ام بوده اند، دستم را گرفتند و راه رفتن را در این وادی زندگی پر فراز و نشیب آموختند. آموزگارانی که
برایم زندگی کردن و انسان بودن را معنا کردند.

تقدیم به وجود بارزشتان...

شکر و قدردانی

از استاد گرامیم جناب آقای دکتر حمید حسن پور بسیار سپاسگزارم چرا که بدون راهنماییهای ایشان تأمین این

پایان نامه بسیار مشکل می نمود.

و

از جناب آقای دکتر منصور فتح سپاسگزارم به دلیل یاری ها و راهنمایی هایشان که بسیاری از سختی ها را برایم آسان تر

نمودند.

تعهد نامه

اینجانب محمد مهدی بخشی دانشجوی دوره کارشناسی ارشد رشته هوش مصنوعی دانشکده کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شاهرود، نویسنده پایان نامه استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه داده، تحت راهنمایی آقای دکتر حمید حسن پور متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های محققان دیگر به مرجع مورداستفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود است و مقالات مستخرج با نام «دانشگاه صنعتی شاهرود» یا «Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آن ها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است. تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود است. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

استخراج ویژگی از تصاویر چهره، طی دو دهه‌ی اخیر در کاربردهای مختلفی از قبیل آشکارسازی چهره، شناسایی چهره، تشخیص حالات چهره و بازسازی چهره مورد استفاده قرار گرفته است. در میان این کاربردها، شناسایی چهره از اهمیت ویژه‌ای برخوردار است. شناسایی چهره با مشکلاتی نظیر کیفیت نامناسب تصویر ورودی، وجود پوشش‌های غیر پیش‌بینی شده روی چهره، حجم زیاد تصاویر پایگاه داده و وجود تنها یک تصویر به ازای هر شخص مواجه بوده است.

در این پایان‌نامه، روشی نوین جهت استخراج ویژگی از تصاویر چهره ارائه شده است. هدف اصلی این پایان‌نامه، بازیابی چهره‌ی ورودی از پایگاه داده حجیم است. با افزایش حجم پایگاه داده، شباهت بین افراد افزایش یافته و قابلیت تفکیک بین تصاویر چهره با مشکل مواجه می‌شود. روش پیشنهادی با استخراج ویژگی‌های مناسب، فاصله بین نمونه‌ها در فضای ویژگی را افزایش می‌دهد و لذا قابلیت تفکیک بین افراد را افزایش می‌دهد. این روش مبتنی بر خاصیت بصری انسان بوده و به صورت ترتیبی، از کل به جزء در راستای مکان از ناحیه چهره ویژگی استخراج می‌کند. برای این منظور، از ویژگی‌های مکان-فرکانس استفاده شده است. در این روش، با اعمال پنجره‌های هم مرکز با ابعاد مختلف روی تصویر چهره، محتوای هر پنجره به فضای فرکانسی منتقل می‌شود. تغییر مؤلفه‌های فرکانسی در پنجره‌های مختلف، فضای ویژگی تصویر را تشکیل می‌دهند. سپس، با استفاده از فیلتر مناسب، تنها مؤلفه‌های فرکانسی با قابلیت جداسازی بالا بین تصاویر چهره، حفظ می‌شوند. در نهایت، با استفاده از معیار فاصله‌ی اقلیدسی، تصویر نهایی از پایگاه داده بازیابی می‌شود.

در اکثر روش‌های موجود شناسایی چهره با لحاظ کردن حجم کم پایگاه داده، نتایج مطلوب حاصل گردیده یا در صورت وجود حجم زیاد پایگاه داده، تعداد تصاویر موجود به ازای هر شخص افزایش پیدا می‌کند. در این پژوهش، به ازای هر شخص تنها یک تصویر در پایگاه داده وجود دارد و از تعداد کلاس‌های به مراتب بیشتر از سایر روش‌ها در پایگاه داده استفاده شده است. در این پایان‌نامه از پایگاه داده FERET استفاده شده است. نرخ شناسایی در مقایسه با بهترین روش قبلی در حجم مشابه با ۲٪ افزایش، به ۹۹٪ ارتقا پیدا کرده است. با افزایش حجم پایگاه داده به ۹۹۰ کلاس، نرخ شناسایی ۹۰/۴٪ حاصل شده است.

کلید واژه‌ها: شناسایی چهره، استخراج ویژگی، حوزه مکان-فرکانس، پایگاه داده حجیم، انتخاب

ویژگی.

فهرست مطالب

فصل اول: مقدمه‌ای بر شناسایی چهره	۱
۱- شناسایی چهره	۲
۲- اهمیت موضوع	۲
۳- تکنیک‌های شناسایی چهره	۳
۳-۱- روش‌های مبتنی بر ویژگی	۴
۳-۲- روش‌های کل‌نگر	۵
۳-۲-۱- روش‌های آماری	۶
۳-۲-۲- روش‌های هوشمند	۷
۳-۳- روش‌های ترکیبی	۷
۴- چالش‌های شناسایی چهره	۸
۵- تعریف مسئله	۹
۶- نوآوری مسئله	۱۰
۷- هدف پایان‌نامه	۱۱
۸- ساختار پایان‌نامه	۱۳
فصل دوم: مروری بر کارهای پیشین	۱۵
۱- مقدمه	۱۶
۲- چهره‌های ویژه	۱۸
۳- تبدیل موجک گابور	۲۱
۳-۱- پیش‌پردازش	۲۲
۳-۲- استخراج ویژگی مبتنی بر فیلتر گابور	۲۵
۳-۳- نتایج	۲۸
۴- الگوریتم SIFT	۲۹

۲۹	۱-۴- استخراج نقاط کلیدی
۳۰	۲-۴- توصیفگر نقاط کلیدی
۳۲	۳-۴- نحوی شباهت سنجی
۳۳	۴-۴- نتایج
۳۴	۵- الگوی باینری محلی
۳۴	۱-۵- استخراج ویژگی
۳۵	۲-۵- معیار فاصله
۳۶	۳-۵- نتایج
۳۷	۶- تبدیل هاف
۴۰	۷- جمع‌بندی
۴۳	فصل سوم: استفاده از ویژگی‌های مکان-فرکانسی جهت شناسایی چهره (روش پیشنهادی)
۴۴	۱- مقدمه
۴۶	۲- آشکارسازی چهره
۴۷	۱-۲- تصویر انتگرال
۴۸	۲-۲- طبقه‌بند آبشاری مبتنی بر آدابوست نامتقارن
۴۹	۳- استخراج ویژگی
۵۳	۴- کاهش ویژگی‌ها
۵۷	۵- معیار فاصله
۶۰	۶- جمع‌بندی
۶۱	فصل چهارم: ارزیابی نتایج
۶۲	۱- مقدمه
۶۲	۲- پایگاه داده
۶۳	۳- آزمایش‌ها

۶۳	۳-۱- نتایج آشکارسازی چهره
۶۳	۳-۲- نتایج کمی شناسایی چهره
۶۶	۳-۳- نتایج کیفی شناسایی چهره
۶۹	۳-۴- نتایج انتخاب ویژگی
۷۰	۴- جمع بندی
۷۳	فصل پنجم: نتیجه گیری و پیشنهادها
۷۴	۱- نتیجه گیری
۷۵	۲- پیشنهادها
۷۶	منابع

فهرست اشکال

- شکل (۱-۱): روال کلی و هدف پایان نامه..... ۱۲
- شکل (۱-۲): یک مجموعه آموزشی از تصاویر چهره و میانگین آن‌ها در روش Eigenface..... ۱۹
- شکل (۲-۲): ۷ چهره ویژه محاسبه شده از تصاویر شکل (۱-۲) بدون حذف پس زمینه..... ۲۰
- شکل (۳-۲): نامگذاری پیکسل‌های همسایه جهت پیشرفت مراحل الگوریتم AD در پیش‌پردازش..... ۲۵
- شکل (۴-۲): فیلترهای گابور مطابق با ۵ مقیاس و ۸ زاویه متفاوت ۲۷
- شکل (۵-۲): نمونه‌ای از اعمال فیلترهای گابور روی تصویر ورودی از پایگاه داده Stirling..... ۲۷
- شکل (۶-۲): استخراج نقاط ویژگی توسط موجک‌های گابور..... ۲۸
- شکل (۷-۲): مقایسه مقادیر پیکسلی در هرم گوسین جهت تشخیص نقاط کلیدی..... ۳۰
- شکل (۸-۲): کاهش زاویه‌های مجاز برای نقاط کلیدی با کلاس‌بندی آنها..... ۳۱
- شکل (۹-۲): ویژگی‌های SIFT روی تصویر چهره..... ۳۲
- شکل (۱۰-۲): ساخت نواحی تصویر و محاسبه شباهت تصاویر چهره در الگوریتم SIFT..... ۳۳
- شکل (۱۱-۲): عملگر LBP..... ۳۴
- شکل (۱۲-۲): همسایگی دایره‌ای در الگوی باینری محلی..... ۳۵
- شکل (۱۳-۲): وزن‌دهی به نواحی چهره بر اساس معادله کای دو وزن دار..... ۳۶
- شکل (۱۴-۲): تغییر پارامترها در شناسایی خطوط مبتنی بر تبدیل هاف..... ۳۸
- شکل (۱۵-۲): تبدیل یک نقطه در حوزه مکان به بی نهایت خط در حوزه پارامتری تبدیل هاف..... ۳۹
- شکل (۱۶-۲): فضای پارامتریک قطبی حاصل از یک نقطه در فضای مکان، حاصل از شکل (۱۵-۲)..... ۳۹
- شکل (۱۷-۲): انتقال تصویر لبه یابی شده به فضای قطبی توسط تبدیل هاف ۳۹

- شکل (۳-۱): تصویر فرآیند کلی الگوریتم ارائه شده..... ۴۶
- شکل (۳-۲): ویژگی‌های مستطیلی نمونه جهت استخراج چهره..... ۴۷
- شکل (۳-۳): محاسبه مجموع مقادیر پیکسلی یک ویژگی مستطیلی، با استفاده از تصویر انتگرال..... ۴۸
- شکل (۳-۴): برخی از ویژگی‌های منتخب توسط طبقه‌بند آدابوست نامتقارن..... ۴۹
- شکل (۳-۵): فرآیند استخراج ویژگی از هر تصویر چهره..... ۵۱
- شکل (۳-۶): مقادیر بهینه دو پارامتر M و L در روش ارائه شده..... ۵۳
- شکل (۳-۷): نمایش فضای ویژگی تصاویر چهره نمونه به صورت برداری در روش پیشنهادی..... ۵۴
- شکل (۳-۸): مولفه‌های فرکانسی در دو پنجره مشخص از سه تصویر نمونه..... ۵۶
- شکل (۳-۹): ماسک طراحی شده جهت فیلتر فضای فرکانسی تصاویر چهره در روش پیشنهادی..... ۵۷
- شکل (۳-۱۰): مقادیر فاصله تصویر آزمون از ۲۰۰ تصویر پایگاه داده..... ۵۹
- شکل (۴-۱): عملکرد سیستم آشکارساز چهره روی چند چهره نمونه..... ۶۳
- شکل (۴-۲): شتاب کاهش نرخ بازیابی با افزایش حجم پایگاه داده..... ۶۶
- شکل (۴-۳): نمونه‌ای از تصاویر پایگاه داده FERET..... ۶۷
- شکل (۴-۴): چند نمونه از تشخیص صحیح الگوریتم پیشنهادی..... ۶۷
- شکل (۴-۵): نمونه‌ای از شناسایی اشتباه در روش پیشنهادی..... ۶۸
- شکل (۴-۶): چند نمونه فیلتر حوزه فرکانس در روش پیشنهادی..... ۶۹
- شکل (۴-۷): کارایی فیلترهای شکل (۴-۶) در روش پیشنهادی..... ۷۰

فهرست جداول

- جدول (۱-۲): نتایج حاصل از اعمال الگوریتم موجک گابور روی پایگاه داده‌های متفاوت ۲۸
- جدول (۲-۲): نتایج حاصل از اعمال الگوریتم SIFT روی پایگاه داده‌های متفاوت ۳۳
- جدول (۳-۲): نتایج حاصل از اعمال الگوریتم LBP روی پایگاه داده FERET ۳۶
- جدول (۴-۲): نتایج حاصل از الگوریتم Hough در شناسایی چهره ۴۰
- جدول (۵-۲): مقایسه روش‌های بررسی شده با یکدیگر ۴۱
- جدول (۱-۳): میزان بهبود نتایج الگوریتم پیشنهادی در صورت استفاده از فیلتر ارائه شده ۵۷
- جدول (۲-۳): مقایسه معیار فاصله اقلیدسی با دو معیار دیگر در روش پیشنهادی ۵۹
- جدول (۱-۴): مقایسه کارایی روش پیشنهادی با الگوریتم‌های مختلف ۶۴
- جدول (۲-۴): نرخ و سرعت بازیابی الگوریتم ارائه شده در حجم‌های متفاوت پایگاه داده ۶۵

فصل اول

مقدمه‌ای بر شناسایی چهره

۱- شناسایی چهره

فن شناسایی چهره (FRT)^۱ در مسئله تشخیص و تایید هویت، به عنوان یک راه حل در راستای دستیابی به بسیاری از نیازهای روز مورد توجه بوده است. این تکنیک، ارتباطی بین انتظار سیستم‌های بیومتریک هوشمند (که جهت تشخیص هویت با خواص انفرادی هر شخص ارائه شده‌اند)، و قابلیت سیستم‌های تجسس بصری انسان برقرار می‌نماید[1].

شناسایی چهره در دهه‌ی اخیر، بخش عظیمی از تحقیقات حوزه شناسایی الگو و بینایی ماشین را به خود اختصاص داده است. انسان‌ها با دقت بالایی شناسایی چهره را انجام می‌دهند در حالی که شناسایی هوشمند چهره توسط ماشین کار دشواری است. روش‌های زیادی برای ارتقای دقت شناسایی هوشمند چهره ارائه شده است. هم اکنون این تکنولوژی به شکل قابل توجهی پیشرفت کرده است[2]. روش‌ها و الگوریتم‌های مختلف در این زمینه با یکدیگر دست و پنجه نرم می‌کنند، ولی تمامی آن‌ها نگرانی‌های مشترکی داشته و دارند. از جمله نگرانی‌هایی که در الگوریتم‌های FRT همواره وجود دارد، دستیابی به دقت بالا، سرعت مناسب و مقابله با حجم زیاد داده‌ها به صورت همزمان، متناسب با زمینه‌های کاری متفاوت و اهداف مشخص است. شرایط ثبت تصویر در پایگاه داده، یکی از مؤثرترین پارامترها در شناسایی چهره است. از جمله این شرایط می‌توان به تعداد تصاویری که باید به ازای هر شخص در پایگاه داده قرار بگیرد، اشاره نمود. همینطور ابعاد، کیفیت و رنگ‌بندی تصاویر ذخیره شده حائز اهمیت است.

۲- اهمیت موضوع

تکنیک‌های بیومتریک به عنوان ابزاری برای تشخیص افراد به صورت هوشمند در سال‌های اخیر، ارائه شده است. لذا به جای تایید و به رسمیت شناختن انسان‌ها و تضمین دسترسی آن‌ها به حوزه‌های مجازی و فیزیکی با استفاده از رمز عبور، کارت‌های هوشمند، نشانه‌ها و کلید، این روش‌ها فیزیولوژی یا مشخصه-

1- Facial recognition technology

های رفتاری افراد را جهت تشخیص هویت آن فرد، مورد بررسی قرار می‌دهند. تکنولوژی‌های مبتنی بر بیومتریک شامل احراز هویت مبتنی بر مشخصات فیزیولوژیکی مثل چهره، اثر انگشت، هندسه انگشت یا دست، رگه‌های دست، کف دست، عنبیه، شبکیه و صدا و همینطور خصیصه‌های رفتاری مثل: طرز راه رفتن، امضا، دینامیک ضربه زدن به کلید، هستند. شناسایی چهره، جهت چیره شدن بر معایب سایر روش‌های بیومتریک دیگر، ارائه شده است. شناسایی چهره می‌تواند بی‌اراده و بدون هیچ عمل یا مشارکتی با کاربر انجام شود. این خود می‌تواند سهولتی جهت بکار بستن این تکنیک در اهداف امنیتی باشد. علاوه بر آن، جمع‌آوری داده برای سایر روش‌های بیومتریک کاری پردردسر است. الگوریتم‌های شناسایی چهره مناسب، به همراه پیش‌پردازش‌های لازم روی تصاویر، می‌تواند نویز و تغییرات در چرخش، مقیاس و روشنایی را تا حد قابل قبولی جبران نماید. لذا با انجام این دو عملیات (پیش‌پردازش مناسب و الگوریتم-های شناسایی مناسب)، می‌توان از این روش به عنوان روشی قدرتمند جهت احراز هویت افراد استفاده نمود[3].

از طرفی سیستم‌های موجود شناسایی چهره، در حجم زیاد تصاویر، کارایی پایین دارند. این موضوع باعث ضعف کاربری‌هایی نظیر کاربردهای امنیتی شده است. لذا با ارائه روشی که بتواند با افزایش حجم پایگاه داده تصاویر، کارایی مناسبی داشته باشد، کاربری‌ها را می‌توان افزایش داد.

۳- تکنیک‌های شناسایی چهره

شناسایی چهره یک مسئله‌ی چالش‌برانگیز در زمینه تحلیل تصویر و بینایی ماشین است و به دلیل کاربردی بودن در حوزه‌های مختلف، توجه زیادی را طی سال‌های اخیر به خود جلب کرده است. لذا تنوع کاربرد در این تکنیک، موجب ایجاد روش‌های متفاوت در به دست آوردن تصاویر چهره شده است. برای مثال، در کاربردهای محافظتی ممکن است بهتر باشد تصاویر به وسیله دوربین‌های ویدئویی گرفته شود، در حالی که در کاربرد دیگر مثل تجسس در پایگاه داده تصاویر، ممکن است نیاز به تصاویر با اندازه ثابت

باشد که توسط دوربین‌های معمولی گرفته شده است. با توجه به روش‌های دریافت داده‌های چهره، تکنیک‌های شناسایی چهره می‌تواند به شکل کلی به سه دسته تقسیم شود: (۱) روش‌هایی که روی تصاویر ساده کار می‌کنند، (۲) آنهایی که با تصاویر ویدئویی کار می‌کنند و (۳) آنهایی که نیازمند داده‌های حسی دیگر نظیر اطلاعات سه بعدی هستند [3]. در این پایان‌نامه دسته اول، یعنی روش‌هایی که با تصاویر ساده سر و کار دارند، جهت رسیدن به هدف، مورد بررسی قرار می‌گیرند.

روش‌های شناسایی چهره برای تصاویر ساده (با در اینجا تصاویر خاکستری دو بعدی) به سه دسته کلی تقسیم می‌شوند [4]: روش‌های مبتنی بر ویژگی، روش‌های کل‌نگر و روش‌های ترکیبی. یک بررسی کلی از برخی تکنیک‌های شناخته شده از این سه دسته در ادامه آمده است.

۳-۱- روش‌های مبتنی بر ویژگی

روش‌های مبتنی بر ویژگی^۱ در ابتدا تصویر ورودی را جهت تعیین نقاط حائز اهمیت تصویر و استخراج و ارزیابی ویژگی‌های تفکیک‌ساز مناسب چهره از این نقاط، مورد پردازش قرار می‌دهند. سپس روابط هندسی بین این ویژگی‌های محلی را محاسبه می‌کنند، به این ترتیب تصویر چهره ورودی را به برداری از ویژگی‌های هندسی کاهش می‌دهند. سپس تکنیک‌های شناسایی الگوی آماری استاندارد جهت مطابقت چهره‌ها، به کار گرفته می‌شوند. از جمله این روش‌ها می‌توان به روش گراف الاستیکی^۲ اشاره نمود [5].

اکثر تکنیک‌های استخراج ویژگی، از روش‌هایی از قبیل تبدیل هاف و فیلترهای گرافی و عملگرهای ریخت شناسی استفاده می‌کنند. درحالی که تمام این تکنیک‌ها به شدت متکی بر الگوریتم‌های اکتشافی هستند؛ به عنوان مثال، محدود کردن زیرفضای جستجو با محدودیت‌های هندسی که باعث از دست دادن اطلاعات مفید موجود در تصویر می‌شود. از این گذشته یک خطای اصلی باید برای این مدل‌ها در نظر گرفته شود، چرا که آن‌ها هرگز نمی‌توانند به شکل عالی و کامل ساختارها را به تصویر نگاشت کنند.

1- Feature - based methods

2- Elastic Bunch Graph

هرچند استفاده از یک مقدار خطای بزرگ منجر به از دست دادن دقت مورد نیاز برای تشخیص صحیح می‌شوند.

مزیت اصلی تکنیک‌های مبتنی بر ویژگی این است که چون استخراج ویژگی از نقاط، مقدم بر انجام تحلیل جهت مقایسه تصویر با شخص شناخته شده می‌باشد، لذا چنین روش‌هایی نسبت به تغییرات موضع در تصویر ورودی، نسبتاً مقاوم هستند. در اصل، ویژگی‌های این رویه‌ها می‌توانند در اندازه، زوایا و روشنایی‌های متفاوت ساخته شوند. مزیت دیگر این رویه‌ها، تراکم نمایش تصاویر چهره و در نتیجه سرعت تطبیق بالا می‌باشد.

ایراد اصلی این روش‌ها سختی در تشخیص خودکار ویژگی و این حقیقت که مجری هر کدام از این تکنیک‌ها مجبور به گرفتن تصمیمات اختیاری در مورد اینکه چه ویژگی‌هایی مهم‌تراند، می‌باشد. بعد از همه این‌ها، اگر مجموعه ویژگی، فاقد توانایی جداسازی مناسب بود، هیچ پردازش ثانویه‌ای نمی‌تواند این نقص درونی را جبران کند. همچنین به دلیل انتخاب محدوده‌ی خاص از چهره به عنوان نقاط حائز اهمیت، در صورت افزایش حجم پایگاه داده و در نتیجه افزایش شباهت ویژگی‌های بین چهره‌ای، کارایی این روش‌ها در حجم زیاد دچار مشکل می‌شود.

۳-۲- روش‌های کل‌نگر

روش‌های کل‌نگر^۱، شناسایی چهره را با استفاده از نمایش‌های عمومی آن انجام می‌دهند. به عنوان مثال توصیف مبتنی بر تصویر اصلی و دست نخورده به جای ویژگی‌های محلی چهره. این رویه‌ها می‌توانند به دو زیر گروه تقسیم شوند: روش‌های آماری و روش‌های هوشمند. توضیح اجمالی از روش‌های این گروه‌ها در زیر آمده است.

1- Holistic methods

۳-۲-۱ - روش‌های آماری

در ساده‌ترین شکل از روش‌های کل‌نگر، تصویر به عنوان یک آرایه‌ی دو بعدی از مقادیر شدتی ارائه می‌شود. سپس فرایند شناسایی با مقایسه همبستگی مستقیم بین تصویر ورودی و تمام تصاویر دیگر در پایگاه داده انجام می‌شود. هرچند نشان داده شده که این روش‌ها تحت شرایط محدود (به عنوان مثال روشنایی، مقیاس و موقعیت یکسان) کار می‌کنند، اما از نظر محاسباتی بسیار گران بوده و از کاستی‌های معمولی روش‌های مبتنی بر همبستگی ساده، نظیر حساسیت به زاویه، اندازه، شرایط روشنایی متفاوت، ناهنجاری پس زمینه و نویز چهره رنج می‌برند. مانع اصلی کارایی روش‌های تطبیق مستقیم این است که آن‌ها در تلاش برای انجام طبقه‌بندی در یک فضای با ابعاد بسیار بالا می‌باشند. جهت مقابله با این مشکل بزرگ (ابعاد بسیار بالا)، رویه‌های متعدد دیگری ارائه گردیده‌اند که روش‌های کاهش ابعاد آماری را جهت محاسبه و همینطور حفظ اکثر ابعاد ویژگی پرمعنا قبل از انجام شناسایی، به کار می‌بندند. برخی از آنها در زیر ذکر گردیده‌اند.

سیروویچ و کیربای^۱، اولین بار تحلیل مولفه اصلی (PCA) را جهت نمایش مقرون بصرفه تصاویر چهره، بکار بردند [6,7]. آنها اثبات کردند که هر چهره منحصر بفرد می‌تواند به شکلی کاربردی روی یک فضای مختصات از تصاویر ویژه، نمایش داده شود و اینکه هر چهره می‌تواند با استفاده از تنها یک تعداد اندکی از تصاویر ویژه و ضرایب شباهت هر تصویر ویژه، بازسازی شود.

تورک و پنتلند^۲، مبتنی بر یافته‌های سیروویچ و کیربای، دریافتند که فرض مطابق با تصاویر ویژه، می‌تواند به عنوان ویژگی‌های کلاس‌بندی جهت شناسایی چهره‌ها، مورد استفاده قرار گیرد [8]. آن‌ها از این استدلال جهت بهبود سیستم شناسایی چهره استفاده کردند که منجر به ساخت چهره‌های ویژه شد، که برابر با بردارهای ویژه مرتبط با مقادیر ویژه برتر ماتریس کواریانس چهره‌های موجود می‌باشد. سپس

1- Sirovich and Kirby

2- Turk and Pentland

چهره‌های خاص (ورودی)، با نگاشت روی چهره‌های ویژه، از بین تصاویر چهره اشخاص شناخته شده، شناسایی می‌شود. تصاویر ویژه یک فضای ویژگی را تعریف می‌کنند که شدیداً ابعاد فضای اصلی را کاهش داده و شناسایی چهره، در این فضای کاهش یافته انجام می‌شود. از دیگر روش‌هایی که در راستای کاهش ابعاد ویژگی ارائه شدند می‌توان به روش‌های تجزیه و تحلیل المان‌های مستقل (ICA) [9] و تجزیه و تحلیل جداساز خطی (LDA) [10] اشاره نمود.

۳-۲-۲- روش‌های هوشمند

روش‌های هوشمند از ابزاری نظیر شبکه‌های عصبی و تکنیک‌های یادگیری ماشین جهت شناسایی چهره‌ها استفاده می‌کنند. مزیت اصلی روش‌های کل‌نگر این است که سعی در حفظ تمام اطلاعات چهره دارند؛ برخلاف روش‌های مبتنی بر ویژگی که متمرکز بر روی نواحی و نقاط جالب توجه محدوده‌ای از چهره می‌شوند. هرچند، همین خاصیت، بزرگترین نقطه ضعف آنها نیز می‌باشد، چرا که اکثر این روش‌ها با این فرض اساسی کار خود را شروع می‌کنند که تمام پیکسل‌های موجود در تصویر مهم هستند. در نتیجه این تکنیک‌ها نه تنها از نظر محاسباتی هزینه بر هستند بلکه نیازمند یک درجه بالا از همبستگی بین تصاویر آموزشی و آزمون می‌باشند و همچنین تحت تغییرات زیاد در موقعیت، مقیاس و روشنایی، موثر واقع نمی‌شوند.

۳-۳- روش‌های ترکیبی

هر یک از دو روش کل‌نگر و مبتنی بر ویژگی، دارای معایبی هستند که اگر بتوان آنها را تا حد امکان کاهش داد، و از مزایای هر کدام استفاده کرد، طبعاً میزان بهبود الگوریتم‌ها افزایش پیدا می‌کند. روش‌های ترکیبی^۱ با استفاده از پیمایش کل تصویر و استخراج ویژگی همزمان از آن (نه لحاظ قرار دادن کل چهره)، علاوه بر استفاده از تمام اطلاعات مفید چهره، از فضای ویژگی کاهش یافته چهره به جای تمام ناحیه

1- Hybrid Methods

چهره استفاده می‌کنند. لذا علاوه بر نرخ شناسایی بالا در حجم زیاد، سرعت بازیابی مطلوبی حاصل می‌گردد. از جمله این روش‌ها می‌توان به پژوهش Pentland و همکاران در [11] اشاره کرد که در آن هم چهره‌های ویژه^۱ عمومی و هم ویژگی‌های ویژه^۲ محلی را مورد استفاده قرار داده است.

۴- چالش‌های شناسایی چهره

شناسایی چهره یک نمونه پیچیده از شناسایی شیء می‌باشد. سختی این مسئله ناشی از این حقیقت است که در رایج‌ترین شکل خود (که تصویر از روبرو می‌باشد) چهره‌ها بسیار شبیه به هم ظاهر می‌شوند و تفاوت‌های بین آنها بسیار ظریف می‌باشد. در نتیجه تصاویر چهره از روبرو، یک تراکم شدید در فضای تصویر شکل می‌دهد که آن را برای تکنیک‌های سنتی شناسایی الگو جهت تمایز دقیق بین آنها با درجه موفقیت بالا، دشوار می‌سازد. بعلاوه چهره انسان یک شیء ثابت و واحد نیست. در واقع فاکتورهای متنوعی که باعث نمایش‌های مختلف چهره می‌شود، وجود دارد. عوامل اختلاف در ظاهر چهره می‌تواند به دو گروه تقسیم شود: فاکتورهای ذاتی (درونی) و فاکتورهای تصادفی (برونی). الف) فاکتورهای درونی صرفاً ناشی از ماهیت فیزیکی چهره می‌باشد و مجزا از ناظر می‌باشد. این فاکتورها مجدداً می‌توانند به دو دسته تقسیم شوند: درون فردی و میان فردی. فاکتورهای درون فردی مبتنی ایجاد تفاوت در ظاهر چهره یک شخص می‌باشد، مثل سن، حالت چهره و متعلقات چهره (شامل موی صورت، عینک، آرایش و ...). در حالی که فاکتورهای بین فردی، مبتنی اختلاف بین ظاهر چهره افراد مختلف می‌باشد. ب) فاکتورهای برونی باعث تغییرات ظاهر چهره در تعامل نور بین چهره و ناظر می‌باشند. این فاکتورها شامل روشنایی، ژست، مقیاس و پارامترهای تصویر برداری (مثل کیفیت، نقطه تمرکز، نویز و ...) می‌باشند.

1 - Eigenfaces

2 - Eigenfeatures

هرچند اکثر سیستم‌های موجود شناسایی چهره، تحت شرایط محدود، به خوبی کار می‌کنند (مثل سناریوهایی که در آنها حداقل تعداد کمی از فاکتورهای درونی و بیرونی مذکور، در تغییر پذیری بین تصاویر چهره، ایفای نقش می‌کنند)، اما تحت شرایطی که هیچ کدام از این فاکتورها تنظیم نشده باشد، کارایی اکثر این سیستم‌ها به شدت کاهش پیدا می‌کند.

علاوه بر مشکلات ذکر شده، عواملی همچون حجم پایگاه داده، وجود تنها یک تصویر به ازای هر شخص در پایگاه داده، ابعاد تصاویر ذخیره شده، کیفیت تصاویر ورودی، تغییر حالات چهره وجود دارد که فرایند شناسایی چهره را با مشکلاتی دوچندان روبرو می‌سازد. با ارائه روشی که بتواند این مشکلات را تا حد قابل توجهی رفع نماید، می‌توان گستره‌ی کاربرد این تکنیک را افزایش داده و سهولت را برای استفاده از این تکنیک بیومتریک به همراه آورد. جهت رفع هر یک از مشکلات، در الگوریتم‌های مختلف روش‌هایی پیشنهاد گردیده است [12].

۵- تعریف مسئله

شناسایی چهره برای دو هدف اصلی بکار می‌رود [3]:

- ۱- تصدیق (مقایسه یکی با یکی): زمانی است که یک تصویر از یک شخص ناشناس با تنها یک ادعای هویت وجود دارد. یا به عبارتی تعیین اینکه آیا شخص همانی می‌باشد که ادعا می‌کند یا خیر.
- ۲- تشخیص هویت (مقایسه یکی با همه): یک تصویر از یک شخص ناشناس داده شده، تعیین هویت شخص با مقایسه (احتمالا کد شده) این تصویر با پایگاه داده‌ای (احتمالا کد شده) از تصاویر اشخاص شناخته شده، مد نظر است.

در این پایان‌نامه هدف اصلی متمرکز بر مورد دوم می‌باشد. مسئله استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه داده، می‌تواند به این شکل عنوان و فرموله بندی شود: تصویری ورودی داده

شده و پایگاه داده‌ای حجیم از تصاویر چهره افراد موجود می‌باشد. تصمیم‌گیری ما در مورد هویت شخص موجود در تصویر ورودی چگونه خواهد بود؟

با استفاده از تکنیک‌های متعدد ارائه شده در روش‌های مختلف شناسایی چهره، و با توجه به خواص تفکیک‌ساز مناسب و ویژگی‌های با ظرفیت بالای^۱ موجود در چهره، در راستای پاسخگویی به سوال فوق، روشی را ارائه کرده‌ایم که با در نظر گرفتن حجم زیاد تصاویر موجود در پایگاه داده، و وجود تنها یک تصویر به ازای هر شخص در آن، بتواند تصویر مورد پرس‌وجو را با عملکردی مناسب، از پایگاه داده بازیابی نماید.

در این پایان‌نامه، فرض اصلی وجود حجم زیاد تصاویر در پایگاه داده می‌باشد که در روش‌های پیشین به این نکته کمتر توجه شده است. با افزایش حجم پایگاه داده، تشابه ویژگی‌ها بین افراد زیاد می‌شود و قابلیت جداسازی آنها در فضای ویژگی و به صورت هوشمند دشوار می‌گردد. روش ارائه شده، ویژگی‌های جداساز با ظرفیت بالا را از تصاویر چهره استخراج می‌نماید تا با وجود حجم زیاد افراد موجود در پایگاه داده، عملیات بازیابی با نرخ و سرعت مناسب انجام شود. روش ارائه شده در این پایان‌نامه، مبتنی بر روش‌های ترکیبی بوده و ویژگی‌های نهایی را با پیمایش کل تصویر چهره استخراج می‌نماید.

۶- نوآوری مسئله

مسئله مطرح شده در این پایان‌نامه، با هدف بهبود برخی چالش‌های روش‌های اخیر در شناسایی چهره ارائه شده است. از جمله بهبودهای حاصل می‌توان به موارد زیر اشاره کرد:

۱- ویژگی‌هایی مدنظر است که بتوانند با توجه به تعداد تصویر مورد مقایسه (تصاویر موجود در پایگاه داده)، تفاوت‌های مناسبی را بین افراد قائل شوند.

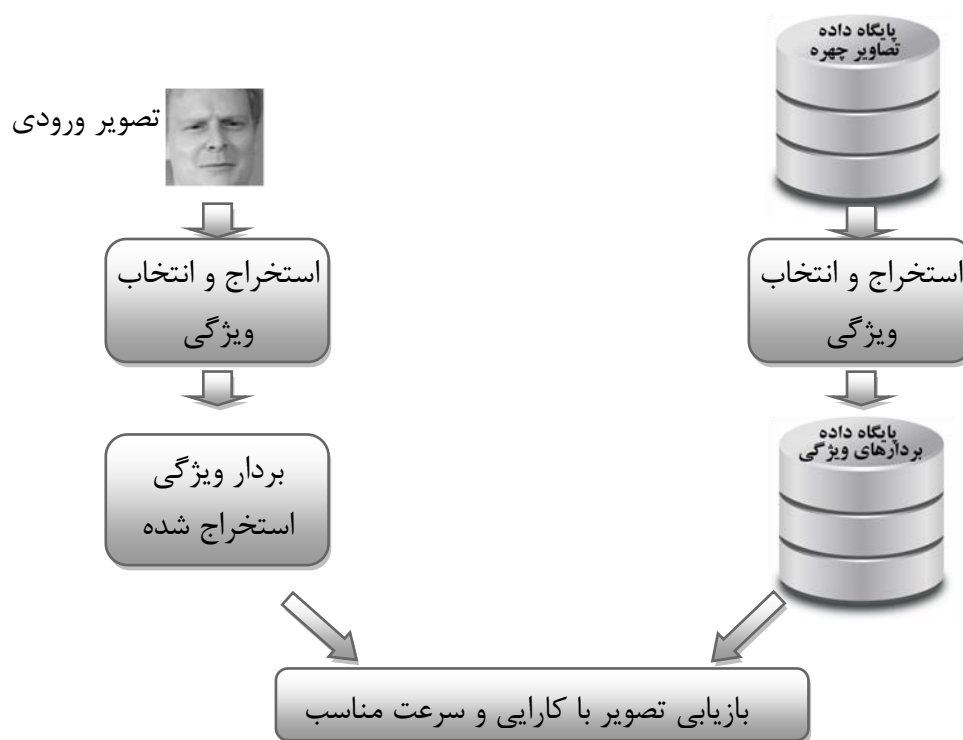
- قابلیت اعمال در پایگاه داده‌های حجیم: روش پیشنهادی ویژگی‌های با ظرفیت بالا را از ناحیه چهره استخراج می‌کند. به همین دلیل، قابلیت جداسازی افراد موجود در پایگاه داده‌ی حجیم فراهم شده و نسبت به روش‌های پیشین بهبود یافته است.
- وجود تنها یک تصویر به ازای هر شخص: از چالش‌های اصلی مسئله شناسایی چهره، نیاز به چند تصویر چهره در پایگاه داده جهت آموزش می‌باشد. روش پیشنهادی تنها با یک تصویر به ازای هر شخص در پایگاه داده، عملیات شناسایی را انجام می‌دهد. از فواید این ویژگی می‌توان به عدم نیاز به آموزش، کاهش فضای ذخیره سازی و سادگی کاربرد اشاره نمود.
- بهبود نرخ شناسایی بدون پیش‌پردازش: در روش پیشنهادی، بدون اعمال پیش‌پردازش روی تصاویر ورودی و تصاویر پایگاه داده، نرخ شناسایی از سایر روش‌ها بیشتر است. با این ویژگی، زمان بازیابی افزایش یافته و پیچیدگی محاسبات کاهش پیدا کرده است.
- دستیابی به سرعت مناسب: روش پیشنهادی با استخراج ویژگی‌های پرمعنا و در عین حال استفاده از الگوریتم‌های سریع جهت استخراج این ویژگی‌ها، سرعت مناسبی را نسبت به سایر روش‌ها ارائه کرده است. با این خاصیت، نشان داده شده است که با افزایش حجم پایگاه داده، میزان کاهش سرعت بسیار مناسب است.

۷- هدف پایان‌نامه

با توجه به تعریفی که از مسئله داشتیم، هدف اصلی، حل مشکل کارایی پایین، و سرعت جستجوی نا-مناسب روش‌های موجود، برای پایگاه داده‌های حجیم تصاویر چهره می‌باشد. لذا علاوه بر تصویر ورودی، باید تمرکز خود را روی پایگاه داده‌ای که تصویر قرار است از آن بازیابی شود نیز دقیق‌تر کنیم. تمرکز اصلی ما بر روی پایگاه داده‌ای می‌باشد که به دلیل حجیم بودن، و در نتیجه شباهت بین خواص طبیعی چهره‌ها در این حجم داده‌ای، بازیابی تصویر با نرخ بالا و سرعت مناسب توسط روش‌های عادی و سنتی

تقریباً غیر ممکن خواهد بود. طبیعتاً برای رفع مشکلات مذکور، باید ویژگی‌هایی را از تصاویر چهره استخراج نماییم، که علاوه بر متراکم کردن بردار ویژگی‌ها، قابلیت تفکیک بین تصاویر با تعداد زیاد را دارا باشد. در مرحله دوم، باید ویژگی‌های مشابه و بلا استفاده و حتی مضر را شناسایی و حذف نماییم تا از شناسایی کاذب در حجم داده زیاد خودداری شود. در نهایت باید روشی را جهت جستجو در پایگاه داده، طبق ویژگی‌های استخراج شده، در نظر بگیریم، که بتواند عمل جستجو را با سرعتی مناسب پیش ببرد، در حالی که به کارایی کلی الگوریتم لطمه‌ای وارد نشود.

در این پایان‌نامه، فرض شده است که تصاویر ورودی و تصاویر موجود در پایگاه داده، تصاویری تمام رخ و کامل از افراد می‌باشند. لذا تمرکز ما روی چگونگی استخراج چهره، یا نحوه محاسبه بهترین چهره از بین چهره‌های دریافتی از ورودی نمی‌باشد. به شکل اجمالی می‌توانیم روال کلی و هدف پایان‌نامه را به صورت دیاگرام نمایش داده شده در شکل (۱-۱) مشاهده نماییم.



شکل (۱-۱): روال کلی و هدف پایان‌نامه

۸- ساختار پایان‌نامه

این پایان‌نامه در ۵ فصل تهیه و تدوین شده است. در فصل دوم، به بررسی برخی از روش‌های استخراج ویژگی و شناسایی چهره پرداخته شده است. در این فصل ابتدا برخی روش‌های مرسوم استخراج ویژگی از تصاویر چهره، به صورت اجمالی بیان شده و سپس به بررسی چند روش منتخب به شکل دقیق‌تر پرداخته شده است. در فصل سوم به بررسی روش پیشنهادی استخراج ویژگی از تصویر چهره پرداخته و الگوریتم پیشنهادی خود را جهت جستجوی تصویر ورودی در پایگاه داده، مبتنی بر ویژگی‌های استخراج شده شرح داده‌ایم. فصل چهارم نتایج حاصل از الگوریتم ارائه شده را مورد بررسی قرار داده و مقایسات لازم این روش با روش‌های پیشین از نظر کارایی و سرعت اجرا، بیان شده است. در ادامه، فصل پنجم با ارائه نتیجه‌گیری و کارهای آینده، پایان بخش مطالب پایان‌نامه خواهد بود.

فصل دوم

مروری بر کارهای پیشین

طی دو دهه اخیر، روش‌های بسیار متنوعی جهت شناسایی چهره انسان‌ها در کاربردهای متفاوت ارائه شده‌اند. هر یک از این روش‌ها جهت رفع یکی از مشکلات شناسایی چهره راهکارهایی ارائه دادند. روش‌های اخیر، به بررسی ویژگی‌هایی پرداخته‌اند که اکثراً در حجم داده کم، عملکرد مناسبی از خود نشان داده‌اند. لذا برای استخراج ویژگی از تصاویر چهره، و سپس به کار بستن آنها جهت رسیدن به اهداف یاد شده در فصل قبل، لازم است ابتدا روش‌هایی که تا کنون در راستای استخراج ویژگی از تصاویر چهره ارائه شده‌اند، به صورت موردی بررسی کرده و سپس به بررسی دقیق چند روش منتخب در این حوزه و مزایا و معایب هر کدام بپردازیم و حوزه کاربرد آن‌ها را مورد بحث قرار دهیم.

ابتدایی‌ترین روش‌های شناسایی چهره، هنوز به عنوان پایه‌ی روش‌های جدید، استفاده می‌شوند. روش چهره‌های ویژه از جمله این روش‌ها است [8]. این روش با محاسبه میانگین و ماتریس کواریانس تصاویر، فضای ویژگی را از تصاویر استخراج می‌کند. همین‌طور در [13] روش الگوی باینری محلی جهت شناسایی چهره بکار گرفته شده است. در این روش ویژگی‌های بافتی و ساختاری با استفاده از هیستوگرام‌های الگوی باینری محلی، جهت نگاشت تصویر چهره به بردار ویژگی‌ها استفاده شده است که هدف اصلی را روی سرعت بالای استخراج ویژگی‌ها قرار داده است. جهت بهبود این روش، در [14] روشی جهت رفع عیب محلی بودن بیش از حد این ویژگی‌ها ارائه شد. این روش با استفاده از چندین بلاک به جای یک پیکسل انفرادی، و در نهایت میانگین‌گیری آنها، محسبات را بسیار بهبود داده است. در [15] روشی مبتنی بر یادگیری ماشین برای شناسایی چهره پیشنهاد شده است. این روش مبتنی بر ویژگی بوده و با استفاده از نواحی محلی چهره، استخراج ویژگی را انجام می‌دهد. چن^۱ روشی برای شناسایی چهره با

1- Chen

سرعت بالا ارائه کرده است [16]. این روش، مبتنی بر تبدیل کسینوسی گسسته و تحلیل خطی فیشر^۱ و شبکه عصبی RBF^۲ است. در مرجع [17] الگوریتم SIFT^۳ در شناسایی چهره به کار گرفته شده است. این الگوریتم مبتنی بر محاسبه دو پارامتر زاویه و اندازه از نقاط کلیدی تصویر است. این نقاط، کوچکترین و بزرگترین نقاط هرم گوسین می‌باشند که از اعمال تابع گوسین در پهناهای متفاوت روی تصویر تشکیل می‌شود. در این پژوهش حجم پایگاه داده انتخابی نسبت به روش‌های دیگر بیشتر است. سپس ترکیبی از دو الگوریتم SIFT و SURF^۴ در [18] ارائه شده است که دقت بالاتری در شناسایی چهره حاصل کرده است، هرچند به حجم دقیق پایگاه داده در این روش اشاره نشده است. ژو^۵ و همکاران در [19] با روش استخراج ویژگی موجک گابور و همینطور انتخاب ویژگی توسط الگوریتم آدابوست^۶، توانستند نتایج نسبتاً مطلوبی را حاصل کنند. در [20] روشی ترکیبی با استفاده از الگوریتم PCA^۷ و موجک گابور ارائه شده است.

در [21] الگوریتمی مبتنی بر تبدیل هاف جهت فائق آمدن بر مشکل تغییرات روشنایی چهره ارائه شده است. همچنین در [22] دو روش پیش پردازش ارائه شده و سپس از فیلتر گابور جهت مرتب‌سازی تصاویر چهره در راستای تغییرات زاویه استفاده شده است. در [23] روشی ترکیبی جهت شناسایی چهره ارائه شده است. در این روش، ابتدا با افزایش فضای ویژگی و سپس با کاهش بُعد همراه با حفظ اطلاعات، ویژگی‌های تصویر استخراج شده‌اند. در [24] با استفاده از ماتریس آنترپی، کواریانس، PCA و در نهایت تعریف ماتریس فاصله آنترپی، عملیات شناسایی چهره انجام شده است. در [25] با استفاده از روش بازسازی محلی مبتنی بر یک چهره واحد، تصاویر را به صورت بلاکی پیمایش کرده و با مقایسه بلاک‌های

1- Fisher
 2- Radial basis function
 3- Scale-invariant feature transform
 4- Speeded Up Robust Features
 5- Zhou
 6- Adaboost
 7- Principal Component Analysis

بازسازی شده‌ی متناظر بین دو تصویر، تشابه تصاویر اندازه‌گیری شده‌اند. در [26-28] نیز روش‌هایی برای بازیابی چهره با استفاده از تشابه نقاط مکانی و ساختاری، شبکه عصبی RBF و دسته بند C- میانگین فازی^۱ ارائه شده‌اند.

در بیشتر روش‌های یاد شده، از تعداد افراد کمتر از ۵۰ نفر در پایگاه داده جهت ارزیابی استفاده شده است. همچنین در اکثر موارد، چندین تصویر به ازای هر شخص جهت فرآیند آموزش به کار گرفته شده است. در ادامه به بررسی چند مورد از روش‌های منتخب به صورت تفصیلی می‌پردازیم.

۲- چهره‌های ویژه

در این بخش به بررسی روشی می‌پردازیم که جهت شناسایی چهره در [8] ارائه شده است. این روش که یکی از دسته روش‌های آماری مبتنی بر روش‌های کل‌نگر می‌باشد، هدف خود را مبتنی بر یافتن اجزای اصلی در توزیعی از چهره‌های موجود در پایگاه داده متمرکز کرده است. برای این هدف، روش مذکور از ماتریس کواریانس تصاویر موجود در پایگاه داده استفاده می‌نماید. این روش، پس از محاسبه ماتریس همبستگی بین تصاویر موجود، با اعمال الگوریتم PCA روی این ماتریس، مولفه‌های اصلی را از آن استخراج می‌نماید. لذا با این کار می‌تواند به توزیعی از چهره‌های موجود دست یابد.

این الگوریتم با انتخاب تنها M مولفه متناظر با مقادیر ویژه بالاتر، و ادغام آنها می‌تواند همواره تقریبی از تصاویر موجود را حاصل نماید. این در صورتی است که با داشتن تمامی مولفه‌های اصلی و اعمال یک ترکیب خطی روی آن‌ها، می‌تواند تمام تصاویر موجود را به شکل دقیق بازسازی نماید. به عبارتی می‌توان این گونه تصور کرد که با استخراج این M مولفه، تفاوت‌های موجود بین تصاویر پایگاه داده، در اندازه‌ها و با ملاک‌های مختلفی، قابل مشاهده خواهد بود.

1- Fuzzy c-means clustering

در نهایت با مقایسه تمام تصاویر موجود، با هر یک از چهره‌های ویژه استخراج شده، یک بردار ویژگی شامل ضرایب همبستگی بین آنها استخراج می‌شود. روال کار به این صورت می‌باشد:

فرض می‌کنیم تصویر ورودی $I(x,y)$ با ابعاد N در N (با مقادیر پیکسلی ۸ بیتی) جهت استخراج ویژگی فراهم شده است. پس از انتقال تصویر به یک بردار مرتب شده به اندازه N^2 و همچنین تشکیل ماتریس کواریانس تصاویر، بردارهای ویژه این ماتریس محاسبه می‌شوند که هر کدام دارای طول N^2 می‌باشند. در ادامه به این فرایند می‌پردازیم.

اگر مجموعه تصاویر آموزشی (تصاویر موجود در پایگاه داده) را $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_R$ در نظر بگیریم،

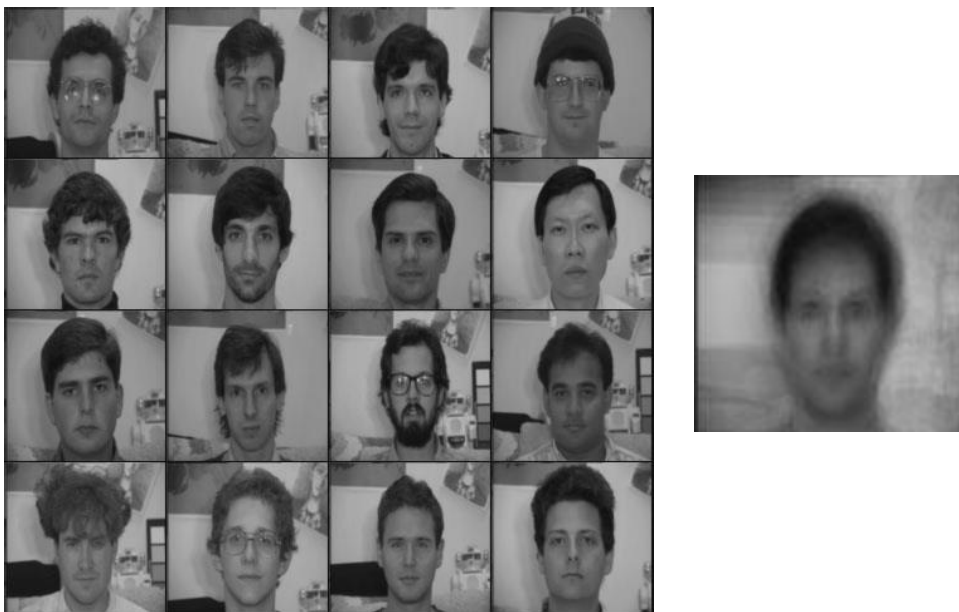
میانگین تصاویر طبق معادله‌ی (۱-۲) محاسبه می‌گردد.

$$\Psi = \frac{1}{R} \sum_{i=1}^R \Gamma_i \quad (1-2)$$

سپس فاصله هر تصویر تا میانگین طبق معادله (۲-۲) محاسبه می‌گردد.

$$\phi_i = \Gamma_i - \Psi \quad (2-2)$$

یک نمونه آموزشی در شکل (۱-۲) نمایش داده شده است.



شکل (۱-۲): یک مجموعه آموزشی از تصاویر چهره و میانگین آنها در روش Eigenface.

سمت چپ: مجموعه آموزشی نمونه. سمت راست: میانگین حاصل تصاویر [8]

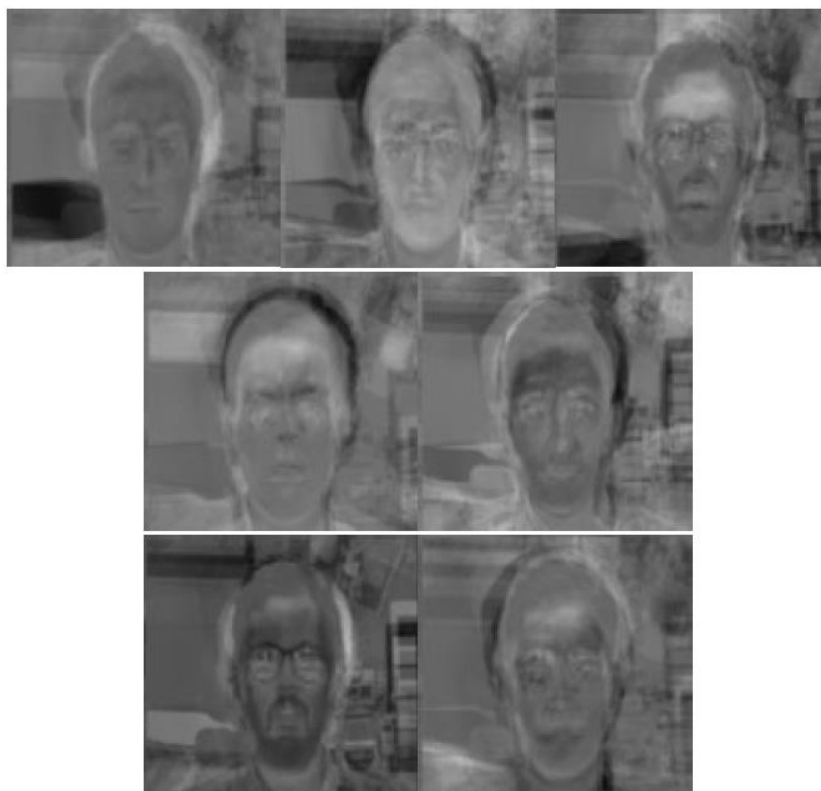
با استفاده از داده‌های فوق، ماتریس کواریانس به شکل معادله (۳-۲) استخراج می‌گردد.

$$C = \frac{1}{R} \sum_{i=1}^R \phi_i \phi_i^T = AA^T \quad (3-2)$$

که در آن: $A = [\phi_1 \ \phi_2 \ \dots \ \phi_R]$.

سپس بردارهای ویژه (μ_i) از ماتریس فوق قابل استخراج خواهند بود. به دلیل ابعاد بالای بردارهای ویژه (N^2) و در نتیجه افزایش زمان اجرا، بردارهای ویژه AA^T را محاسبه نموده و سپس طبق تبدیل زیر، آنها را به بردارهای ویژه AA^T تبدیل می‌نماییم:

اگر فرض کنیم v_i بردارهای ویژه AA^T باشند به طوری که $AA^T v_i = \mu_i v_i$ با ضرب طرفین در A داریم $AA^T A v_i = \mu_i A v_i$ لذا می‌توان $A v_i$ را بردارهای ویژه ماتریس AA^T در نظر گرفت. که در این صورت بعد N^2 به R تبدیل خواهد شد. ۷ تصویر ویژه حاصل از تصاویر فوق در شکل (۲-۲) قابل مشاهده است [8].



شکل (۲-۲): ۷ چهره ویژه محاسبه شده از تصاویر شکل (۱-۲) بدون حذف پس زمینه [8]

با افزودن تصویر جدید به پایگاه داده، روال استخراج تصاویر ویژه از سر گرفته می‌شود. اما این فرایند می‌تواند در زمانی که سیستم در حال استراحت یا خارج از کار می‌باشد انجام شود و به این صورت با موازی سازی محاسبات، از کاهش سرعت سیستم جلوگیری نمود.

این روش با استفاده از پایگاه داده دست ساز شامل ۲۵۰۰ تصویر از ۱۶ شخص تحت تمام ترکیب‌های حالات مختلف شامل سه زاویه‌ی سر، سه اندازه‌ی سر یا مقیاس، و سه شرایط روشنایی و رزولوشن متفاوت تست شده است. نرخ شناسایی ۹۶٪، ۸۵٪ و ۶۴٪ برای اختلاف روشنایی، زاویه، و مقیاس گزارش شده است. هرچند به نظر می‌رسد این روش قدرت نسبتاً خوبی با تغییرات روشنایی دارد، اما کارایی آن با تغییرات مقیاس تنزل پیدا می‌کند [8].

۳- تبدیل موجک گابور

فیلترهای معروف گابور^۱، اوایل جهت تشخیص و رفع نویز در سیگنال‌ها مورد استفاده قرار می‌گرفتند [29]. بعدها مدل‌هایی جهت نگاشت این فیلترها در فضای دو بعدی مثل فضای تصویر، تعریف و به کار گرفته شد. در حوزه شناسایی شیء در تصاویر، فیلترهای گابور مختلفی مورد استفاده قرار گرفته است. در این بخش روش ارائه شده در [22] را مورد بررسی قرار می‌دهیم که جهت شناسایی چهره از این فیلتر استفاده می‌نماید.

این روش از چند مرحله اصلی تشکیل شده است. مرحله اول بر پیش پردازش متمرکز است. در این مرحله از الگوریتم AD^2 جهت استخراج لبه‌ها و نرم کردن تصویر به عنوان پیش پردازش استفاده شده است. مرحله دوم از فیلتر گابور جهت استخراج ویژگی در راستای زاویه‌های خاص استفاده می‌شود. در

1- Gabor Wavelet Transform
2- Anisotropic Diffusion

مرحله نهایی با جستجوی فضای ویژگی و انتخاب ویژگی‌های بهینه با استفاده از الگوریتم^۱ BPSO، فضای ویژگی کاهش پیدا می‌کند.

۳-۱- پیش پردازش

در مرحله اول با استفاده از تبدیل AD تصویر اصلی مورد پیش پردازش قرار می‌گیرد. این تبدیل دو خاصیت اصلی دارد: (۱) تصویر اصلی را در فضای مقیاس هموارسازی^۲ می‌کند. (۲) لبه‌ها و خطوط و خواص مهم دیگر تصویر را استخراج می‌کند. ویژگی این الگوریتم به همراه داشتن هر دو این خواص با هم می‌باشد.

در پردازش تصویر و ماشین بینایی، AD یک تکنیک جهت کاهش نویز تصویر بدون حذف بخش‌های مهم از محتویات تصویر از جمله لبه‌ها و خطوط می‌باشد. AD شبیه به فرایند تبدیل یک تصویر به مجموعه‌ای از تصاویر بلور شده به صورت ترتیبی در فضای مقیاس، مبتنی بر یک فرایند انتشار^۳ می‌باشد. فرایند انتشار در حالت عادی، یک تبدیل مقیاس ثابت و خطی از تصویر اصلی می‌باشد. شدت بلور این تصاویر به شکل مرحله‌ای افزایش پیدا می‌کند. در حالت عادی فرایند هموارسازی، می‌توان هر کدام از تصاویر حاصل از فرایند را نتیجه‌ی کانالو بین تصویر اصلی و یک فیلتر گوسین دو بعدی تصور کرد که انحراف معیار فیلتر به شکل مرحله‌ای کاهش پیدا می‌کند. AD یک حالت کلی از این فرایند انتشار می‌باشد که مجموعه‌ای از تصاویر را تولید می‌کند اما هر تصویر حاصل، یک ترکیب بین تصویر اصلی و یک فیلتر است که این فیلتر وابسته به محتویات محلی تصویر اصلی می‌باشد. بنابراین AD یک تبدیل مقیاس ثابت و غیر خطی از تصویر اصلی می‌باشد.

1- Binary Particle Swarm Optimization
2- Smoothing
3- Diffusion process

تکنیک‌های انتشار غیر خطی که نمونه‌ای از روش‌های هموارسازی تطبیقی^۱ هستند، مبتنی بر اعمال فرایندی روی تصویر می‌باشند که خود وابسته به خواص تصویر است. این تکنیک‌ها برای چیره شدن بر مشکلات انتشار خطی ارائه شدند. انتشار خطی دارای یک نفوذ همگن در تصویر با یک ضریب هدایت ثابت می‌باشد. در این فرایند انتشار، نویز و لبه‌ها در تصویر هموار می‌شوند. این نوع انتشار را انتشار همسانگرد^۲ می‌گویند. جهت جلوگیری از هموار شدن لبه‌ها طی عملیات حذف نویز، باید از انتشار غیر خطی استفاده شود. در این روش نویز هموار و لبه‌ها برجسته می‌شوند. فرایند این انتشار وابسته به اطلاعات محلی تصویر است. این دسته انتشارها را انتشار ناهمسانگرد^۳ می‌گویند [30]. هرچند این مفهوم در جامعه پردازش تصویر شناخته شده است (برای مروری کلی به [31] و ارجاعات آن مراجعه شود)، اما یک فرمول‌بندی مرتبط، اولین بار توسط Malik و Perona در سال ۱۹۸۷ ارائه گردید. Malik و Perona در [32]، یک تعریف جدید از فضای مقیاس انجام دادند و الگوریتمی را که با استفاده از یک فرایند انتشار محقق می‌شود، ارائه کردند [33]. این الگوریتم شامل اعمال عملگرهای محلی و پایه به صورت تکراری روی تصویر می‌باشد که پیاده‌سازی‌های سخت‌افزاری موازی را امکان پذیر می‌نماید. معادله انتشار در این الگوریتم، جهت هموارسازی و حذف نویز از لاپلاس تصویر استفاده می‌کند. این معادله انتشار ناهمسانگرد در معادله (۴-۲) نشان داده شده است.

$$I_t = \text{div}(c(x, y; t) \nabla I) = c(x, y; t) \Delta I + \nabla c \cdot \nabla I \quad (4-2)$$

که در آن t پارامتر زمان (مقیاس t ام)، $I(x, y, 0)$ تصویر اصلی، ∇I گرادیان تصویر، ΔI تبدیل لاپلاس تصویر و c ضریب انتقال می‌باشند.

در معادله‌ی فوق با ثابت بودن c ، معادله به شکل انتشار همسانگرد و خطی یعنی $I_t = c \Delta I$ تبدیل خواهد شد.

1- Adaptive smoothing
2- Isotropic
3- Anisotropic

می‌خواهیم هموارسازی نواحی یکنواخت را نسبت به هموارسازی مرزها اولویت دهیم. برای این منظور باید ضریب انتقال جوری تنظیم شود که داخل نواحی مقدار ۱ و روی مرز نواحی مقدار صفر به خود اختصاص دهد. به بیان ریاضی، ضریب انتقال باید تابعی از مرزهای تصویر در هر مرحله باشد. در این صورت عملیات بلورسازی نواحی را از هم مجزا کرده و تداخل بین نواحی ایجاد نمی‌کند. به عبارتی مرز بین نواحی را برجسته می‌نماید. در این الگوریتم، از گرادیان به عنوان بهترین تابع تقریب برای یافتن مرز نواحی در هر مرحله استفاده شده است.

بنابراین c باید تابعی از گرادیان تصویر باشد، از طرفی طبق شرایط فوق، این تابع باید یکنوا نزولی و غیر صفر باشد. با فرض $g(\cdot)$ به عنوان این تابع انتقال، c به شکل معادله‌ی (۲-۵) محاسبه می‌گردد.

$$c = g(\|\nabla I(x, y, t)\|) \quad (۲-۵)$$

این تابع انتقال طوری انتخاب می‌شود که $\lim_{x \rightarrow 0} g(x) = 1$ که یعنی انتشار در نواحی یکنواخت بیشینه می‌شود و $\lim_{x \rightarrow \infty} g(x) = 0$ که یعنی انتشار در راستای لبه‌ها متوقف می‌شود. این خاصیت، در راستای دستیابی به دو ویژگی تبدیل AD می‌باشند. با این مشخصات، دو تابع توسط Perona و Malik ارائه گردید که در معادله (۲-۶) و (۲-۷) نشان داده شده است.

$$g_1(\nabla I) = e^{-\left(\frac{\|\nabla I\|}{K}\right)^2} \quad (۲-۶)$$

$$g_2(\nabla I) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{K}\right)^2} \quad (۲-۷)$$

که در آنها K پارامتر حد آستانه اندازه گرادیان است که نرخ انتشار را کنترل می‌کند. فضای مقیاس تولید شده توسط این دو تابع متفاوت‌اند: تابع اول، به لبه‌های با کنتراست بالا، امتیاز بیشتر می‌دهد، تابع دوم، به نواحی وسیع‌تر امتیاز بیشتری نسبت به نواحی کوچک‌تر می‌دهد.

با توجه به موارد فوق، جهت پیش‌روی تبدیل فوق، گسسته‌سازی چهار نزدیک‌ترین همسایه در عملگر لاپلاس به شکل معادله (۲-۸) انجام می‌شود.

$$I_{i,j}^{t+1} = I_{i,j}^t + \lambda [c_N \cdot \Lambda_N I + c_S \cdot \Lambda_S I + c_E \cdot \Lambda_E I + c_W \cdot \Lambda_W I]_{i,j}^t \quad (۸-۲)$$

که در آن $1 < \lambda < \frac{1}{4}$ به عنوان ضریب عددی جهت پایداری، N و S و E و W نشان‌دهنده‌ی همسایه‌های پیکسلی بالا، پایین، راست و چپ پیکسل فعلی می‌باشند. عملگر Λ نشان‌دهنده‌ی اختلاف فاصله بین نزدیک‌ترین پیکسل همسایه در جهت مشخص می‌باشد که به همراه ضریب انتقال c به شکل معادلات (۲-۹) تا (۱۲-۲) تعریف می‌شوند.

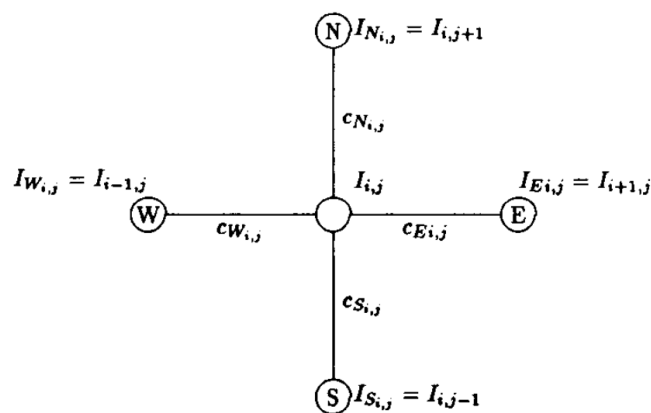
$$c_{N_{i,j}}^t = g(|\Lambda_N I_{i,j}^t|) \quad \Lambda_N I_{i,j} = I_{i-1,j} - I_{i,j} \quad (۹-۲)$$

$$c_{S_{i,j}}^t = g(|\Lambda_S I_{i,j}^t|) \quad \Lambda_S I_{i,j} = I_{i+1,j} - I_{i,j} \quad (۱۰-۲)$$

$$c_{E_{i,j}}^t = g(|\Lambda_E I_{i,j}^t|) \quad \Lambda_E I_{i,j} = I_{i,j+1} - I_{i,j} \quad (۱۱-۲)$$

$$c_{W_{i,j}}^t = g(|\Lambda_W I_{i,j}^t|) \quad \Lambda_W I_{i,j} = I_{i,j-1} - I_{i,j} \quad (۱۲-۲)$$

نام‌گذاری پیکسل‌های همسایه در شکل (۲-۳) نشان داده شده است.



شکل (۲-۳): نام‌گذاری پیکسل‌های همسایه جهت پیشرفت مراحل الگوریتم AD در پیش‌پردازش [22]

۲-۳- استخراج ویژگی مبتنی بر فیلتر گابور

یکی از مهمترین فیلترهایی که جهت نگاشت روی تصویر چهره انسان، از آن بهره گرفته شده است، توسط Daugman طراحی گردید [34]. در این روش، اثبات گردیده که ویژگی‌های بصری چهره انسان را می‌توان با نگاشت یک تابع توزیع گوسین مدل شده با تابع نمایی، از چهره استخراج نمود. این فیلتر، که

مشابه سایر فیلترهای گابور، دارای یک تابع موجک دوبعدی می‌باشد، به شکل معادله (۲-۱۳) تعریف و فرموله بندی شده است [22].

$$\psi_{f,\theta}(x,y) = e^{(-\frac{1}{2})(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2})} e^{i 2\pi f x'} \quad (2-13)$$

که در آن $x' = x \cos \theta + y \sin \theta$ و $y' = -x \sin \theta + y \cos \theta$ و f فرکانس موج سینوسی (مقیاس) در هر مرحله بوده و θ زاویه تابع گابور می‌باشد.

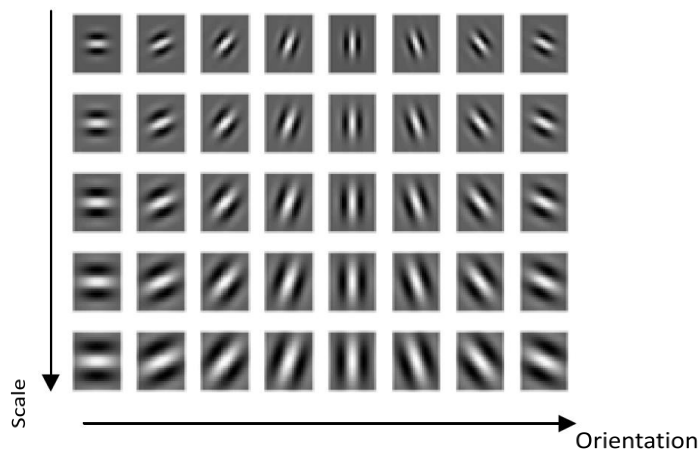
با توجه به قابلیت نمایش زاویه و مقیاس به صورت همزمان در توابع نمایی، از این تابع جهت استفاده در توزیع گوسین که روی چهره انسان نگاشت می‌شود، استفاده گردیده است [35].

با توجه به اینکه موجک‌ها در کاربردهای تک بعدی مثل سیگنال‌ها، در مقیاس‌های متعدد روی سیگنال اصلی شیف‌ت داده شده، و ضرایب همبستگی موجک با شیف‌ت‌های متعدد سیگنال، به عنوان بردار ویژگی استخراج می‌شود، در حالت دو بعدی نیز، این موجک مادر، در مقیاس‌ها و بعلاوه زاویه‌های متعدد، روی تصویر اصلی شیف‌ت داده شده و خواصی مبنی بر قسمت‌هایی از تصویر که با موجک مذکور، دارای توزیع گوسی مشابه می‌باشند، به عنوان بردار ویژگی از تصویر استخراج می‌شود. لذا با کانالو کردن موجک صفحه‌ای، روی تصویر چهره، با مقیاس و زاویه‌های متفاوت، برداری از ویژگی‌ها را از آن چهره خواهیم داشت.

نحوه اعمال این موجک روی تصویر به شکل معادله (۲-۱۴) می‌باشد که در حالت پیوسته ارائه گردیده است.

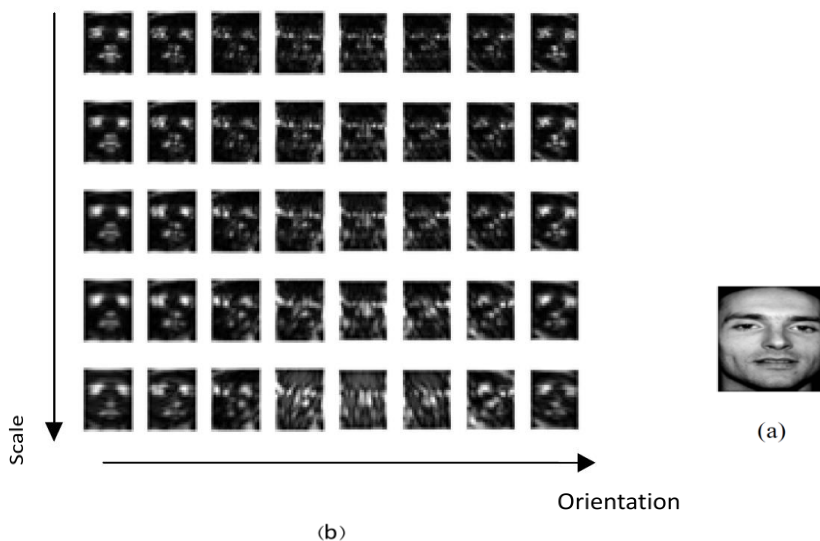
$$g(x,y) = I(x,y) \otimes \psi_{f,\theta}(x,y) \quad (2-14)$$

که در آن I تصویر ورودی به این فیلتر می‌باشد. نمونه‌ای از اعمال این موجک روی یک تصویر ورودی، در شکل (۲-۴) آمده است.



شکل (۲-۴): فیلترهای گابور مطابق با ۵ مقیاس و ۸ زاویه متفاوت [35]

با اعمال هر یک از فیلترهای گابور فوق، روی تصویر و کانالو کردن آن‌ها در تصویر ورودی، توزیعی از تصویر حاصل می‌گردد. نمونه‌ای از نتیجه این کانولوشن در شکل (۲-۵) مشاهده می‌شود.



شکل (۲-۵): نمونه‌ای از اعمال فیلترهای گابور روی تصویر ورودی از پایگاه داده Stirling. (a) تصویر اصلی (b) حاصل فیلتر [35]

در آخرین قدم، استخراج ویژگی‌ها از توزیع‌های حاصل روی تصویر چهره، انجام می‌گردد. می‌دانیم که هر توزیع گوسین دارای دو مشخصه میانی و واریانس می‌باشد. از این دو مشخصه جهت تشکیل بردار ویژگی‌های این روش استفاده می‌گردد. لذا به ازای هر توزیع، دو

ویژگی خواهیم داشت. از طرفی در این مثال ۴۰ توزیع گوسی داریم که در مجموع، بردار ویژگی شامل ۸۰ ویژگی خواهد بود. از طرفی اگر نقطه بیشینه هر توزیع را به عنوان نقطه ویژگی آن توزیع استخراج کنیم، می‌توانیم نقاط حساس که به عنوان ویژگی از تصویر ورودی استخراج شده‌اند را استخراج نماییم. شکل (۶-۲) این عملکرد را نشان می‌دهد.



شکل (۶-۲): استخراج نقاط ویژگی توسط موجک‌های گابور. سمت راست: تصویر اصلی، سمت چپ: نقاط استخراج شده. [35]

۳-۳- نتایج

عیب اصلی این روش این است که به یک پایه متعامد منجر نمی‌شود. یعنی تبدیل موجک گابور، دارای افزونگی است. لذا باعث کندی محاسبات مخصوصاً زمانی که پایگاه داده‌ای حجیم داشته باشیم، خواهد شد. از طرفی زمان استخراج ویژگی‌ها بسیار زیاد است که این امر باعث کاهش سرعت بازیابی می‌شود. [36,37]

نتایج حاصل از اعمال این روش روی پایگاه داده‌های متفاوت در جدول (۱-۲) نشان داده شده است.

جدول (۱-۲): نتایج حاصل از اعمال الگوریتم موجک گابور روی پایگاه داده‌های متفاوت در [22]

پایگاه داده	تعداد افراد	تعداد تصاویر به ازای هر شخص	نرخ شناسایی %
ORL	۴۰	۱۰	۹۴/۴۱
FERET	۳۵	۲۰	۸۳/۳
Cropped YALE B	۳۵	۲۰	۸۵/۱۲
FEI	۲۰۰	۱۰	۹۰/۴۱

۴- الگوریتم SIFT

در این بخش به بررسی عملکرد الگوریتم تبدیل ویژگی مقیاس ثابت^۱ ارائه شده در [17] پرداخته‌ایم. این الگوریتم از سه بخش اصلی شامل: (۱) استخراج نقاط کلیدی (۲) توصیف‌گر نقاط کلیدی و (۳) شباهت-سنجی تشکیل شده است. SIFT، برای استخراج ویژگی‌های متمایز از تصاویر جهت شناسایی یک شیء در حالت‌های مختلف ارائه شده است.

۴-۱- استخراج نقاط کلیدی

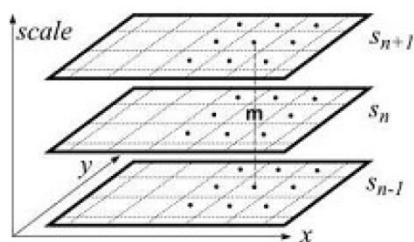
SIFT از فاصله تبدیل گوسین‌های متوالی حاصل از تصویر اصلی در فضای مقیاس، جهت شناسایی نقاط جالب توجه استفاده می‌کند. با لحاظ کردن یک تصویر ورودی I فضای مقیاس به عنوان یک تابع $L(x,y,\sigma)$ تعریف می‌شود که از کانولوشن تصویر اصلی با تابع گوسین $G(x,y,\sigma)$ در مقیاس‌های متفاوت حاصل می‌شود. سپس تابع اختلاف گوسین^۲ $D(x,y,\sigma)$ از اختلاف بین دو گوسین مجاور هم با اختلاف ضریب k به شکل معادله (۲-۱۵) محاسبه می‌شود:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (۲-۱۵)$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \text{که در آن}$$

در اختلاف گوسین‌های حاصل، هر پیکسل با ۸ همسایه در گوسین موجود و ۹ پیکسل در گوسین بالا و پایین خود، مقایسه می‌شود. این مقایسه در شکل (۲-۷) نشان داده شده است.

1- Scale invariant feature transform
2- Difference of Gaussian



شکل (۷-۲): مقایسه مقادیر پیکسلی در هرم گوسین جهت تشخیص نقاط کلیدی

در هر مقایسه، اگر مقدار پیکسل مورد نظر کمینه یا بیشینه بود، آن نقطه به عنوان نقطه کلیدی لحاظ می‌گردد.

۲-۴ - توصیف گر نقاط کلیدی

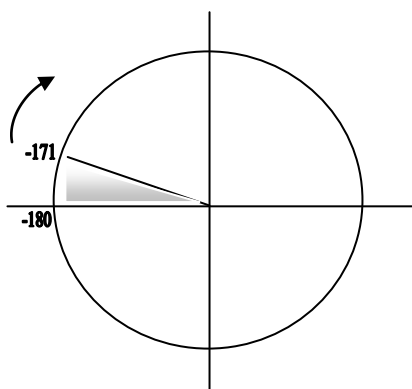
برای توصیف هر نقطه کلیدی که در مرحله قبل استخراج گردید، به ازای تمام پیکسل‌های موجود در تمام سطوح D ، اندازه گرادیان $M(x,y)$ و زاویه $\theta(x,y)$ با استفاده از معادله (۱۶-۲) و (۱۷-۲) محاسبه می‌شوند.

$$M(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (16-2)$$

$$\theta(x,y) = \tan^{-1} \frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \quad (17-2)$$

در مرحله بعد، زاویه نقطه کلیدی مبتنی بر یک ناحیه همسایگی اطراف آن تغییر می‌کند. برای این منظور ۳۶۰ زاویه ممکن را به ۳۶ کلاس تقسیم‌بندی می‌کنیم که هر کدام ۱۰ زاویه را در بر می‌گیرند. نماینده‌ی هر کلاس، زاویه‌ی میانی آن کلاس می‌باشد. یک پنجره 5×5 اطراف هر نقطه کلیدی انتخاب می‌نماییم. برای هر کدام از عناصر پنجره، کلاس مربوط به زاویه‌ی آن عنصر را مشخص می‌کنیم. مقدار نماینده‌ی آن کلاس جایگزین زاویه‌ی عنصر فعلی می‌شود

به این صورت تعداد زوایای مجاز کاهش پیدا می‌کنند. یک کلاس فرضی بین زاویه ۱۷۱- و ۱۸۰- درجه در شکل (۸-۲) نمایش داده شده است.



شکل (۸-۲): کاهش زاویه‌های مجاز برای نقاط کلیدی با کلاس بندی آن‌ها

سپس در هر پنجره، M های تمام خانه‌هایی که دارای θ های یکسان‌اند، با هم جمع می‌شوند. به عنوان مثال اگر در پنجره ۲۵ تایی اول، ۵ خانه دارای زاویه بین ۱۷۱ و ۱۸۰ بودند، M های این خانه‌ها با هم جمع می‌شود و به عنوان قدرت زاویه ۱۷۵، ثبت می‌شود.

با این کار، در کل ۳۶ منطقه زاویه‌ای داریم که هر کدام ۱۰ زاویه‌ی پشت سر هم را شامل می‌شوند. (۳۶۰ زاویه به ۳۶ زاویه کاهش پیدا کرد).

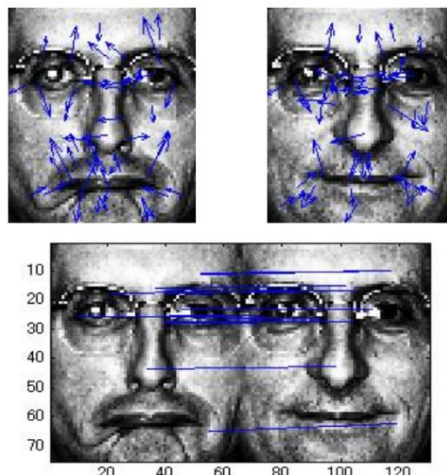
در نهایت تنها زاویه‌ای که بیشترین قدرت را جمع کرده است، نگه داشته شده و به آن نقطه کلیدی (نقطه مرکزی پنجره) اختصاص داده می‌شود.

در مرحله نهایی همسایگی‌های ۱۶×۱۶ هر نقطه کلیدی را دریافت کرده و مراحل زیر انجام می‌شوند:

- پنجره ۱۶×۱۶ به پنجره‌های ۴×۴ تقسیم می‌شوند.
- برای هر پنجره ۴×۴ مراحل قبل (کاهش زاویه‌ها) اینبار با قطعات ۴۵ درجه‌ای، تکرار می‌شوند و M ها را برای هر محدوده زاویه‌ای با هم جمع می‌کنیم. لذا در هر چرخه ۸ تا M استخراج می‌شود. (به ازای هر خانه ۴×۴).

لذا برای هر نقطه کلیدی ۱۲۸ توصیف‌گر لحاظ می‌گردد. که به ترتیب برای زوایای ۴۵ تایی پنجره‌ی ۴×۴ اول تا پنجره‌ی ۴×۴ آخر می‌باشند.

نمونه‌ای از نقاط کلیدی استخراج شده توسط این الگوریتم و مقایسه‌ی این نقاط در دو تصویر از یک شخص با تغییر حالات چهره در شکل (۲-۹) نشان داده شده است:



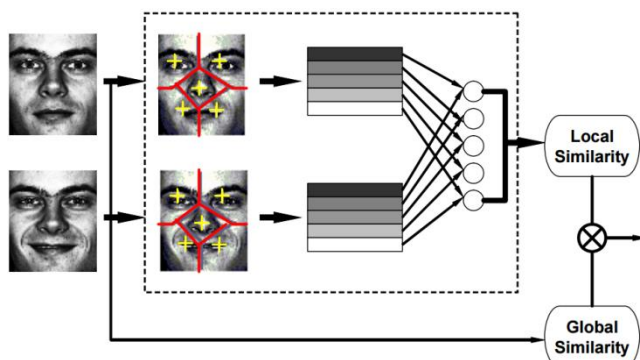
شکل (۲-۹): ویژگی‌های SIFT روی تصویر چهره و مقایسه ویژگی‌ها در دو تصویر از یک شخص با تفاوت حالات چهره [17]

۴-۳- نحوه‌ی شباهت سنجی

پس از استخراج نقاط کلیدی و اعمال توصیف‌گر روی آنها جهت ساخت بردار ویژگی‌ها، از فرایند زیر جهت اندازه‌گیری دو بردار ویژگی استفاده می‌شود:

(۱) تصویر به k ناحیه‌ی اولیه با مراکز تصادفی اولیه تقسیم‌بندی می‌شود. (۲) برای هر نقطه‌ی کلیدی در تصویر نزدیک‌ترین ناحیه را با استفاده از فاصله‌ی اقلیدسی لحاظ کرده و سپس مقادیر هر مرکز دسته را جهت بازسازی ناحیه‌ها به روز رسانی می‌کنیم. (۳) اگر مراکز جدید، به شکل قبلی باقی ماندند، کلاس‌بندی متوقف شده و k ناحیه‌ی باقی‌مانده به عنوان نواحی نگاشت برای مقایسه ثبت می‌شوند. (۴) پس از ساخت نواحی روی تصویر چهره، زمان آزمون چهره‌ی جدید، تمامی ویژگی‌های استخراج شده از تصویر، مبتنی بر موقعیت مکانی، به نواحی متناظر اختصاص داده می‌شوند.

نمونه‌ای از ناحیه‌بندی چهره در این روش در شکل (۲-۱۰) نشان داده شده است.



شکل (۲-۱۰): ساخت نواحی تصویری و محاسبه شباهت تصاویر چهره در الگوریتم SIFT [17]
 مبتنی بر نواحی ساخته شده و جهت شناسایی چهره، استراتژی نگاشت عمومی و محلی به صورت ترکیبی استفاده می‌شوند. با محاسبه دو معیار فاصله‌ی عمومی و محلی تعریف شده در [17]، بین دو تصویر و ضرب نهایی آن‌ها، فاصله‌ی نهایی دو تصویر از هم محاسبه می‌شود که عدد بزرگتر نشان‌دهنده‌ی شباهت بیشتر می‌باشد.

۴-۴- نتایج

این الگوریتم نسبت به سایر روش‌ها تا کنون نتایج قابل قبول‌تری حاصل کرده است. این روش با استفاده از نقاط کلیدی چهره و انتقال آن‌ها به فضای مقیاس ثابت و زاویه ثابت، قابلیت شناسایی را بیشتر کرده است. هرچند به دلیل محلی بودن نقاط کلیدی، عملکرد این الگوریتم با افزایش حجم پایگاه داده و در نتیجه ایجاد تشابه در چهره‌ها، ضعیف‌تر می‌شود. در جدول (۲-۲)، نتایج اعمال این الگوریتم بر دو پایگاه داده FERET و CAS-PEAL نمایش داده شده است.

جدول (۲-۲): نتایج حاصل از اعمال الگوریتم SIFT روی پایگاه داده‌های متفاوت

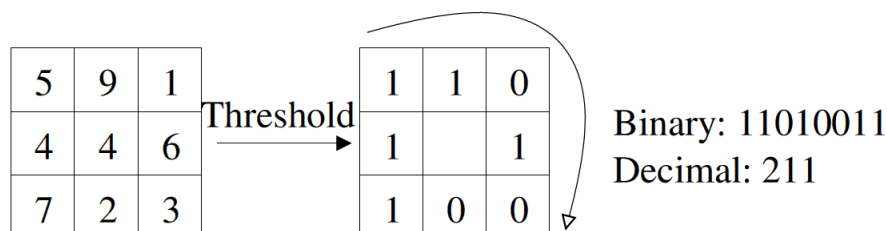
پایگاه داده	تعداد افراد	تعداد تصاویر به ازای هر شخص	نرخ شناسایی %
FERET	۲۰۰	۱	۹۷
CAS-PEAL	۴۰۰	۱	۹۴/۷

۵- الگوی باینری محلی

در این بخش روش ارائه شده در [13] را مورد بررسی قرار داده‌ایم. در این روش هم اطلاعات شکل ظاهری و هم بافت تصویر جهت نمایش تصویر استفاده می‌شوند. ابتدا ناحیه چهره به نواحی کوچک‌تر تقسیم می‌شود که از هر ناحیه ویژگی‌های الگوی باینری محلی^۱ استخراج شده و در نهایت به صورت یک هیستوگرام واحد به هم الحاق می‌شوند و این هیستوگرام، تصویر چهره را توصیف می‌کند. بافت نواحی تصویر به شکل محلی توسط الگوی باینری محلی کد می‌شوند در حالی که تمام ناحیه چهره با بازسازی هیستوگرام چهره کد شده، نمایش داده می‌شود.

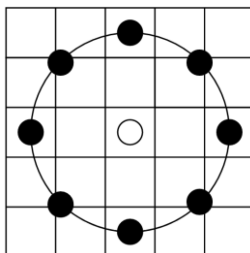
۵-۱- استخراج ویژگی

برای محاسبه الگوی باینری محلی تصویر به شکل ساده، برای هر موقعیت پیکسل، ۸ همسایه آن را در نظر می‌گیریم. سپس فاصله بین پیکسل مرکزی و هر یک از ۸ پیکسل همسایه‌ی 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 پیکسل مرکزی با هر پیکسل همسایه از d_{av} بیشتر یا مساوی باشند، به آن موقعیت، ۱ و در غیر این صورت صفر اختصاص می‌دهیم. لذا برای هر موقعیت پیکسل یک رشته دودویی ۸ رقمی بدست می‌آید که با تبدیل آن به معادل دهدهی، مقداری بین ۰ تا ۲۵۵ برای آن پیکسل محاسبه خواهد شد. نحوه این کدگذاری در شکل (۲-۱۱) مشاهده می‌شود.



شکل (۲-۱۱): عملگر LBP [13]

در این روش از همسایگی دایره‌ای استفاده شده است. مقادیر هر پیکسل روی شعاع همسایگی با استفاده از درون یابی پیکسل‌های اطراف محاسبه می‌شود. این همسایگی در شکل (۲-۱۲) نشان داده شده است.



شکل (۲-۱۲): همسایگی دایره‌ای در الگوی باینری محلی [13]

هیستوگرام تصویر جدید حاصل از الگوی باینری محلی، به عنوان ویژگی‌های بافتی تصویر استفاده می‌شود. این هیستوگرام دارای اطلاعاتی در مورد توزیع الگوهای محلی چهره مثل لبه‌ها، نقاط و نواحی هموار روی کل چهره است. برای نمایش مؤثرتر چهره، اطلاعات مکانی نیز به آن اضافه شده است. برای این منظور تصویر به چندین ناحیه تقسیم می‌شود. هیستوگرام هر ناحیه با استفاده از الگوی باینری محلی محاسبه شده، و با الحاق آنها به یکدیگر، هیستوگرام نهایی به عنوان بردار ویژگی تصویر بدست می‌آید.

۲-۵- معیار فاصله

در این روش جهت مقایسه دو بردار ویژگی از فاصله کای دو^۱ استفاده شده است. این معیار فاصله در معادله (۲-۱۸) نشان داده شده است.

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \quad (2-18)$$

زمانی که تصویر به نحوی متفاوت تقسیم‌بندی می‌شود، انتظار می‌رود در راستای تمایز بین افراد مختلف، برخی نواحی دارای اطلاعات مفیدتری نسبت به نواحی دیگر باشند. به همین دلیل به ازای هر

1- Chi Square

ناحیه یک وزن اختصاص داده شده است که متناسب با اهمیت اطلاعاتی است که آن ناحیه دارد. لذا معادله (۱۸-۲) به شکل معادله (۱۹-۲) استفاده می‌شود.

$$\chi^2_{\omega}(S, M) = \sum_{i,j} \omega_j \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}} \quad (۱۹-۲)$$

نمونه‌ای از این ناحیه‌بندی و وزن‌دهی در شکل (۱۳-۲) نشان داده شده است.



شکل (۱۳-۲): وزن‌دهی به نواحی چهره بر اساس معادله کای دو وزن دار. سمت چپ: یک نمونه تصویر چهره تقسیم شده به ۷×۷ پنجره. سمت راست: مجموعه وزن‌های معادله کای دو وزن دار. مربع‌های سیاه دارای وزن صفر، خاکستری پررنگ دارای وزن ۱، خاکستری کم رنگ دارای وزن ۲ و سفید دارای وزن ۴ می‌باشند [13].

۳-۵- نتایج

این روش نتایج خود را روی پایگاه داده FERET در دو حالت بیان کرده است. ابتدا نتایج بازیابی با وزن‌دهی به نواحی مختلف حاصل گردیده است. سپس نتایج بدون وزن‌دهی گزارش شده است. جدول (۳-۲) این نتایج را نشان می‌دهد.

جدول (۳-۲): نتایج حاصل از اعمال الگوریتم LBP روی پایگاه داده FERET

الگوریتم	تعداد افراد	تعداد تصاویر به ازای هر شخص	نرخ شناسایی %
LBP وزن دار	۱۶۰	۴	۹۷
LBP بدون وزن	۱۶۰	۴	۹۳

۶- تبدیل هاف

در این بخش به الگوریتم تبدیل هاف^۱ در [21] پرداخته‌ایم. در این روش با تقسیم‌بندی تصویر به بلاک‌های مشخص و استخراج تعداد مشخص از رأس‌های^۲ تبدیل هاف به عنوان ویژگی‌های تصویر، عملیات شناسایی انجام می‌شود. در این روش، از استخراج لبه با عملگر سوبل^۳ به عنوان پیش پردازش و از تبدیل هاف به عنوان روش استخراج ویژگی استفاده شده است. جهت انتخاب ویژگی‌ها نیز از الگوریتم BPSO^۴ استفاده شده است [38].

در این روش، ابتدا تصویر لبه یابی شده به تعداد بلاک‌های کوچک‌تر تقسیم‌بندی می‌شود. برای هر بخش تبدیل هاف اعمال شده و تعداد مشخصی از نقاط ویژگی استخراج می‌شوند. برای تعداد N بلاک، تصویر شامل $\sqrt{N}$ بلاک در راستای سطرها و $\sqrt{N}$ بلاک در راستای ستون‌ها می‌باشد. ابعاد بلاک به شکل معادلات (۲۰-۲) و (۲۱-۲) بدست می‌آید:

$$Blok_height(Bh) = \frac{Image_height}{\sqrt{N}} \quad (20-2)$$

$$Blok_width(Bw) = \frac{Image_Width}{\sqrt{N}} \quad (21-2)$$

که در آن Bh طول بلاک و Bw عرض بلاک می‌باشند.

تبدیل هاف در تعیین اشکال دلخواه از تصویر مثل خطوط، دایره‌ها و بیضی استفاده می‌شود. این فرایند مبتنی بر یک طرح رأی‌گیری است که برای بدست آوردن بهترین مقادیر رأس‌ها در فضای پارامتری انجام می‌شود. در روش ارائه شده در این مقاله این فرایند مبتنی بر تبدیل هاف استاندارد می‌باشد که جهت شناسایی خطوط مستقیم استفاده شده است. یک خط راست می‌تواند مطابق معادله (۲۲-۲) نشان داده شود.

1- Hough Transform

2- Peaks

3- Sobel

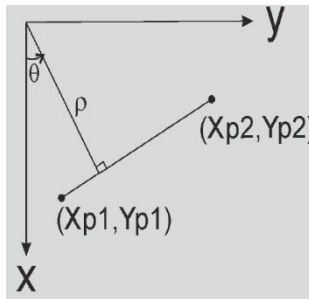
4- Bimary Particle Swarm Optimization

$$y = mx + c \quad (22-2)$$

که در آن m شیب خط و c محل قطع محور y ها توسط خط می‌باشد. برای فرایند محاسبه قوی‌تر Hart و Doda فضای پارامتری جدیدی را جهت شناسایی بهتر خطوط توسط تبدیل هاف ارائه دادند که با ρ و θ تعریف می‌شوند [39]. در فضای پارامتری جدید معادله خط مطابق معادله (23-2) در می‌آید.

$$y = \left(-\frac{\cos \theta}{\sin \theta}\right)x + \left(\frac{\rho}{\sin \theta}\right) \quad (23-2)$$

که در آن ρ فاصله متعامد خط تا مبدا و θ زاویه خط متعامد بر آن با یکی از محورها (معمولا محور x -ها) می‌باشد. این تبدیل پارامتری در شکل (2-14) نشان داده شده است.

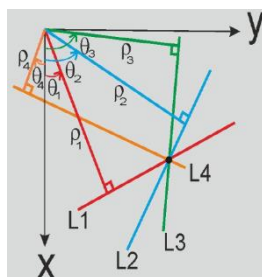


شکل (2-14): تغییر پارامترها در شناسایی خطوط مبتنی بر تبدیل هاف [21]

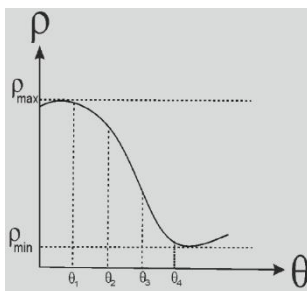
لذا معادله فوق به شکل معادله (2-24) می‌تواند تغییر یابد.

$$\rho = x \cos \theta + y \sin \theta \quad (24-2)$$

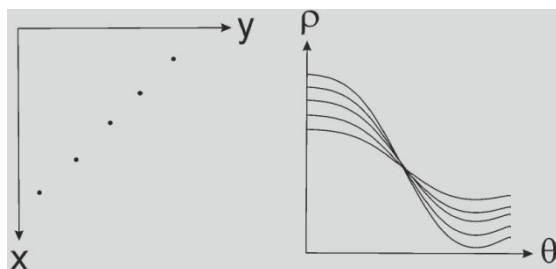
هدف اصلی تبدیل هاف را می‌توان اینگونه توصیف کرد: یک نقطه (x, y) در فضای مکان حاصل از فرایند لبه یابی تصویر فرض می‌کنم. هر نقطه در تصویر لبه‌یابی شده توسط بی‌نهایت خط با زاویه‌های متفاوت با راستای محور x ها و فاصله از مبدأهای متفاوت قط شده است. این مفهوم در شکل (2-15) نشان داده شده است.



شکل (۲-۱۵): تبدیل یک نقطه در حوزه مکان به بی نهایت خط در حوزه پارامتری تبدیل هاف [21]
 هر خط در تصویر لبه یابی شده شامل تعداد زیادی نقطه می باشد که به آن پاره خط رای می دهند. جهت تعیین خطوط مهم در تصویر لبه یابی شده در فضای مکان، هر نقطه را به فضای قطبی انتقال می دهیم. هر نقطه در فضای مکان، معادل یک موج سینوسی در فضای قطبی است که در شکل (۲-۱۶) نشان داده شده است.



شکل (۲-۱۶): فضای پارامتریک قطبی حاصل از یک نقطه در فضای مکان - حاصل از شکل (۲-۱۵) [21]
 بنابراین هر نقطه در فضای مکان، به مجموعه ای از پارامترهای (ρ, θ) در فضای قطبی رای می دهند. در نهایت نقطه ای که در فضای قطبی بیشترین رای را به خود اختصاص دهد، انتخاب می شود. این نقطه معادل یک پاره خط در فضای مکان می باشد. این فرایند در شکل (۲-۱۷) نشان داده شده است.



شکل (۲-۱۷): انتقال تصویر لبه یابی شده به فضای قطبی توسط تبدیل هاف و انتخاب نقطه با بیشترین رای به عنوان لبه ی برگزیده در فضای مکان [21]

تعداد رأسهای حاصل از هر بلاک باید برای تمامی تصاویر پایگاه داده لحاظ شود. این تعداد بسیار مهم است. تعداد زیاد منجر به استخراج ویژگی‌های ناخواسته شده که باعث کاهش نرخ شناسایی می‌شود. مقدار کم منجر به از دست دادن اطلاعات می‌شود. یک حد مشخص برای استخراج رأسهای تبدیل هاف طبق معادله (۲-۲۵) محاسبه می‌شود.

$$Peaks = \frac{(SAR) \times (BAR)}{Nb / 100} \times \alpha \quad (2-25)$$

که در آن SAR نسبت ابعاد تصویر است که به مقدار $\frac{4}{3}$ لحاظ شده است و BAR نسبت ابعاد بلاک‌ها می‌باشد که برابر $\frac{Bw}{Bh}$ می‌باشد. Nb تعداد بلاک‌هایی است که تصویر توسط آن‌ها بخش‌بندی شده است. α ثابت بهینه‌سازی است که در آزمایش‌ها مقدار ۱۶ برای آن حاصل شده است. لذا معادله فوق به شکل معادله (۲-۲۶) بازنویسی می‌شود.

$$Peaks = \frac{4 \times (Bw) \times 16 \times 100}{3 \times (Bh) \times (Nb)} \quad (2-26)$$

نتایج حاصل از این الگوریتم در جدول (۲-۴) مشاهده می‌شود.

جدول (۲-۴): نتایج حاصل از الگوریتم Hough در شناسایی چهره

پایگاه داده	Color FERET	CMU PIE (ILLUM)	CMU PIE (original)	Extended Yale B	CAS-PEAL
نرخ بازیابی	80.81	95.16	32.41	92.86	71.75

۷- جمع‌بندی

در این فصل به ۵ الگوریتم شناسایی چهره به صورت تفصیلی پرداخته شد. ویژگی‌های استخراجی از الگوریتم چهره‌های ویژه به دلیل حجم بالا و وابستگی مستقیم به کل چهره، در حجم زیاد پایگاه داده دچار مشکل خواهد شد. الگوریتم موجک گابور به دلیل افزونگی داده‌ها، دارای کاهش نرخ شناسایی در حجم زیاد پایگاه داده می‌باشد. الگوریتم الگوی باینری محلی، دارای حجم محاسبات زیادی می‌باشد. هرچند در این الگوریتم سعی شده است به بعد مکان نیز توجه شود، اما از قادر به پیمایش تمام چهره

جهت استخراج ویژگی نمی باشد. الگوریتم تبدیل هاف نیز به دلیل محلی بودن و استخراج لبه‌های تصویر، در حجم زیاد پایگاه داده و تشابه ویژگی‌ها دچار تشخیص اشتباه می‌شود. در بین الگوریتم‌های مطرح شده، الگوریتم SIFT قابلیت بالا تری داشته و نسبت به سایر الگوریتم‌ها در حجم قابل توجهی از پایگاه داده توانسته نرخ مناسب را کسب کند. اکثر روشهای پیشین، به دلیل مد نظر قرار دادن چهره به صورت محلی و استخراج ویژگی از آن، قادر به بازیابی تصویر در حجم زیاد تصاویر را ندارند. لذا نتایج حاصل از الگوریتمها، در حجم زیر ۲۰۰ چهره اعلام شده اند.

مقایسه کلی از این روشها در جدول (۵-۲) ارائه شده است.

جدول (۵-۲): مقایسه روش‌های بررسی شده با یکدیگر

روش	مقاومت در برابر تغییرات مقیاس و چرخش	مقاومت در برابر افزایش حجم پایگاه داده	سرعت بازیابی	کل نگر	مبتنی بر ویژگی
Eigenface	ضعیف	ضعیف	ضعیف	✓	-
Gabor	مقاوم	ضعیف	ضعیف	-	✓
SIFT	مقاوم	متوسط	مناسب	-	✓
LBP	مقاوم	متوسط	متوسط	-	✓
Hough	متوسط	متوسط	ضعیف	-	✓

فصل سوم

استفاده از ویژگی‌های مکان-فرکانسی جهت

شناسایی چهره (روش پیشنهادی)

۱- مقدمه

تمرکز اصلی روش پیشنهادی، بازیابی تصویر چهره از پایگاه داده حجیم از تصاویر می‌باشد که در آن به ازای هر شخص تنها یک تصویر چهره ذخیره شده است. در روش پیشنهادی، تعداد تصاویر پایگاه داده به مراتب بیشتر از روش‌های پیشین می‌باشد. بنابراین مشکل حجم پایگاه داده و وجود تنها یک تصویر به ازای هر شخص در پایگاه داده، تا حد قابل توجهی رفع شده است. همچنین مرحله آزمون با استفاده از تنها یک تصویر به ازای هر شخص انجام می‌شود.

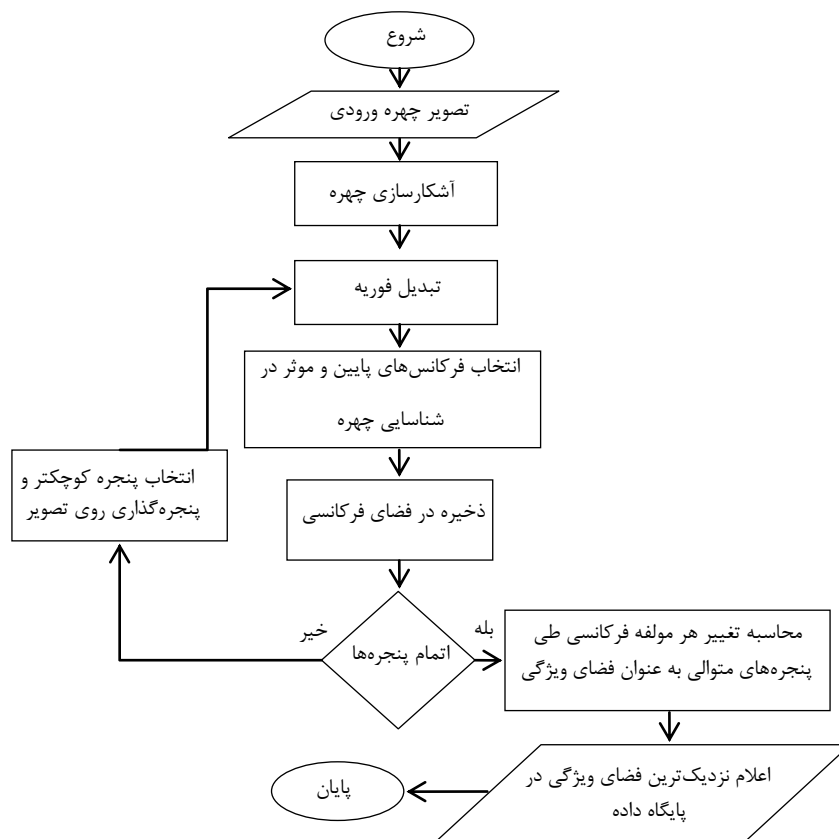
تلاش اصلی در سیستم‌های شناسایی چهره، نزدیک کردن ویژگی‌های استخراج شده به ویژگی‌های بصری انسان است. روش پیشنهادی نیز در راستای این هدف، طراحی شده است. ایده اصلی این روش مبتنی بر خاصیت شناسایی اشخاص توسط قوه بصری انسان است. انسان‌ها هنگام شناسایی شخص خاص در بین انبوهی از افراد از قوه بصری استفاده می‌کنند و با پیمایش تصویر از کل به جزء و بررسی چهره در هر مرحله، به شناسایی شخص می‌پردازند. در هر مرحله از پیمایش، استخراج معنایی تصویر چهره انجام می‌شود. در ادامه، ویژگی‌های تصویر مورد نظر با ویژگی‌های موجود در ذهن، مقایسه می‌شود و جهت اطمینان، جستجو ادامه می‌یابد. الگوریتم پیشنهادی نیز با استفاده از این منطق طبیعی، شناسایی چهره انسان را انجام داده است. برای استفاده از این منطق توسط ماشین‌های شناسایی هوشمند چهره، از اصول ریاضیات کاربردی استفاده شده است. به بیان زبان ماشین می‌توان گفت روش پیشنهادی به صورت ترکیبی از روش‌های مبتنی بر ویژگی و روش‌های کل‌نگر عمل می‌کند. در این روش با لحاظ کردن کل نقاط چهره، از ناحیه چهره استخراج ویژگی می‌شود و به این ترتیب با لحاظ کردن کل ناحیه چهره در ویژگی‌های استخراج شده، قابلیت مقایسه بین تعداد تصاویر بیشتری را فراهم می‌نماید. همچنین در روش پیشنهادی از تکنیکی جهت کاهش ابعاد ویژگی استفاده شده است. روش استخراج ویژگی به نحوی

انتخاب گردیده است که سرعت بازیابی مناسبی داشته باشد تا بتواند در مجموع نسبت به سایر روش‌ها کارایی مناسب‌تری داشته باشد.

روش پیشنهادی فاقد پیش پردازش می‌باشد. لذا تصاویر استفاده شده در این روش، به شکل اولیه و بدون تغییرات وارد فرآیند شناسایی می‌شوند. در بسیاری از روش‌های دیگر از فرایند پیش پردازش جهت بهبود نتایج استفاده شده است. این ویژگی یکی از فواید روش پیشنهادی می‌باشد. بدون پیش پردازش، سرعت بازیابی و موارد کاربرد الگوریتم افزایش پیدا می‌کند. در این فصل، از ۲۰۰ تصویر جهت داده‌های اعتبارسنجی استفاده شده است و پارامترهای بهینه روی این تصاویر حاصل گردیده است.

تمرکز اصلی روش پیشنهادی بر شناسایی چهره می‌باشد. از این رو برای آشکارسازی ناحیه چهره از الگوریتم Viola-Jones و توابع از پیش تعریف شده در نرم‌افزار متلب استفاده شده است [40]. روش پیشنهادی در بخش شناسایی چهره، شامل سه مرحله اصلی است: (۱) استخراج ویژگی از ناحیه چهره هر شخص (۲) کاهش فضای ویژگی (۳) اندازه‌گیری فاصله تصویر آزمون و تصاویر پایگاه داده.

پس از استخراج ناحیه چهره، از تصویر ورودی طی پنجره‌گذاری‌های متوالی استخراج معنا می‌شود. با استفاده از سطوح معنایی استخراج شده، فضای ویژگی تصویر ایجاد می‌شود. با استفاده از فیلتر مناسب، ویژگی‌ها را کاهش داده و سپس با استفاده از معیار فاصله مناسب، فاصله‌ی فضای ویژگی تصویر آزمون با فضای ویژگی تصاویر پایگاه داده محاسبه می‌شود. در نهایت، نزدیک‌ترین تصویر به تصویر آزمون، به عنوان جواب نهایی انتخاب می‌شود. فرایند کلی الگوریتم پیشنهادی در شکل (۳-۱) به تصویر کشیده شده است.



شکل (۳-۱): تصویر فرآیند کلی الگوریتم ارائه شده

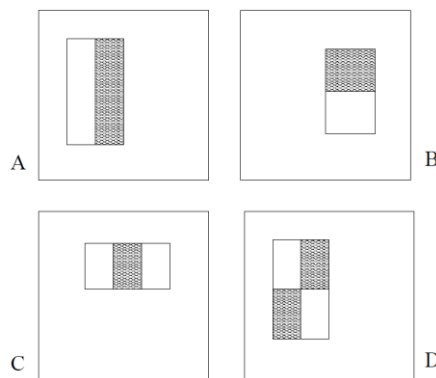
۲- آشکارسازی چهره

برای شناسایی ناحیه چهره از الگوریتم Viola-Jones در [38] استفاده شده است. در این روش، یک الگوریتم یادگیری ماشین جهت شناسایی اشیاء ارائه شده است. این الگوریتم سه مرحله است. ابتدا یک نمایش جدید از تصویر حاصل می‌شود که تصویر انتگرال نامیده شده است. این کار باعث می‌شود ویژگی‌های مورد استفاده برای الگوریتم، با سرعت بیشتری استخراج شوند. مرحله دوم شامل یک الگوریتم یادگیری مبتنی بر آدابوست^۱ می‌باشد که یک مجموعه اندک از ویژگی‌های بصری مهم را از مجموعه بزرگتر انتخاب می‌کند که باعث افزایش کارآمدی رده‌بند می‌شود. مرحله سوم استفاده از یک ترکیب از

1- Adaboost

چند کلاس‌بند به صورت آبشاری^۱ می‌باشد که باعث می‌شود نواحی پس زمینه تصویر که موجب افزایش حجم محاسبات می‌شوند، حذف گردند.

برای استخراج ویژگی‌ها از سه نوع ویژگی استفاده می‌شود. مقدار ویژگی مستطیلی دوبرخی، اختلاف بین مجموع مقادیر پیکسلی هر بخش از بخش دیگر می‌باشد. دو ناحیه در هر مستطیل، هم اندازه بوده و به صورت افقی و یا عمودی مجاور می‌باشند. در حالت کلی ویژگی‌های مستطیلی در شکل (۲-۳) نشان داده شده‌اند.



شکل (۲-۳): ویژگی‌های مستطیلی نمونه جهت استخراج چهره. مجموعه پیکسل‌های مستطیل‌های خاکستری از مستطیل‌های سفید کم می‌شود. A و B ویژگی‌های مستطیلی دوبرخی و C ویژگی‌های مستطیلی سه بخشی و D ویژگی‌های مستطیلی چهار بخشی را نشان می‌دهند.

ویژگی مستطیلی سه بخشی، مجموعه مقادیر پیکسلی دو مستطیل خارجی را از مجموع مقادیر مستطیل داخلی کم می‌کند. ویژگی‌های مستطیلی چهار بخشی نیز اختلاف بین جفت‌های قطری را محاسبه می‌کند. ابعاد پنجره‌های ویژگی، 24×24 می‌باشد.

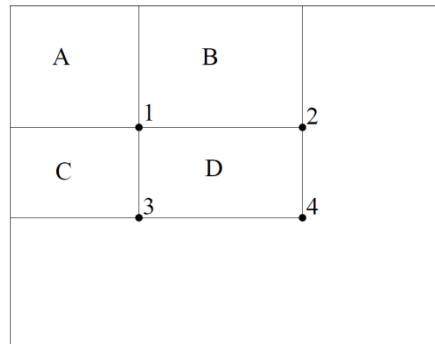
۲-۱- تصویر انتگرال

ویژگی‌های مستطیلی با اعمال روی یک نمایش میانی از تصویر می‌توانند بسیار سریع‌تر عمل کنند. این نمایش میانی را تصویر انتگرال می‌نامیم. تصویر انتگرال در موقعیت (x,y) شامل مجموع مقادیر پیکسلی سمت چپ و بالای پیکسل (x,y) است. این محاسبه در معادله (۳-۱) بیان شده است.

1- Cascade

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \quad (۱-۳)$$

که در آن $ii(x,y)$ تصویر انتگرال، و $i(x,y)$ تصویر اصلی است. با استفاده از تصویر انتگرال، هر مجموعه مستطیلی می‌تواند توسط چهار آرایه مرجع محاسبه شود و به شکل واضح‌تر، اختلاف بین دو ناحیه از ویژگی مستطیلی دو بخشی توسط ۶ آرایه‌ی از قبل محاسبه شده (در تصویر انتگرال) بدست می‌آید. این فرایند در شکل (۳-۳) نشان داده شده است.



شکل (۳-۳): محاسبه مجموع مقادیر پیکسلی یک ویژگی مستطیلی، با استفاده از چهار آرایه در تصویر انتگرال. مقدار تصویر انتگرال در موقعیت ۱ برابر با مجموع مقادیر پیکسلی ناحیه A می‌باشد. این مقدار در موقعیت ۲ برابر با $A+B$ می‌باشد و در موقعیت ۳ و ۴ به ترتیب برابر با $A+C$ و $A+B+C+D$ می‌باشد. مجموع مقادیر پیکسلی در ناحیه D می‌تواند اینگونه محاسبه شود: $4+1-(2+3)$

۲-۲ - طبقه‌بند آبخاری مبتنی بر آدابوست نامتقارن

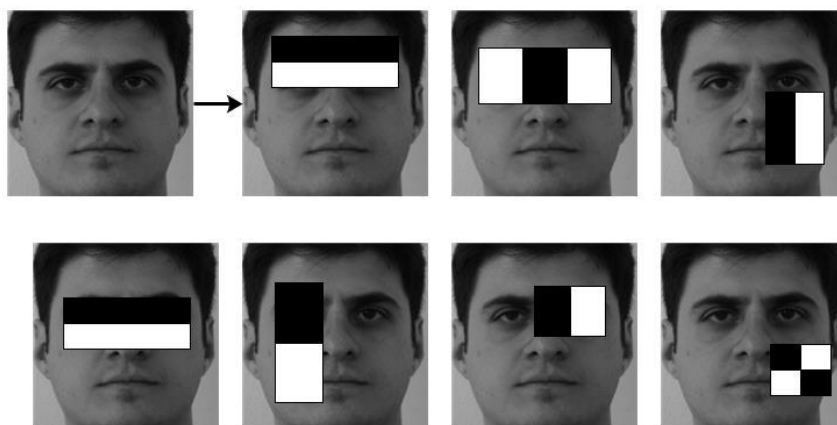
با داشتن مجموعه‌ای عظیم از ویژگی‌ها که باید بر روی تک تک تصاویر چهره و غیرچهره اعمال شوند، پروسه‌ای بسیار سنگین و زمان‌بر پیش رو می‌باشد. به صورت تجربی، سیستم نهایی توسط تعداد کمی از این ویژگی‌ها تشکیل می‌شود. این تعداد باید از بین چنین مجموعه بزرگی استخراج شوند. به منظور طبقه‌بندی از سیستم آدابوست نامتقارن استفاده شده است. این طبقه‌بند دارای دو ویژگی کلیدی می‌باشد:

- سرعت بالا در آموزش

- انتخاب ویژگی‌های برتر در حین آموزش

فرایند این الگوریتم در [38] به طور کامل شرح داده شده است. هدف از فرآیند آموزش، انتخاب مجموعه‌ای از بهترین طبقه‌بندهای ضعیف و تشکیل یک طبقه‌بند قوی^۱ است. مجموعه‌ای از طبقه‌بندهای قوی، سیستم نهایی را تشکیل می‌دهند.

ویژگی‌های انتخاب شده توسط آدابوست نامتقارن، معنی دار و به راحتی قابل تفسیر می‌باشند. برای نمونه شکل (۳-۴) چند مورد از ویژگی‌های انتخاب شده توسط آدابوست نامتقارن را نشان می‌دهد. ویژگی اول به اختلاف ناحیه چشم‌ها و پیشانی و ویژگی دوم به اختلاف چشم‌ها با ناحیه وسط دو چشم اشاره دارد.



شکل (۳-۴): برخی از ویژگی‌های منتخب توسط طبقه بند آدابوست نامتقارن

۳- استخراج ویژگی

بخش اصلی سیستم شناسایی چهره، استخراج ویژگی از تصویر چهره است. تمرکز اصلی الگوریتم پیشنهادی نیز معطوف به این قسمت است. استخراج ویژگی در این بخش، به صورت اعمال ترتیبی از

1- Strong Classifier

پنجره‌های هم مرکز با اندازه‌های متفاوت روی تصویر چهره انجام می‌شود. این ویژگی‌ها مبتنی بر تبدیل فوریه دو بعدی هستند. در مراحل متوالی، تصویر را از حاشیه به سمت مرکز تصویر پنجره‌گذاری می‌کنیم. نتیجه این فرایند در مرحله p پنجره‌گذاری، به شکل ریاضی در رابطه (۲-۳) نشان داده شده است.

$$h_{N \times N}^p(i, j) = I(p + i - 1, p + j - 1) \quad (2-3)$$

در رابطه‌ی (۲-۳)، I تصویر چهره‌ی ورودی با ابعاد $L \times L$ و h^p تصویر حاصل از پنجره‌گذاری در مرحله p است. در این رابطه همواره $2 < N < L$ بوده و در هر مرحله $N = L - 2(p - 1)$ می‌باشد.

میزان تغییر اندازه پنجره‌ها، متناسب با تعداد پیکسل‌های تعیین شده توسط کاربر است. با اعمال تبدیل فوریه دوبعدی روی تصویر حاصل از هر مرحله پنجره‌گذاری، اندازه‌ی مولفه‌های فرکانسی استخراج می‌شوند. این تبدیل در رابطه (۳-۳) نشان داده شده است.

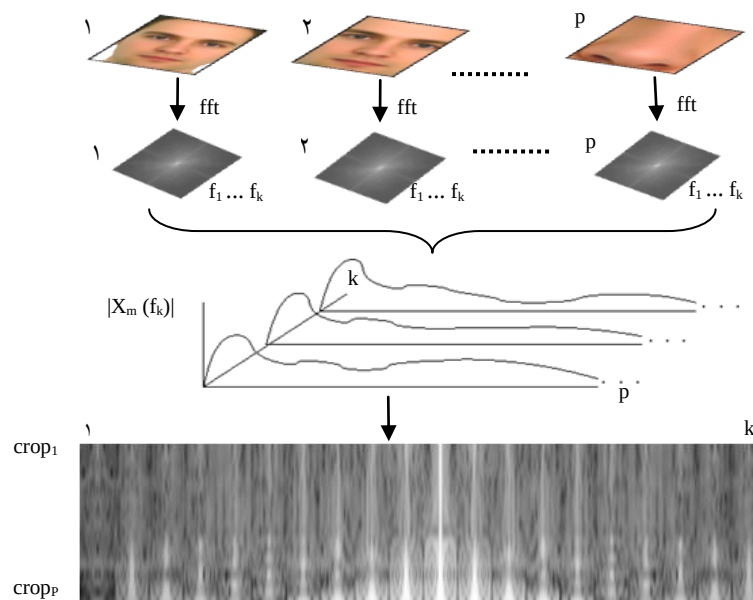
$$H(k, l) = \sum_{n=0}^{M-1} \sum_{m=0}^{M-1} e^{-i(kn+lm)} h(n, m) \quad (3-3)$$

در رابطه‌ی (۳-۳)، h تصویر ورودی و H اندازه تبدیل فوریه تصویر است. n و m معرف موقعیت پیکسل تصویر ورودی، k و l بیانگر موقعیت مولفه فرکانسی تصویر ورودی در صفحه تبدیل یافته‌ی فرکانسی هستند.

ویژگی‌های نهایی مبتنی بر تغییرات اندازه‌ی هر مولفه فرکانسی طی پنجره‌های متوالی، تعیین می‌شوند. ویژگی‌های نهایی جهت بهبود مقیاس و قابلیت مقایسه دقیق‌تر، به فضای دسیبل^۱ انتقال می‌یابند. منظور از مؤلفه‌های فرکانسی، پیکسل‌های تبدیل یافته تصویر در فضای فوریه دو بعدی تصویر است. این تغییرات به عنوان فضای ویژگی چهره‌ی هر شخص معرفی می‌شوند.

دلایل استفاده از ویژگی‌های فوریه، قدرت بالای استخراج خواص معنایی تصاویر، ثابت بودن نسبت به چرخش و سرعت قابل قبول آن است. در شناسایی چهره، سرعت مناسب و مقاومت نسبت به چرخش از ویژگی‌های بسیار مهم بشمار می‌رود.

در سیگنال‌های یک بعدی، تبدیل فوریه زمان کوتاه^۱ جهت استخراج معنایی در کاربردهای فراوانی استفاده می‌شود که مشابه فرآیند فوق است. از جمله این کاربردها، می‌توان به درک معنایی گفتار اشاره کرد [41]. در این تحقیق، ویژگی‌های تغییرات فرکانسی در راستای مکان، در شناسایی چهره استفاده شده‌اند که بیانگر مفاهیمی بسیار سودمند در تصاویر هستند. در شکل (۳-۵) فرایند استخراج ویژگی نشان داده شده و ماتریس ویژگی یک تصویر نمونه استخراج شده است.



شکل (۳-۵): فرآیند استخراج ویژگی از هر تصویر چهره.
 p : تعداد پنجره‌گذاری‌های تصویر. k : تعداد مولفه‌های فرکانسی انتخاب شده. X_m : تبدیل یافته تصویر در حوزه فرکانس.

روش پیشنهادی شامل پارامترهایی می‌باشد که در زیر به تعریف آن‌ها پرداخته‌ایم:

(۱) $L \times L$: ابعاد تصویر ورودی چهره است.

(۲) μ : در هر مرحله، اندازه پنجره به اندازه‌ی تعداد پیکسلی مشخص، کاهش پیدا می‌کند و این فرایند به سمت مرکز تصویر ادامه پیدا می‌کند. این پارامتر تعداد پیکسل مذکور را مشخص می‌کند.

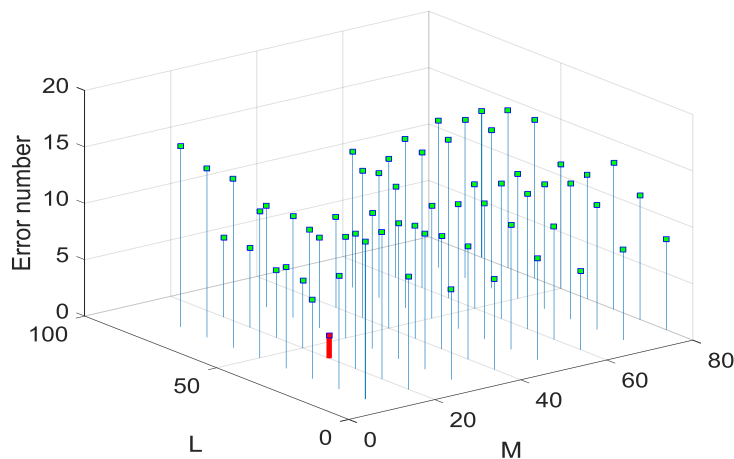
(۳) $M \times M$: ابعاد تبدیل فوری دوبعدی مورد استفاده در تمامی مراحل الگوریتم یا به عبارتی M^2 تعداد مولفه‌های فرکانسی تعریف شده است.

تعداد دفعات پنجره‌گذاری روی تصویر از حاشیه به سمت مرکز تصویر، با استفاده از پارامترهای فوق طبق رابطه (۳-۴) محاسبه می‌شود.

$$P = \frac{\frac{L}{2} - \mu}{\mu} \quad (4-3)$$

با افزایش پارامتر μ ، سرعت شناسایی چهره افزایش می‌یابد و با کاهش این پارامتر دقت الگوریتم افزون می‌شود. در دقیق‌ترین حالت، مقدار μ برابر با ۱ است.

تغییر دو پارامتر M و L نیز، نرخ شناسایی و سرعت بازیابی تصویر را دستخوش تغییر می‌کنند. افزایش این دو پارامتر، فضای ویژگی را افزون و سرعت بازیابی را کاهش می‌دهند؛ لذا باید مقداری بهینه برای این دو پارامتر انتخاب شود تا به ازای آن‌ها نرخ بازیابی بیشینه شود. مقدار بهینه این دو پارامتر طی آزمایش‌های متوالی در پایگاه داده ۲۰۰ تایی به عنوان داده‌های اعتبارسنجی محاسبه شد که نتایج برخی از این آزمایش‌ها در محدوده نقاط مشخص در شکل (۳-۸) نمایش داده شده است. چنانچه در این شکل مشاهده می‌شود، مقادیر بهینه‌ی یافت شده برای این دو پارامتر، $L=40$ و $M=20$ است.



شکل (۳-۶): مقادیر بهینه دو پارامتر M و L در روش ارائه شده

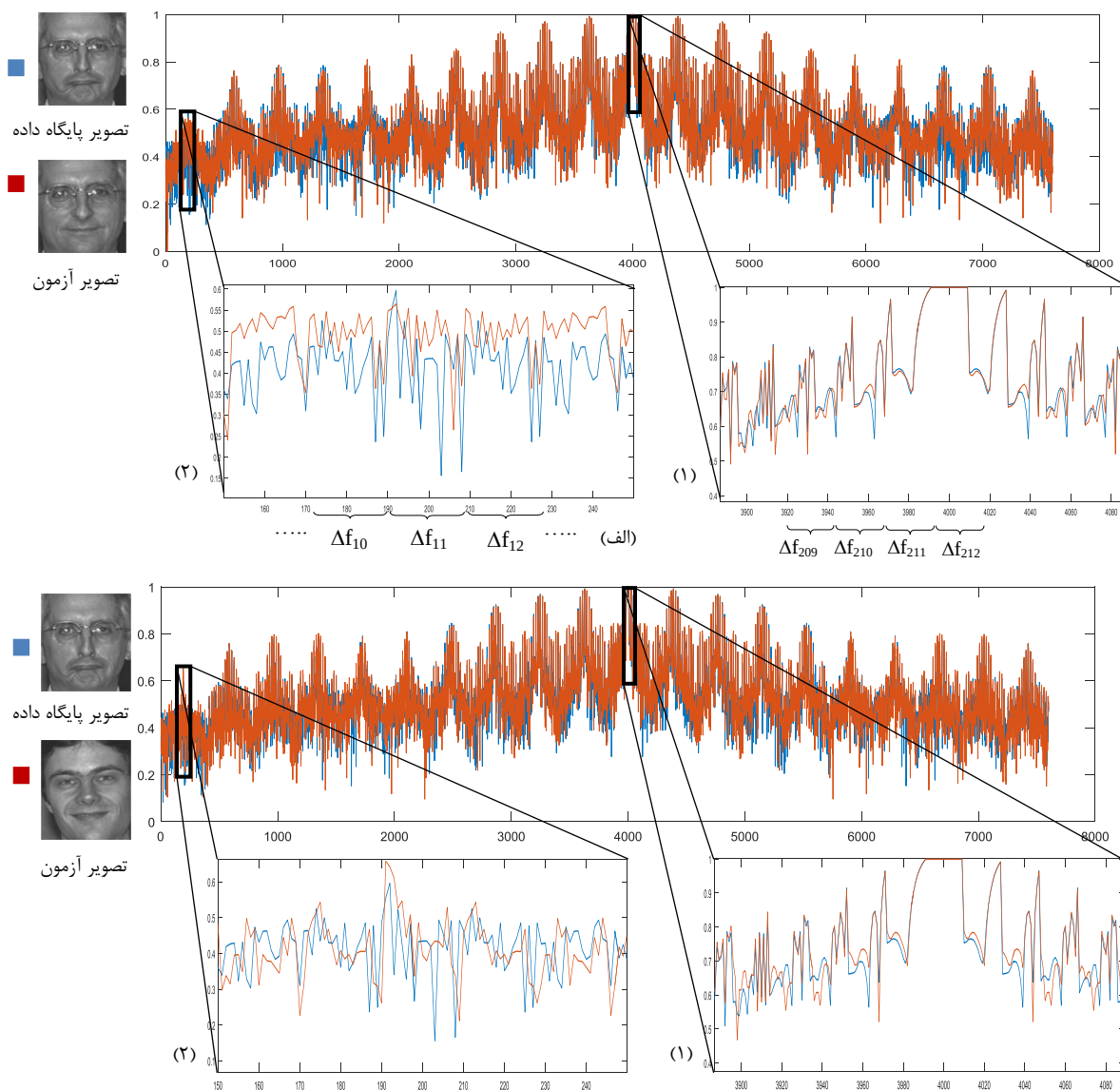
چنانچه مشخص است، با افزایش یا کاهش بیش از حد این پارامترها، میزان خطا به مراتب افزایش می-یابد. لذا محدوده‌ی کمتر از ۸۰ و بیشتر از ۱۰ نمایش داده نشده‌اند. در ادامه‌ی پایان‌نامه، از این مقادیر بهینه در آزمایش‌ها استفاده شده است.

۴- کاهش ویژگی‌ها

پس از استخراج ویژگی، نیاز به انتخاب ویژگی‌های سودمند و صرف نظر کردن از ویژگی‌های غیرمفید است. در روش پیشنهادی نیز جهت کاهش ابعاد ویژگی، افزایش سرعت و همینطور افزایش دقت بازیابی، فضای ویژگی کاهش داده شده است. به همین منظور مولفه‌های فرکانسی غیر سودمند استخراج شده از تصویر حذف می‌شوند تا علاوه بر افزایش دقت بازیابی، سرعت بازیابی افزایش یابد. برای این منظور، بررسی‌های فراوانی بر روی فضای ویژگی تصویر یک شخص در حالت‌های مختلف و فضای ویژگی دو شخص متفاوت انجام شده است. در این بررسی‌ها مشخص شد که در روش پیشنهادی، فرکانس‌های پایین اهمیت بیشتری نسبت به فرکانس‌های بالا دارند. فرکانس‌های بالا نشان‌دهنده‌ی بافت تصویر می‌باشند که

در انسان‌های مختلف به هم نزدیک هستند. اما فرکانس‌های پایین وابسته به موقعیت و اندازه‌ی اجزاء تشکیل دهنده‌ی چهره نظیر بینی، لب و چشم هستند که در افراد مختلف متفاوتند.

در شکل (۷-۳) ماتریس ویژگی‌های استخراج شده برای افراد نمونه، به صورت یک بردار واحد به تصویر

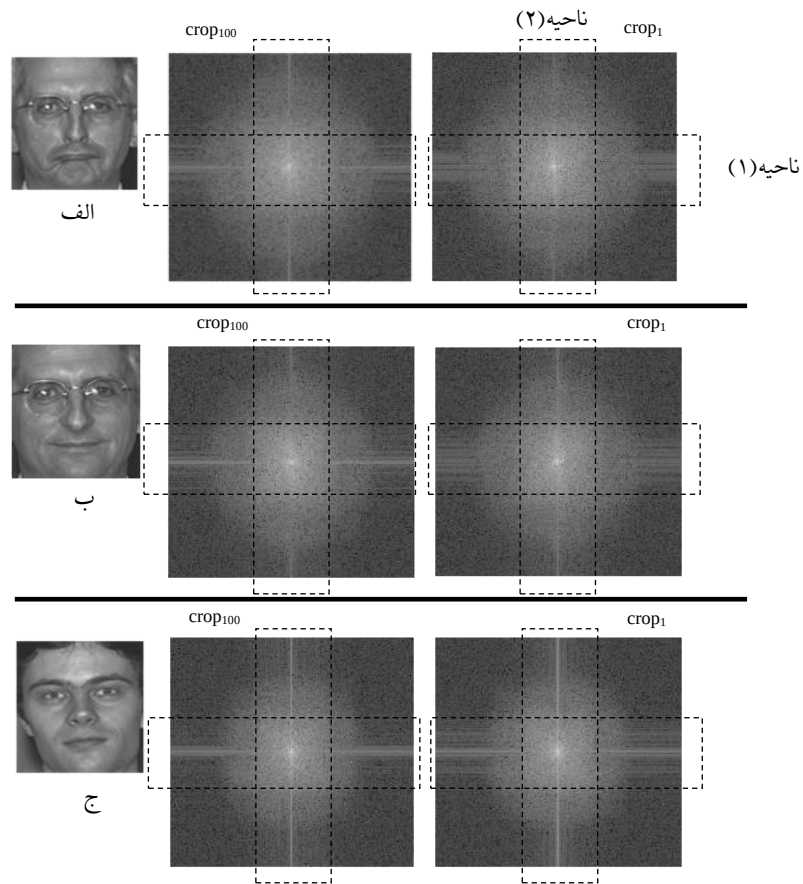


شکل (۷-۳): نمایش فضای ویژگی تصاویر چهره نمونه به صورت برداری در روش پیشنهادی (ب) پیمایش ستونی ماتریس ویژگی‌های استخراج شده. (الف) مقایسه دو تصویر هم‌منوع. (ب) مقایسه دو تصویر غیرهم‌منوع. Δf_i نشان دهنده تغییرات مولفه فرکانسی i ام در برش‌های متوالی تصویر می‌باشد.

کشیده شده است. همانگونه که مشاهده می‌شود، تغییرات مولفه‌های فرکانس پایین در دو تصویر از یک شخص، مشابه (شکل (۷-۳)-الف-۱) و در دو تصویر از دو شخص متفاوت مغایر هم هستند (شکل (۷-۳)-ب-۱). اما تغییرات مولفه‌های فرکانس بالا در دو تصویر از یک شخص، مغایر هم هستند (شکل (۷-۳)-الف-۲). به عبارت دیگر در روش پیشنهادی، فرکانس‌های پایین موجب تشخیص صحیح و اکثر فرکانس‌های بالا موجب تشخیص ناصحیح می‌شوند. لذا برای افزایش نرخ بازیابی در روش پیشنهادی باید ویژگی‌های فرکانس پایین انتخاب شوند. انتخاب این ویژگی‌ها با اعمال فیلتری بر روی تمام فضای فرکانسی ویژگی‌ها، میسر می‌شود.

فیلترهای رایج حوزه‌ی فرکانس، تصویر را تغییر می‌دهند و ویژگی‌های مقایسه‌ای آن حفظ نمی‌شود. اما در این تحقیق، فیلتر باید به گونه‌ای طراحی شود که ویژگی‌های منحصر به فرد تصویر ورودی، جهت مقایسه، حفظ شوند. از این‌رو، فیلترهای حوزه فرکانس متداول، در این کاربرد کارا نیستند. برای منحصر به فرد کردن تصاویر طی پنجره‌گذاری‌های متوالی، باید نواحی را که دارای تغییرات فرکانسی زیادی هستند، از تصاویر استخراج و نواحی دارای تغییرات فرکانسی اندک، فیلتر شوند. طی آزمایش‌های فراوان، نواحی حساس و تغییرپذیر در فضای فوریه تصویر چهره مشخص شدند. این نواحی در شکل (۸-۳) نشان داده شده است. سایر قسمت‌های تصویر دارای اطلاعات غیر ضروری در شناسایی چهره هستند و باعث شناسایی کاذب می‌شوند. بنابراین در گام اول نواحی غیر ضروری فیلتر می‌شوند و در گام دوم فیلتری مناسب نیز برای نواحی ضروری طراحی می‌شود.

در شکل (۸-۳) فضای فوریه سه تصویر در نواحی ضروری، یعنی دو ناحیه (۱) و (۲) در فضای فوریه، مورد مقایسه قرار گرفته‌اند. چنانچه در این شکل مشاهده می‌شود، تغییرات مولفه‌های ناحیه (۱) طی پنجره‌گذاری ۱ تا پنجره‌گذاری ۱۰۰ در هر سه تصویر ناهمگون و غیر قابل تفکیک هستند. در حالیکه تغییرات مولفه‌های ناحیه (۲) برای دو تصویر الف و ب از یک شخص، از تصویر سوم قابل تفکیک هستند.



شکل (۳-۸): مولفه‌های فرکانسی در دو پنجره مشخص از سه تصویر نمونه (تغییرات اصلی در راستای پنجره‌ها، در دو ناحیه دسته بندی شده‌اند). الف: تصویر پایگاه داده. ب: تصویر آزمون هم‌نوع. ج: تصویر آزمون غیر هم‌نوع

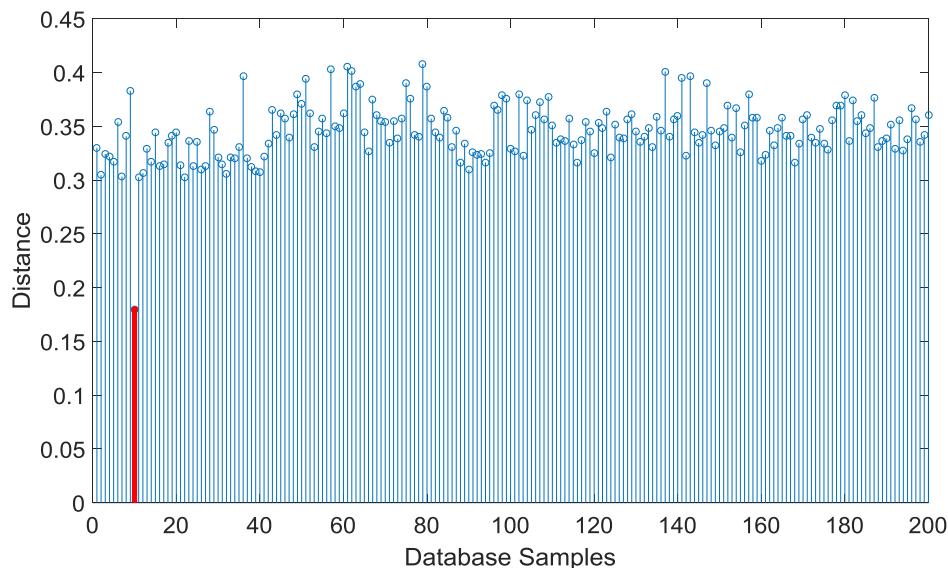
همانگونه که اشاره شد، محدوده‌ای از فرکانس حائز اهمیت است که تغییرات آن طی پنجره‌گذاری‌های متوالی بین دو تصویر از یک شخص و دو شخص متفاوت قابل تفکیک باشد. لذا اهمیت ناحیه (۲) به مراتب بیشتر از ناحیه (۱) است. بنابراین در فیلتر نهایی، اهمیت ناحیه (۲) نیز باید لحاظ شود. ماسک طراحی شده در قالب شکل (۳-۹) خواص مذکور را دارد و در این پایان‌نامه برای طراحی فیلتر، از آن

می‌شود. این فضاها و ویژگی، k مولفه فرکانسی دارند و هر مولفه فرکانسی، P مقدار طی پنجره‌های متوالی از تصویر دارد. در واقع هر بردار ویژگی، $k \times P$ ویژگی دارد. در ادامه، فاصله اقلیدسی بردار تغییرات هر مولفه فرکانسی در I_1 ، با بردار تغییرات مولفه فرکانسی متناظر خود در I_2 محاسبه می‌شود. در واقع هر بردار ویژگی، به k دسته P تایی تقسیم می‌شود و فاصله هر دسته، از دسته متناظر در تصویر دوم محاسبه می‌شود. با داشتن k دسته، در نهایت k مقدار فاصله حاصل می‌شود. با میانگین‌گیری این k مقدار، فاصله دو تصویر محاسبه می‌شود. به این ترتیب فاصله تصویر آزمون با تمام تصاویر موجود در پایگاه داده محاسبه شده و نزدیک‌ترین فاصله به عنوان محتمل‌ترین جواب در نظر گرفته می‌شود. در رابطه‌ی (۳-۶) این فاصله نمایش داده شده است.

$$distance(I_1, I_2) = \frac{\sum_{i=1}^k \sqrt{\sum_{j=1}^p (I_{1ij} - I_{2ij})^2}}{k} \quad (۳-۶)$$

در رابطه‌ی بالا، k تعداد مولفه‌های فرکانسی، p تعداد پنجره‌گذاری تصویر و I_1 و I_2 تبدیل یافته دو تصویر به فضای ویژگی در حوزه فرکانس هستند.

آزمایش‌ها نشان می‌دهد که تصویر آزمون، از تصویر هم برچسب خود در پایگاه داده، فاصله اندکی دارد و شناسایی چهره با دقت بالایی انجام می‌شود. شکل (۳-۱۲) نمونه‌ای از فاصله یک تصویر آزمون با تمامی تصاویر موجود در پایگاه داده را نمایش می‌دهد. در این مثال حجم پایگاه داده ۲۰۰ تصویر چهره متمایز در نظر گرفته شده است. در این نمونه، فاصله‌ی تصویر آزمون از تصویر هم‌برچسب خود اندک است و با دومین کم‌ترین فاصله، اختلاف قابل توجهی دارد. در واقع، فاصله تصویر آزمون تا تصویر هم‌برچسب خود در پایگاه داده، به طرز قابل توجهی قابلیت جداسازی از سایر فاصله‌ها را دارد.



شکل (۳-۱۰): مقادیر فاصله تصویر آزمون از ۲۰۰ تصویر پایگاه داده. تشخیص صحیح دهمین داده به عنوان نزدیک‌ترین تصویر چهره به تصویر آزمون.

در این تحقیق، چندین معیار فاصله دیگر نیز بررسی شده‌اند. نتایج بررسی‌ها نشان داده است که جواب بهینه، متعلق به معیار فاصله اقلیدسی است. کارایی دو معیار فاصله‌ی دیگر در جدول (۳-۲) نسبت به فاصله اقلیدسی نشان داده شده است. این دو معیار بهترین عملکرد را در بین سایر معیارهای فاصله داشته‌اند. مقایسه‌ی این معیارها روی پایگاه داده با تعداد ۲۰۰ چهره متمایز انجام گرفته است. تعاریف هر معیار در معادلات (۳-۷) و (۳-۸) نشان داده شده است.

$$conberra(p, q) = \sum_{i=1}^n \frac{|p_i - q_i|}{|p_i| + |q_i|} \quad (۳-۷)$$

$$Jensen(p, q) = \sum_{i=1}^n \frac{\log(p_i)}{2} + \frac{\log(q_i)}{2} - \log\left(\frac{p_i + q_i}{2}\right) \quad (۳-۸)$$

جدول (۳-۲): مقایسه معیار فاصله اقلیدسی با دو معیار دیگر در روش پیشنهادی

میزان افزایش زمان بازیابی نسبت به فاصله اقلیدسی	میزان کاهش نرخ بازیابی نسبت به فاصله اقلیدسی	معیار فاصله
۱۳۲,۷۵ms	٪۱	Canberra
۸۱۲,۵ms	٪۵	Jensen

۶- جمع‌بندی

در این فصل الگوریتم پیشنهادی جهت شناسایی چهره ارائه گردید. ابتدا با استفاده از الگوریتم Viola-Jones ناحیه چهره استخراج شده و سپس با الگوریتم پیشنهادی، ویژگی‌های مکان - فرکانسی از چهره استخراج شده است. در این روش، جهت کاهش فضای ویژگی و در نتیجه افزایش سرعت بازیابی و همینطور افزایش نرخ بازیابی، از فیلتری در حوزه فرکانس استفاده گردیده است. با این فیلتر ویژگی‌های پر تاثیر در شناسایی چهره استخراج شده و نرخ شناسایی افزایش پیدا کرده است. با استفاده از فضای ویژگی استخراج شده از چهره ها، با معیار فاصله مناسب نزدیک ترین تصویر به تصویر آزمون به عنوان جواب نهایی اعلام می‌شود.

در روش پیشنهادی از پارامترهای تنظیم کننده ای استفاده شده است که مقادیر بهینه این پارامترها با استفاده از مجموعه داده‌های اعتبار سنجی بدست آمده است.

فصل چهارم

ارزیابی نتایج

۱ - مقدمه

در این فصل به بررسی نتایج آزمایشی روش پیشنهادی پرداخته شده است. روش پیشنهادی در چهار بخش مورد ارزیابی قرار گرفته است. (۱) آشکارسازی چهره (۲) نتایج کمی روش شناسایی چهره (۳) نتایج کیفی روش شناسایی چهره (۴) نتایج روش انتخاب ویژگی. با ارائه نتایج روش پیشنهادی و مقایسه نتایج با سایر روش‌ها، عملکرد سیستم پیشنهادی در مقابل سایر سیستم‌های شناسایی چهره نشان داده شده است. جهت مقایسه نتایج روش پیشنهادی با روش‌های پیشین، سعی شده است بهترین حالت که قابلیت مقایسه را داشته باشد انتخاب شود. برای این منظور، حجم پایگاه داده روش پیشنهادی را جهت مقایسه با روش‌های دیگر، مشابه با سایر روش‌ها لحاظ کرده‌ایم. سپس عملکرد سیستم پیشنهادی را در حجم‌های بیشتر مورد بررسی قرار داده‌ایم.

۲ - پایگاه داده

جهت ارزیابی الگوریتم پیشنهادی، از پایگاه داده FERET استفاده شده است. این پایگاه داده شامل ۹۹۰ فرد متمایز و چندین چهره برای هر فرد است. در این پایگاه‌نامه، از زیرمجموعه‌های fa و fb این پایگاه داده استفاده شده است. این دو زیرمجموعه شامل تصاویر چهره از روبرو هستند. fa شامل تصاویر ساده و fb تصاویر متناظر با fa و همراه با ایجاد تغییرات حالت چهره است. در الگوریتم پیشنهادی، از زیرمجموعه fa به عنوان تصاویر پایگاه داده و از زیرمجموعه fb به عنوان تصاویر آزمون استفاده شده است. از هر زیر مجموعه ۲۰۰ تصویر برای داده‌های اعتبارسنجی (جهت تشخیص پارامترهای بهینه) و از مابقی تصاویر به عنوان داده‌های آزمون استفاده شده است.

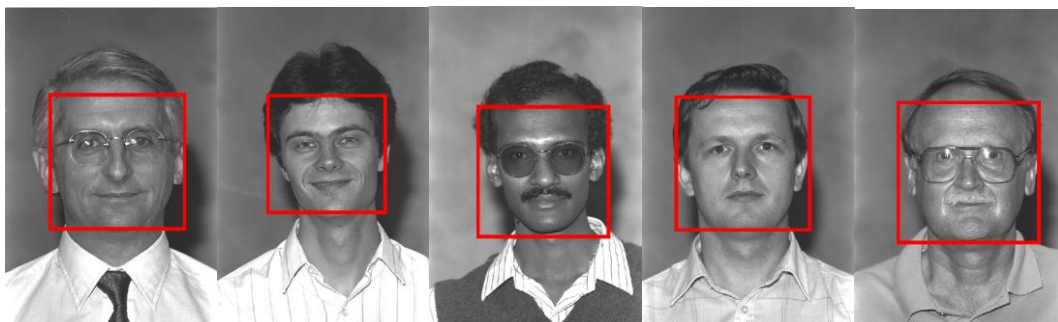
۳- آزمایش‌ها

۳-۱- نتایج آشکارسازی چهره

چنانچه در فصل پیش اشاره شد، ناحیه چهره با استفاده از ویژگی‌های مستطیلی استخراج می‌شود.

نتایج حاصل از الگوریتم Viola-Jonse روی چند نمونه از تصاویر پایگاه داده FERET در شکل (۴-۱)

نشان داده شده است.



شکل (۴-۱): عملکرد سیستم آشکارساز چهره روی چند چهره نمونه

۳-۲- نتایج کمی شناسایی چهره

در این بخش، روش پیشنهادی در دو مرحله ارزیابی می‌شود. ابتدا نتایج حاصل از الگوریتم پیشنهادی،

با روش‌های اخیر مقایسه می‌شود. سپس توانایی الگوریتم پیشنهادی در ارزیابی چهره، بر اساس تغییرات

حجم پایگاه داده ارزیابی می‌شود.

هدف اصلی الگوریتم ارائه شده، شناسایی چهره در حجم زیاد پایگاه داده و وجود تنها یک تصویر چهره

به ازای هر شخص است. از بین روش‌های اخیر در شناسایی چهره، تنها دو روش از حجم نسبتاً زیاد پایگاه

داده استفاده کردند که نتایج قابل قبول‌تری نسبت به سایرین حاصل کرده اند. اولین روش، الگوریتم

SIFT در [17] است. این روش، از پایگاه داده FERET با ۲۰۰ چهره انتخابی استفاده و نتایج خود را ثبت

کرده است. الگوریتم دوم، الگوریتم LBP در [13] است. در این الگوریتم، ۱۶۰ چهره از پایگاه داده

FERET انتخاب شده است. در این روش، از هر چهره ۴ تصویر در اختیار بوده است که ۱ تصویر در پایگاه داده جهت بازیابی و ۳ تصویر جهت آزمون مورد استفاده قرار گرفته‌اند.

همچنین روش پیشنهادی با روش‌های دیگر که دارای حجم پایگاه داده کمتر هستند مقایسه شده است. این روش‌ها در فصل مربوط به روش‌های پیشین، مورد تجزیه و تحلیل قرار گرفتند. در جدول (۴-۱) نتایج ارزیابی الگوریتم ارائه شده با روش‌های اخیر نشان داده شده است.

جدول (۴-۱): مقایسه کارایی روش پیشنهادی با الگوریتم‌های مختلف

الگوریتم	پایگاه داده	تعداد افراد پایگاه داده	تعداد تصاویر به ازای هر فرد	نرخ بازیابی
DCP[26]	AR	۱۲۶	۲۶	۹۴/۶۴
Hough[21]	FERET	۳۵	۲۰	٪۸۰/۸۱
Gabor[22]	FERET	۳۵	۲۰	٪۸۳/۳
SIFT[17]	FERET	۲۰۰	۱	٪۹۷
LBP[13]	FERET	۱۶۰	۴	٪۹۷
روش پیشنهادی	FERET	۲۰۰	۱	٪۹۹

در بین روش‌های مورد مقایسه، الگوریتم SIFT عملکرد مناسب‌تری از خود نشان داده است. علت این امر استفاده از ویژگی‌های چرخش ثابت و مقیاس ثابت و همین‌طور استفاده از کلاس‌بند مناسب در مرحله شباهت سنجی است. هر یک از الگوریتم‌های فوق، از روش‌های مبتنی بر ویژگی و یا روش‌های کل‌نگر استفاده کرده‌اند. روش پیشنهادی با استخراج ویژگی‌های ترکیبی توانست عملکرد بهتری را از خود نشان دهد.

مزیت اصلی الگوریتم پیشنهادی، استفاده از تمام ناحیه چهره جهت استخراج ویژگی است. این در حالی است که به دلیل افزایش ابعاد ویژگی‌ها، سرعت بازیابی کاهش چشم‌گیری پیدا نمی‌کند. اکثر روش‌های شناسایی چهره، به دلیل کاهش زمان بازیابی، از الگوریتم‌های مدل‌گرا استفاده می‌کنند که در نتیجه، به دلیل مد نظر قرار دادن نواحی خاص از چهره، نرخ شناسایی کاهش پیدا می‌کند. برخی روش‌ها نیز

مجبور به استفاده از چندین تصویر به ازای هر شخص می‌شوند تا بتوانند مشکل کمبود ویژگی‌ها را حل نمایند که خود باعث افزایش حجم پایگاه داده و کاهش کاربرد می‌شود. در الگوریتم ارائه شده، به ازای هر شخص تنها یک تصویر چهره در پایگاه داده وجود دارد و سرعت بازیابی آن، بسیار بالا است.

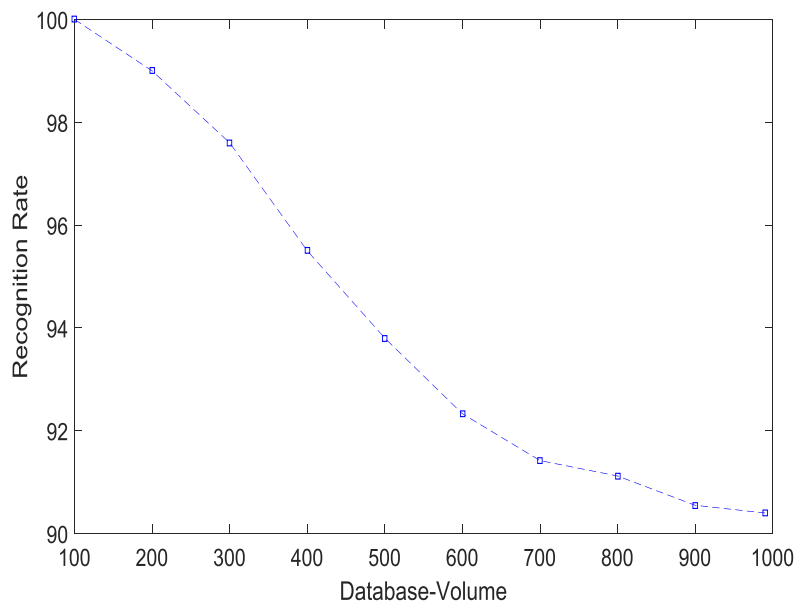
در جدول (۲-۴) نتایج الگوریتم پیشنهادی در حجم‌های متفاوت پایگاه داده، آورده شده است.

جدول (۲-۴): نرخ و سرعت بازیابی الگوریتم ارائه شده در حجم‌های متفاوت پایگاه داده

نرخ بازیابی	سرعت بازیابی (ms)	حجم پایگاه داده	مشخصات
٪۱۰۰	۹۹/۵	۱۰۰	CPU: Intel Core i7, 2.0 GHz Memory: 4 GB
٪۹۹	۱۸۶	۲۰۰	
٪۹۷/۶۶	۲۵۶	۳۰۰	
٪۹۵/۵	۳۴۳	۴۰۰	
٪۹۳/۸	۴۳۸	۵۰۰	
٪۹۲/۳۳	۵۲۲	۶۰۰	
٪۹۱/۴۳	۶۳۶	۷۰۰	
٪۹۱/۱۲	۷۴۰	۸۰۰	
٪۹۰/۵۵	۷۹۰	۹۰۰	
٪۹۰/۴	۸۶۰	۹۹۰	

در این بررسی، هر بار به صورت تصادفی تعداد مختلفی تصویر از مجموعه داده‌های آزمون پایگاه داده انتخاب می‌شوند. روش‌های اخیر، تعداد کلاس‌ها را کمتر از ۲۰۰ لحاظ کرده‌اند. با توجه به آزمایش‌های انجام شده با افزایش تعداد کلاس‌ها، بازدهی این روش‌ها به شدت کاهش می‌یابد.

با توجه به تاکید این پایان‌نامه بر حجم نمونه‌ها، چنانچه در شکل (۲-۴) نشان داده شده است، شتاب کاهش نرخ بازیابی تصویر چهره با افزایش تعداد تصاویر پایگاه داده، بسیار کم است که کارایی مناسب روش پیشنهادی در حجم زیاد نمونه‌ها را نشان می‌دهد.



شکل (۴-۲): شتاب کاهش نرخ بازیابی با افزایش حجم پایگاه داده

۳-۳- نتایج کیفی شناسایی چهره

در این بخش نمونه‌ای از تصاویر پایگاه داده و عملکرد سیستم شناسایی چهره ارائه شده در این پایان‌نامه در بازیابی چهره نشان داده شده است. شکل (۴-۳) نمونه‌ای از تصاویر پایگاه داده FERET را نشان می‌دهد. تصاویر انتخاب شده در این شکل، از مجموعه تصاویر آزمون پایگاه داده انتخاب شده‌اند. یکی از ویژگی‌های پایگاه داده FERET وجود چهره‌هایی از ملیت‌های متفاوت و سنین مختلف می‌باشد. این ویژگی به واقعی بودن این پایگاه داده کمک کرده است و لذا نتایج حاصل، از نظر تفاوت سنی و ملیت‌های متفاوت و همین‌طور تفاوت حالات چهره، به دنیای واقعی نزدیک است.



شکل (۴-۳): نمونه‌ای از تصاویر پایگاه داده FERET

در روش پیشنهادی از تغییرات حالات چهره به عنوان تفاوت بین تصاویر پایگاه داده و تصاویر آزمون ورودی استفاده شده است. در شکل (۴-۴) چند نمونه از تشخیص‌های صحیح الگوریتم پیشنهادی در شناسایی چهره نشان داده شده است.



(پایگاه داده) (آزمون) (پایگاه داده) (آزمون) (پایگاه داده) (آزمون)

شکل (۴-۴): چند نمونه از تشخیص صحیح الگوریتم پیشنهادی

تشخیص‌های اشتباه نیز در پایگاه داده با حجم ۲۰۰ چهره متمایز در شکل (۴-۵) نشان داده شده است.

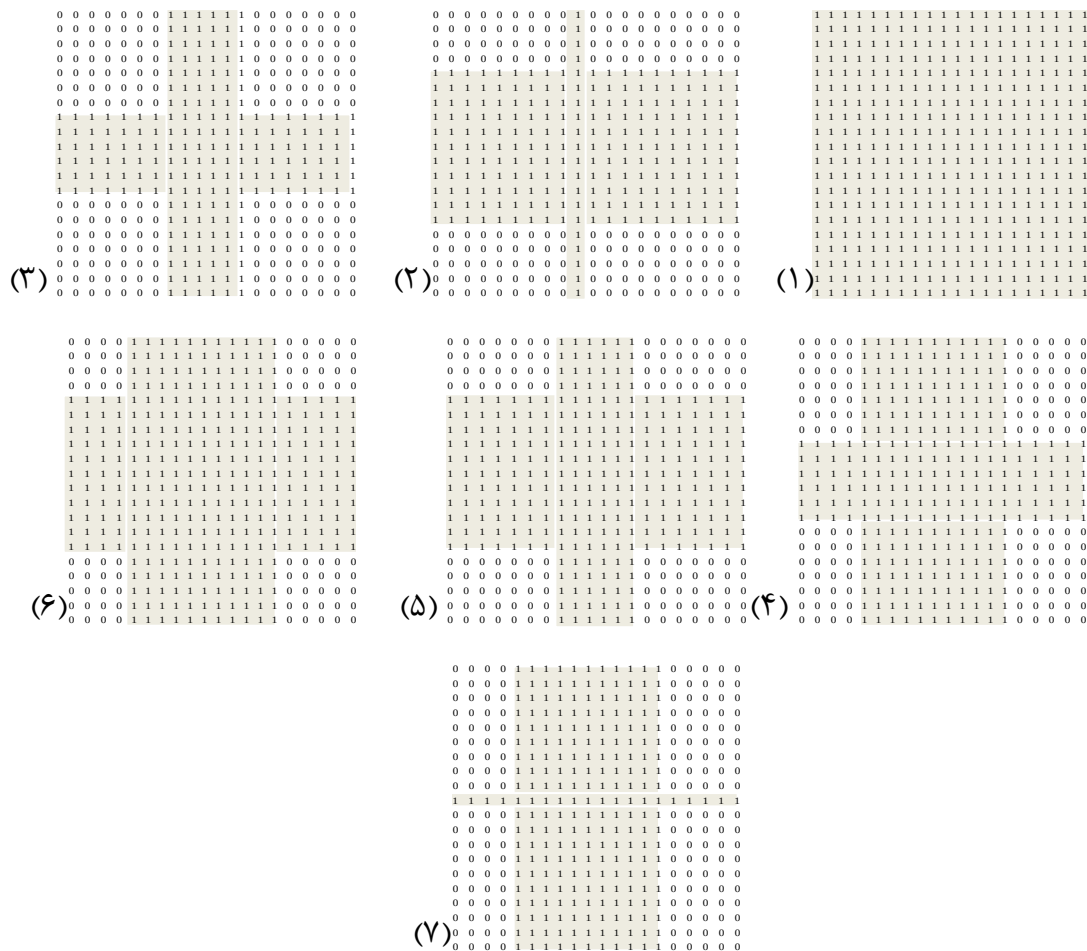


شکل (۴-۵): نمونه‌ای از شناسایی اشتباه در روش پیشنهادی
در تصویر فوق دو اتفاق به چشم می‌خورد. تصویر سمت راست، به دلیل وجود چهره‌ی بسیار مشابه در پایگاه داده، و از طرفی وجود تغییر حالت چهره زیاد بین دو تصویر از یک شخص، سیستم دچار شناسایی اشتباه شده است.

در تصویر سمت چپ، ناحیه چهره تصویر ورودی به شکل نامناسب استخراج شده است. لذا به دلیل از بین رفتن اطلاعاتی از چهره که در تصویر متناظر خود در پایگاه داده بسیار اهمیت دارد، سیستم دچار شناسایی اشتباه شده و نتوانسته از اطلاعات چهره استفاده مفید نماید.

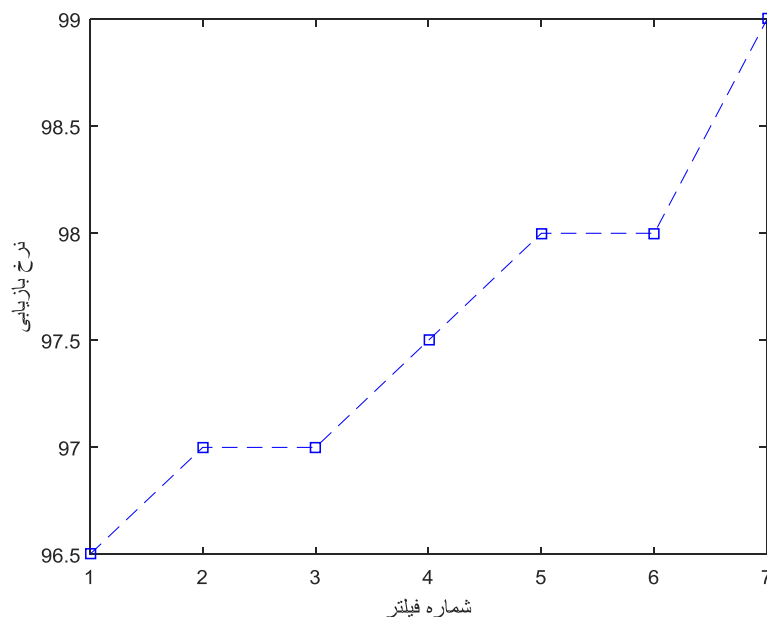
۴-۳- نتایج انتخاب ویژگی

جهت انتخاب ویژگی‌های جداساز مناسب، از فیلتری در حوزه فرکانس استفاده شده است. این فیلتر باید فرکانس‌های پایین را به شکلی انتخاب نماید که نرخ بازیابی بهینه شود. دلیل انتخاب فرکانس‌های پایین نیز این است که فرکانس‌های پایین وابسته به موقعیت و اندازه‌ی اجزاء تشکیل دهنده‌ی چهره نظیر بینی، لب و چشم هستند که در افراد مختلف متفاوتند. فیلتر در حالات مختلف جهت استخراج فرکانس-های پایین روی داده‌های اعتبارسنجی اعمال گردیده است. چند نمونه از حالات مختلف این فیلتر در شکل (۴-۶) نشان داده شده است.



شکل (۴-۶): چند نمونه فیلتر حوزه فرکانس در روش پیشنهادی. شماره (۱) بدون اعمال فیلتر می‌باشد.

با اعمال پنجره‌های فیلتر متفاوت در حوزه فرکانس تصاویر، نتایج متفاوتی دریافت گردید که در شکل (۷-۴) نشان داده شده است.



شکل (۷-۴): کارایی فیلترهای شکل (۶-۴) در روش پیشنهادی

در این شکل نتایج اعمال فیلترهای مختلف شکل (۶-۴) روی فضای ویژگی داده‌های اعتبارسنجی در روش پیشنهادی نشان داده شده است. چنانچه مشاهده می‌شود، فیلتر شماره ۷ بهترین نرخ بازیابی را از خود نشان داده است. چنانچه در فصل گذشته به دلایل این امر پرداختیم، دلیل این است که تغییرات مولفه‌های فرکانس پایین در محدوده این فیلتر، در پنجره گذاری‌های متوالی روی دو تصویر از یک شخص نسبت به سایر فیلترها مشابه تر می‌باشد.

۴- جمع‌بندی

در این فصل به بررسی نتایج آزمایشی روش پیشنهادی پرداخته شد. تحقیقات متوالی نشان داد که استفاده از الگوریتم Viola-Jonse جهت آشکارسازی چهره نسبت به سایر الگوریتم‌ها در کاربرد پیش رو

مناسبتر بوده و نتایج آن نشان داده شد. در بخش شناسایی چهره، نتایج کمی و کیفی نشان دهنده این موضوع هستند که الگوریتم پیشنهادی شناسایی را با نرخ بسیار مطلوبی انجام می‌دهد. شناسایی‌های اشتباه در دو حالت رخ می‌دهند: (۱) شباهت بیش از حد دو چهره متفاوت به یکدیگر. (۲) عدم توانایی در استخراج ویژگی‌ها به دلیل وجود انسدادهایی بر چهره ورودی و یا پایگاه داده. شتاب اندک کاهش نرخ بازیابی با افزایش حجم پایگاه داده، نشان دهنده کارآمدی روش پیشنهادی در حجم زیاد پایگاه داده است. همچنین در بخش انتخاب ویژگی‌ها، نحوه انتخاب بهترین فیلتر جهت حذف ویژگی‌های ناکارآمد شرح داده شد و بهترین فیلتر بدست آمد.

در حالت کلی در این فصل نشان داده شد که الگوریتم پیشنهادی نسبت به سایر الگوریتم‌های شناسایی چهره که روی پایگاه داده FERET پیاده سازی شده اند، بهتر عمل می‌کند.

فصل پنجم

نتیجه گیری و پیشنهادها

۱- نتیجه گیری

در این پایان نامه روشی جهت شناسایی چهره ارائه گردید. با بررسی تکنیک‌های شناسایی چهره، ویژگی‌های روش پیشنهادی مورد بررسی قرار گرفت و نتایج آن با چند روش اخیر مقایسه شد.

تکنیک‌های شناسایی چهره همواره با مشکلاتی در بازیابی تصویر از پایگاه داده مواجه بودند. روش‌های متعدد جهت رفع هر یک از این مشکلات ارائه گردیدند. جهت رفع مشکل حجم پایگاه داده و وجود تنها یک تصویر به ازای هر شخص، در این پایان نامه روشی نوین برای شناسایی چهره با استفاده از ویژگی‌های مکان-فرکانس ارائه شد. در این روش، ویژگی‌های مناسبی از تصویر چهره در حوزه فرکانس و با محاسبه تغییرات مولفه‌های فرکانسی در راستای مکان، استخراج شد. تغییرات راستای مکان متناسب با فرایند شناسایی چهره توسط قوهی بصری انسان بوده و به شکل کل به جزء فضای چهره را پیمایش می‌کند.

فضای ویژگی استخراج شده در این روش، با پیمایش تمام ناحیه چهره و استخراج همزمان ویژگی در هر مرحله از پیمایش، قابلیت بسیار بالایی در تفکیک بین کلاس‌های مختلف دارد. همچنین در این پایان نامه از فیلتری در حوزه فرکانس، جهت کاهش ابعاد ویژگی و بهبود نرخ بازیابی استفاده شده است. روش استخراج ویژگی، جدید بوده و تا کنون از آن جهت بازیابی چهره استفاده نشده است. در نهایت با استفاده از معیار فاصله اقلیدسی، نزدیک‌ترین تصویر پایگاه داده شناسایی شده است. الگوریتم پیشنهادی، در کاربردهایی با حجم پایگاه داده انبوه، نسبت به سایر الگوریتم‌ها بهتر ایفای نقش می‌کند. همچنین این الگوریتم در پایگاه داده‌هایی با تنها یک تصویر به ازای هر شخص، بازدهی بسیار مناسبی دارد. در این تحقیق، نتایج حاصل با روش‌های پیشین مورد مقایسه قرار گرفته و نشان داده شده است که نرخ بازیابی از روش‌های قبلی بیشتر بوده و با ۲٪ افزایش نسبت به بهترین روش ارائه شده تا کنون، به ۹۹٪ ارتقا پیدا کرده است. همچنین با افزایش حجم پایگاه داده تا ۹۹۰ تصویر چهره، نرخ بازیابی مناسبی حاصل گردیده

و نشان داده شد که شتاب کاهش نرخ بازیابی با افزایش حجم پایگاه داده بسیار کم است. این خاصیت بیانگر نرخ بازیابی مناسب در افزایش چندین برابری حجم پایگاه داده، است.

۲- پیشنهادها

به دلیل نوین بودن روش ارائه شده در این پایان‌نامه، جهت افزایش کارایی روش پیشنهادی و همین‌طور انتخاب زمینه‌های کاری مرتبط با این موضوع، در این بخش به مواردی اشاره کرده‌ایم.

فضای ویژگی در روش پیشنهادی، از فضای فرکانسی تصویر طی پنجره‌گذاری‌های متوالی روی تصویر چهره حاصل گردیده است. نتایج آزمایش‌ها نشان داد که با استخراج ویژگی مجدد از فضای ویژگی فرکانسی فعلی، می‌توان ویژگی‌های تفکیک‌ساز مناسب‌تری از چهره استخراج نمود.

در روش پیشنهادی با اعمال پنجره‌های هم‌مرکز روی تصویر و استخراج فضای فرکانسی در هر مرحله، فضای ویژگی حاصل گردیده است. با انتخاب روال پنجره‌گذاری متفاوت روی تصویر چهره، می‌توان نتایج دقیق‌تری را حاصل نمود.

روش پیشنهادی با استفاده از جستجوی ترتیبی، فضای ویژگی تصویر آزمون را با فضای ویژگی تمام تصاویر پایگاه داده مقایسه می‌کند. با ارائه الگوریتم خوشه‌بندی مناسب روی پایگاه داده مبتنی بر ویژگی‌های استخراج شده، می‌توان سرعت بازیابی الگوریتم را افزایش داد.

همچنین می‌توان از روشهای مبتنی بر موجک نیز جهت استخراج ویژگی‌های فرکانسی در تصویر چهره استفاده نمود و آن را جایگزین تبدیل فوریه کرد.

- [1] Lucas D. Introna, Lancaster University and Helen Nissenbaum, New York University, "Facial Recognition Technology A Survey of Policy and Implementation Issues", The Center for Catastrophe Preparedness & Response, New York University.
- [2] Li,Stan Z.,Jain,Anil K., "Handbook of Face Recognition", Edition 2, Springer-Verlag LondonPublishing, 2011.
- [3] Rabia Jafri and Hamid R. Arabnia, "A Survey of Face Recognition Techniques", Journal of Information Processing Systems, Vol.5, No.2, pp 41-56, 2009.
- [4] Zhao, Wenyi, Rama Chellappa, P. Jonathon Phillips and Azriel Rosenfeld, "Face recognition: A literature survey", ACM computing surveys (CSUR), Vol. 35, no. 4, pp 399-458, 2003.
- [5] Wiskott, Laurenz, J-M. Fellous, N. Kuiger, and Christoph Von Der Malsburg, "Face recognition by elastic bunch graph matching", IEEE Transactions on pattern analysis and machine intelligence, Vol. 19, No. 7, pp 775-779, 1997.
- [6] Kirby, Michael, and Lawrence Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE Transactions on Pattern analysis and Machine intelligence, Vol. 12, No. 1, pp 103-108, 1990.
- [7] Sirovich, Lawrence, and Michael Kirby, "Low-dimensional procedure for the characterization of human faces", Vol. 4, No. 3, pp 519-524, 1987.
- [8] Matthew Turk And Alex Pentland, "Eigenfaces for Recognition", Vision and Modeling Group, The Media laboratory, Massachusetts Institute of Technology, Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp 71-86, 1991.
- [9] Bartlett, Marian Stewart, "Independent component representations for face recognition", In Face Image Analysis by Unsupervised Learning, pp 39-67, 2001.
- [10] Belhumeur, Peter N., João P. Hespanha, and David J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", IEEE Transactions on pattern analysis and machine intelligence, Vol. 19, No. 7, pp 711-720, 1997.
- [11] Pentland, Alex, Baback Moghaddam, and Thad Starner, "View-based and modular eigenspaces for face recognition", Computer Vision and Pattern Recognition, Proceedings CVPR'94., IEEE Computer Society Conference on, pp 84-91, 1994.
- [12] L. Lenc and P. Kral, "Face Recognition under Real-world Conditions", ICAART(2), pp 250-256, 2013.
- [13] T. Ahonen, A. Hadid and M. Pietikainen, "Face recognition with local binary patterns", Computer Vision - ECCV, Vol. 302, pp 469-481, 2004.
- [14] S. Liao, X. Zhu, Z. Lei, L. Zhang and Stan Z. Li, "Learning multi-scale block local binary patterns for face recognition", Advances in Biometrics, Vol. 4642, pp 828-837, 2007.
- [15] D. H. Wang and P. Conilione, "Machine learning approach for face image retrieval", Neural Computing and Applications, Vol. 21, No. 4, pp 683-694, 2012.
- [16] M. J. Er, W. Chen and S. Wu, "High-speed face recognition based on discrete cosine transform and RBF neural networks", IEEE Transactions on Neural Networks , Vol. 16, No. 3, pp 679-691, 2005.

- [17] J. Luo, Y. Ma, E. Takikawa , S. Lao, M. Kawade and B. Lu , “Person-specific SIFT features for face recognition”, IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2, 2007.
- [18] L. Lenc and P. Král , “A combined SIFT/SURF descriptor for automatic face recognition”, Sixth International Conference on Machine Vision, Vol. 9067, 2013.
- [19] M. Zhou , H. Wei and S. Maybank, “Gabor wavelets and AdaBoost in feature selection for face verification”, Workshop in application of computer vision, pp 101-109, 2006.
- [20] H. Cho, R. Roberts, B. Jung, O. Choi and S. Moon, “An efficient hybrid face recognition algorithm using PCA and GABOR wavelets”, International Journal of Advanced Robotic Systems, Vol. 11, 2014.
- [21] R. Varun, Y. VivekanandKini, K. Manikantan and S. Ramachandran, “Face recognition using hough transform based feature extraction”, Procedia Computer Science, Vol. 46, pp 1491-1500, 2015.
- [22] T. M. Abhishree, J. Latha, K. Manikantan and S. Ramachandran, “Face recognition using gabor filter based feature extraction with anisotropic diffusion as a pre-processing technique”, Procedia Computer Science, Vol. 45, pp 312-321, 2015.
- [23] C. Ma, J. Y. Jung, S. W. Kim and J. S. Ko, “Random projection-based partial feature extraction for robust face recognition”, Neurocomputing, Vol. 149, pp. 1232-1244, 2015.
- [24] S. Wang and P.Liu, “A new feature extraction method based on the information fusion of entropy matrix and covariance matrix and its application in face recognition”, entropy, Vol. 17, No. 7, pp 4664-4683, 2015.
- [25] X. Xie and K. M. Lam, “Face recognition using elastic local reconstruction based on a single face image”, Pattern Recognition, Vol. 41, No. 1, pp 406–417, 2008.
- [26] Y. Gao, Y. Qi, “Robust visual similarity retrieval in single model face databases”, Pattern Recognition, Vol. 38, No. 7, pp 1009-1020, 2005.
- [27] J. Haddadnia, K. Faez and M. Ahmadi, “A fuzzy hybrid learning algorithm for radial basis function neural network with application in human face recognition”, Pattern Recognition, Vol. 36, No. 5, pp 1187–1202, 2003.
- [28] J. Lu, X. Yuan and T. Yahagi, “A method of face recognition based on fuzzy c-means clustering and associated Sub-NNs”, IEEE Transactions On Neural Networks, Vol. 18, No. 1, pp 150 – 160, 2007.
- [29] Gabor, Dennis. “Theory of communication. Part 1: The analysis of information”, Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution, Vol. 93, No. 26, pp 429-441, 1946.
- [30] Liu, Cui Yin, Chun Yu Zhang, Hong Zhao Yuan, and Xi Long Qu. “The Comparison of Isotropic and Anisotropic Diffusion Techniques for Image Denoising” In Applied Mechanics and Materials, Vol. 34, pp 557-561. Trans Tech Publications, 2010.
- [31] P. Saint-Marc, J.S. Chen, G. Medioni, “Adaptive smoothing: a general tool for early vision”, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 13, pp 514–529, 1990.
- [32] Weickert, Joachim, “Anisotropic diffusion in image processing”, Vol. 1. Stuttgart: Teubner, 1998.
- [33] P. Perona, J. Malik, “Scale space and edge detection using anisotropic diffusion”, Proc. IEEE Comp. Soc. Workshop on Computer Vision, IEEE Computer Society Press, Washington, pp 16–22, 1987.

- [34] Daugman, John G. "Two-dimensional spectral analysis of cortical receptive field profiles", *Vision research*, Vol. 20, No. 10, pp 847-856, 1980.
- [35] Kepenekci, Burcu. "Face recognition using gabor wavelet transform." PhD diss., Middle Eest Technical University, 2001.
- [36] Ferrari, Ricardo J., Rangaraj M. Rangayyan, JE Leo Desautels, and Annie F. Frere. "Analysis of Bilateral Asymmetry in Mammograms via Directional Filtering with Gabor Wavelets", *Journal of Electronic Imaging*, Vol. 16, no.1, 2007.
- [37] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters", *Pattern Recognition*, Vol. 24, No. 12, pp 1167–1186, 1991.
- [38] J. Kennedy and R. C. Eberhart, "A Discrete Binary Version of the Particle Swarm Algorithm", In: *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, pp 4104–4108, 1997.
- [39] Richard O. Duda and Peter E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures", *Technical Note 36, Communications of the ACM*, Vol. 15, No.1, pp 11–15, 1971.
- [40] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
- [41] K. K. Paliwal and L. D. Alsteris, "On The usefulness of STFT phase spectrom in human listening tests", *Speech Communication*, Vol. 45, No. 2, pp 153-170, 2004.

Abstract

Feature extraction from facial images is used in various applications Over the past two decades, e.g. face detection, face recognition, facial expression and face reconstruction. Among these applications, Face recognition has a particular importance.Face recognition has problems such as poor quality of the input image, unanticipated coatings on the face, high volume database and being only one picture per person.

In this research, a new method is presented to feature extraction from facial images. The main purpose of this research is retrieval of probe image from a big database. By increasing the size of the database, the similarities between people increases and the separation capability decreases.

The proposed method increases the distance between people in feature space by extracting appropriate features. This method is based on properties of the human vision system and sequentially extracts features in top-down manner. For this purpose, spatial-frequency features are used. In this method, by applying concentric windows in different size on the facial image, the content of each window are mapped to frequency space. The change of frequency components in different windows forms the feature space of image. Then frequency component with high separation capability between face images is remained by appropriate filter. In the end, the final image is retrieved from database by Euclidean distance criterion. In most existing facial recognition methods, desirable results are achived considering the small size of the database or in big database, the number of image per person increases.

In this research only one image per person is considered in database and used more subjects in database against another methods. In this research the FERET database is used. Recognition rate compared with the best current method in similar size of database with 2% improvement, reached to 99%. By increasing database size to 990 classes, 90/4% of recognition rate is achieved.

Key words: Face recognition, Feature extraction, Spatial-Frequency domain, Big database, Feature selection.



Faculty of computer Engineering
MSc Thesis in Artificial Intelligence Engineering

Feature Extraction from Facial Images for Image Retrieval from Database

By:

Mohammad.M Bakhshi

Supervisor:

Prof. H. Hassanpour

Advisor

Dr. M. Fateh

Feb 2017