

شناسایی حالت چهره با استفاده از نرمالیزاسیون هندسی و تبدیل موجک Dual-Tree مختلط

حمید صادقی^۱، دانشجوی دکتری، فضائل آیت‌اللهی^۲، کارشناسی ارشد، ابوالقاسم اسدالله راعی^۳، دانشیار

۱، ۲ و ۳ - دانشکده مهندسی برق - دانشگاه صنعتی امیرکبیر - تهران - ایران
hamid.sadeghi@aut.ac.ir^۱, fazael_ayat@yahoo.com^۲, raie@aut.ac.ir^۳

چکیده: در سال‌های اخیر شناسایی حالت چهره به دلیل کاربردهای فراوان مورد توجه زیادی در تحقیقات بینایی ماشین قرار گرفته است. به دلیل تنوع و تغییرپذیری حالت‌ها، شناسایی حالت چهره با دقت بالا هنوز هم به‌عنوان یک مسئله چالش‌برانگیز مطرح است. در این مقاله روش نوینی ارائه می‌شود که هم‌زمان با حذف تنوع هندسی در چهره، از استخراج ویژگی‌های دقیق بهره می‌برد. برای این کار از مدل هندسی میانگین برای نرمالیزاسیون و حذف تنوع هندسی تصاویر چهره و در مرحله بعد از تبدیل موجک مختلط Dual-Tree برای استخراج ویژگی‌های ظاهری استفاده شده است. این تبدیل موجک یک تبدیل جهت‌دار با قابلیت تشخیص لبه‌ها در زوایای مختلف و دارای خاصیت تغییرناپذیری نسبی در برابر جابجایی است. این نوآوری، یعنی استفاده هم‌زمان از نرمالیزاسیون هندسی و تبدیل موجک مذکور، دقت الگوریتم را در برابر استفاده تک‌تک از این روش‌ها بیش از ۵٪ افزایش می‌دهد؛ به طوری که با این روش، نرخ شناسایی هفت حالت چهره بر روی پایگاه داده CK+ برابر با ۹۳/۷۸٪ است، که در برابر معتبرترین منابع تحقیقاتی اخیر در این پایگاه داده دقت قابل توجهی است.

واژه‌های کلیدی: شناسایی حالت چهره، تبدیل موجک مختلط Dual-Tree، نرمالیزاسیون هندسی، شکل میانگین.

Facial Expression Recognition using Geometric Normalization and Dual-Tree Complex Wavelet Transform

H. Sadeghi¹, F. Ayatollahi², A. Asaddollah Raie³

1, 2, 3-Electrical Engineering Department-Amirkabir University of Technology-Tehran-Iran

Abstract: Due to its various applications, facial expression recognition has been attracted much attention in machine vision research in recent years. Till now, facial expression recognition with high accuracy remains as a challenging problem due to the variability of facial expressions. In this paper, a novel approach is presented, which utilizes an accurate feature extraction method along with the elimination of geometric variability. For this purpose, the mean geometric model is used for normalization and elimination of geometric variability in facial images. Then, Dual-Tree Complex Wavelet Transform is used for appearance feature extraction. It is an approximately shift-invariant and directional transform which can detect edges in different angles. Utilizing the geometric normalization and the mentioned wavelet transform together improves the accuracy more than 5% respects to utilizing each method separately. From the experiments, average recognition rate on CK+ dataset is 93.78% being a noticeable result compared to the existing studies on this dataset.

Keywords: Facial expression recognition, Dual-Tree Complex Wavelet Transform (DT-CWT), geometric normalization, mean shape.

تاریخ ارسال مقاله: ۱۳۹۲/۱۲/۲۷

تاریخ اصلاح مقاله: ۱۳۹۳/۰۴/۲۱

تاریخ پذیرش مقاله: ۱۳۹۳/۰۵/۲۰

نام نویسنده مسئول: ابوالقاسم اسدالله راعی

نشانی نویسنده مسئول: ایران - تهران - خیابان حافظ - دانشگاه صنعتی امیرکبیر - دانشکده مهندسی برق

۱- مقدمه

احساس را ابراز می‌کنند، به طوری که جایگاه و نحوه قرار گرفتن اجزای چهره در آن‌ها برای ابراز یک احساس مشابه، می‌تواند کاملاً متفاوت باشد. یکدیگر باشد. به عنوان مثال می‌توان به شکل (۱) مراجعه نمود. در این شکل که از تصاویر یک پایگاه داده [۱۵] انتخاب شده‌اند، تمامی افراد احساس غمگینی را از خود ابراز می‌کنند، اما نحوه ابراز این احساس دارای تفاوت‌هایی در افراد مختلف است. حالت‌های دیگر چهره نیز وضع تقریباً مشابهی دارند. روش‌های هندسی مستقیماً تحت تأثیر این تنوع قرار می‌گیرند و در روش‌های ظاهری نیز تنوع هندسی و جایگاه قرار گرفتن اجزای مختلف چهره نتایج را تحت تأثیر خود قرار می‌دهد. در [۱۶] سعی شده است که با انجام یک نرمالیزاسیون هندسی، تنوع هندسی در چهره افراد برطرف شود. برای این کار از مدل هندسی میانگین در الگوریتم مدل ظاهری فعال^۳ (AAM) استفاده شده و تصاویر تمام افراد به این مدل نگاشته شده‌اند. اما در این کار از توصیف دقیق بافت تصویر چهره استفاده نشده است و نتایج به دست آمده برای شناسایی حالت چهره چند شخص از پیش تعیین شده، به طور میانگین حدود ۷۵٪ است.

همان‌طور که گفته شد، در تحقیقاتی که تاکنون انجام شده است، پس از نرمالیزاسیون هندسی از مقادیر مستقیم پیکسل‌ها برای استخراج ویژگی استفاده شده است. در این مقاله برای نخستین بار بعد از نرمالیزاسیون هندسی با مدل میانگین، از یک تبدیل مویک برای استخراج ویژگی دقیق از چهره استفاده شده است. برای دسته‌بندی ویژگی‌ها نیز از دسته‌بند ماشین بردار پشتیبان (SVM) [۱۷] استفاده شده است. بلوک دیاگرام الگوریتم پیشنهادی در شکل (۲) رسم شده است.



شکل (۱): ابراز حالت غمگینی در چند شخص از پایگاه داده CK+ [۱۵]

به طور خلاصه در این مقاله به مسائل زیر پرداخته خواهد شد:

- در مرحله اول به منظور رفع مشکل تنوع هندسی در چهره افراد و همچنین ابراز احساس، نرمالیزاسیون هندسی چهره انجام می‌شود، در این نرمالیزاسیون اجزای مختلف چهره در جایگاه از پیش تعریف شده قرار می‌گیرند. برای این کار از مثلث‌بندی دلانی^۴ و تبدیل آفاین قطعه‌ای^۵ استفاده شده است. برای انجام نرمالیزاسیون هندسی، مدل هندسی میانگین و یا شکل میانگین مورد استفاده قرار گرفته است.
- به جهت توصیف دقیق بافت چهره و استخراج ویژگی‌های ظاهری چهره از تبدیل مویک مختلط^۶ Dual-Tree استفاده شده است.

چندین سال است که شناسایی حالت چهره یا تشخیص احساس ابراز شده در چهره، صرف‌نظر از هویت شخص، مورد توجه فراوان در کاربردهای پردازش تصویر و بینایی ماشین قرار گرفته است [۱، ۲]. از کاربردهای مهم شناسایی حالت چهره می‌توان به ساخت دوربین‌های عکاسی هوشمند، ربات‌های دارای احساس، انیمیشن‌های دارای ابراز احساس مشابه انسان اشاره نمود. اغلب مطالعات انجام شده، احساس‌های اصلی معرفی شده در [۳] یعنی شش حالت احساسی خشم، تنفر، ترس، خوشحالی، غمگینی و تعجب^۱ که در تعاملات اجتماعی و تمام جوامع بشری فراگیر هستند [۴]، را بررسی کرده‌اند [۵-۱۰]. به دلیل اهمیت این احساسات، این مقاله نیز ابتدا به بررسی این حالت‌ها می‌پردازد و در مرحله بعد حالت تحقیر^۲ را نیز به این حالت‌ها اضافه و آزمایش می‌کند، که این امر عمل شناسایی را مشکل‌تر می‌سازد.

شناسایی حالت چهره شامل دو مرحله اصلی است: استخراج ویژگی و دسته‌بندی. استخراج ویژگی‌های مناسب نقش مهمی در دقت نتایج دارد، ویژگی‌هایی که در شناسایی حالت چهره استفاده می‌شود را می‌توان به دودسته ویژگی‌های هندسی و ویژگی‌های ظاهری تقسیم نمود [۲]. برای ارائه تصویر چهره می‌توان ویژگی‌های هندسی، یا ویژگی‌های ظاهری و یا ترکیبی از آن‌ها را به کار برد. ویژگی‌های هندسی شکل و موقعیت اجزای چهره مانند دهان، چشم‌ها، پیشانی و غیره را ارائه می‌کنند و ویژگی‌های ظاهری اطلاعات مختلف بافت تصویر چهره از قبیل برآمدگی و فرورفتگی‌ها، خطوط و چین‌وچروک‌ها را از تمام یا نواحی خاصی از تصویر چهره ارائه می‌کنند [۲].

در [۵، ۶] یک مدل هندسی چهره متشکل از ۳۰ نقطه چهره پیشنهاد شده و فواصل مشخصی از نقاط به عنوان ویژگی مورد استفاده قرار گرفته است. در [۱۱] با محاسبه توابع همبستگی از چند ویژگی هندسی چهره مانند موقعیت گوشه‌های لب نسبت به مرکز آن، وضعیت دهان و نسبت ارتفاع و عرض آن، چند حالت اصلی شناسایی شده است. در روش‌های مبتنی بر ویژگی‌های ظاهری، فیلترهای تصاویر، مانند مویک‌های گابور [۱۲]، به تمام یا نواحی خاصی از تصویر چهره اعمال می‌شود. به علت نتایج مناسب به دست آمده، در این مقاله نیز از ویژگی‌های ظاهری برای شناسایی حالت چهره استفاده می‌شود.

از آنجایی که هندسه چهره و جایگاه قرار گرفتن اجزای مختلف چهره اهمیت بالایی در نتایج دارد، قبل از استخراج ویژگی به بررسی این مسئله پرداخته می‌شود. تنوع هندسی در چهره حالت عادی افراد یک مشکل اساسی در شناسایی حالت چهره است که باعث اختلاط مسئله شناسایی حالت چهره با تشخیص هویت می‌شود [۱۳].

این تنوع به صورت گسترده‌تر در ابراز احساس‌های مختلف وجود دارد که مشکلی بنیادی در شناسایی حالت چهره است [۱۴]. در اثر تنوع هندسی در ابراز احساس، افراد مختلف به شکل‌های متفاوت یک

نرمالیزاسیون هندسی چهره پرداخته می‌شود که تاکنون در الگوریتم AAM مورد استفاده قرار گرفته است [۱۶، ۱۸، ۱۹].

برای محاسبه شکل یا مدل میانگین، از نقاط مشخص شده روی تصاویر چهره استفاده می‌شود. برای این کار، ابتدا شکل هندسی تمام تصاویر به میانگین صفر و واریانس واحد نرمالیزه می‌شوند و در مرحله بعد، میانگین شکل هندسی تصاویر از رابطه ۱ محاسبه می‌گردد.

$$mean\ shape = \frac{1}{m} \sum_{i=1}^m shape_i \quad (1)$$

که $shape_i$ شکل هندسی نرمالیزه شده متشکل از نقاط مشخصه تصویر i ام، m تعداد تصاویر ورودی و $mean\ shape$ مدل یا شکل میانگین نرمالیزه شده است. مدل میانگین محاسبه شده در این مقاله در شکل (۳) قابل مشاهده است.

نحوه نرمالیزاسیون هندسی به این ترتیب است که ۶۰ نقطه مشخصه چهره به مختصات ثابت مدل میانگین نگاشته می‌شوند. برای این کار از مثلث بندی دلانی و تبدیل آفاین قطعه‌ای [۲۰] همراه با درون‌یابی دوخطی^{۱۲} استفاده شده است. شکل (۴) نمونه‌ای از نرمالیزاسیون هندسی را برای یک تصویر چهره نشان می‌دهد.

نرمالیزاسیون هندسی چهره علاوه بر برطرف کردن تنوع هندسی در ابراز احساس، مزیت‌های دیگری نیز دارد که از این قبیل می‌توان به حذف چرخش سر (در حد قابل قبول)، یکسان نمودن اندازه چهره تمام افراد و حذف پس‌زمینه در تصویر اشاره نمود. تمام این عوامل در دقت شناسایی حالت چهره مؤثر هستند. لازم به ذکر است، در [۱۴] نشان داده شده است که اندازه 110×150 پیکسل برای استخراج ویژگی‌های بافت تصویر چهره مناسب است؛ بنابراین در این مقاله از اندازه مذکور برای یکسان کردن اندازه تمام تصاویر استفاده شده است.

۲-۲- استخراج ویژگی

در این مقاله برای استخراج ویژگی از تبدیل موجک مختلط Dual-Tree استفاده شده است؛ به همین جهت ابتدا توضیح مختصری درباره ساختار و خصوصیات این تبدیل ارائه می‌شود و سپس به چگونگی استخراج ویژگی به کمک آن خواهیم پرداخت.

این تبدیل یک تبدیل جهت‌دار است که قابلیت تشخیص لبه‌ها در شش زاویه $\pm 15^\circ$ ، $\pm 45^\circ$ و $\pm 75^\circ$ را دارا است. علاوه بر این DT-CWT به صورت تقریبی نسبت به جابجایی تغییرناپذیر است که این خواص به استخراج ویژگی‌های تقریباً یکسان در تصاویر مربوط به یک حالت چهره کمک می‌کند.

برای ارزیابی الگوریتم روش‌های مختلف آزمایش شامل ارزیابی متقاطع ۱۰ لایه^۷ و leave-one-subject-out به صورت مستقل از شخص به کار گرفته شده‌اند. برای انجام آزمایش‌ها از دسته‌بند قدرتمند ماشین بردار پشتیبان^۸ استفاده شده است. جهت ارزیابی دقیق‌تر، در قسمت آزمایش‌ها دسته‌بند K- نزدیک‌ترین همسایه^۹ نیز برای دسته‌بندی ویژگی‌های استخراج شده استفاده شده است.

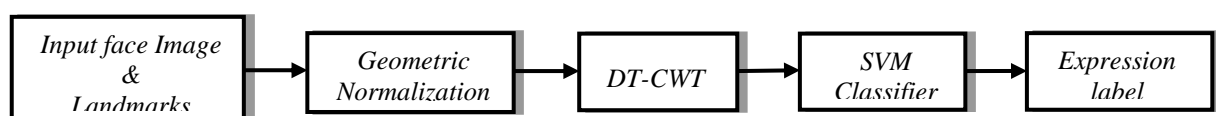
در بخش بعد نحوه نرمالیزاسیون هندسی و استخراج ویژگی استفاده از موجک مختلط Dual-Tree (DT-CWT) در روش پیشنهادی تشریح می‌شود. بخش ۳ به ارزیابی الگوریتم بر روی پایگاه داده و ارائه نتایج و مقایسه با کارهای انجام شده اختصاص دارد. نهایتاً بخش چهارم مربوط به نتیجه‌گیری و جمع‌بندی این مقاله است.

۲- روش پیشنهادی

در این بخش به معرفی و تشریح الگوریتم پیشنهادی برای شناسایی حالت چهره پرداخته می‌شود. به طور خلاصه در این بخش الف) نرمالیزاسیون هندسی چهره و ب) استخراج ویژگی، مورد بررسی قرار می‌گیرد. برای دسته‌بندی ویژگی‌های استخراج شده، از دسته‌بند ماشین بردار پشتیبان با هسته خطی استفاده شده است. از آنجایی که ماشین بردار پشتیبان یک دسته‌بند دودویی^{۱۱} است، روش دسته‌بندی یک در مقابل کل^{۱۱} برای دسته‌بندی چند کلاسه استفاده شده است.

۲-۱- نرمالیزاسیون هندسی تصاویر چهره

تنوع در حالت‌های چهره، مشکل اصلی در مسئله شناسایی حالت چهره است [۱۴]. در ظاهر به نظر می‌رسد، تنوع ابراز احساس بیشتر در موقعیت و هندسه اجزای چهره (مانند دهان، چشم‌ها و ابروها) رخ می‌دهد و چین و چروک‌های ناشی از ابراز یک احساس دارای تنوع کمتری نسبت به تنوع وضعیت هندسی چهره است. بنابراین نرمالیزاسیون هندسی چهره به یک شکل ثابت می‌تواند راهگشای مسئله باشد. از این رو، در این مقاله ابتدا هندسه چهره به یک مدل ثابت نرمالیزه می‌شود، سپس با استفاده از ویژگی‌های ظاهری، حالت چهره شناسایی می‌شود. در این بخش به تشریح مدل هندسی میانگین برای



شکل (۲): بلوک دیاگرام الگوریتم پیشنهادی

$$\psi^1(x, y) = \varphi(x)\psi(y) \quad (2)$$

$$\psi^2(x, y) = \psi(x)\psi(y) \quad (3)$$

$$\psi^3(x, y) = \psi(x)\varphi(y) \quad (4)$$

$$\psi^4(x, y) = \psi(x)\overline{\varphi(y)} \quad (5)$$

$$\psi^5(x, y) = \psi(x)\overline{\psi(y)} \quad (6)$$

$$\psi^6(x, y) = \varphi(x)\overline{\psi(y)} \quad (7)$$

که در آن‌ها:

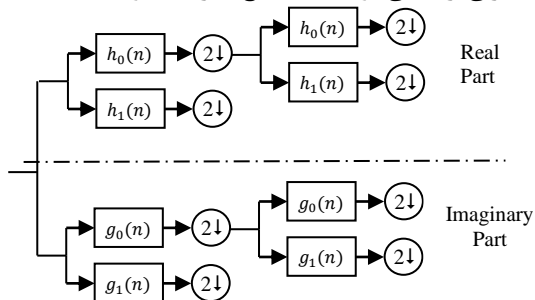
$$\varphi(x) \triangleq \varphi_i(x) + j\varphi_j(x) \quad (8)$$

$$\psi(x) \triangleq \psi_i(x) + j\psi_j(x) \quad (9)$$

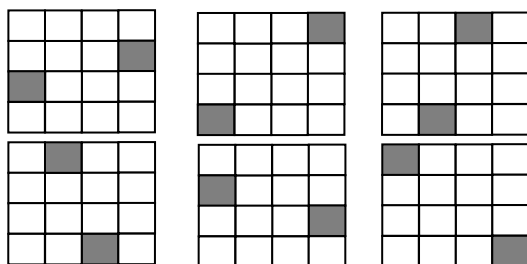
DT-CWT دوبعدی تحلیلی است و نیازی به تقارن در محورها ندارد بنابراین این تبدیل یک تبدیل جهت‌دار است و قابلیت تشخیص لبه‌ها در شش زاویه $\pm 15^\circ$ ، $\pm 45^\circ$ و $\pm 75^\circ$ را دارا است (شکل ۶). از دیگر خصوصیات DT-CWT می‌توان به تغییرناپذیری تقریبی نسبت به جابجایی [۲۱] و همچنین خاصیت کاهش داده‌های زائد اشاره کرد [۲۲].

۲-۳-۲- استفاده از DT-CWT به‌عنوان ویژگی

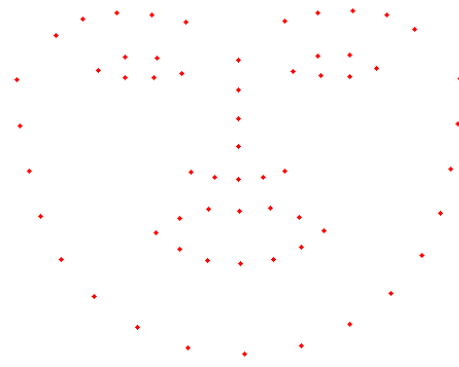
به‌منظور اعمال DT-CWT دوبعدی به تصاویر چهره، اندازه تصاویر می‌بایست تغییر یابد. زیرا در هر سطح تجزیه توسط این تبدیل موجک، طول و عرض تصاویر نصف می‌شود. به همین سبب ابعاد تصاویر چهره با اضافه شدن پیکسل‌های با مقدار ثابت به مرزهای تصاویر، گسترش یافته و مضربی از 2^l می‌شوند که J سطح تجزیه مطلوب است.



شکل (۵): ساختار بانک فیلتر 2D DT-CWT [۲۱]



شکل (۶): شش زاویه مختلف تبدیل موجک مختلط Dual-Tree در صفحه شطرنجی فرکانسی [۲۱]



شکل (۳): شکل یا مدل میانگین محاسبه‌شده



شکل (۴): یک نمونه نرمالیزاسیون هندسی چهره

۲-۳-۳- استخراج ویژگی

در این مقاله برای استخراج ویژگی از تبدیل موجک مختلط Dual-Tree استفاده شده است به همین جهت ابتدا توضیح مختصری درباره ساختار و خصوصیات این تبدیل ارائه می‌شود و سپس به چگونگی استخراج ویژگی به کمک آن خواهیم پرداخت.

۲-۳-۱- تبدیل موجک مختلط (DT-CWT) Dual-Tree

DT-CWT به‌منظور بهبود مشکلات موجود در تبدیل موجک گسسته (DWT) به وجود آمد. پیاده‌سازی بانک فیلتر این تبدیل موجک به‌سادگی از روی ساختار بانک فیلتر DWT امکان‌پذیر است؛ در نتیجه پیاده‌سازی آن به‌سادگی و بدون صرف هزینه ریاضی اضافی قابل انجام است. در DT-CWT یک‌بعدی این بانک فیلتر از دو شاخه مجزا DWT تشکیل شده است، که به تبدیل موجک، این ویژگی را می‌دهد که بخش حقیقی و موهومی آن به‌صورت تقریبی تبدیل هیلبرت یک‌دیگر باشند، در نتیجه تبدیلی تقریباً تحلیلی حاصل شود. تحلیلی بودن DT-CWT امکان جهت‌یابی مناسب را که DWT از آن بی‌بهره بوده است به DT-CWT اعطا می‌کند [۲۱] (شکل ۵).

همان‌گونه که در DWT دوبعدی با ضرب توابع بالاگذر ψ و پایین‌گذر φ در جهات x, y سه تابع موجک LH, HL و HH تشکیل می‌شوند، در DT-CWT نیز همین اتفاق می‌افتد با این تفاوت که توابع پایین‌گذر φ و بالاگذر ψ خود توابعی مختلط هستند، در نتیجه شش تابع موجک روابط ۲ تا ۷ را حاصل می‌کنند.

سپس برای ارزیابی بهتر الگوریتم، تصاویر دارای برچسب حالت تحقیر نیز به داده‌ها افزوده و مورد آزمایش قرار گرفته‌اند. همچنین تمامی آزمایش‌ها به صورت مستقل از شخص انجام می‌شود، به طوری که از افراد مرحله تعلیم هیچ تصویری در مرحله آزمایش وجود ندارد.



شکل (۷): نمونه‌هایی از تصاویر پایگاه داده CK+؛ به ترتیب از بالا چپ به سمت پایین راست: تنفر، خوشحالی، تعجب، ترس، خشم، تحقیر، ناراحتی و عادی [۱۵]

۲-۳- نتایج آزمایش‌ها

برای انجام آزمایش‌ها ابتدا ویژگی‌های تصاویر انتخاب شده از پایگاه داده که در قسمت قبل ذکر شد، طبق بخش ۲ استخراج شده است. سپس ویژگی‌های استخراج شده به روش ارزیابی leave-one-subject-out (LOSO) مورد ارزیابی قرار گرفته‌اند. در این نوع ارزیابی، تمام تصاویر مربوط به یک شخص، در مجموعه آزمایش قرار می‌گیرد و تصاویر مربوط به اشخاص باقی‌مانده در مجموعه آموزش قرار می‌گیرد. این روند برای تمام اشخاص موجود در پایگاه داده تکرار می‌شود و در نهایت میانگین نرخ شناسایی تمام اشخاص موجود در پایگاه داده (با احتساب تعداد تصاویر)، به عنوان نتیجه نهایی محاسبه می‌شود.

برای دسته‌بندی ویژگی‌های استخراج شده، از دسته‌بند ماشین بردار پشتیبان با هسته خطی استفاده شده است. نتایج به دست آمده برای شناسایی شش حالت چهره در قالب یک ماتریس ابهام^{۱۳} در جدول (۱) نشان داده شده است. همچنین ماتریس ابهام شناسایی هفت حالت چهره در جدول (۲) قابل مشاهده است.

همان‌طور که از جدول (۲) قابل مشاهده است، نرخ شناسایی به دست آمده در حالت‌های ترس (Fe.) و تحقیر (Co.) نسبت به دیگر حالت‌های چهره کمتر است. یکی از دلایل این امر می‌تواند تعداد کم تصاویر در این دو حالت چهره در پایگاه داده باشد که باعث می‌شود تعداد داده‌های آموزش در این دسته‌ها به اندازه کافی نباشد.

جدول (۱): ماتریس ابهام شناسایی شش حالت چهره

	An. (%)	Di. (%)	Fe. (%)	Ha. (%)	Sa. (%)	Su. (%)
An.	۹۱/۸۵	۵/۱۹	۰	۰	۱/۴۸	۱/۴۸
Di.	۳/۹۵	۹۴/۹۲	۰	۰	۱/۱۳	۰
Fe.	۰	۰	۶۹/۴۴	۱۶/۶۷	۰	۱۳/۸۹
Ha.	۰	۰	۰	۱۰۰	۰	۰
Sa.	۳/۵۷	۰	۰	۰	۹۴/۰۵	۲/۳۸
Su.	۰	۱/۲	۰	۰	۱/۶۱	۹۷/۱۹
Avg.	٪۹۴/۱۶					

برای استخراج ویژگی، تصاویر چهره گسترش یافته، به کمک DT-CWT دوبعدی تا سطح J تجزیه می‌شوند تا یکسری زیرباند با مقیاس‌های متفاوت متشکل از ضرایب موجک مختلط تولید شوند. حال تنها زیرباندهای سطح J حفظ شده و سایر زیرباندها با فرکانس بالاتر دور ریخته می‌شوند.

از آنجایی که اندازه ضرایب مختلط معیار دقیقی از انرژی طیفی را بیان می‌کنند و به علت دارا بودن خواص ذکر شده در بخش ۲-۲-۱، این معیار به عنوان ویژگی لحاظ می‌شود. بدین ترتیب هشت زیر تصویر پس از تجزیه هر تصویر چهره تولید می‌شود، که شامل شش زیر تصویر میان‌گذر و دو زیر تصویر پایین‌گذر است. زیر تصاویر میان‌گذر نشان‌دهنده جزئیات و زیر تصاویر پایین‌گذر نشان‌دهنده کلیات تصویر هستند. این زیر تصاویر پس از نرمالیزاسیون به میانگین صفر و واریانس واحد، به صورت بردار درآمده و پشت سر هم چیده می‌شوند تا بردار ویژگی نهایی را ایجاد کنند.

استفاده از یک سطح موجک در این روش باعث کاهش ابعاد بردار ویژگی می‌شود و نیاز به استفاده از روش‌های کاهش ابعاد را مرتفع می‌سازد.

۳- ارزیابی الگوریتم

در این بخش روش پیشنهادی مورد ارزیابی قرار می‌گیرد. برای این کار از یکی از معتبرترین پایگاه‌های داده در زمینه شناسایی حالت چهره استفاده شده است، که ابتدا به معرفی آن پرداخته می‌شود. پس از معرفی پایگاه داده، نتایج به دست آمده بر روی آن نشان داده می‌شود. سپس این نتایج مورد بررسی قرار می‌گیرد.

۳-۱- پایگاه داده

برای ارزیابی الگوریتم پیشنهادی از پایگاه داده CK+ [۱۵] استفاده شده است که نمونه‌های از تصاویر آن در شکل (۱) نشان داده شد. پایگاه داده CK+ نسخه کامل شده پایگاه داده Cohn-Kanade [۲۳] بوده و برای رفع محدودیت‌های آن معرفی شد. به طوری که تعداد رشته‌ها در پایگاه داده ۲۲٪ و تعداد افراد ۲۷٪ افزایش داده شد و برای برخی از حالات نیز حالت غیرارادی به داده‌ها افزوده شد. این پایگاه داده متشکل از ۱۲۳ شخص و دارای تمام احساسات اصلی است. در این نسخه حالت تحقیر نیز به پایگاه داده افزوده شده است. شکل (۷) نمونه‌هایی از تصاویر پایگاه داده CK+ را نشان می‌دهد.

هر داده دنباله‌ای از تصاویر است که از حالت عادی شروع و به حداکثر حالت احساسی ختم می‌شود. همچنین مختصات ۶۸ نقطه از چهره افراد، توسط الگوریتم AAM مشخص و در پایگاه داده قرار داده شده است؛ که در این مقاله از همین نقاط برای موقعیت‌یابی نقاط کلیدی و نرمالیزاسیون هندسی چهره استفاده شده است.

در این مقاله سه فریم انتهایی تمام دنباله‌هایی که برچسب آن‌ها یکی از شش حالت اصلی است، برای انجام آزمایش‌ها انتخاب شده‌اند.



طبق ماتریس‌های ابهام به‌دست‌آمده در جدول (۱) و (۲)، بیشترین اشتباه بین حالت ترس با خوشحالی رخ داده است. همان‌طور که ذکر یکی از دلایل این امر می‌تواند تعداد ناکافی داده‌های آموزشی باشد. دلیل دیگر این تداخل می‌تواند شباهت ظاهری این دو حالت چهره باشد. به‌عنوان مثال، جدول (۳) نتیجه شناسایی حالت‌های خوشحالی و ترس یک شخص را نشان می‌دهد. در این جدول حالت ترس به‌اشتباه حالت خوشحالی تشخیص داده شده است.

یکی از چالش‌های مهم در شناسایی حالت چهره این است که الگوریتم ارائه‌شده بتواند در صورت استفاده از قسمت‌های خاصی از چهره به‌جای تمام چهره به‌درستی عمل کند. از مزایای این امر می‌توان به کارایی الگوریتم در شرایط خاصی همچون وجود انسداد در نواحی مختلف چهره اشاره نمود. همچنین با انتخاب برخی از نواحی مهم چهره به‌جای کل آن، می‌توان هزینه محاسباتی را نیز کاهش داد. حذف برخی از نواحی چهره، که در شناسایی حالت چهره تأثیر زیادی ندارند، می‌تواند باعث افزایش دقت الگوریتم نیز بشود.





برای ارزیابی الگوریتم پیشنهادشده از نظر استفاده از قسمت‌های مختلف چهره، در اینجا نواحی چشم‌ها و دهان و ترکیب آن‌ها انتخاب و مورد آزمایش قرار گرفته‌اند. نتایج به‌دست‌آمده در این مورد در جدول (۴) نشان داده شده است. در این آزمایش‌ها از روش ارزیابی متقاطع ۱۰ لایه به‌صورت مستقل از شخص استفاده شده است و میانگین نرخ شناسایی شش حالت اصلی چهره به نمایش گذاشته شده است.

جهت انجام یک ارزیابی دقیق‌تر از الگوریتم پیشنهادی، در این قسمت از دسته‌بندی K-نزدیک‌ترین همسایه با معیار فاصله spearman برای دسته‌بندی ویژگی‌های استخراج‌شده استفاده شده است. برای این کار، نواحی مختلف چهره که در جدول (۴) نشان داده شد، مورد آزمایش قرار گرفته است. نتایج دسته‌بندی شش حالت اصلی چهره به ازای استفاده از نواحی مختلف چهره برحسب مقادیر مختلف K در نمودار شکل (۹) نشان داده شده است. در این شکل از روش ارزیابی متقاطع ۱۰ لایه به‌صورت مستقل از شخص استفاده شده است.

جدول (۳): مثالی از دسته‌بندی اشتباه در حالت ترس

		
Ground truth	Happiness	Fear
Classification result	Happiness	Happiness

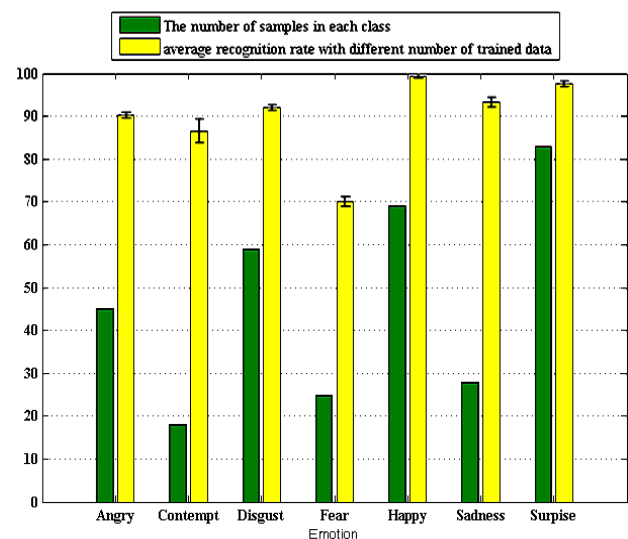
جدول (۴): نتایج حذف برخی از نواحی چهره

Input image example				
Accuracy (%)	۹۳/۲۹	۹۳/۸۳	۹۱/۲۳	۷۶/۳

جدول (۲): ماتریس ابهام شناسایی هفت حالت چهره

	An. (%)	Co. (%)	Di. (%)	Fe. (%)	Ha. (%)	Sa. (%)	Su. (%)
An.	۹۱/۱۱	۲/۲۲	۵/۹۳	.	.	.	۰/۷۴
Co.	۱/۸۵	۸۷/۰۴	.	.	.	۱۱/۱۱	.
Di.	۴/۵۲	.	۹۳/۷۹	.	.	۱/۱۳	۰/۵۷
Fe.	.	.	.	۷۰/۸۳	۱۶/۶۷	.	۱۲/۵
Ha.	۱۰۰	.	.
Sa.	۳/۵۷	۲/۳۸	.	.	.	۹۴/۰۵	.
Su.	.	۱/۲۱	.	.	.	۰/۴	۹۸/۳۹
Avg.	٪ ۹۳/۷۸						

نمودار نشان داده‌شده در شکل (۸) گواهی مناسب از نسبت تعداد داده‌ها و نرخ شناسایی است. همان‌طور که ملاحظه می‌شود، تقریباً می‌توان گفت میانگین نرخ شناسایی در دسته‌هایی که تعداد داده‌ها در آن‌ها بیشتر بوده، بالاتر است. در این شکل میانگین و واریانس دقت شناسایی در هر حالت چهره به ازای تعداد مختلف داده آموزشی ترسیم شده است. برای انجام این آزمایش، به‌طور تقریبی ۴۰، ۵۰، ۶۰، ۷۰، ۸۰، ۹۰ و ۹۵ درصد از کل تصاویر هر حالت به‌عنوان مجموعه آموزشی و مابقی تصاویر به‌عنوان داده‌های آزمایشی استفاده شده است. هر یک از آزمایش‌ها چند مرتبه با انتخاب تصادفی داده‌ها انجام شده است. همان‌طور که ملاحظه می‌شود، حساسیت الگوریتم به تعداد داده‌های آموزشی تقریباً کم است. باین‌حال واریانس نتایج در برخی از حالت‌ها (مانند تحقیر و ترس) نسبت به دیگر حالت‌ها بیشتر است. البته بزرگ بودن واریانس در این دسته‌ها لزوماً به معنای حساسیت مذکور در آن‌ها نیست. بخشی از این واریانس می‌تواند به دلیل تعداد بسیار کم تصاویر پایگاه داده در این دسته‌ها باشد که در شکل نیز قابل‌مشاهده است.



شکل (۸): میانگین و واریانس نرخ شناسایی هر دسته به ازای تعداد داده‌های آموزشی مختلف (میله‌های زرد)؛ تعداد تصاویر پایگاه داده در هر دسته (میله‌های سبز)

استخراج شده استفاده شده است. برای ارزیابی الگوریتم ارائه شده در [۱۵]، از تصاویر پایگاه داده معرفی شده در آن (پایگاه داده CK+) شامل هفت حالت چهره به روش LOSO، که در ابتدای بخش ۳-۲ توضیح داده شد، استفاده شده است. همان طور که از جدول (۶) ملاحظه می شود، استفاده از ترکیب این نوع موجک با نرمالیزاسیون هندسی توسط مدل میانگین، دقت شناسایی حالت چهره را نسبت به استفاده تنها از مدل میانگین، بیش از ۵٪ افزایش می دهد. همچنین نتایج به دست آمده از هر حالت احساسی به صورت جداگانه در جدول (۷) با [۱۵] مقایسه شده است.

جدول (۶): مقایسه نتایج به دست آمده با الگوریتم AAM [۱۵]

dataset	Classes	Measure	Recognition rate (%)
This paper	CK+	γ	۹۳/۷۸
AAM [۱۵]	CK+	γ	۸۸/۳۳

جدول (۷): مقایسه دقت به دست آمده برای هر حالت با [۱۵]

	An. (%)	Co. (%)	Di. (%)	Fe. (%)	Ha. (%)	Sa. (%)	Su. (%)
This paper	۹۱/۱	۸۷/۰	۹۳/۸	۷۰/۸	۱۰۰	۹۴/۱	۹۸/۴
AAM [۱۵]	۷۵/۰	۸۴/۴	۹۴/۷	۶۵/۲	۱۰۰	۶۸/۰	۹۶/۰

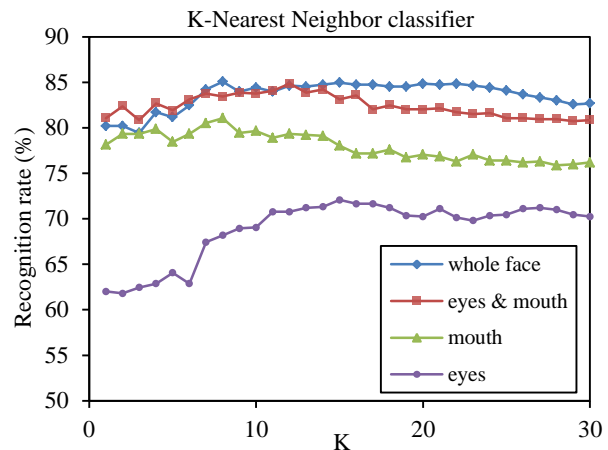
برای بررسی اثر نرمالیزاسیون هندسی بر دقت الگوریتم ارائه شده، در اینجا شناسایی هفت حالت چهره با استفاده از ویژگی های موجک تشریح شده در بخش ۲-۲ و بدون استفاده از نرمالیزاسیون هندسی انجام شده است. نتایج به دست آمده از این آزمایش در جدول (۸) آورده شده است. با توجه به نتایج به دست آمده در این جدول می توان گفت که انجام نرمالیزاسیون هندسی باعث بهبود بیش از ۵٪ در دقت الگوریتم نسبت به زمانی که بدون نرمالیزاسیون اقدام به استخراج ویژگی می نماییم، شده است.

جدول (۸): تأثیر استفاده از نرمالیزاسیون در برابر عدم استفاده از آن

dataset	Classes	Measure	Recognition rate (%)
This paper	CK+	γ	۹۳/۲۹
DT-CWT	CK+	γ	۸۷/۹۹

به عنوان یک نتیجه مهم از جدول (۶) و (۸) می توان گفت، هنگامی که از ترکیب نرمالیزاسیون هندسی (با مدل میانگین) و تبدیل موجک تشریح شده در بخش ۲-۲ استفاده شود، نرخ شناسایی حالت چهره نسبت به استفاده مجزا از هر یک از این دو، بیش از ۵٪ افزایش یافته و از حدود ۸۸٪ به بیش از ۹۳٪ می رسد. البته واضح است که هزینه محاسباتی استفاده هم زمان از این دو روش، هزینه محاسباتی را افزایش می دهد.

جهت انجام مقایسه ای مناسب با تحقیقات انجام شده اخیر، در شرایط برابر از نظر تعداد تصاویر، نوع پایگاه داده و نحوه ارزیابی در این



شکل (۹): نرخ شناسایی شش حالت چهره با استفاده از دسته بند K- نزدیک ترین همسایه با استفاده از نواحی مختلف چهره به ازای مقادیر مختلف K

همان طور که از نمودار شکل (۹) قابل مشاهده است، تقریباً می توان گفت که با افزایش تعداد همسایگی K تا حدود ۱۵، نرخ شناسایی افزایش یافته و پس از آن کاهش می یابد. اما نکته مهمی که از شکل (۹) می توان دریافت، میزان اهمیت نواحی مختلف چهره در شناسایی حالت چهره است. طبق این شکل، ناحیه دهان از اهمیت بیشتری نسبت به ناحیه چشمها برخوردار است. همچنین می توان گفت که زمانی که از ترکیب نواحی چشمها و دهان استفاده می شود، نرخ شناسایی تقریباً با استفاده از کل چهره برابر است؛ که می توان از این روش برای کاهش هزینه محاسباتی استفاده نمود.

جدول (۵) جهت انجام یک مقایسه مناسب با نتایج گزارش شده در [۱۳] آورده شده است. اعداد موجود در ردیف اول این جدول از نتایج به دست آمده در جدول (۴) اخذ شده است. ردیف دوم مربوط به بهترین نتایج گزارش شده در [۱۳] به ازای استفاده از نواحی مختلف تصویر چهره است.

جدول (۵): مقایسه نتایج به دست آمده در نواحی مختلف چهره

method	Input	whole face	Best region selection	Mouth	eyes
	This paper (%)	۹۳/۲۹	۹۳/۸۳	۹۱/۲۳	۷۶/۳
[۱۳] (%)	۸۸	۹۱/۴	۸۱/۰	۶۴/۸	

۳-۳- مقایسه نتایج به دست آمده با دیگر کارها

در جدول (۶) نتایج به دست آمده در این مقاله، با نتایج گزارش شده در [۱۵] که از الگوریتم AAM استفاده کرده است، مقایسه می شود. الگوریتم AAM مستقیماً از مقادیر پیکسلها برای ارائه بافت تصاویر نرمالیزه شده چهره به مدل میانگین استفاده می کند. در الگوریتم AAM، برای استخراج ویژگی از ترکیب اطلاعات مربوط به ظاهری (یا بافت تصویر) و اطلاعات هندسی استفاده می شود. در [۱۵] بعد از استخراج ویژگی از دسته بند SVM برای دسته بندی ویژگی های

مراجع

- [1] J. Fasel and J. Luetten, "Automatic facial expression analysis: a survey," *Pattern Recognition*, vol. 36, pp. 259-275, 2003.
- [2] Y. Tian, T. Kanade and J. Cohn, *Handbook of Face Recognition*, Springer, 2005 (Chapter 11. Facial Expression Analysis).
- [3] P. Ekman and W.V. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality Social Psychology*, vol. 17, no. 2, pp. 124-129, 1971.
- [4] P. Ekman and W. Friesen, *Pictures of Facial Affect*, Consulting Psychologists, 1976.
- [5] H. Kobayashi and F. Hara. "Recognition of six basic facial expression and their strength by neural network," in *Proceedings of International Workshop Robot and Human Communication*, pp. 381-386, 1992.
- [6] H. Kobayashi and F. Hara. "Recognition of mixed facial expression by neural network," *IEEE International Workshop Robot and Human Communication*, pp. 387-391, 1992.
- [7] L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," *IEEE Transaction on Affective Computing*, vol. 2, no. 4, pp. 219-229, 2011.
- [8] R. Xiao, Q. Zhao, D. Zhang and P. Shi, "Facial expression Recognition on multiple manifolds," *Pattern Recognition*, vol. 44, no. 1, pp. 107-116, 2011.
- [9] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 915-928, 2007.
- [10] G. Zhao and M. Pietikäinen, "Boosted multi-resolution spatiotemporal descriptors for facial expression recognition," *Pattern Recognition Letters*, vol. 30, no. 12, pp. 1117-1127, 2009.
- [11] N-S. Pai and S-P. Chang, "An embedded system for real-time facial expression recognition based on the extension theory," *Computers and Mathematics with Applications*, vol. 61, no. 8, pp. 2101-2106, 2011.
- [12] J. Daugmen, "Complete discrete 2d Gabor transforms by neutral networks for image analysis and compression," *IEEE Transaction on Acoustic, Speech and Signal Processing*, vol. 36, no. 7, pp. 1169-1179, 1988.
- [13] R. Ptucha and A. Savakis, "Manifold based sparse representation for facial understanding in natural images," *Image and Vision Computing*, vol. 31, no. 5, pp. 365-378, 2013.
- [14] C. Shan, Sh. Gong and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.
- [15] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete expression dataset for action unit and emotion-specified expression," in *Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010)*, San Francisco, USA, pp. 94-101, 2010.
- [16] P. Martins and J. Batista, "Identity and expression recognition on low dimensional manifolds," in *16th IEEE International Conference on Image Processing (ICIP)*, pp. 3341 - 3344, 2009.
- [17] Ch. Ch. Chang and Ch. J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2012. available at <http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf> >.
- [18] A. Lanitis, Ch. J. Taylor and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 743-756, 1997.
- [19] C. Martin, U. Werner and H.-M. Gross, "A real-time facial expression recognition system based on active appearance models using gray images and edge images," in *8th IEEE International Conference on Face and Gesture Recognition*, pp. 1-6, 2008.
- [20] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135-164, 2004.

قسمت نتایج گزارش شده چند مورد از تحقیقات انجام شده بر روی پایگاه داده CK+ آورده شده است. با توجه به اینکه در این تحقیقات، آزمایش‌ها به صورت‌های مختلف انجام شده است، نتایج به دست آمده با روش‌های ارزیابی و شرایط مشابه در مقایسه با نتایج گزارش شده در [۱۳، ۱۵، ۲۶-۲۴] بر روی پایگاه داده CK+، در جدول (۹) آورده شده است. به عنوان مثال برای انجام مقایسه‌ای مناسب با مطالعاتی که بر روی ۷ حالت چهره انجام شده است، به صورت مشابه، حالت چهره هفتم، تحت عنوان حالت تحقیر نیز در آزمایش‌ها در نظر گرفته شده و نتایج به دست آمده در جدول ذکر شده است.

جدول (۹): مقایسه نتایج به دست آمده با کارهای انجام شده بر روی پایگاه داده

CK+			
	Classes	Measure	Recognition rate (%)
This paper	۶	LOSO	۹۴/۱۶
This paper	۶	10-fold	۹۳/۲۹
This paper	۷	LOSO	۹۳/۷۸
This paper	۷	10-fold	۹۳/۱۹
[۱۳]	۷	LOSO	۹۱/۴
[۱۵]	۷	LOSO	۸۸/۳۳
[۲۴]	۷	10-fold	۹۰/۱
[۲۵]	۷	LOSO	۸۲/۶
[۲۶]	۷	10-fold	۸۹/۳

۴- نتیجه گیری

در این مقاله با استفاده از مدل میانگین به حذف تنوع هندسی در چهره اشخاص مختلف پرداخته شد. برای این کار از تبدیل آفاین قطعه‌ای و مثلث بندی دلانی استفاده شده است. سپس برای استخراج ویژگی از تصاویر چهره، تبدیل موجک مختلط Dual-Tree به کار برده شده است. خواص این تبدیل، نظیر جهت یابی مناسب برای تشخیص لبه‌ها در زوایای گوناگون و همچنین تغییرناپذیری نسبت به جابجایی، استخراج ویژگی‌های یکسان در تصاویر با حالت‌های چهره یکسان و شدت‌های مختلف را تضمین می‌کند. استفاده هم‌زمان از نرمالیزاسیون هندسی با مدل میانگین و استخراج ویژگی با استفاده از تبدیل موجک مذکور که برای نخستین بار انجام شده است، باعث افزایش درصد شناسایی حالت چهره می‌شود. نتایج حاصل شده در شناسایی شش حالت اصلی چهره، به روش leave-one-subject-out به دقت ۹۴/۱۶٪ و در شناسایی هفت حالت چهره، به روش leave-one-subject-out به دقت ۹۳/۷۸٪ دست یافته است؛ که نسبت به روش‌های موجود بهبود مناسبی داشته است. نتایج آزمایش‌ها مؤید این نکته است که به کارگیری نواحی دهان و چشم‌ها با دستیابی به دقت مشابه، بردار ویژگی کوتاه‌تر و در نتیجه هزینه محاسباتی کمتری را دربر خواهد داشت.

- [21] W. Selesnick, R. G. Baraniuk and N. G. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123-151, 2005.
- [22] C. C. Liu and D. Q. Dai, "Face recognition using dual-tree complex wavelet features," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2593-2599, 2009.
- [23] T. Kanade, J. F. Cohn and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, Grenoble, France, pp. 46-53, 2000.
- [24] M. S. Islam and S. Auwatanamongkol, "A novel feature extraction technique for facial expression recognition," *International Journal of Computer Science Issues*, vol. 10, no. 3, pp. 9-14, 2013.
- [25] S. Yang and B. Bhanu, "Understanding discrete facial expressions in video using an emotion avatar image," *IEEE Transaction on Systems Man Cybernetics, Part B, Cybernetics*, vol. 42, no. 4, pp. 980-992, 2012.
- [26] A. R. Rivera, J. R. Castillo and O. Chae, "Local directional number pattern for face analysis: face and expression recognition," *IEEE Transaction on Image Processing*, vol. 22, no. 5, pp. 1740 - 1752, 2013.

زیر نویس ها

- ¹ Anger, Disgust, Fear, Happiness, Sadness, and Surprise
- ² Contempt
- ³ Active Appearance Model (AAM)
- ⁴ Delaunay Triangulation
- ⁵ Piecewise Affine Warp
- ⁶ Dual-Tree Complex Wavelet Transform
- ⁷ 10-fold Cross Validation
- ⁸ Support Vector Machine (SVM)
- ⁹ K-Nearest Neighbor
- ¹⁰ binary
- ¹¹ One Versus The Rest
- ¹² Bilinear Interpolation
- ¹³ Confusion Matrix