دانشگاه صنعتی امیرکبیر
(پلی‌تکنیک تهران)

دانشکده مهندسی برق
گروه مخابرات - سیستم

پایان نامه کارشناسی ارشد

مکان‌یابی چهره و انطباق یک مستطیل به دور آن در تصاویر رنگی به منظور فشرده سازی با دقت بالا جهت بایگانی

نگارش:
علیرضا بایسته تاشک

استاد راهنما:
دکتر ابوالقاسم صیادیان

خرداد ماه ۱۳۸۷

دانشگاه صنعتی امیرکبیر
(پلی‌تکنیک تهران)

بسمه تعالی

فرم اطلاعات پایان‌نامه
کارشناسی- ارشد و دکترا

تاریخ:
شماره:

معاونت پژوهشی
فرم پروژه تحصیلات تکمیلی ۷

مشخصات دانشجو:

نام و نام خانوادگی: علیرضا بایسته ناشک دانشجوی آزاد ● بورسیه ○ سمادل ○

شماره دانشجویی: ۸۴۱۲۳۰۳۷ دانشکده: مهندسی برق رشته تحصیلی: مخابرات گروه: سیستم

مشخصات استاد راهنما:

نام و نام خانوادگی: دکتر ابوالقاسم صیادیان درجه و رتبه: دانشیار

نام و نام خانوادگی: درجه و رتبه:

مشخصات استاد مشاور:

نام و نام خانوادگی: درجه و رتبه:

نام و نام خانوادگی: درجه و رتبه:

عنوان پایان‌نامه به فارسی : مکان‌یابی چهره و انطباق یک مستطیل به دور آن در تصاویر رنگی به منظور فشرده سازی با دقت بالا جهت بایگانی

عنوان پایان‌نامه به انگلیسی: Optimum Face Rectangular Fitting and High Precision Compression in Color Image for Archiving

نوع پروژه: کارشناسی ○ ارشد ● دکترا ○ سال تحصیلی: ۱۳۸۴

کاربردی ○ بنیادی ○ توسعه‌ای ○ نظری ○

تاریخ شروع: ۱۳۸۵/۷/۱ تاریخ خاتمه: ۱۳۸۷/۳/۲۷ تعداد واحد: ۶ سازمان تأمین کننده اعتبار:

واژه‌های کلیدی به فارسی: شناسایی چهره،طبقه بندی کننده آبشاری،انتخاب ویژگی،ماشین بردار رابطه ای

واژه‌های کلیدی به انگلیسی: Face Detection, Cascade Classifier, Feature Selection, RVM

مشخصات ظاهری	تعداد صفحات: ۷۹	تصویر ● جدول ● نمودار ● نقشه ○ واژه‌نامه ○	تعداد مراجع: ۳۹	تعداد صفحات ضمائم: ۱
زبان متن	فارسی ●	انگلیسی ○	چکیده	فارسی ● انگلیسی ○
یادداشت				

نظرها و پیشنهادها به منظور بهبود فعالیت‌های پژوهشی دانشگاه

استاد:

دانشجو:

امضاء استاد راهنما: تاریخ:

چکیده:

در این پایان نامه سیستمی برای تشخیص سریع چهره در تصاویر رنگی ارائه می شود. هدف این سیستم پیدا کردن محل چهره‌های موجود در یک تصویر در حداقل زمان ممکن و دقت مطلوب می‌باشد. سیستم پیشنهادی در این پایان‌نامه مبتنی بر یک روش دو مرحله‌ای می‌باشد. در مرحله‌ی نخست، به کمک اطلاعات رنگ فضای جستجو کاهش می یابد. به این ترتیب نواحی که ممکن است چهره در آن وجود داشته باشد،مشخص می شوند و لازم نیست کل تصویر را برای یافتن چهره جستجو نماییم. در مرحله‌ی بعد هریک از نواحی جستجو به‌وسیله‌ی یک طبقه‌بندی‌کننده بررسی می‌شود و چهره‌های موجود در آن‌ها استخراج می‌گردند. مزیت عمده‌ی این ساختار دو مرحله‌ای این می باشد که قابل اعمال به تصاویرخاکستری می باشد.

در طراحی طبقه‌بندی‌کننده از ویژگی هار دوبعدی غیر استاندارد استفاده شده است. تعداد این ویژگی ها برای یک تصویر بسیار زیاد بوده و در وهله اول این مطلب که کدام ویژگی ها به بهترین نتایج طبقه‌بندی منجر می‌شوند مشخص نمی باشد. به همین دلیل در این سیستم بهترین مؤلفه ها توسط یک الگوریتم انتخاب ویژگی، انتخاب شده و از آنها برای آموزش طبقه‌بندی کننده استفاده می‌شود. در این سیستم برای تشخیص چهره در یک ناحیه، کل ناحیه توسط یک پنجره کوچک با اندازه‌ی ۲۰×۲۰در چند مقیاس جستجو می شود. از آنجایی که در هر موقعیت از تصویر می بایست عمل استخراج ویژگی و طبقه بندی انجام شود، برای بالا بردن سرعت جستجو، دو راهکار عمده دراین سیستم مد نظر قرار گرفته است. اولاً، یک روش بسیار سریع برای محاسبه ویژگی استفاده شده است . ثانیاً، در این سیستم برای طبقه‌بندی از یک طبقه‌بندی کننده با ساختار آبشاری استفاده شده است که سرعت طبقه بندی را بشدت افزایش می‌دهد.

نتایج حاصله نشان می‌دهند که نرخ شناسایی سیستم ارائه شده برای مجموعه‌ی تصاویر رنگی گردآوری شده۹۳/۷۷ می باشد. نرخ شناسایی برای پایگاه داده MIT- CMU ۸۹ با ۲۰ خطا می باشد.

زمان لازم برای یافتن چهره ها در یک تصویر به نحوی می باشد که سیستم قادر است در نرخ ۲۵ fps باوضوح تصویر ۲٤۰×۳۲۰ به صورت بلادرنگ عمل نماید.

فهرست مطالب

فصل اول ۱

۱-۱- مقدمه ۲

۱-۲- روشهای تشخیص چهره و دسته‌بندی آنها ۳

۱-۲-۱- روشهای مبتنی بر آگاهی ۳

۱-۲-۲- روشهای مبتنی بر ویژگیهای نامتغیر چهره ٤

۱-۲-۳- روشهای مبتنی بر تطبیق الگو ٥

۱-۲-٤- روشهای مبتنی بر ظاهر ٦

۱-۳- معرفی سیستم ارائه شده ۷

فصل دوم ۹

۲-۱- مقدمه ۱۰

۲-۲- فضاهای رنگ ۱۱

۲-۲-۱- فضاهای رنگ پایه: ۱۱

۲-۲-۲-فضاهای رنگ اداراکی ۱۲

۲-۲-۳- فضاهای رنگ یکنواخت ادراکی ۱۲

۲-۲-٤- فضاهای رنگ دیگر ۱۳

۲-۳- بررسی عملکرد فضاهای رنگ ۱۳

۲-٤- طبقه‌بندی کنندههای رنگ ۱۳

۲-٤-۱- آستانه‌گیری ۱۳

۲-٤-۲- مدل هیستوگرام و طبقه‌بندی‌کننده بیز ١٤

۲-٤-۳- مدل گوسی تک ١٦

۲-٤-٤- مدل مخلوط گوسی ١٦

۲-٤-٥- مدل محدوده‌ی بیضوی ١٧

۲-٤-٦- طبقه‌بندی‌کننده پرسپترون چند لایه ١٨

۲-٥- یافتن نواحی کاندیدای چهره ٢١

۲-٥-۱- جبران تغییرات روشنایی ٢١

۲-٥-۲- استخراج پوست ٢٢

فصل سوم ٢٧

۳-۱- مقدمه ٢٨

۳-۲- روش پیشنهادی ٣٥

۳-۳-جمع بندی ٣٧

فصل چهارم ٣٨

٤-۱- مقدمه ٣٩

٤-۲-عدم توازن در تشخیص چهره ٤٠

٤-۳- ماشین بردار رابطه ای ٤٣

٤-۳-۱- رگرسیون RVM ٤٤

٤-۳-۱-۱- توزیع پیشین وزن ها ٤٦

٤-۳-۱-۲- استنتاج بر اساس قواعد بیزین ٤٧

٤-۳-۲- طبقه بندی کننده RVM ٤٩

٤-٤- پالایش داده ها ٥٢

٤-٥- روش پیشنهادی برای آموزش هر گره ٥٢

٤-٥-١- انگیزه ٥٢

٤-٥-٢- RVM نامتقارن ٥٣

٤-٥-٣- نتایج بررسی روش پیشنهادی ٥٥

فصل پنجم ٥٧

٥-١- بررسی روش انتخاب ویژگی پیشنهاد شده ٥٨

٥-٢- بررسی ساختار آبشاری ٦٠

٥-٣-سیستم کلی ارائه شده ٦٦

فصل ششم ٦٨

نتیجه گیری و پیشنهادات ٦٩

مراجع ٧٠

ضمیمه‌ی یک ٧٣

فصل اول

مکان‌یابی چهره و معرفی روش‌های موجود

۱-۱- مقدمه

از دیرباز بشر به دنبال طراحی سیستمی‌هایی بوده است که بتوانند با انسان به صورت خودکار تعامل داشته باشند؛ در این راستا علوم مختلفی چون رباتیک، بینایی ماشین، هوش مصنوعی و ... رشد بسیار نموده‌اند. گام اول در رسیدن به این هدف درک حالات و احساسات انسان به وسیله‌ی کامپیوتر است. چهره‌ها دسته‌ای مهم از اشیاء هستند که نقش ویژه‌ای در بیان حالات و احساسات انسان ایفا می‌نمایند. لذا ادراک چهره به وسیله کامپیوتر از اهمیت ویژه‌ای برخوردار است که علاوه بر تعامل ماشین و انسان ، کاربردهای متنوعی از قبیل، مالتی مدیا، سیستم‌های نظارتی و امنیتی، مخابرات تصویری، متحرک‌سازی و غیره شامل می‌شود. لذا، در طی سال‌های اخیر روش‌ها و الگوریتم‌های متنوعی برای ادراک خود کار چهره بوجود آمده‌اند. درک حالت چهره، وضعیت و موقعیت چهره، تشخیص هویت از روی چهره، ردیابی چهره و ... در حوزه کلی ادراک چهره قرار می‌گیرند. تشخیص چهره به عنوان یک مسئله طبقه بندی دو کلاسی، وجود یا عدم وجود چهره و محل آن در تصویر ورودی را، مشخص می‌سازد. توصیف کلاس چهره به دلیل وجود مولفه‌های ساختاری از قبیل چشم‌ها، بینی و رابطه مکانی آن‌ها، نسبت به کلاس غیرچهره آسان‌تر است. علاوه بر این ، تنوع بسیار زیاد نمونه‌های کلاس غیرچهره، جمع آوری یک مجموعه آموزشی جامع برای این کلاس را مشکل می‌سازد، که این امر به نوبه خود توصیف قابل اطمینان این کلاس را، پیچیده می‌کند. چالش های مرتبط با مسئله آشکار سازی چهره را می توان به صورت زیر دسته بندی کرد[۱]:

- **وضعیت چهره:** وضعیت چهره در تصویر بسته به وضع نسبی چهره- دوربین (تمام رخ ، ۴۵ درجه، نیم رخ، وارونه) می‌تواند تغییر کند و ممکن است برخی از اجزای چهره مانند چشم، بینی یا ... به صورت جزئی یا کلی از دید پنهان شود.

- **وجود یا عدم وجود مؤلفه‌های ساختاری:** برخی مولفه‌های چهره مانند ریش، سبیل و عینک ممکن است که وجود داشته یا نداشته نباشند و شکل ، رنگ و اندازه این مؤلفه ها می‌تواند تغییرات قابل ملاحظه ای در نمود چهره ایجاد کند.

- **حالت چهره:** حالت های چهره از قبیل شادی،غم و ..... ظاهر چهره را تغییر می‌دهند.

- **پوشانندگی**: برخی از اجزا چهره ممکن است توسط اشیاء دیگر پوشیده شوند. مثلا در یک تصویر گروهی ممکن است برخی چهره ها ، چهره های دیگر را به صورت جزئی بپوشانند.

-**جهت تصویر**: چهره می‌تواند حول محور اپتیکی دوربین چرخیده باشد.

-**شرایط تصویربرداری**: فاکتورهایی مانند روشنایی(طیف، توزیع منابع، شدت) و خصوصیات دوربین (پاسخ سنسورهاو لنزها) بر ظاهر یک چهره در حین تصویربرداری اثر گذارند.

مشخصاً، تشخیص چهره اولین قدم در سیستمهای خودکار طراحی شده برای دستیابی به اهداف ادراکی سطح بالاتر فوق الذکر می‌باشد.

مسائل متعددی با مسئله تشخیص چهره در ارتباط نزدیک هستند [۱] از جمله:

– **مکان یابی چهره**: که عبارت است از مسئله مشخص کردن محل چهره در تصویری که فقط حاوی یک چهره است. در واقع این مسئله، یک حالت ساده شده از مسئله تشخیص چهره است که در آن فرض بر وجود فقط یک چهره در تصویر می‌باشد.

– **تشخیص اجزاء چهره**: که هدف از آن تشخیص وجود وموقعیت اجزاء چهره، مانند چشم‌ها، بینی، ابرو، دهان و لب‌ها، گوش و غیره، با فرض وجود فقط یک چهره در تصویر می‌باشد.

– **تشخیص هویت از روی چهره**: که در آن تصویرچهره ورودی با تصاویر موجود در یک پایگاه داده مقایسه شده و در صورت وقوع تطابق، هویت فرد اعلام می‌شود.

– **ردیابی چهره**: که در آن ردیابی پیوسته موقعیت و جهت چهره در یک سلسله تصاویر (ویدئو) بصورت بلادرنگ هدف مطلوب سیستم می‌باشد.

– تشخیص حالت چهره: که در آن مشخص کردن حالت چهره اعم از شاد بودن، غمگین و... مد نظر می‌باشد.

۱-۲- روش‌های تشخیص چهره و دسته‌بندی آنها

باتوجه به ملاحظاتی که در انتخاب یک استراتژی مناسب برای تشخیص چهره دخیل هستند از قبیل نحوهٔ ارائه و انتخاب روش تصمیم‌گیری یا طبقه بندی، راهکارهای متنوعی برای مسئله تشخیص چهره ارائه شده است [۱]. در این بخش روش‌هایی که تاکنون برای تشخیص چهره ارائه شده‌اند را باتوجه به نوع برخورد آن‌ها با مسئله، دسته‌بندی کرده و مزایا و معایب هریک را تا حد ممکن بررسی می‌کنیم تا درک جامعی از روش‌های ارائه شده حاصل گردد و جهت انتخاب روش مناسب بصیرت کافی حاصل شود.

۱-۲-۱- روش‌های مبتنی بر آگاهی[1]

در این روش‌ها، باتوجه به آگاهی محقق درباره‌ی چهره انسان، قوانینی که توصیف‌کننده‌ی چهره‌ی انسان هستند، بدست می‌آیند. این قوانین عمدتاً اجزاء چهره‌ی انسان و روابط آن‌ها را

[1] Knowledge-Based

توصیف می‌نمایند. بعنوان مثال یک چهره همواره در تصویر با دو چشم که نسبت به یکدیگر به‌صورت متقارن قرار دارند، یک بینی و یک دهان ظاهر می‌شود. رابطهٔ بین این‌اجزاء با استفاده از فاصله و موقعیت نسبی آن‌ها قابل تعریف می‌باشد. در این روش‌ها که عمدتاً برای تصاویری که فقط شامل یک چهره هستند استفاده می‌شوند، ابتدا اجزاء چهره‌استخراج شده و سپس چهره‌های احتمالی با توجه به قوانین بدست آمده، مکان‌یابی‌می‌شوند. سپس این کاندیداها وارد یک مرحله‌ی تایید می‌شوند تا میزان تشخیص‌های غلط‌کاهش یابد. یک مشکل عمده که برای این روش‌ها متصور است این است که مسئله تشخیص شئ به مسئله‌ی تشخیص چند شئ (اجزاء مختلف چهره) که مسئله‌ای پیچیده‌تر است می‌تواند گسترش پیدا کند و بعلاوه اجزاء چهره می‌بایست کاملاً در تصویر مشخص باشند که این در مورد مسئله مطلوب این پایان‌نامه صادق نمی‌باشد. همچنین مشکل دیگری که در این روش‌ها وجوددارد، دشوار بودن استخراج قوانینی موثر و کارآمد با استفاده از آگاهی انسان می‌باشد. اگراین قوانین بسیار دقیق و سختگیرانه باشند، ممکن است بسیاری از چهره‌ها توسط این قوانین رد شوند و اگر این قوانین کلی و سهل باشند، تعداد تشخیص‌های غلط بالا می‌رود. در واقع جز آزمایش راه دیگری برای آزمودن این قوانین که عمدتاً بصورت شهودی وابتکاری بدست می‌آیند وجود ندارد.

بعلاوه تعمیم این روش‌ها به وضعیت‌های مختلف چهره نیز دشوار می‌باشد چرا که برای تمام وضعیت‌های ممکن باید قوانین متفاوتی وضع شود. با این حال این روش‌ها ساده هستند ودر مواردی که تصویر حاوی یک چهره تمام رخ بوده و پس زمینه نیز ساده و قابل‌پیش‌بینی باشد، می‌توانند موثر واقع شوند.

۱-۲-۲- روش‌های مبتنی بر ویژگی‌های نامتغیر چهره

باتوجه به اینکه انسان‌ها به راحتی توانایی تشخیص چهره‌ها در وضعیت‌ها و شرایط‌نورپردازی مختلف را، دارا هستند، لذا می‌بایست ویژگی‌ها یا خصوصیاتی درچهره وجودداشته باشند که در شرایط متغیر، تقریباً ثابت باقی می‌مانند. با در نظر گرفتن این‌فرضیه، روش‌هایی بوجود آمدند که سعی در پیدا کردن چنین ویژگی‌هایی دارند، ویژگی‌های مانند لبه‌های چهره، نسبت شدت روشنایی در مناطق مختلف چهره، شکل اجزاء، بافت و رنگ قسمت‌های مختلف چهره و غیره. سپس با استفاده از این ویژگی‌ها مدل‌هایی که توصیف کننده ارتباط این ویژگی‌ها می‌باشند، بدست می‌آیند و یا این ویژگی‌ها به صورت مفهومی دسته‌بندی می‌شوند[1] تا وجود چهره تشخیص داده شود.

[1] Perceptual Grouping

عمده‌ترین مشکلی که برای این روش‌ها وجود دارد، تشخیص چهره در یک پس‌زمینه‌ی پیچیده[1] می‌باشد. چرا که در پس زمینه‌های پیچیده جزئیات فراوانی وجود دارند که روش‌های دسته‌بندی را، بشدت دچار مشکل می‌کنند. لذا این روش‌ها نیز عمدتاً برای تصاویر حاوی یک چهره و پس زمینه‌های ساده مناسب هستند. مشکل دیگر، بعلت تغییرات بسیار متنوعی که در چهره‌ها وجود دارد و در بخش قبلی به آن‌ها اشاره شد، بوجود می‌آید. ممکن است ویژگی‌های مطلوب بر اثر نورپردازی نامناسب، نویز و پوشیدگی، مختل شده و یا محدوده‌ی ویژگی‌ها کاملاً مشخص و قابل تفکیک نباشند؛ چنین شرایط نامساعدی این روش‌ها را با مشکلات جدی روبرو می‌نمایند. ولی بهرحال در شرایط کنترل‌شده و پس زمینه‌های ساده، می‌توان از این روش‌ها استفاده نمود.

۱-۲-۳- روش‌های مبتنی بر تطبیق الگو

در روش‌های تطبیق الگو، یک الگوی مشخص برای چهره (اغلب تمام رخ) تعریف شده و یا بصورت یک مدل پارامتریک مشخص می‌شود. این مدل‌ها معمولاً برپایه لبه‌ها و خطوط موجود در صورت و یا مشخصات نواحی صورت از قبیل میزان روشنایی، تعریف می‌شوند. سپس میزان همبستگی این الگو با نواحی مختلف تصویر محاسبه شده و نواحی با میزان همبستگی بالا بعنوان چهره مشخص می‌گردند. این قبیل الگوها از آنجایی که بصورت اصطلاحاً دستی طراحی می‌شوند، درمواجهه با تغییرات متنوع کلاس چهره‌ها ممکن است دچار مشکل شوند و می‌بایست تحت شرایط کنترل شده مورد استفاده قرار گیرند. البته برای مقابله با این مسئله، الگوهای انعطاف پذیر ارائه شده‌اند که در آنها نواحی مختلف تصویر به یک مدل انعطاف‌پذیر پارامتریک، با کمینه‌سازی یک تابع انرژی، برازانده می‌شوند[۲]. این روش‌ها برای اجزاء متغیر چهره خوب عمل می‌کنند ولی یک عیب عمده‌ی آنها این است که می‌بایست مدل مربوطه به‌گونه‌ای مقداردهی اولیه شود که به حالت فعلی شئ مورد تشخیص نزدیک باشد و در غیر این صورت عمل برازش دچار مشکل می‌شود. بطورکلی روش‌های مبتنی برتطبیق الگو برای تشخیص چهره‌های متعدد در تصویر با توجه به تغییرات متنوع موجود در کلاس چهره‌ها، و همینطور از لحاظ سرعت پردازش، برای کاربرد مطلوب این پایان نامه مناسب به نظر نمی‌رسند.

[1] Cluttered

۱-۲-٤- روش‌های مبتنی بر ظاهر[1]

در روش‌های مبتنی بر ظاهر به جای اینکه قوانین یا الگوها از پیش تعریف شوند، این قوانین یا الگوها از روی داده‌های ارائه کننده کلاس مورد تشخیص (و کلاس متمم) آن، آموخته می‌شوند. بطور کلی، در روشهای مبتنی برظاهر، از روش‌های تجزیه و تحلیل‌آماری و آموزش ماشین برای پیدا کردن مشخصاتی که به خوبی توصیفگر شیء مورد نظر باشند، استفاده می‌شود. این مشخصات که از روی داده‌ها آموخته می‌شوند اغلب به‌صورت مدلهای توزیع و یا توابع تمایز هستند که نهایتاً از آن‌ها برای تشخیص چهره استفاده می‌شود. اغلب از روش‌های کاهش ابعاد برای پایین آوردن حجم محاسبات و بالابردن دقت تقریب‌ها (باتوجه به حجم محدود داده‌های آموزشی)، قبل از آموزش استفاده می‌شود که در بخش قبل در این مورد صحبت شد. بسیاری از روش‌های مبتنی بر ظاهر در قالب توابع احتمال قابل بررسی هستند. دراین رویکرد، تصویر یا بردار ویژگی استخراج شده از آن، به عنوان یک متغیر تصادفی x درنظر گرفته می‌شود و این متغیر تصادفی با استفاده از توابع احتمال شرطی کلاس‌های چهره و غیرچهره ، P(x|face) و P(x|non-face)، توصیف می‌شود. سپس با استفاده ازطبقه‌بندی کننده‌ی بیز یا حداکثر احتمال پسینه، عمل‌کلاس‌بندی یک تصویر نمونه انجام می‌شود. متأسفانه، پیاده سازی طبقه‌بندی کننده‌ی بیز براحتی امکانپذیر نمی‌باشد، چرا که P(x|face) و P(x|non-face) ، اغلب بصورت چند قله‌ای[2] بوده و فرم پارامتریک مشخصی برای آنها متصور نیست. بعلاوه بدلیل ابعاد بالای x ، تخمین آن‌ها با داده‌های آموزشی محدود دشوار می باشد. لذا در این رویکرد در گذشته، عمده‌ی فعالیت‌ها روی روش‌های مختلف پیدا کردن تقریب‌های پارامتریک یا غیرپارامتریک برای این توابع توزیع احتمال، متمرکز بوده است. در فصل ٤ در این مورد با عمق بیشتری صحبت خواهیم کرد. رویکرد دیگری که در روش‌های مبتنی بر ظاهر وجود دارد (که رویکرد منتخب این‌پایان‌نامه نیز می‌باشد) یافتن یک تابع تمایز برای جداسازی کلاس‌های چهره و غیره،می‌باشد. در این رویکرد، اغلب، الگوهای تصویری، به یک فضای با ابعاد کمتر (مثلاً با استفاده از استخراج ویژگی) تصویر می‌شوند و سپس یک تابع تمایز ازروی داده‌های آموزشی، تولید (آموخته) می‌شود. این تابع تمایز یا به فرم یک تابع خطی (خط، صفحه و ابرصفحه) بوده و یا به صورت یک تابع غیرخطی (منحنی یا ابرصفحهٔ غیر مسطح) مانند آنچه در شبکه‌های عصبی چند لایه وجود دارد، می‌باشد.

اخیراً از ماشین های بردار رابطه ای[3] ها یا روشهای مبتنی بر توابع هسته[1] نیز برای کاربرد

---

[1] Appearance Based
[2] Multimodal
[3] Relevance Vector Machines

تشخیص شیء استفاده شده است. این روش‌ها برخلاف روش‌های متداول شبکه عصبی که در آنها آموزش، با کمینه‌سازی تابع خطای تجربی انجام می‌شود، بر اساس روابط بیزین وفرض یک توزیع اولیه مناسب بر روی پارامترها بدست می‌آیند. در مورد روش‌های ذکر شده، در فصل ٤ مفصل‌تر صحبت خواهیم کرد. روش‌های مبتنی بر ظاهر به خاطر ماهیت آموزش بیننده‌ای که دارند، علیرغم منابع متنوع تغییر که در کلاس چهره‌ها حضور دارند، به خوبی عمل می‌کنند و اغلب روش-هایی که اخیراً برای تشخیص چهره مطرح شده‌اند در این قالب بوده و نتایج بهتر ومطمئن‌تری نسبت به سه روش قبلی مخصوصاً در مواردی که پس زمینه پیچیده باشد، ازخود نشان داده‌اند ([٤] تا [٨]). لذا در این پایان‌نامه، این رویکرد برای تشخیص چهره انتخاب شده‌است.

۱-۳- معرفی سیستم ارائه شده

در این بخش به معرفی عملکرد سیستم ارائه شده و محتوای فصول بعدی پایان‌نامه می‌پردازیم. همانطور که در بخش‌های گذشته مطرح شد، بعلت وجود مشخصات متغیر بسیار در کلاس‌چهره‌ها، مدل کردن چهره‌ها بصورت تحلیلی بسیار دشوار بوده و معمولاً از روشهای آماری برای مدل کردن چهره‌ها استفاده می‌شود. این روش‌ها که در طبقه‌بندی بخش قبلی در گروه روشهای مبتنی بر ظاهر قرار می‌گیرند، مدل آماری از روی یک مجموعه‌ی داده‌های کلاس موردنظر (وکلاس متمم آن) تولید کرده یا اصطلاحاً آموزش می‌بیند.

در این سیستم ما ابتدا به کمک اطلاعات رنگ ناحیه‌ی جستجو را به مقدار قابل ملاحظه‌ای کاهش می دهیم. این کار در سرعت سیستم تاثیر بسزایی ایفا می کند. سپس نواحی جستجو را به سطوح خاکستری تبدیل می‌نماییم و به استخراج ویژگی برای آن می‌پردازیم.

ماسک هار به علت سادگی و سهولت محاسبه مورد توجه قرارگرفته اند.این ویژگیها به عنوان ساده‌ترین موجک ها قادرند تغییرات روشنایی در جهت های افقی ،عمودی و قطری کد کنند و ابزاری مناسب را، برای توصیف یک جسم فراهم نمایند.

با توجه به بزرگ بودن حجم پایگاه داده تصاویر،جهت افزایش سرعت طبقه‌بندی از ساختار آبشاری برای طبقه‌بندی کننده استفاده شده است. یک طبقه‌بندی کننده‌ی آبشاری از چندین طبقه‌بندی کننده معمولی که در طبقات پی در پی قرارگرفته‌اند، تشکیل شده است. طبقات ابتدایی دارای طبقه-بندی کننده‌های ساده‌ای بوده و پیچیدگی طبقه‌بندی کننده‌ها درطبقات بالاتر بیشتر می‌شود. به هنگام جستجوی تصویر برای پیدا کردن چهره‌ها بسیاری‌از مناطق تصویر توسط طبقه‌بندی کننده‌های ساده‌ی

[1] Kernel Methods

طبقات ابتدایی رد[1] شده و فقط مناطقی که به چهره شباهت دارند به طبقات بالاتر راه پیدا کرده و مورد پردازش بیشتر قرار می‌گیرند و بهمین دلیل سرعت جستجو افزایش پیدا می‌کند. هر منطقه‌ای از تصویر که بتواند از تمامی طبقات عبور کند بعنوان چهره ثبت می‌گردد فصل ٤ به معرفی آموزش آماری و طبقه‌بندی کننده انتخاب شده و مسائل‌مرتبط با آموزش آن اختصاص داده شده است.

فصل سوم به مسئله انتخاب بهترین مولفه‌های ممکن جهت طبقه‌بندی، اختصاص دارد. مسئله انتخاب بهترین ویژگی‌ها از دوجهت حائز اهمیت است. اولاً) ویژگیهای اضافی[2] و یا نامربوط[3] می‌توانند آموزش را مختل نمایند. ثانیاً) همانطور که قبلاً نیز گفته شد، طبقات ابتدایی طبقه‌بندی کننده‌ی آبشاری می‌بایست ساده‌ترین حالت ممکن را داشته باشند لذا می‌بایست با تعداد کمی ویژگی به دقت طبقه‌بندی موردنظر دست پیدا کنند. بنابراین مسئله انتخاب بهترین ویژگی‌ها یا مولفه‌ها، مخصوصاً برای طبقات ابتدایی که بیشترین مناطق تصویر را مورد پردازش قرار می‌دهند، بسیار مهم بوده و در سرعت عملکرد سیستم نقش کلیدی دارد. در فصل پنجم پیاده‌سازی سیستم کلی و نتایج شبیه سازی روش پیشنهاد شده در این پایان نامه به تفصیل بیان گردیده است.در فصل پایانی جمع-بندی و پیشنهادات ارائه شده است. در ضمیمه‌ی روشی برای فشرده سازی تصویر چهره بررسی می-گردد. این روش بر اساس الگوریتم VQ تصاویر چهره را در دو مرحله چندی سازی می‌نماید.

1 Reject
2 Redundant
3 Irrelevant

فصل دوم

رنگ و کاربرد آن در مکان‌یابی چهره

۲-۱- مقدمه

یافتن پوست، در بسیاری از کاربردهای پردازش تصویر مانند شناسایی چهره[۱]، ردیابی چهره[۲] و تحلیل بافت[۳] و سیستم‌های بازیابی تصاویر نقش مهمی ایفا می کند. اخیرا روش‌هایی که از رنگ، برای یافتن پوست استفاده می‌کنند، توجه زیادی را به خود جلب کرده است. مقاومت نسبت به عواملی چون چرخش[٤]، تغییر اندازه[٥] و پوشانندگیِ[٦] جزئی از دلایل اصلی استفاده از رنگ می‌باشد. علاوه بر این، از رنگ پوست در کنار سایرشاخص‌ها مانند شکل و هندسه، می‌توان برای ساختن سیستم‌های شناسایی چهره‌ی دقیق استفاده کرد.

تشخیص رنگ پوست اغلب به عنوان یک مرحله اولیه در سیستم‌های تشخیص چهره، ردیابی چهره و ... استفاده می گردد. از اطلاعات رنگ پوست می توان به عنوان ابزاری موثر برای یافتن نواحی کاندیدای چهره استفاده کرد. تلاش ها در راستای یافتن پوست بیشتر در حوزه تصویر-برداری طیف مرئی متمرکز است.عوامل ذیل توصیف رنگ پوست را دشوار می سازند:

- **روشنایی**[۷]: تغییرات در توزیع منابع نور و شدت روشنایی،نمود رنگ پوست را در تصویر به شدت تحت تاثیر قرار می دهد. تغییرات روشنایی به شدت کارایی روش‌های یافتن رنگ پوست را کاهش دهد.

- **خصوصیات دوربین**[۸]: حتی در شرایط نورپردازی یکسان، توزیع رنگ پوست یک فرد از یک دوربین به دوربین دیگر متفاوت است.

- **نژاد**[۹]: رنگ پوست در اقوام مختلف تغییرات فراوانی می کند برای مثال رنگ پوست درمیان اقوام آسیایی،آفریقایی و اروپایی نمود کاملا متفاوتی دارند.

- **خصیصه فردی**[۱۰]: سن، جنسیت و... می توانند از عواملی باشند که رنگ پوست را تحت تاثیر قرار دهند.

عواملی دیگری چون گریم، عینک، رنگ‌های زمینه و حرکت می توانند نمود رنگ پوست را تغییر دهد. این گونه مشکلات را براحتی می توان به کمک طیف نامرئی یا تصویر برداری چند طیفی

---

[1] Face Detection
[2] Face Trakcing
[3] Texture Analysis
[4] Texture Analysis
[5] Scale
[6] Occlusion
[7] Illumination
[8] Camera Characteristic
[9] Ethnicity
[10] Individual Characteristics

حل نمود. طیف نامرئی رنگِ پوست نسبت به شرایط نورپردازی ، نژاد ، سایه و گریم مقاوم است. اگر چه تجهیزات گران این گونه تصویر برداری کاربرد روش های مبتنی بر طیف نامرئی را بیسار محدود می‌کند.

مسئله تشخیص رنگ پوست یک مسئله طبقه بندی دو کلاسی می باشد. اولین گام در راستا، استفاده از اطلاعات رنگ در قالب یک فضای رنگ مناسب می باشد. گام بعدی مدل کردن پیکسل-های پوست و غیر پوست به وسیله یک توزیع مناسب است ودر نهایت طبقه‌بندی پیکسل‌ها به‌وسیله مدل‌های بدست آمده می باشد(در همه‌ی روش‌ها نمی توان بخش‌های مدل سازی و طبقه بندی را به عنوان دو مرحله مجزا در نظر گرفت). انتخاب مناسب فضای رنگ، همپوشانی بین دو کلاس پوست و غیر پوست را کاهش می دهد و به این وسیله به طبقه‌بندی کردن پوست و ایجاد پارامترهای مقاوم در برابرتغییرات شرایط نورپردازی کمک می کند. به بیانی دیگر مدل سازی و طراحی طبقه‌بندی-کننده ساده ترمی‌گردد.

چندین فضای رنگ برای تشخیص رنگ پوست پیشنهاد و استفاده شده است. فضای رنگ RGB بعنوان فضای پیش فرض اغلب فرمت‌های تصویرشناخته می‌شود و معمولاً هر فضای رنگ دیگر به وسیله‌ی یک تبدیل(خطی یا غیرخطی) از فضای RGB توصیف می‌گردد.

بررسی ها نشان می دهد که تفاوت عمده رنگ پوست افراد درمولفه‌ی شدت روشنایی می‌باشد. بنابراین استفاده از مولفه‌ی روشنایی برای طبقه‌بندی توصیه نمی‌گردد .و غالباً روش‌ها از مولفه‌های رنگینگیِ[1]پوست، که در افراد مختلف شباهت زیادی با یکدیگر دارند، استفاده می‌کنند[۱۳]. در ادامه ما چند فضای رنگ متداول را توضیح می‌دهیم.

۲-۲- فضاهای رنگ

۲-۲-۱- فضاهای رنگ پایه:

از فضای رنگ RGB برای ذخیره سازی و نمایش تصاویر رقمی استفاده می شود.چشم انسان رنگ را به صورت ترکیب این سه مولفه توصیف می نماید.فضای RGB از سه مولفه رنگ اولیه(قرمز،آبی،سبز)تشکیل می شود.

به منظور کاهش وابستگی به شرایط نورپردازی مولفه های RGB را نرمالیزه می کنند طوری که جمع مولفه های حاصل یک گردد.این وابستگی خطی سبب می گردد که یکی از مولفه های حامل

[1] Chromaticity

اطلاعات اضافه نباشد و بتوان از آن صرف نظر کرد.فضای حاصل را RGB نرمالیزه شده(rgb) می-نامند. در این فضا تفاوت های ناشی از شرایط نورپردازی و نژاد به صورت قابل ملاحظه ای کاهش می یابد. خوشه های پوست در فضای rgb واریانس کمتری نسبت به فضای RGB دارند. از این رو مدل سازی و تشخیص در این فضا راحت است.به علت مزایای ذکر شده فضای rgb یکی از فضاهای پر کاربرد در تشخیص پوست می‌باشد.

سیستم رنگ CIE از یک مولفه Y روشنایی و دو مولفه Z،X تشکیل می‌گردد مولفه‌های این فضا طوری طراحی گردیده‌اند که با ویژگی‌های انطباق رنگ درسیستم بصری انسان مطابقت دارند.

۲-۲-۲-فضاهای رنگ اداراکی

توصیف ویژگی‌های ادراکی[1] رنگ مانند اصل رنگ[2](H)اشباع[3](S)وشدت روشنایی(I) به صورت تبدیل خطی از RGB ممکن نمی باشد.تبدیل های غیرخطی فراوانی برای نگاشت فضای RGB به فضای ویژگی های ادراکی پیشنهاد شده است.

فضای رنگ HSV رنگ را به صورت اصل رنگ (خاصیتی از رنگ که با عبور از قرمز به سبز تغییر می کند) اشباع (خاصیتی از رنگ که با عبور از قرمز به صورتی تغییر می کند) و روشنایی(خاصیتی که با عبور از سیاه به سفید تغییر می کند) توصیف می کند.تبدیل RGB به HSV نسبت نور شدید[4] در منبع سفید،نور محیطی و سوگیری سطح نسبت به منبع نور مقاوم است از این رو انتخاب خوب برای روش های تشخیص پوست می باشد.فضاهای مشابه دیگری مانند YIQ– YES،YUV،YCgCr،YCbCr در این دسته قرار می گیرند.

۲-۲-۳- فضاهای رنگ یکنواخت ادراکی

یکنواختی ادراکی نشان می‌دهد که چطور دو رنگ برای انسان در ظاهر متفاوت می باشد. فضای رنگ یکنواخت به گونه ای تعریف می گردند که همه رنگ بوسیله تفاوت های ادراکی مرتب گردند.در این فضای رنگ مولفه روشنایی[5] L و مولفه رنگینگی[6](UV یا ab) بوسیله یک نگاشت غیر خطی مختصات XYZ بدست می آیند.

---

[1] Perceptual Features
[2] Hue
[3] Saturation
[4] Highlight
[5] Luminance
[6] Chroma

۲-۲-٤- فضاهای رنگ دیگر

معمولاً مولفه ی قرمزدررنگ پوست انسان مولفه‌ی غالب می باشد. از این رو برخی محققین از نسبت‌های رنگ (R/G و غیره) برای شناسایی بهره جستند. برخی هم از ترکیبی از مولفه‌های رنگ استفاده کردند (H,R/G,E) و(R/G+R/B+G/B) .

۲-۳- بررسی عملکرد فضاهای رنگ

از معیارهای مختلفی برای ارزیابی عملکرد فضاهای رنگ می‌توان بهره جست. بعنوان مثال [۱۹] عملکرد ۱٦ فضای رنگ مختلف را به وسیله‌ی ۵ معیار متفاوت بررسی کرده است. معیار هایی چون آنتروپی توزیع پوست(E)، اشتراک هیستوگرام‌ها(HI)، HCE[1]، $tr(S_w)/tr(S_B)$ و $tr(S_w^{-1}S_B)$ دراین بررسی استفاده شده‌اند. فضاهای رنگ به صورت دو و سه بعدی با یکدیگر مقایسه شده-اند[2].نتایج این بررسی را می‌توان در جدول ۲-۱ مشاهده نمود.

۲-٤- طبقه‌بندی کننده‌های رنگ

از دیدگاه طبقه بندی، تشخیص رنگ پوست متمایز ساختن دو کلاس پوست وغیر پوست از یکدیگر می باشد. بسته به فضای رنگ انتخاب شده و پایگاه داده‌ی استفاده شده، روش‌های مختلفی برای حل مسئله اتخاذ گردیده است. بنابراین مرور مختصری از کارهای انجام شده در این زمینه، می‌تواند مارا در انتخاب یک راهکار مناسب برای حل این مسئله یاری نماید

۲-٤-۱- آستانه‌گیری[3]

رنگ پوست انسان در محیط هایی با نورپردازی کنترل شده، نواحی کوچکی در فضای رنگ اشغال می‌کند. از این رو یکی از متداول ترین روش‌ها تعریف محدوده‌ای برای مولفه‌های مختلف رنگ می-باشد. یک یا چندین بازه برای هر مولفه‌ی رنگ مشخص می گردد و سپس پیکسل‌هایی که در این محدوده ها قرار می‌گیرند به عنوان پوست شناخته می شوند. [١٤]ازیک بازه ثابت برای مولفه I در فضای YIQ برای تشخیص رنگ پوست بهره جست. همه پیکسل‌های تصویر که در بازه‌ی $R_I = [0\ 50]$

[1] Histogram Chi-Error

[2] فضا های دو بعدی از حذف مولفه ی شدت روشنایی بدست می آیند. برای RGB مولفه G و برای KLT مولفه Tحذف گردیده اند.

[3] Thersholding

جدول:۲-۱ :عملکرد فضاهای رنگ [۱۹]

	E		HI		HCE		$tr[S_w]/tr[S_B]$		$tr[S_w^{-1}S_B]$	
	3D	2D	3D	2D	3D	2D	3D	2D	3D	2D
RGB	12.54	9.932	0.344	0.416	1.030	0.841	24.14	17.74	0.331	0.305
XYZ	12.32	9.686	0.346	0.423	1.024	0.827	37.61	38.45	0.331	0.288
YCbCr	12.48	7.779	0.345	0.419	1.025	0.834	20.35	4.471	0.331	0.285
YIQ	12.09	7.394	0.347	0.418	1.021	0.836	21.37	4.349	0.331	0.285
YES	12.17	7.586	0.347	0.414	1.022	0.847	20.61	4.427	0.330	0.284
KLT	12.22	7.713	0.346	0.375	1.022	0.926	20.21	3.769	0.331	0.331
L*a*b*	11.34	7.570	0.349	0.473	1.015	0.709	19.58	13.30	0.201	0.110
L*u*v*	11.49	7.848	0.348	0.526	1.017	0.601	21.14	13.31	0.215	0.105
HSI	12.20	8.690	0.345	0.452	1.025	0.754	62.52	96.85	0.076	0.024
HSV	12.47	8.867	0.344	0.452	1.027	0.757	20.54	49.29	0.188	0.038
HSL	12.33	8.712	0.345	0.452	1.026	0.754	49.22	79.52	0.106	0.027
TSL	12.12	8.103	0.346	0.457	1.023	0.746	63.32	872.15	0.063	0.003
NCC rgb	/	7.420	/	0.462	/	0.731	/	23.61	/	0.048
NCC xyz	/	6.340	/	0.468	/	0.719	/	25.23	/	0.041
E-R/G-H	9.094	/	0.362	/	0.979	/	29.49	/	0.315	/
H-GY-Wr	9.260	/	0.369	/	0.966	/	9.495	/	0.226	/

قرار می‌گیرند به عنوان پوست طبقه‌بندی می گردند. البته این روش اساساً برای تصاویر افراد زرد پوست طراحی شده است.

[۳۷] از بازه های ثابتی در فضای HS استفاده کرده‌اند. در این روش پیکسل‌هایی که در بازه R_H=[0 50] و R_S=[0.23 0.68] قرار می‌گیرند به عنوان پوست شناخته می شوند. این محدوده ها، برای پایگاه داده M2VIS بدست آمده اند.

[۱٤] یک الگوریتم قطعه بندی چهره پیشنهاد داده است که در آن از بازه‌های ثابت در فضای C_bC_r استفاده شده است. پیکسل‌هایی که مقادیر مولفه C_r و C_b در بازه‌های R_{cr} =[133 173] و R_{cb}=[77 127] قرار می‌گیرند به عنوان پوست در نظر گرفته می شوند. این روش از پایگاه داده ECU برای یافتن بازه های خود استفاده کرده است.

[۱٤] ناحیه پوست را به کمک ۸ صفحه در فضای YCbCr جدا ساختند. [۱٤]از مقدار آستانه در فضای rg وhsv استفاده کردند. این مقادیر عبارتند از: R_r =[0.35 1]، R_g =[0.28, 0.363]، R_H =[0, 50]، R_S =[0.20, 0.68]. برای جزئیات بیشتر به مرجع [۱٤] مراجعه شود.

۲-٤-۲- مدل هیستوگرام و طبقه‌بندی‌کننده بیز[1]

در این روش از یک مدل دو بعدی یا سه بعدی هیستوگرام برای توصیف رنگ پوست استفاده می شود. . از هیستوگرام رنگ می‌توان برای متمایز ساختن اشیاء و ردیابی اهداف بهره جست[۱۵]. از

[1] Bayes Classifiers