مدل‌سازی چهره با استفاده از میانگین‌گیری بر پایه دگردیسی تصویر و تجزیه مرتبه پایین

سینا اخلاقی[1] حمید حسن‌پور[2] وحید ابوالقاسمی[3]

۱-دانش آموخته کارشناسی ارشد- دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود- ایران
sinaakhlaqi@shahroodut.ac.ir

۲- استاد- دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود- ایران
h.hassanpour@shahroodut.ac.ir

۳- استادیار- دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود- ایران
vabolghasemi@shahroodut.ac.ir

چکیده: در نظارت ویدئویی، زاویه دید چهره نسبت به دوربین (انسداد زاویه‌ای) توانایی سیستم در شناسایی چهره را محدود می‌سازد. در این مقاله روشی برای رفع انسداد زاویه‌ای در تصاویر چهره پیشنهاد می‌شود که مبتنی بر دگردیسی تصویر بوده و با دریافت تعدادی تصویر زاویه‌دار از چهره یک شخص، تصویری از روبه‌رو مدل می‌سازد. روش پیشنهادی با تکنیک‌هایی از قبیل انتقال مکانی پیکسل‌های تصاویر ورودی با استفاده از توابع انتقال، درونیابی خطی و میانگین‌گیری از شدت روشنایی آنها، مدل چهره از روبه‌رو را بدست می‌آورد. برای بهبود تصویر مدل، روش پیشنهادی بر روی تصاویر ورودی واگرا (تصاویر چهره دارای زاویه قرینه نسبت به‌هم) ، اعمال می‌شود. سپس تصاویر مدل شده، با تجزیه مرتبه پایین، نسبت به‌هم تراز می‌شوند. شبکه عصبی تابع پایه شعاعی به عنوان تابع انتقال برای این روش در نظر گرفته شده است. از مزایای این روش عدم نیاز به اطلاعات عمق در تصویر، کالیبره بودن تصاویر و اطلاعات زاویه چهره است که معمولا در مدل‌سازی سه بعدی استفاده می‌شود. عملکرد الگوریتم روی تصاویر پایگاه دادهPRIMAمورد بررسی قرار گرفته است. با فرض اینکه تصاویر چهره تنها دارای انسداد زاویه‌ای باشند، نتایج آزمایشات نشان می‌دهد که روش پیشنهادی با دقت مناسبی تصویر نمای روبه‌روی چهره را مدل‌سازی می‌نماید.

کلمات کلیدی: انسداد زاویه‌ای، دگردیسی، تابع پیچش، ترازسازی، شبکه عصبی تابع پایه شعاعی.

تاریخ ارسال مقاله: ۱۳۹۵/۱۱/۷

تاریخ پذیرش مشروط مقاله: ۱۳۹۶/۰۸/۲۳

تاریخ پذیرش مقاله: ۱۳۹۶/۱۰/۱۶

نام نویسنده‌ی مسئول: دکتر وحید ابوالقاسمی

نشانی نویسنده‌ی مسئول: ایران- شاهرود- دانشگاه صنعتی شاهرود- دانشکده مهندسی برق و رباتیک

[Downloaded from jiaeee.com on 2026-11-20]

[DOR: 20.1001.1.26765810.1397.15.2.6.6]

Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018

۱- مقدمه

امروزه به دلیل کاربردهای گسترده از جمله تأمین امنیت و ثبت ورود و خروج کارکنان در سازمان‌ها، نیاز به شناسایی هویت بیش از پیش احساس می‌شود. از بین روش‌های شناسایی هویت، شناسایی چهره به دلیل دخالت کمتر انسان، نسبت به سایر روش‌ها بیشتر مورد توجه می‌باشد[۱].

هر سیستم شناسایی چهره را می‌توان متشکل از دو بخش استخراج چهره و شناسایی چهره دانست. علی‌رغم پیشرفت‌های قابل توجه، این سیستم‌ها همچنان با مشکلاتی روبرو هستند. زاویه دید، تغییر شدت روشنایی محیط، حالت‌های مختلف چهره، تغییرات مقیاس، تاری تصاویر و انسداد[۱] از جمله این چالش‌ها هستند[۲].

طی دو دهه اخیر، روش‌های مختلفی برای شناسایی چهره ارائه‌شده‌اند. Kondo و همکارانش در[۳]، روشی پنج مرحله‌ای برای سیستم شناسایی چهره ارائه داده‌اند که در برابر شدت روشنایی و پس‌زمینه‌های پیچیده مقاوم می‌باشد. مراحل این روش شامل تبدیل موجک با توابع پایه هار[۲] تصویر چهره، لبه‌یابی بخش‌های صورت، پیدا کردن محور تقارن آن، تشخیص چهره و در نهایت شناسایی آن است. اینگونه سیستم‌هابه جهت داشتن فرضیات زیاد،کاربردی محدود دارند.

در سیر تکاملی سیستم‌ها، شاهد ظهور روش‌های مبتنی بر مدل ظاهر- فعال (AAM)[۳]هستیم[۴]. این روش‌ها قادر به استخراج ویژگی برای شناسایی با دریافت یک تصویر چهره، تحت شرایط نوری و زوایای مختلف می‌باشد. اما پیچیدگی مدل و تنظیم پارامترهای زیاد از مشکلاتی است که روش‌های بر پایه AAM از آن رنج می‌برند[۴].

برای اخذ نتایج بهتر در بخش شناسایی، معمولا پیش‌پردازش و پس‌پردازش‌هایی به ترتیب بر روی تصویر ورودی یا خروجی سیستم استخراج چهره صورت می‌گیرد [۵]. این روش‌ها نرمال‌سازی تصاویر چهره نامیده می‌شوند. روش‌های نرمال‌سازی جهت ارائه تصویری واحد از مجموعه تصاویر چهره شخص، به منظور رفع محدودیت‌هایی مانند تغییرات شدت روشنایی، وجود انسداد و حالات مختلف چهره ارائه می‌شوند. هدف این روش‌ها ساده‌سازی سیستم شناسایی و حذف فرضیات متعدد آن است. برای مثال Guo و همکارانش در [۶]، در ابتدا با استفاده از کلاس‌بندی برپایه بازنمایی[۴] تنک، یک واژه‌نامه[۵] را برای شناسایی چهره آموزش می‌دهد. پس از استخراج چهره، از اختلاف میان چهره بازسازی شده توسط واژه‌نامه و چهره اصلی که به آن باقیمانده بازسازی[۶] گفته می‌شود ماسک دودویی برای از بین بردن انسداد در تصویر چهره استخراج می‌نماید.

شدت روشنایی غیریکنواخت در تصویر چهره، باعث پنهان شدن ویژگی‌های چهره می‌شود و نرخ شناسایی را کاهش می‌دهد. Shan و همکارانش در[۷]، با استفاده از تکنیک‌های برابر سازی هیستوگرام، تصحیح شدت روشنایی گاما (GIC)[۷] و بازسازی باقی‌مانده شدت روشنایی (QIR)[۸] به نرمال‌سازی شدت روشنایی می‌پردازد. با بهبود روش نرمال‌سازی شدت روشنایی، Juefei و همکارانش در [۸]، ابتدا تصاویر چهره را به حوزه الگوی دودویی محلی[۹](LBP) نگاشت می‌کنند. سپس در طی یک نگاشت معکوس از فضای LBP به فضای پیکسل، که از طریق یادگیری واژه‌نامه توامان میان دو حوزه مذکور بدست می‌آید، نرمال‌سازی انجام می‌شود. حالات مختلف چهره نیز فاکتوری تاثیرگذار در نرخ شناسایی چهره است. Arad و همکارانش در[۹]، با استفاده از تکنیک پیچش تصویر به یکسان‌سازی حالات چهره می‌پردازد.

از دیگر نرمال‌سازی‌ها که قادر به تصحیح حالات چهره، انسداد و شدت روشنایی‌هایی متفاوت است، می‌توان به روش ترازسازی مقاوم با تجزیه مرتبه پایین (RASL)[۱۰] اشاره نمود[۱۰]. روشRASL به ترازسازی دسته‌ای از تصاویر می‌پردازد که با هم همبستگی خطی دارند. در این روش، مدل $I_i = \left(I_i^0 + e_i\right) \circ T_i^{-1}$ به تصاویر نسبت داده شده است که در آن، I_i تصویر در دنیای واقعی، I_i^0تصویر با مرتبه پایین[۱۱]، e_i خطای تجزیه تنک[۱۲] و T_i^{-1} تابع انتقال است. تابع هزینه[۱۳] در این مدل به صورت بهینه‌سازی توامان نامحدب[۱۴] با قیدی غیر خطی[۱۵] بر روی مجموعه انتقالات و خطای تنک هر تصویر تعریف می‌شود. برای یافتن پاسخ (تصاویر I_i^0)، چارچوبی مبتنی بر الگوریتم‌های تقریبی و تجزیه مرتبه پایین، جهت بهینه‌سازی تابع هزینهRASL ارائه شده است[۱۱,۱۰].

در روش‌های نرمال‌سازی، تصحیح زاویه برای چهره کمتر مشاهده می‌شود. در این مقاله روش پیش‌پردازشی برای سیستم شناسایی چهره ارائه شده است که با دریافت چند تصویر زاویه‌دار از چهره یک شخص، تصویر چهره روبه‌رو ازآنرا مدل نماید. معمولا در تصاویر ویدئویی کنترلی بر روی زاویه چهره فرد نسبت به دوربین وجود ندارد. اما تعداد زیادی از این تصاویر زاویه‌دار، توسط سیستم ویدئوی قابل استخراج است.

در تهیه تصاویر مختلف از یک صحنه، میان اشیای درون صحنه و همچنین صفحه تصویر ممکن است سه نوع زاویه ایجاد شود. در نوع اول، شیء موجود در تصویر ساختار خود را حفظ نموده و نسبت بین اجزای آن ثابت باقی می‌ماند. تصحیح این نوع زاویه تحت تبدیل نسبی[۱۶] امکان‌پذیر است [۱۳,۱۲]. نوع دوم هنگامی رخ می‌دهد که صحنه تصویر هم‌راستا با صفحه دوربین نباشد. در این نوع از زاویه، هنگام نگاشت تصویر از فضای سه بعدی به دو بعدی، ساختار اشیا حفظ می‌شود، اما ممکن است نسبت اجزای شیء به‌هم و اندازه آنها ثابت باقی نماند. این نوع زاویه تحت تبدیل معکوس پرسپکتیو[۱۷] قابل تصحیح است[۱۴]. روش RASL قادر به تصحیح تبدیلات نسبی و پرسپکتیو است. نوع سوم زاویه با چرخش شی در محیط به وجود می‌آید. چون بخشی از شی در تصویر پنهان می‌شود، تصحیح نوع سوم زاویه دشوار است. شکل (۱) انواع زاویه‌های مطرح شده را نشان می‌دهد.

Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer2018
[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]



الف ب پ

شکل(۱): انواع زاویه‌های موجود در تصویر. الف: زاویه نسبی. ب: زاویه پرسپکتیو. ج: زاویه چهره حاصل از چرخش سر.

در تصاویر ویدئویی، امکان ایجاد زاویه از نوع اول و سوم در چهره، به دلیل چرخش سر وجود دارد. اگر دوربین تصویربرداری قادر به کالیبره‌سازی[۱۸] تصاویر باشد یا تصاویر استریو[۱۹]موجود باشد، روش‌هایی وجود دارند که با استفاده از مدل‌های سه‌بعدی قادر به مدل‌سازی تصویر از روبه‌رو برای چهره هستند [۱۶,۱۵]. اما به دلیل هزینه‌بر بودن تامین دوربین‌هایی که قادر به تولید این نوع از تصاویر هستند، کاربرد روش‌های مبتنی بر مدل‌سازی سه‌بعدی، در نظارت ویدئویی محدود می‌شود. لازم به ذکر است که روش‌های مبتنی بر مدل سه‌بعدی معمولا بلادرنگ نبوده و در مدل‌سازی آنها، امکان بازسازی بخش‌هایی از چهره وجود دارد که از تصاویر ورودی قابل استخراج نیستند. بنابراین ممکن است که مدل چهره ارائه شده توسط این روش‌ها با چهره شخص تفاوت‌هایی داشته باشد.

روش پیشنهادی در این مقاله، با مدل‌سازی چهره از تصاویر زاویه‌دار، به تصحیح زاویه‌های نوع اول و سومِ چهره در تصاویر ویدئویی خواهد پرداخت. این روش، بدون در نظر گرفتن فرض‌های سه‌بعدی و تنها با استفاده از تصاویر ساده دوبعدی، مدل‌سازی چهره از روبه‌رو را بهبود می‌بخشد.

در ادامه مقاله در بخش دوم به مفاهیم و الگوریتم‌های مورد نیاز در مقاله می‌پردازیم. در بخش سوم، الگوریتم پیشنهادی مطرح می‌شود. بخش چهارم آزمایشات و نتایج کمی و کیفی روش پیشنهادی نمایش داده می‌شوند. در بخش پنجم نیز به نتیجه‌گیری می‌پردازیم.

۲-۱- مفاهیم و الگوریتم ها

در این بخش، الگوریتم‌های مورد نیاز برای روش پیشنهادی در این مقاله، مورد بحث قرار می‌گیرند. روش پیشنهادی از دو تکنیک استفاده می‌کند. تکنیک اول با انتقال مکان پیکسل‌های تصویر، بدون تغییر در شدت روشنایی آنها، تصویر را تغییر می‌دهد. این تکنیک پیچش[۲۰] تصویر نام دارد. تکنیک دوم،با دریافت دو تصویر ورودی به نام‌های تصویر مبدا و مقصد، تصویر مبدا را به صورت تدریجی به سمت تصویر مقصد تغییر می‌دهد که حاصل آن دنباله‌ای از تصاویر میانی است و به دگردیسی[۲۱] معروف است. تکنیک پیچش، بخشی از فرایند دگردیسی است. در ادامه به صورت خلاصه به توصیف این دو تکنیک می‌پردازیم.

۱-۱- پیچش تصویر

فرض کنید تصویری مانند $I \in R^{w \times h}$ به عنوان تصویر ورودی داریم و $w \in W = \{T_i(x,y) : \mathbb{R}^2 \to \mathbb{R}, i \in \{u,v\}\}$ مجموعه توابع انتقال بر روی پیکسل‌های تصویر باشد. فرایند پیچش تصویر با رابطه (۱) تعریف می‌شود.

$$\hat{I}(u,v) = w \circ I(x,y) = I(T_u(x,y), T_v(x,y)) \quad (۱)$$

که $\hat{I}$ تصویر انتقال یافته و عملگر $\circ$ نشان‌دهنده انتقال مکانی پیکسل‌های تصویر توسط مجموعه توابع انتقال w است.

در پیچش تصویر ممکن است که مکان چند پیکسل به یک نقطه نگاشت شود، یا یک پیکسل به خارج از محدوده تصویر انتقال یابد. با توجه به محدود بودن تعداد پیکسل‌های تصویر، ممکن است که به تعدادی از پیکسل‌ها مقداری به عنوان شدت روشنایی تعلق نگیرد. این پیکسل‌ها باعث به وجود آمدن مصنوعاتی در تصویر انتقال یافته می‌شوند[۲۲]. برای حذف مصنوعات به وجود آمده، با معکوس نمودن توابع پیچش (با فرض یک‌به‌یک بون آنها) و مکان‌یابی پیسکل‌های تصویر خروجی در تصویر ورودی، انتقال به صورت معکوس انجام می‌شود. در انتقال معکوس، برای پیکسل‌هایی که به بیرون از صفحه تصویر نگاشت می‌شوند، می‌توان از مکاشفه[۲۳] ساده نگاشت به نزدیکترین پیکسل استفاده نمود. همچنین می‌توان با درونیابی از شدت روشنایی پیکسل‌های همسایه، به پیکسل‌هایی شدت روشنایی به آنها تعلق نگرفته، عددی نسبت داد و با این روش اثر مصنوعات را از بین برد [۱۴]. از دیگر کاربردهای پیچش تصویر می‌توان به نرمال‌سازی، و تطبیق تصاویر اشاره نمود[۱۷]. در اکثر انتقال‌ها، نقاطی بر روی تصاویر انتخاب می‌شوند و تابع انتقالِ پیچش از این نقاط استخراج می‌شود. یکی از روش‌های استخراج توابع انتقال از نقاط، درونیابی می‌باشد[۹].

۲-۱- دگردیسی تصویر

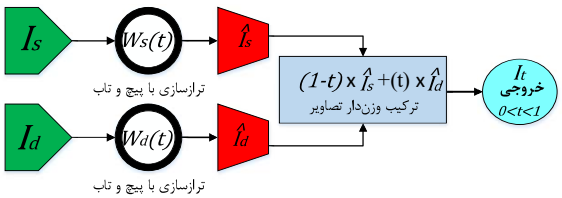
فرایند دگردیسی متشکل از سه مرحله می‌باشد: ۱) استخراج توابع انتقال از نقاط ایجاد شده. ۲) اعمال توابع انتقال بر روی تصاویر مبدا و مقصد برای ترازسازی مکان نقاط. ۳) ترکیب وزن‌دار تصاویر تراز شده مبدا و مقصد نسبت به پارامتر زمان [۲۰-۱۸].

بخش مهم دگردیسی، انتخاب مناسب تابع انتقال برای پیچش است. اگر w مجموعه توابع پیچش، t پارامتر زمان و I_s و I_d به ترتیب تصاویر مبدا و مقصد باشند، دگردیسی این تصاویر با I_t نشان داده شده و به صورت دیاگرام (۲) می‌باشد.

[Downloaded from jiaeee.com on 2026-11-20]

[DOR: 20.1001.1.26765810.1397.15.2.6.6]

Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018



شکل (۲): فرایند دگردیسی بر روی تصاویر مبدا و مقصد.

۳- روش پیشنهادی

روش پیشنهادی مطرح شده در این مقاله مبتنی بر دگردیسی بوده و از چند تصویر چهره (مثلا ۲۴ عدد) به عنوان تصاویر مبدا استفاده می‌نماید. در ابتدا تعدادی نقطهاز بخش‌های چهره در تصاویر مختلف ورودی انتخاب می‌شوند. فرض کنید تعداد این نقاط n عدد در هر تصویر باشد و m تصویر ورودی نیز در اختیار باشد. برای ترازسازی تصاویر مبدا نسبت به‌هم، نیاز است تا نقاط انتخاب شده در تصاویر مبدا باهم دارای تطابق باشند. به این معنی که مجموعه نقاط نشان دهنده یک بخش خاص از تصویر چهره فرد (مثلا چشم راست) درتمامیتصاویر مبدا، در یک بردار سازمان‌دهی شوند. طول این بردارها برابر با m (تعداد تصاویر ورودی) می‌باشد و تعداد این بردارها بستگی به انتخاب نقاط در بخش‌های مختلف چهره دارد. به نقاط این مجموعه، نقاط مطابق[۲۴] می‌گوییم. نقاط مطابق در تصویر مقصد، از میانگین بردارهای نقاط مطابق تصاویر مبدا بدست می‌آیند. سپس توابع انتقال میان تصویر مقصد و هر تصویر مبدا محاسبه می‌شوند (به ازای هر تصویر مبدا یک مجموعه تابع انتقال). در نهایت با استفاده از انتقال معکوس، مکان تمامی پیکسل‌های تصویر مقصد، در تصویر مبدا محاسبه شده و مقدار شدت روشنایی آنها با درونیابی خطیبدست می‌آید. این روند برای هر تصویر ورودی انجام می‌شود و هر تصویر به اندازه $1/m$ به شدت روشنایی پیکسل‌های تصویر مقصد می‌افزاید. شکل(۳) شمای کلی روش پیشنهادی را برای چهار تصویر ورودی و سه نقطه مطابق، در سه گام خلاصه می‌نماید.

اکنون دقیق‌تر به فرضیات مساله و روش پیشنهادی می‌پردازیم. زاویه چهره را با جفت $(\pm y^{\circ} \pm x^{\circ})$ نشان می‌دهیم که در آن x و y به ترتیب میزان چرخش چهره در راستای افقی و عمودی می‌باشد. زاویه تصویر چهره را با $I^{\pm y^{\circ} \pm x^{\circ}}$ نشان می‌دهیم. مجموعه‌ای از تصاویر با اندازه یکسان به صورت $C = \left\{ i \in I^{\pm y^{\circ} \pm x^{\circ}}, x \text{ or } y \neq 0 \right\}$ به عنوان ورودی داریم. می‌خواهیم با تصاویر مجموعه C، تصویر روبه‌روی چهره شخص را مدل کنیم. در شکل (۴) یک نمونه از این مجموعه نشان داده شده است. فرض می‌کنیم که تصاویر مجموعه C فقط دارای زاویه چهره باشند. همچنین مجموعه C، خاصیت $\left\{ \forall i : i^{-y^{\circ}-x^{\circ}} \in C \Rightarrow i^{+y^{\circ}+x^{\circ}} \in C \right\}$ (قرینگی نسبت به زاویه) را دارد.

همانطور که پیش‌تر بیان شد یکی از روش‌های استخراج توابع انتقال از مجموعه نقاط مطابق، درونیابی می‌باشد. شبکه‌های عصبی تابع پایه شعاعی[۲۵] (RBF) ابزاری قدرتمند جهت درونیابی داده‌های چند متغیره پراکنده[۲۶] به شمار می‌روند. برای تابع انتقال، از معماری شبکه عصبی تابع پایه شعاعی معرفی شده در [۹] استفاده می‌کنیم. این تابع با رابطه (۲) نشان داده شده است.

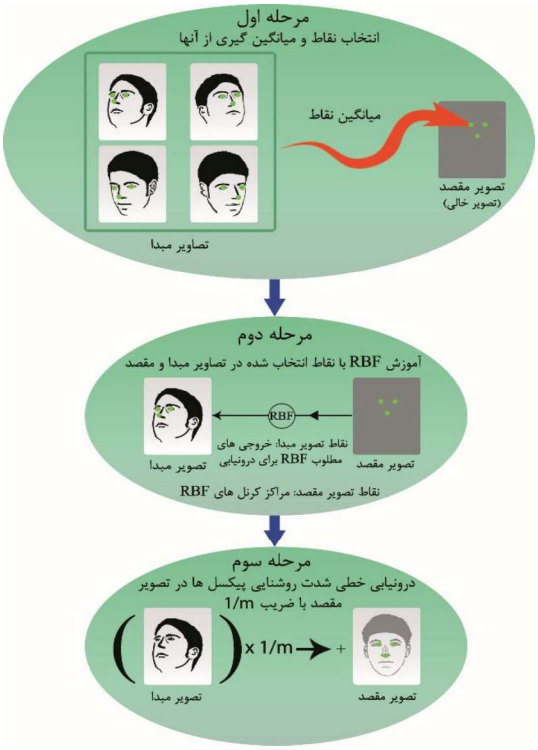
$$RBF(x,y): \mathbb{R}^2 \to \mathbb{R} = \sum_{i=1}^{n} \omega_i k \left\| c_i - \rho \right\|_2 + P_m(x,y),$$

$$P_m(x,y) \in \sum_{i=0}^{m} \sum_{j=0}^{m-i} a_{ij} x^i y^j \quad (۲)$$

با شرط ارضای قید زیر

$$\sum_{i=1}^{n} \omega_i = \sum_{i=1}^{n} \omega_i x_i = \ldots = \sum_{i=1}^{n} \omega_i y_i = \ldots = 0$$

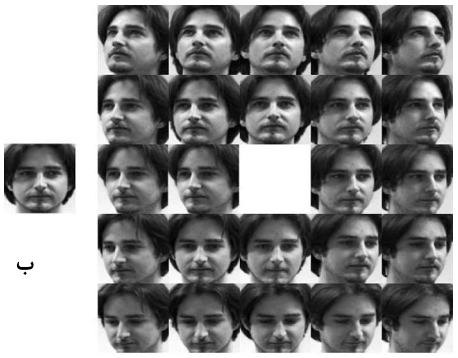
که در آن RBF شبکه عصبی، ω_i و a_{ij} وزن‌های شبکه عصبی، P_m تابع تعمیم چندجمله‌ای، m درجه تابع تعمیم، c_i مراکز هسته‌ها، $\rho = \left(\rho^x, \rho^y \right)$ مختصات پیکسل مورد نظر برای درونیابی و k تابع هسته می‌باشند. $\|.\|_2$ فاصله اقلیدسی است.



شکل(۳): خلاصه روش پیشنهادی در سه گام. گام اول: ایجاد مجموعه نقاط مطابق و میانگین‌گیری از آنها. گام دوم: آموزش تابع انتقال و استفاده از انتقال معکوس برای مکان‌یابی پیکسل‌های تصویر مدل در تصاویر مبدا. گام سوم: بدست آوردن مقدار شدت روشنایی پیکسل‌های تصویر مدل.

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]

آموزش شبکه عصبی رابطه (۲)، با عبارت تعمیم آن، باعث می‌شود که اگر فاصله نقاطی که با این شبکه درونیابی می‌شوند تا مراکز شبکه زیاد باشد، مقداری که شبکه به عنوان خروجی برای آنهاتولید می‌نماید از یک چندجمله‌ای پیروی نمایند[۹]. در روش پیشنهادی مطلوب است که تابع انتقال تاثیرات سراسری برحرکت پیکسل‌ها داشته باشد و حداقل خمیدگی را به صفحه تصویر وارد کند [۱۹,۹]. تحقیقات انجام شده توسط Arad و همکارانش، انتخاب تابع پایه شعاعی $k(x)=x^2\log x$، با $k(0)=0$ (TPS)[۲۷] به عنوان هسته شبکه عصبی و تنظیم درجه تابع P_m به ۱را برای رسیدن بهاین هدف لازم می‌داند[۹]. پس مجموعه توابع انتقال با رابطه (۳) مشخص می‌شوند.



ب

الف

شکل(۴):تصاویر مبدا و تصویر مقصد. الف:تصاویر مبدا، تصویرگوشهبالاسمتچپدارایزاویه(۳۰+۳۰+) است. میزان چرخش چهره در راستای افقی و عمودی تصاویر متوالی به اندازه ۱۵- است. ب: تصویر مقصد که در شکل (الف) جای آن خالی گذاشته شده و دارای زاویه(۰+۰+)است.

$$
\begin{aligned}
w &= \{(T_u,T_v)\in RBF(x,y)\}:\mathbb{R}^2\to\mathbb{R}^2 = \\
&(T_u(\rho)=\sum_{i=1}^{n}\omega_i k\left\|c_i-\rho\right\|_2 \\
&\quad +a_{00}+a_{01}\rho^y+a_{10}\rho^x, \qquad (۳)\\
&T_v(\rho)=\sum_{i=1}^{n}\alpha_i k\left\|c_i-\rho\right\|_2 \\
&\quad +b_{00}+b_{01}\rho^y+b_{10}\rho^x)
\end{aligned}
$$

برای بدست آوردن توابع انتقال، مجموعه بردارهای نقاط تصاویر مبدا را با $S=\{s_i^d \mid d\in\{x,y\}, i=1,2,...,m\}$ در نظر می‌گیریم. نقاط مقصد، D، نیز همین‌طور نمایش داده می‌شوند.

با تشکیل ماتریس S از بردارهای نقاط مطابق، به صورت (۴)

$$
s^x=\begin{bmatrix} s_{11}^x & \cdots & s_{1m}^x \\ \vdots & \ddots & \vdots \\ s_{n1}^x & \cdots & s_{nm}^x \end{bmatrix}, s^x=\begin{bmatrix} s_{11}^y & \cdots & s_{1m}^y \\ \vdots & \ddots & \vdots \\ s_{n1}^y & \cdots & s_{nm}^y \end{bmatrix} \quad (۴)
$$

نقاط مطابق در تصویر مقصد را با بردارهای (۵) تعریف می‌کنیم.

$$
D^x=\begin{bmatrix} \frac{\sum_{j=1}^{m} s_{1,j}^x}{m} \\ \vdots \\ \frac{\sum_{j=1}^{m} s_{n,j}^x}{m} \end{bmatrix}, D^y=\begin{bmatrix} \frac{\sum_{j=1}^{m} s_{1,j}^y}{m} \\ \vdots \\ \frac{\sum_{j=1}^{m} s_{n,j}^y}{m} \end{bmatrix} \quad (۵)
$$

فرض کنید $s_1=\{(s_{i,1}^x,s_{i,1}^y)\}_{i=1}^n$ نقاط انتخاب شده یکی از تصاویر مبدا باشد. تفاوتی در ترتیب ورود تصاویر مبدا به الگوریتم وجود ندارد. علت آن در ادامه مقاله ذکر خواهد شد. با در نظر گرفتن مجموعه D به عنوان مراکز هسته‌های توابع انتقال و با درونیابی دقیق نقاط s_1، توابع انتقال w تشکیل می‌شوند. اگر از انتقال معکوس استفاده کنیم رابطه (۶) را خواهیم داشت:

$$
\begin{aligned}
&w=(T(D))= \\
&\{(T_u(d_i),T_v(d_i))\}_{i=1}^n=\{s_i^x,s_i^y\}_{i=1}^n \qquad (۶)
\end{aligned}
$$

مطابق رابطه (۶) برای راستای x پیکسل‌های تصویر مدل سیستم خطی (۷) را خواهیم داشت:

$$
\{s_{i,1}^x,d\}_{i=1}^n=\begin{bmatrix}
T_u(d_1)=\sum_{i=1}^{n}\omega_i k\left\|d_i-d_1\right\|_2 \\
+a_{00}+a_{01}d_1^y+a_{10}d_1^x=s_{1,1}^x \\
\vdots \\
T_u(d_n)=\sum_{i=1}^{n}\omega_i k\left\|d_i-d_n\right\|_2+ \\
a_{00}+a_{01}d_n^y+a_{10}d_n^x=s_{n,1}^x \\
d_1^x\omega_1+...+d_n^x\omega_n=0 \\
d_1^y\omega_1+...+d_n^y\omega_n=0 \\
\omega_1+...+\omega_n=0
\end{bmatrix} \quad (۷)
$$

به صورتی مشابه می‌توان برای مجموعه $\{s_{i,1}^y,d\}_{i=1}^n$ رابطه (۷) را نوشت. پس با داشتن n نقطه متناظر، برای ساخت توابع انتقال، نیاز به حل سیستم خطی با $n+3$ مجهول برای هر راستای مختصات داریم. برای اینکه سیستم فوق دارای پاسخ باشد، تعداد نقاط انتخابی باید حداقل سه عدد باشد.

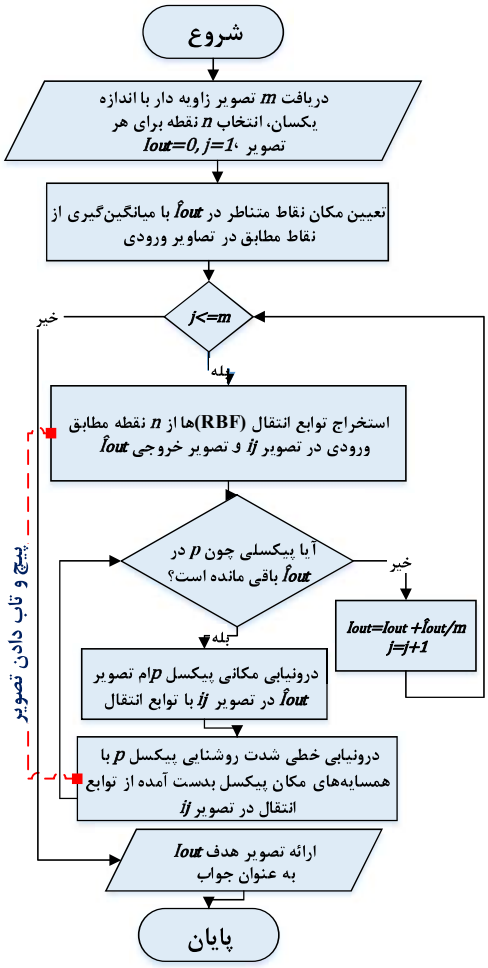
بعد از حل سیستم‌های خطی رابطه (۷) توابع انتقال بدست خواهند آمد. این الگوریتم میانگین‌گیری بر پایه دگردیسی نامیده می‌شود. شکل (۵) نمودار گردشی میانگین‌گیری دگردیسی را نمایش می‌دهد. برای دو تصویر ورودی، $m=2$، خروجی الگوریتم برابر است با دگردیسی تصویر با همان توابع انتقال در $t=0.5$. جهت ایجاد نقاط مطابق، از روش نقطه‌گذاری استفاده می‌کنیم. با استفاده از

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]
Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018

روش‌هایی که از یک مدل عمومی چهره استفاده می‌کنند، امکان خودکارسازی انتخاب نقاط وجود دارد[۲۱].

شکل (۶) نمونه‌ای از خروجی الگوریتم را برای تصاویر شکل (۴-الف) نشان می‌دهد.

برای تحلیل اثر این الگوریتم، $I^{+y^{\circ}+x^{\circ}}$ و $I^{-y^{\circ}-x^{\circ}}$ را از مجموعه تصاویر C (مجموعه ورودی که در ابتدای بخش معرفی شد) در نظر بگیرید. این تصاویر را واگرا می‌نامیم.



شکل(۵): نمودار گردشی الگوریتم پیشنهادی، میانگین‌گیری بر پایه دگردیسی.

بردارهای D در رابطه (۵) مطابق با مکان نقاط در تصویر روبه‌روی چهره می‌باشند. شکل (۷) این مطلب را نشان می‌دهد. در شکل (۷)، بر روی چهار تصویر واگرا، با زوایای (۳۰+،۳۰+)،(۳۰-،۳۰+)،(۳۰+،۳۰-)و (۳۰-،۳۰-)، یک نقطه بر روی بینی ایجاد شده است. پس مجموعه D دارای یک نقطه است.

اگر محور مختصات را بر روی این نقطه قرار دهیم، نقاط مطابق در مبدا محور مختصات تصاویر واگرا، نسبت به‌هم قرینه خواهند بود. اگر دو تصویر واگرا را به عنوان ورودی الگوریتم در نظر بگیریم، در گذر از یک تصویر واگرا به تصویر واگرای دیگر، یک تصویر چهره از روبه‌رو که حداکثر دارای مقداری زاویه از نوع اول مانند شکل (۱-الف) است، ایجاد می‌شود [۲۲]. این تصویر روبه‌رو در میانگین مکان نقاط مطابق دو تصویر واقع می‌شود. چون جابه‌جایی در میانگین‌گیری تاثیری ندارد، می‌توانیم به ورودی‌های الگوریتم به صورت دوبه‌دو واگرا نگاه کنیم. در این صورت اثر الگوریتم میانگین‌گیری دگردیسی بر روی مجموعه تصاویر ورودی، به میانگین مجموعه تصاویر روبه‌روی بدست آمده از تصاویر مبدا و مقصد واگرا، توجیه می‌شود.

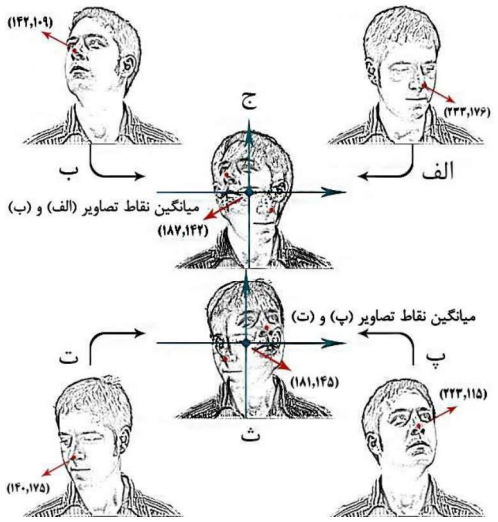


شکل (۶): نتایج حاصل ازاعمال الگوریتم پیشنهادی. میانگین‌گیری برروی تمامی تصاویرشکل(۴-الف).دراین آزمایش درهرتصویر چهره شخص با ایجاد ۲۰ نقطه متناظر (مجموعا ۴۸۰ نقطه) مجموعه D تشکیلشد.

چون زاویه نوع سوم در چهره مانند شکل (۱-پ) از چرخش سر حول محور گردن به وجود می‌آید، نحوه جابه‌جایی نقاط مختلف چهره با توجه به فاصله آنها تا محور گردن مشخص می‌شود. در میان مجموعه نقاط مطابق، آنهایی که دارای فاصله بیشتری از هم هستند، بر روی مداری دورتر از محور گردن خواهند بود (برای مثال در چرخش سر بر محور گردن در راستای افقی و عمودی، میزان جابه-جایی نوک بینی بیشتر از چشم یا گوش می‌باشد). پس رابطه‌ای مستقیم میان فواصل نقاط مطابق و فاصله آنها با محور گردن وجود دارد. در انتقال با استفاده از شبکه عصبی، فاصله بین نقاط مطابق، رابطه مستقیم با اندازه حرکت نقطه در تصویر مدل دارد.به عبارت دیگر در تصویر مدل، نقاط دور از هم، بیشتر از نقاط نزدیک به‌هم تغییر مکان می‌دهند. اگر تابع انتقال به خوبی عمل درونیابی را تحت تابعی هموار یاد بگیرد به منزله این است که الگوی اندازه جابه‌جایی نقاط مطابق نسبت به‌هم را آموزش دیده است. همچنین چون این شبکه عصبی، تابعی پیوسته می‌باشد، تصاویرمیانی که تولید می‌کند، دارای زاویه‌های مختلف نسبت به دوربین می‌باشند.برای نشان دادن نحوه حرکت نقاط، یک فرایند دگردیسی با دیاگرام (۲) و استفاده از توابع انتقال رابطه (۲) تشکیل دادیم و دو تصویر چهره با زوایای(۰-،۳۰-) و (۰+،۳۰+) را به ترتیب به عنوان تصاویر مبدا و مقصد دگردیسی در نظر گرفتیم.شکل (۸) نحوه حرکت نقاط در تصاویر میانی تولید شده توسط فرایند دگردیسی را نشان می‌دهد. در شکل (۸) تصاویر (پ۱)، (پ۲) و (پ۳) بدست آمده از دگردیسی به ترتیب در $t=1/4$، $t=1/2$ و $t=3/4$ نشان داده شده است. با توجه به نحوه انتقال توابع پیچش رابطه (۲) انتظار می‌رود تصاویر تولید شده به ترتیب دارای زوایای (۱۵-،۰+)، (۰+،۰+) و (۱۵+،۰+) باشند. به همین خاطر تصاویر چهره شخص از پایگاه داده با همان زاویه‌ها، برای مقایسه در شکل (۸) آورده شده‌اند. جدول (۱) مختصات دو نقطه مطابق قرار

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]
Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer2018

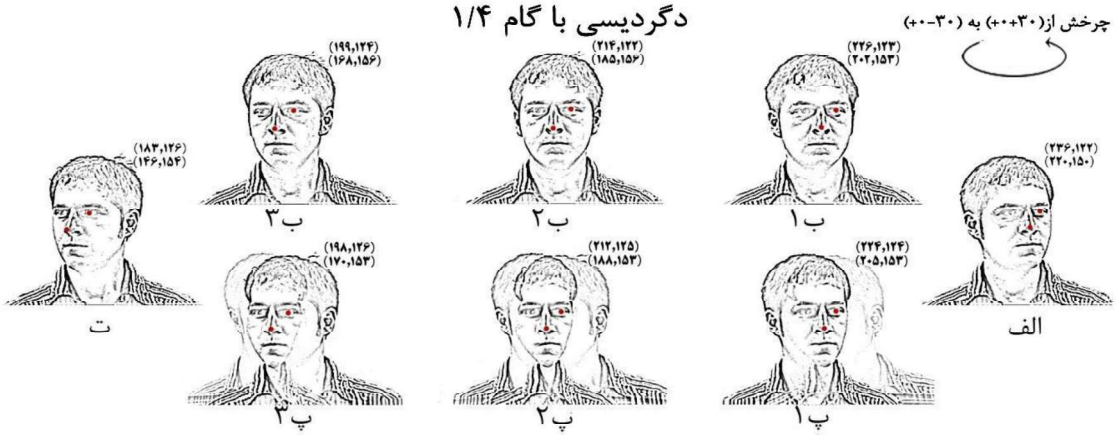
داده شده بر روی بینی و چشم راست شخص در تصاویر شکل (۸) را نشان می‌دهد.



شکل (۷): چهار تصویر الف، ب، پو ت به ترتیب با زوایای (۳۰-۳۰-)، (۳۰+۳۰)، (۳۰-۳۰+) و (۳۰+۳۰-). تصاویر الف وب جفت واگرای هم هستند. تصاویر پ و ت نیز همین‌طور. از روی مختصات بینی در تصاویر ث و ج که به ترتیب مکان نگاشت نقطه در تصویر مدل برای جفت‌های (الف و ب) و (پ و ت) را نشان می‌دهد، مشخص است که بینی در خروجی به صورت تقریبی، به یک نقطه نگاشت می‌شود و حذف دو تصویر واگرا از مجموعه باعث جابه‌جایی نقاط مقصد، D، نخواهد شد.

به دلیل استفاده از مکاشفه ساده برای پیکسل‌هایی که به بیرون از صفحه تصویر نگاشت می‌شوند، عدم اطلاع از تعداد نقاط مورد نیاز برای استخراج توابع انتقال دقیق‌تر و اینکه نمی‌توان در بخش‌های پوشیده شده، نقاط مطابق ایجاد نمود، به وجود آمدن اثرات روح در تصویر خروجی اجتناب ناپذیر است[۱۸].

ایجاد نقاط مطابق بر روی تعداد زیادی از تصاویر مبدا نیز (به علت عدم دقت در انتخاب مکان آنها) باعث به وجود آمدن مقدار زیادی خطای مکانی می‌شود. برای بهبود تصویر مدل، اگر تصویر مدل را از تصاویر واگرا بدست آوریم، مجموعه‌ای از تصاویر روبه‌رو خواهیم داشت. در این حالت نیزتاثیر درونیابی میانگین‌گیری دگردیسی، مشابه حالات مختلف چهره بوده و دارای اثرات روح نیز هست. در بخش مقدمه راجع به روش ترازسازی RASL بحث شد[۱۰]. روش RASL مقاومتی در مقابل زاویه چهره ندارد. شکل (۹) نمایانگر خروجی این الگوریتم برای تصاویر ورودی شکل (۴-الف) است. اگر تصاویر بدست آمده به صورت مطرح شده در بالا را با روش RASL نسبت به‌هم تراز کنیم، اثرات به وجود آمده در خروجی میانگین دگردیسی، خطا محسوب شده و تصاویر مرتبه پایینی از روبه‌رو بدون اثرات هاله‌ای و حالات مختلف بدست می‌آید. شکل (۱۰) نشان‌دهنده نتایج میانگین دگردیسی بر روی تصاویر واگرا و سپس ترازسازی آنها با روش RASL نسبت به‌هم است. پارامترهای روش RASL برای تولید تصاویر شکل (۹) و (۱۰) یکسان تنظیم شده است. با این روش، انسداد زاویه‌ای چهره به حالت‌های چهره و خطاهایی با انرژی پایین تبدیل شده و قابلیت تصحیح با روش RASLرا دارد.شکل(۱۱) نتیجه ترازسازی با روش RASL، میانگیندگردیسی و ترکیب میانگین دگردیسی و ترازسازی با روش RASL را به همراه تصویر روبه‌روی پایگاه داده شخص برای مقایسه کیفی نمایش می‌دهد.



شکل (۸): نمایش نحوه حرکت نقاط در تصاویر میانی دگردیسی با تابع انتقال (۳). الف و ت به ترتیب تصاویر مبدا و مقصد هستند. در این شکل ب۱ ،ب۲و ب۳به ترتیب تصاویر با زوایای (۱۵-۰+)، (۰+۰+) و (۱۵+۰+) از پایگاه داده هستند. انتظار داریم در تصاویر ب۱ ،ب۲و ب۳ و تصاویر متناظر آنها یعنی پ۱ ،پ۲و پ۳ که به ترتیب حاصل دگردیسی در $t=1/4$، $t=1/2$، $t=3/4$هستند، نقاط اختلاف مکانی ناچیزی داشته باشند.

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]
Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018

تصاویر	ب۱ / پ۱	فاصله میان نقاط به پیکسل	ب۲ / پ۲	فاصله میان نقاط به پیکسل	ب۳ / پ۳	فاصله میان نقاط به پیکسل
مختصات چشم	(۲۲۶،۱۲۳) (۲۲۴،۱۲۴)	(۲،۱)	(۲۱۴،۱۲۲) (۲۱۲،۱۲۵)	(۲،۳)	(۱۹۹،۱۲۴) (۱۹۸،۱۲۶)	(۱،۲)
مختصات بینی	(۲۰۲،۱۵۳) (۲۰۵،۱۵۳)	(۳،۰)	(۱۸۵،۱۵۶) (۱۸۸،۱۵۳)	(۳،۳)	(۱۶۸،۱۵۶) (۱۷۰،۱۵۳)	(۲،۳)
فاصله اقلیدسی میان نقاط بینی	۲،۲۳۶		۳،۶۰۶		۲،۲۳۶	
فاصله اقلیدسی میان نقاط چشم	۳		۴،۲۴۲		۳،۶۰۶	

جدول(۱). مقایسه مکان نقاط در تصاویر متناظر. در این جدول، مختصات نقاط تصاویر متناظر، در شکل(۸) برای مقایسه در یک ستون درج شده است. در این جدول اختلاف نقاط دو راستای محور مختصات و همچنین فاصله اقلیدسی آنها درج شده است.

۴- آزمایش

در این بخش در ابتدا پایگاه داده استفاده شده در مقاله مورد بررسی قرار می‌گیرد. سپس به پیکربندی روش پیشنهادی پرداخته می‌شود. در آخر نحوه آزمایشات بیان می‌شود و مقایسات کمی جهت ارزیابی روش نشان داده می‌شوند.



ب

الف

شکل (۹): نتایج حاصل ازاعمال الگوریتم RASL برروی تمامی تصاویرشکل (۴-الف). الف:تصاویرمرتبه پایین (I_i^0)، ب: میانگین تصاویرالف.



ب الف

شکل (۱۰): تصاویر خروجی میانگین‌گیری دگردیسی و تصاویر مرتبه پایین الگوریتم RASL. الف: تصاویر مبدا شکل (۴-الف)، دوبه‌دو میانگین دگردیسی آنها تولید شده و حاصل آن ۱۲ تصویر است. به عنوان مثال تصویر بالا سمت چپ حاصل میانگین دگردیسی تصاویر واگرای (۳۰+۳۰+) و (۳۰-۳۰-) است. ب: تصاویر مرتبه پایین بعد از ترازسازی با الگوریتم RASL.



شکل (۱۱): مقایسه‌ای میان خروجی روش‌های مطرح شده. تصویر بالا سمت چپ میانگین تصاویر تراز شده با الگوریتم RASL با تصاویر ورودی شکل (۴-الف) است. تصویر بالا سمت راست خروجی میانگین‌گیری دگردیسی بر روی همان تصاویر ورودی است. تصویر پایین سمت چپ: میانگین تصاویر شکل (۱۰-ب) است. تصویر پایین سمت راست تصویر روبه‌روی پایگاه داده است.

۴-۱- پایگاه داده

در این مقاله از پایگاه دادهPRIMA استفاده شده است[۲۳]. این مجموعه داده شامل ۲۷۹۰ تصویر چهرهاز پانزده نفر است. از هر شخص ۱۸۶ تصویر با تغییرات زاویه‌ای در دو راستای x و y در محدوده (۹۰+،۹۰-) و با گام ۱۵درجه در هر دو راستا، موجود می‌باشد. مرکز چهره به همراه طول و عرض آن در پایگاه داده ارائه شده است.در

[Downloaded from jiaeee.com on 2026-11-20]

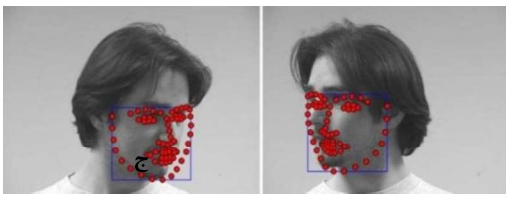
[DOR: 20.1001.1.26765810.1397.15.2.6.6]

Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer2018

این پایگاه داده ذکر شده که اندازه‌گیری زاویه‌های چهره دارای مقداری خطا بوده و دقیق نمی‌باشد.

۴-۲- تنظیم پارامترها برای روش پیشنهادی

در ابتدا به پیکربندی الگوریتم RASLمی‌پردازیم. در روش RASLپارامتر λ مصالحه‌ای میان مرتبه ماتریس و میزان تنکی خطا برقرار می‌کند.Peng و همکارانش برای این پارامتر مقدار $1/\sqrt{m}$ که m تعداد پیکسل‌های تصویر است را پیشنهاد دادند [۱۰]. در این پژوهش نیز پارامتر λ به همان مقدار تنظیم شده است. اندازه تصاویر چهره پس از استخراج به 100×100تنظیم شدند. روش RASLترازسازی تصاویر را نسبت به دو نقطه کانونی و فاصله میان آنها، در یک قاب کانونی انجام می‌دهد. برای انجام آزمایش، ابعاد قاب کانونی را100×100قرار دادیم. نقاط کانونی را به گوشه‌های خارجی چشم تنظیم نموده و فاصله کانونی این نقاط را ۸۳ پیکسل قرار دادیم. ازدیگر پارامترهای روشRASL تعیین نوع تابع انتقال است. از مجموعه انتقال‌های نسبی، تشابه[۲۸] و هموگرافی صفحه‌ای[۲۹] که به عنوان گروه‌های انتقال برای روش RASLدر نظر گرفته شده است، تابع انتقال روش RASLبه انتقال نسبی تنظیم شد. تعداد تصاویر ورودی الگوریتم برای تمامی اشخاص، ۲۴ تصویر با تغییر زاویه از (۳۰-۳۰-) تا (۳۰+۳۰+) مانند شکل(۴-الف) در نظر گرفته شد.



شکل (۱۲):دو نمونه از چهره که روش‌های استخراج چهره [۲۴] و استخراج ویژگی‌های چهره [۲۵] دارای خطای قابل ملاحضه‌ای درخروجی‌های خود می‌باشند. کادر آبی نشان‌دهنده نتیجه استخراج با روش [۲۴]. چهره و نقاط قرمز نشان‌دهنده روش [۲۵]برای استخراج ویژگی‌های چهره می‌باشند. این دو چهره از راست به چپ دارای زوایای (۴۵+۰+) و (۴۵-۳۰-) می‌باشند.

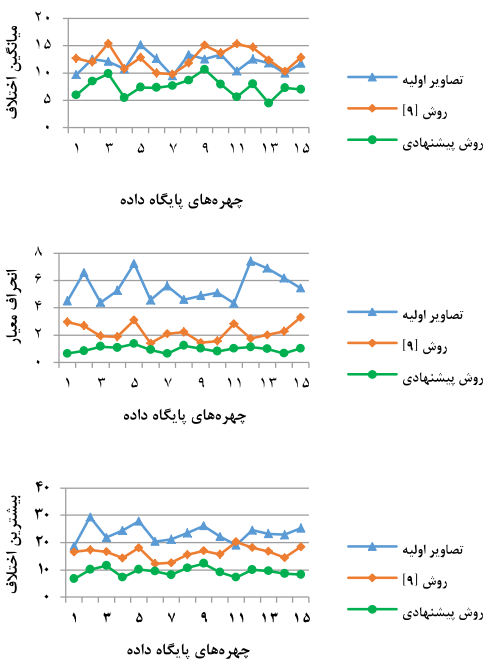
لازم به ذکر است که روش ارائه شده در این مقاله قابلیت خودکار سازی را دارد. از جمله عوامل تاثیرگذار در تعیین بازه زاویه چهره به صورت مطرح شده، محدودیت روش‌های استخراج چهره و همچنین عدم توانایی روش‌های استخراج ویژگی از چهره هنگام مواجه با چالش زاویه چهره می‌باشد. به عنوان مثال در شکل (۱۲) از روش[۲۴] برای استخراج چهره که در حضور چهره‌های زاویه‌دار عملکردی مناسب داشته، استفاده شده است. همچنین از روش [۲۵] برای استخراج ویژگی استفاده شده است. همانطور که مشاهده می‌شود این روش‌ها در مواجه با چهره‌های دارای زاویه خارج از محدوده مطرح شده، نتایجی همراه باخطای زیاد را در بر دارند.در این مقالهبرای استخراج چهره، از مرکز چهره با طول و عرض آن استفاده شد. ایجاد نقاط مطابق در تصاویر مبدا، به صورت دستی انجام شد و از نواحی گوشه‌های چشم، ابروها، گوشه‌های دهان، بینی و حاشیه صورت استفاده نمودیم.



شکل (۱۳):مقایسه بصری خروجی تصاویر با روش‌های مختلف وتصاویر روبه‌رو. الف: میانگین معمولی تصاویر. ب: نتیجه ترازسازی با استفاده از روشRASL.[۱۰] ج: تصاویر بدست آمده از روش پیشنهادی. د: تصاویر روبه‌روی پایگاه داده.

[Downloaded from jiaeee.com on 2026-11-20]

[DOR: 20.1001.1.26765810.1397.15.2.6.6]



شکل (۱۴): نمودارهای میانگین، انحراف معیار و بیشترین اختلاف از چهره روبه‌روی هر فرد و خروجی روش‌های مختلف، برای نقاط گوشه‌های دهان و نوک بینی.

۴-۳- چارچوب آزمایش و نتایج کمی

۴-۳-۱- مقایسه کمی با روش‌های مختلف

در این بخش، در ابتدا تصاویر مدل شده با روش پیشنهادی و همچنین روش RASL مورد مقایسه کیفی و کمی قرار می‌گیرد[۳۰]. جهت انجام مقایسات، از چارچوب‌های مطرح شده در [۱۰] استفاده می‌کنیم. شکل (۱۳) نتایج سه روش میانگین تصاویر، ترازسازی با[RASL۱۰] و روش پیشنهادی را نشان می‌دهد. شکل (۱۴) نیز آماره‌های میانگین، انحراف معیار و بیشترین اختلاف باچهره روبه‌روی هر فرد را برای سه نقطه گوشه‌های دهان و نوک بینی، برای تصاویر ورودی، تصاویر تراز شده با روش[RASL۱۰] و همچنین روش پیشنهادی این مقاله نشان می‌دهد. از فاصله اقلیدسی جهت سنجش فاصله بین نقاط استفاده شده است. در شکل (۱۴) نمودارها نشان‌دهنده این نکته هستند که روش پیشنهادی در تمامی چهره‌ها کاریی به مراتب بهتر از میانگین معمولی و روش [۱۰] دارد. همچنین اختلاف کمی که میان نتایج روش [۱۰] و میانگین معمولی تصاویر وجود دارد حاکی از این است که این روش قادر به رفع انسداد زاویه‌ای در تصاویر چهره نیست.

۴-۳-۲- سنجش زاویه چهره

روش‌های زیادی برای تخمین زاویه چهره وجود دارند [۲۷,۲۶]. برای تخمین زاویه چهره از روش مبتنی بر رگرسیون احتمالی با ابعاد بالا استفاده نمودیم[۲۶].پس از آموزش این روش، بر روی پایگاه داده PRIMA، زاویه تصاویر مدل شده سنجیده شد. جدول (۲) نشان‌دهنده زاویه چهره اشخاص است که با این روش تخمین زده شده است. همانطور که از جدول (۲) مشخص است زاویه چهره برای اشخاص اول، دوازدهم و پانزدهم به ترتیب (۵-،۰+)، (۵+،۰+)، (۵-،۰+) و برای باقی اشخاص پایگاه داده (۰+،۰+) تخمین زده شده است.

۴-۳-۳- مقایسه تصویر مدل با چهره روبه رو

در این آزمایش، میزان شباهت میان تصویر مدل شده به‌همراه تعدادی تصاویر چهره از همان شخص با تصویر روبه‌رو سنجیده می‌شود.

$$E=(I_1^{w\times l},I_2^{w\times l})=\frac{1}{w\times l}\sum_{i=1}^{w}\sum_{j=1}^{l}\left|I_{1_{i,j}}-I_{2_{i,j}}\right| \quad (۸)$$

جدول(۲): زوایای بدست آمده برای چهره‌های مدل با روش[۲۶]

شخص	زاویه	شخص	زاویه	شخص	زاویه
اول	(۵-،۰+)	ششم	(۰+،۰+)	یازدهم	(۰+،۰+)
دوم	(۰+،۰+)	هفتم	(۰+،۰+)	دوازدهم	(۵+،۰+)
سوم	(۰+،۰+)	هشتم	(۰+،۰+)	سیزدهم	(۰+،۰+)
چهارم	(۰+،۰+)	نهم	(۰+،۰+)	چهاردهم	(۰+،۰+)
پنجم	(۰+،۰+)	دهم	(۰+،۰+)	پانزدهم	(۵-،۰+)

برای انجام مقایسات، از سیستم شناسایی چهره مطرح در[۲۸] با معیار فاصله نرم ۱(رابطه ۸) استفاده می‌کنیم.

که I_1 و I_2 به ترتیب تصویر پرس و جو و تصویر پایگاه داده بوده، همچنین l و w نیز به ترتیب طول و عرض دو تصویر بوده و در نهایت $I_{1_{i,j}}$ و $I_{2_{i,j}}$ به ترتیب مقدار شدت روشنایی پیکسل‌ها در تصاویر پرس و جو و تصویر پایگاه داده است.

جهت سنجش مقاومت الگوریتم،از ۲۴ تصویر ورودی، ۶ شبیه‌ترین تصویر به تصویر روبه‌رو برای هر شخص، با زوایای (۰+،۱۵+)، (۰+،۱۵-)، (۱۵+،۰+)، (۱۵-،۰+)، (۳۰+،۰+) و (۳۰-،۰+) را از ورودی الگوریتم حذف نمودیم.از ۱۸ تصویر ورودی باقی مانده، ۴ شبیه‌ترین تصویر به تصویر چهره روبه‌رو با زوایای (۱۵+،۱۵+)،(۱۵-،۱۵+)، (۱۵+،۱۵-)و(۱۵+،۱۵+) شناسایی شدند. در جدول (۳) میزان شباهت هر یک از این چهار تصویر و تصویر مدل شده، با تصویر روبه‌روی پایگاه داده گزارش شده است.

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]

جدول (۳): جدول مقایسات چهار شبیه‌ترین تصویر با زوایای مطرح شده به همراه تصویر مدل شده با تصویر روبه‌روی پایگاه داده. هر سطر نشان‌دهنده نتیجه مقایسات یک شخص است. نشان‌دهنده شبیه‌ترین به تصویر روبه‌روی پایگاه داده است. نشان‌دهنده اختلافی کمتر از ٪۱ میان تصویر پرس‌وجو و تصویر پایگاه داده است. همچنین نشان‌دهنده اختلافی کمتر از ٪۱۰ و نشان‌دهنده اختلافی بیشتر از ٪۱۰ است. برای تصویر هر شخص، رتبه آن بر اساس شباهت با تصویر روبه‌رو مشخص شده است. A دارای بیشترین شباهت و E دارای کمترین شباهت است.

زاویه / شخص	(+۱۵+۱۵)		(−۱۵+۱۵)		(+۱۵−۱۵)		(−۱۵−۱۵)		تصویر مدل شده	
اول	۰,۱۱۲۱	A	۰,۱۴۰۰	D	۰,۱۲۹۸	C	۰,۱۶۵۱	E	۰,۱۱۳۲	B
دوم	۰,۱۲۳۷	D	۰,۱۲۳۰	C	۰,۱۰۲۷	A	۰,۱۴۱۵	E	۰,۱۱۳۷	B
سوم	۰,۰۸۵۰	A	۰,۰۹۲۸	C	۰,۱۳۶۴	D	۰,۰۹۱۲	B	۰,۲۰۷۰	E
چهارم	۰,۱۲۲۳	C	۰,۱۵۲۵	E	۰,۱۴۷۵	D	۰,۱۰۰۸	A	۰,۱۲۰۶	B
پنجم	۰,۱۱۰۲	C	۰,۱۲۸۵	E	۰,۱۲۴۶	D	۰,۱۰۹۴	B	۰,۱۰۳۳	A
ششم	۰,۱۲۴۱	B	۰,۱۳۸۲	D	۰,۱۷۳۱	E	۰,۱۲۶۷	C	۰,۰۷۳۷	A
هفتم	۰,۱۳۱۹	C	۰,۱۳۴۶	D	۰,۱۴۷۵	E	۰,۰۸۶۹	A	۰,۰۸۹۹	B
هشتم	۰,۱۱۰۶	A	۰,۱۲۱۸	B	۰,۱۷۱۹	E	۰,۱۲۶۸	C	۰,۱۳۰۹	D
نهم	۰,۰۸۸۰	B	۰,۱۰۹۵	D	۰,۰۶۹۳	A	۰,۱۱۱۴	C	۰,۰۹۳۴	E
دهم	۰,۱۰۴۸	A	۰,۱۳۳۶	C	۰,۱۵۶۳	E	۰,۱۳۵۰	D	۰,۱۰۵۸	B
یازدهم	۰,۰۹۳۳	A	۰,۱۰۴۶	C	۰,۱۳۴۰	E	۰,۱۱۰۳	D	۰,۱۰۳۹	B
دوازدهم	۰,۰۸۱۵	A	۰,۱۱۹۶	C	۰,۱۳۰۱	E	۰,۱۲۶۰	D	۰,۱۰۱۶	B
سیزدهم	۰,۰۹۱۴	B	۰,۱۰۷۹	C	۰,۱۲۹۶	E	۰,۱۱۵۰	D	۰,۰۸۹۶	A
چهاردهم	۰,۱۲۲۳	D	۰,۱۲۹۴	E	۰,۱۱۰۱	C	۰,۱۰۰۵	B	۰,۰۹۸۳	A
پانزدهم	۰,۱۳۵۹	D	۰,۱۱۲۷	B	۰,۱۱۹۰	C	۰,۱۰۵۵	A	۰,۱۴۱۳	E

در جدول (۳) رنگ زرد نشان‌دهنده تفاوتی کمتر از ۱ درصد، رنگ خاکستری نشان‌دهنده تفاوت بین ۱ تا ۱۰ درصد و رنگ قرمز نشان‌دهنده تفاوت بیش از ۱۰ درصد است از جدول (۳) می‌توان دریافت که تصویر مدل شده در ٪ ۲۶,۶۷ مواقع، شبیه‌ترین تصویر به تصویر روبه‌رو می‌باشد. همچنین در ٪۷۳,۳۳ مواقع، تصویر مدل دارای اختلافی کمتر از ٪۱ و در ٪۹۳,۳۳ مواقع اختلافی کمتر از ٪۵ با تصویر روبه‌رو داشته است. جدول (۴) این موضوع را نشان می‌دهد.

جدول (۴): خلاصه نتایج استخراج شده از جدول (۳).

پذیرش خطای کمتر از ۵ درصد با شبیه‌ترین تصویر	داشتن رتبه Aیا B	شبیه‌ترین تصویر به تصویر روبه‌رو
٪۹۳,۳۳	٪۷۳,۳۳	٪۲۶,۶۷

بیشترین اختلافی که تصویر مدل شده با شبیه‌ترین تصویر به تصویر روبه‌رو دارد، بعد از شخص سوم، مربوط به شخص چهاردهم است که در حدود ۳,۵۸ درصد است. دلیل ساخت مدلی ناکارآمد برای

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]
Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018

چهره شخص سوم، وجود مقدار زیادی انسداد (موی شخص که کل پیشانی آنرا پوشش داده) در تمامی تصاویر ورودی این شخص است.

۴-۴- بار محاسباتی روش پیشنهادی

با توجه به اینکه هنوز بخش‌هایی از روش پیشنهادی خودکار نشده، در این بخش برای اندازه‌گیری بار محاسباتی، مجموع زمان پردازش انتقال توسط شبکه‌های عصبی و ترازسازی با روش RASL اندازه‌گیری می‌شود. در [۱۰] راجع به پیچیدگی زمانی روش RASL بحثی نشده و تنها مدت زمان اجرای روش اندازه‌گیری شده است. مجموع زمان پردازش برای هر شخص با ۲۴ تصویر و وضوح 100×100برای هر تصویر چهره بر روی یک لپتاپ HPبا پردازنده COREi5 2.8 GHZ و6GB حافظه اصلی، در نرم‌افزار MATLAB R2014b، با حداقل ۱۵ثانیه و حداکثر ۱۹ثانیه گزارش شد که ۱۰ ثانیه از آن مربوط به عملیات انتقال ۲۴ تصویر و باقی مربوط به ترازسازی است. علت این اختلاف، همگرایی روش RASLدر دفعات تکرار مختلف برای چهره‌های متفاوت اشخاص می باشد. این پیچیدگی زمانی، پیشرفت قابل مقایسه‌ای بر روش‌های مبتنی بر مدل سه‌بعدی همچون [۳۰] دارد[۳۱].

۴-۵- جمع‌بندی

پایگاه داده‌های زیادی همچون FERETوجود دارند که برای آزمایشات روش پیشنهادی این مقاله مناسب می‌باشند. از میان این پایگاه داده‌ها، PRIMA دارای دقت کمتری در اندازه‌گیری زاویه چهره نسبت به سایر پایگاه‌های داده بود و تنها حاوی تصاویر دو بعدی ساده است[۲۴,۲۳]. با توجه به عملکرد مناسب روش پیشنهادی بر روی پایگاه داده PRIMA، می‌توان به نتایج این روش در تصاویر ویدئویی اخذ شده در درگاه‌های ورودی اتکا نموده و همچنین از تصاویر تقریبا واگرا(تصاویری که درگذر از یکی به دیگری یک تصویر چهره روبه‌رو وجود دارد) نیز استفاده نمود.

با استفاده از چارچوب ساخت تابع (۲) قابل اثبات است که از یک تابع چندجمله‌ای بدست آمده با حداقل مربعات نیز می‌توان به عنوان تابع انتقال استفاده نمود.سعی بر آن داریم تا در آینده این مساله حداقل مربعات را با تابع هزینه الگوریتم RASLدر قالب یک معادله بهینه‌سازی توامان، یکپارچه‌سازی نماییم و با به‌کارگیری روش‌های استخراج ویژگی چهره همچون [۳۱,۳۰,۲۵] برای ایجاد نقاط مطابق، روش پیشنهادی را کاملا خودکار سازیم.

۲- نتیجه‌گیری

در این مقاله روشی برای مدل‌سازی تصویر چهره شخص از روبه‌رو، با استفاده از مجموعه تصاویر زاویه‌دار ارائه شد. این روش میانگین‌گیری دگردیسی نامیده شد و نشان داد که قادر به نرمال‌سازی تصاویر چهره نسبت به زاویه می‌باشد. همچنین نشان داده شد که با استفاده از تصاویر دوبه‌دوی واگرا و با میانگین‌گیری دگردیسی، قادریم که انسداد زاویه‌ای را به حالات چهره تبدیل کنیم. با تبدیل خطای تصویر، امکان تصحیح آن با استفاده از تجزیه مرتبه پایین RASLفراهم شد. نتایج کمی و کیفی نشان می‌دهند که روش پیشنهادی به صورتی مناسب تصویر روبه‌رو را از تصاویر زاویه‌دار می‌سازد.این روش از اطلاعات زاویه چهره استفاده نمی‌کند و قادر به مدل‌سازی تصویر چهره از تصاویر ساده‌ی دوبعدی که ۳۰ درجه زاویه در هر دو جهت دارند، می‌باشد.

مراجع

[1] Rashidi S, Fallah A, Tohidkhah F. Dynamic Signature Verification Based on Extraction of Stability Dominant Points and Segmentation of Pattern. Journal of Iranian Association of Electrical and Electronics Engineers. Vol. 9, No. 1, pp. 23-32, 2012.

[2] Hjelmås, E. and Low, B.K., Face detection: A survey. Computer vision and image understanding, Vol. 83, No. 3, pp.236-274, 2001.

[3] Kondo, T. and Yan, H., Automatic human face detection and recognition under non-uniform illumination. Pattern Recognition, Vol. 32, No. 10, pp.1707-1718, 1999.

[4] Haghighat, M., Abdel-Mottaleb, M. and Alhalabi, W., 2016. Fully automatic face normalization and single sample face recognition in unconstrained environments. Expert Systems with Applications, Vol. 47, pp.23-34.

[5] Arslan, M., Mirza, M.B., Shabbir, B., Bokhari, S.T.F. and Azam, F., Face Recognition: A Survey. 2014.

[6] Guo, T., Tan, X.H. and Chao, C.X, January, Robust Face Recognition via Occlusion Detection and Masking, In MATEC Web of Conferences Vol. 44, EDP Sciences, 2016.

[7] Shan, S.,Gao, W., Cao, B. and Zhao, D., Illumination normalization for robust face recognition against varying lighting conditions, In Analysis and Modeling of Faces and Gesture, IEEE International Workshop on, pp. 157-164, 2003.

[8] Juefei-Xu, F. and Savvides, M., Encoding and decoding local binary patterns for harsh face illumination normalization, In Image Processing (ICIP), IEEE International Conference on, pp. 3220-3224, 2015.

[9] Arad, N., Dyn, N., Reisfeld, D. and Yeshurun, Y., Image warping by radial basis functions: Application to facial expressions. CVGIP: Graphical models and image processing, Vol.56, No. 2, pp.161-172, 1994.

[10] Peng, Y., Ganesh, A., Wright, J., Xu, W. and Ma, Y., RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 11, pp. 2233-2246, 2012.

[11] Zhang, Z., Xu, Y., Yang, J., Li, X. and Zhang, D., A survey of sparse representation: algorithms and applications. IEEE Access, Vol. 3, pp.490-530, 2015.

[12] Baker, S. and Matthews, I., Lucas-kanade 20 years on: A unifying framework, International journal of computer vision, Vol.56, No. 3, pp.221-255, 2004.

[13] Brown, L.G., A survey of image registration techniques, ACM computing surveys (CSUR), Vol. 24, No. 4, pp. 325-376, 1992.

[14] Wolberg, G., Geometric transformation techniques for digital images: a survey. Department of Computer Science, Columbia University, 1988. Analysis and

[Downloaded from jiaeee.com on 2026-11-20]

[DOR: 20.1001.1.26765810.1397.15.2.6.6]

زیرنویس‌ها

[1]Occlusion
[2] Haar-Wavelet
[3]Active Appearance Model
[4]Sparse Representation based Classification
[5]Dictionary
[6]Reconstruction Residual
[7]Gamma Intensity Correction
[8]Quotient Illumination Relighting
[9]Local Binary Pattern
[10]**R**obust **A**lignment by **S**parse and **L**ow rank decomposition
[11]Low Rank
[12]Sparse error
[13]Cost function
[14]Non convex optimization
[15]Nonlinear Constraint
[16]Affine
[17]Perspective

[18]Calibration: در تصاویر کالیبره، امکان تخمین فاصله اشیا تا دوربین تحت پارامتری به‌نام فاکتور کالیبره که دوربین تصویربرداری در اختیار می‌گذارد وجود دارد.

[19]Stereo Images: جفت تصاویری هستند که توسط دوربین‌های ثابت شده بر اساس طول یک خط پایه اخذ می‌شود. این تصاویر حاوی اطلاعاتی مانند عمق و زاویه صفحه تصویر نسبت به دوربین می‌باشند که برای سه‌بعدی سازی استفاده می‌شود.

[20]Warp
[21]Morphing
[22]Aliasing
[23]Heuristic
[24]Correspond points
[25]Radial Basis Function
[26]Scattered data interpolation
[27]Thin Plate Spline
[28]Similarity
[29]PlanarHomography

[30]علت مقایسه تنها با روش RASL کارایی چشم‌گیر آن نسبت به روش‌های مشابه است. خواننده می‌تواند برای اطلاعات بیشتر به [۹] مراجعه نماید.

[31]به دلیل اینکه روش‌های مبتنی بر مدل سه‌بعدی از تصاویر کالیبره و استریو استفاده می‌نمایند و همچنین پایگاه داده PRIMA حاوی این اطلاعات نمی باشد، در مقایسات کمی، روش پیشنهادی با این روش‌ها مقایسه نشد.

Machine Intelligence, Vol. 34, No. 11, pp. 2233-2246, 2012.

[15] Fanelli, G., Dantone, M., Gall, J., Fossati, A. and Van Gool, L., Random forests for real time 3D face analysis, International Journal of Computer Vision, Vol. 101, No. 3, pp .437-458, 2013.

[16] Jain, H., Hellwich, O. and Anand, R.S., Improving 3D face geometry by adapting reconstruction from stereo image pair to generic Morphable Model, In Information Fusion (FUSION), 2016 19th International Conference on, pp. 1720-1727, 2016.

[17] Zitova, B. and Flusser, J., Image registration methods: a survey Image and vision computing, Vol. 21, No. 11, pp. 977-1000, 2003.

[18] Beier, T. and Neely, S., July. Feature-based image metamorphosis, In ACM SIGGRAPH Computer Graphics, Vol. 26, No. 2, pp. 35-42, 1992.

[19] Lee, S.Y., Chwa, K.Y., Hahn, J.K. and Shin, S.Y., Image morphing using deformation techniques, Journal of Visualization and Computer Animation, Vol.7, No. 1, pp. 3-23, 1996.

[20] Prince, T., Fragoso, Y. and Gaye, O., Applications of Image Morphing Techniques to Analyze Changes in Our Environment, Open Access Library Journal, Vol. 3, No. 9, p. 1, 2016.

[21] Zanella V, Fuentes O., An approach to automatic morphing of face images in frontal view, InMexican International Conference on Artificial Intelligence, pp. 679-687. Springer Berlin Heidelberg, 2004.

[22] Seitz, S.M. and Dyer, C.R., View morphing: Uniquely predicting scene appearance from basis images, In Proc. Image Understanding Workshop, pp. 881-887, 1997.

[23] "PRIMADataset",[Online].Available: http://www prima.inrialpes.fr/perso/Gourier/Faces/HPDatabase.html,[Accessed Monday, July 6, 2015]

[24] Maturana, D., Domingo M., and Alvaro S..Face recognition with decision tree-based local binary patterns. In Asian Conference on Computer Vision, pp. 618-629. Springer, Berlin, Heidelberg, 2010.

[25] Zhang, Zhanpeng, Ping Luo, Chen Change Loy, and Xiaoou Tang. Facial landmark detection by deep multi-task learning. In European Conference on Computer Vision, pp. 94-108. Springer, Cham, 2014.

[26] Drouard, V., Ba, S., Evangelidis, G., Deleforge, A. and Horaud, R., Head pose estimation via probabilistic high-dimensional regression, In Image Processing (ICIP), IEEE International Conference on, pp. 4624-4628, 2015.

[27] Zhu, X. and Ramanan, D., Face detection, pose estimation, and landmark localization in the wild, In Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, pp. 2879-2886, 2012.

[28] Jain AK, Li SZ.Handbook of face recognition. NewYork: springer; 2011.

[29] Hwang, J., Sunjin Yu, Joongrock K., and Sangyoun Lee.3D face modeling using the multi-deformable method. Sensors 12, no. 10: 12870-12889, 2012.

[30] Asthana, A., Zafeiriou, S., Tzimiropoulos, G., Cheng, S. and Pantic, M., From pixels to response maps: Discriminative image filtering for face alignment in the wild, IEEE transactions on pattern analysis and machine intelligence, Vol. 37, No. 6, pp. 1312-1320, 2015.

[31] Seyedarabi H, Aghagolzadeh A, Khanmohammadi S, Kabir E. Analysis and Synthesis of Facial Expressions by Feature-Points Tracking and Deformable Model, Journal of Iranian Association of Electrical and Electronics Engineers. Vol. 4, No. 1, pp. 11-19, 2007.

[Downloaded from jiaeee.com on 2026-11-20]
[DOR: 20.1001.1.26765810.1397.15.2.6.6]
Journal of Iranian Association of Electrical and Electronics Engineers - Vol.15- No.2- Summer 2018