

## سیستم خبره شناسایی احساس از روی تصویر ویدیویی چهره

محرم منصوری زاده      نصرآ... مقدم چرکری      احسان اله کبیر

دانشگاه تربیت مدرس، تهران، ایران

### چکیده

در این مقاله، سیستم خبره‌ای برای شناسایی احساس از روی تصاویر چهره ارائه می‌شود. چنین سیستمی را می‌توان برای ارتباط انسان و ماشین بکار گرفت. ورودی سیستم فریمهای متوالی نمای تمام رخ چهره شخص است. فرض می‌شود که سر حرکت محسوسی ندارد و توزیع روشنایی فریمها یکنواخت است. سیستم، نقاط مشخص شده اعضای صورت را در فریمهای متوالی دنبال کرده و تغییرات آنها را در قالب مجموعه اولیه حرکات چهره شناسایی می‌کند و سپس این حرکات را به مجموعه‌ای از احساس‌های شادی، تعجب، غم، بی‌اعتنایی، خشم، تنفر و ترس تفسیر می‌کند. در این سیستم برای هر واحد حرکت و همچنین هر احساس، معیار شدت ظهور تعریف شده است. شدت ظهور واحد حرکت برای ادغام واحدهای مشابه و شدت ظهور احساس نیز برای ادغام احساس‌های مشابه استفاده می‌شود. سیستم قادر به شناسایی احساس‌های مرکب است و تناقضهای احتمالی (مثلا احساس شادی همراه با غم) را با حذف احساسی که شدت ظهور پایین‌تری دارد، رفع می‌کند. دقت کلی شناسایی احساس ۸۰ درصد است.

**واژه های کلیدی:** شناسایی احساس، سیستم خبره، حالات چهره، ویدیو

### ۱- مقدمه

کشف کرده و در قالب کدهای FACS<sup>۲</sup> دسته بندی و به احساس‌هایی با شدت ظهور<sup>۳</sup> معین تفسیر می‌کند. نوآوری سیستم پیشنهادی، تعریف کمیت<sup>۴</sup> شدت ظهور احساس<sup>۵</sup> و استفاده از آن برای تشخیص و مدیریت احساس‌های مرکب است. برای نیل به این هدف، این سیستم فرایند بازشناسی را با استفاده از سیستمهای خبره<sup>۶</sup> انجام می‌دهد. هدف از بکارگیری سیستم خبره در این مبحث، استفاده از مدلی مناسب برای بیان نحوه ادراک انسان از حالات چهره است. همچنین استقلال کنترل از دانش در سیستم خبره، امکان تغییر و توسعه دانش سیستم را در طول حیات آن فراهم می‌کند.

در جهان واقعی معمولاً مجموعه‌ای از چند احساس به صورت مرکب در چهره انسانها ظاهر می‌شوند. در این تحقیق بطور ویژه روی این حقیقت تاکید شده و کمیت شدت ظهور احساس را برای شناسایی مجموعه احساس‌های موجود در چهره و شدت و ضعف هریک از آنها، تعریف و استفاده کرده ایم. این کمیت به ویژه زمانی که ممکن است احساس‌های متضاد از مجموعه حرکات چهره استنتاج شود، برای رفع تناقضهای احتمالی بکار می‌رود.

ساختار بقیه مقاله چنین است: در بخش ۲ نگاهی کلی به پیش زمینه موضوع خواهیم داشت و دو نمونه از کارهایی را که از سیستم خبره برای شناسایی استفاده کرده‌اند، بررسی می‌کنیم. در بخش ۳ سیستم پیشنهادی توضیح داده می‌شود. بخش ۴ نتایج حاصل از آزمایشهای انجام شده و بررسی خطاها را در بر می‌گیرد. بخش ۵ به جمع بندی و نتیجه گیری می‌پردازد.

چهره مهمترین ابزار ارتباط غیر کلامی است. از این نظر متخصصان کامپیوتر همواره سعی کرده اند به نحوی از این طریق نیز ارتباط انسان و ماشین را برقرار کنند. با انجام چنین کاری، ضمن تسهیل ارتباط انسان و ماشین به ویژه برای افراد ناتوان جسمی، افقهای جدیدی برای کاربرد کامپیوتر نیز گشوده می‌شود. اتوماسیون فرایند آموزش، گرافیک کامپیوتری، بینایی روباتها، تشخیص و معالجه بیماریهای روانی از این جمله اند. یکی از اصول اولیه ماشینی کردن یک فرایند، بیان کمی نحوه انجام آن است. گرچه روانشناسان مطالعات زیادی روی نحوه ادراک انسان از حالات چهره اشخاص انجام داده اند، تحقیقات کمی در زمینه بیان کمی حرکات چهره انجام شده است [۱]. در این میان کار Ekman و Friesen (۱۹۷۸) برای تعریف سیستم کدگذاری حرکات چهره بسیارمشهور و مورد توجه است. تغییر حالت صورت در نتیجه فعالیت عضلات و ماهیچه‌های متعدد صورت است و هر تغییر کلی در صورت، نتیجه انقباض یا انبساط مجموعه‌ای از این ماهیچه‌هاست. در این کد گذاری، هر واحد حرکت<sup>۱</sup>، تغییری در چهره است که به صورت مستقل و غیر قابل تقسیم انجام می‌گیرد. بنابراین هر نوع تغییر حالت چهره از جمله بروز احساس‌های مختلف را می‌توان در قالب ترکیبی از حرکات اولیه مذکور بیان کرد [2].

سیستم پیشنهادی این مقاله، با ردیابی خودکار اعضای چهره، تغییرات آنها را

## ۲- بررسی پیشینه موضوع

شناسایی حالات چهره، از دیرباز مورد توجه روانشناسان بوده است. داروین در سال ۱۸۷۸ یکسان بودن نحوه بروز احساس در چهره افراد مختلف را بررسی کرد. یک قرن بعد، در سال ۱۹۷۸ اکمن، مطالعه مشابهی روی افراد بدوی در گینه نو انجام داد [3]. او در ادامه کارهایش FACS را تدوین کرد [2] و همچنین با استفاده از افراد جوامع مختلف، مشابه بودن شکل بروز احساس در صورت را بررسی کرد. در FACS، هر واحد حرکت به تغییر در صورت اطلاق می‌شود که اولاً به تنهایی قابل اجرا باشد و ثانیاً قابل تقسیم نباشد. مثلاً حرکت "باز کردن دهان همراه با بالا انداختن ابروها" گرچه یکباره انجام می‌گیرد، به دو حرکت "باز کردن دهان" و "بالا انداختن ابرو" تقسیم می‌شود که مستقل از هم می‌توانند انجام گیرند. محدودیت این سیستم، بیان واحدهای حرکت تنها بر اساس مشخصات مکانی است. با این وجود، این سیستم مبنای اغلب کارهای شناسایی خودکار یا نیمه خودکار حالات چهره و احساس بوده است [۱]. نسخه اول FACS ۴۲ واحد حرکت را تعریف کرده است. نمونه‌هایی از این واحدها را در جدول ۱ می‌بینید.

جدول ۱. تعریف برخی از واحدهای حرکت چهره در FACS

کد	توصیف
۱	انتهای درونی ابرو بالا رفته است.
۲	انتهای خارجی ابرو بالا رفته است.
۴	ابرو پایین تر آمده است.
۵	پلک بالایی، بالاتر رفته است.
۶	گونه بالاتر رفته است.
۷	پلک، تنگتر شده است.
۹	ناحیه بین دو چشم چروک شده است.
۱۰	لب بالا، بالاتر رفته است.
۱۱	بینی فشرده تر شده است.
۱۲	طول لب بزرگتر شده است.
۱۶	لب پایین منقبض شده است.
۱۷	چانه بر آمده است.
۱۸	لبها جمع و گرد شده اند.
۲۰	طول دهان بزرگتر شده است.
۲۳	لبها بهم فشرده اند.
۲۴	گوشه لب بالاتر رفته است.
۲۵	لبها از هم جدا شده اند.
۲۶	فک پایینتر آمده است.
۲۷	دهان کاملاً باز شده است.
۴۳	چشم بسته است.

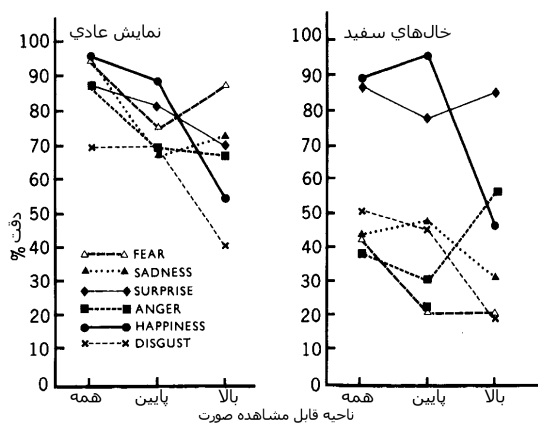
\*در این جدول صفات بالاتر، پایین‌تر و بزرگتر نسبت به حالت عادی شخص مقایسه شده‌اند.

Scherer نیز در سال ۱۹۸۸ طی تحقیقاتی دریافت هنگام بروز احساس، تغییر حالت چهره نابینایان مادرزادی، مشابه افراد بینای عادی است [4]. هرچند نوزادان بینا در تعامل با محیط مهارت‌های ارتباطی خود را بهبود می‌بخشند [5]، شکل بروز احساس‌های پایه در آنها تغییرات زیادی نمی‌کند [6].

بررسی نحوه تغییر چهره هنگام بروز احساس مورد توجه روانشناسان بوده و سوالاتی از این دست موضوع تحقیقات زیادی بوده است: نقش چشمها، و بینی و دهان در بروز یک احساس چیست؟ آیا می‌توان تنها با نگاه کردن به یک عکس بر اساس تغییرات صورت گرفته شده، احساس بروز یافته را تشخیص داد؟

آیا حرکت و تغییرات نواحی خاصی از صورت برای تشخیص احساس کافی است یا باید تغییرات کل صورت را مشاهده کرد؟

در سال ۱۹۷۹، Bassili [6] برای پاسخ به این سوالات مطالعاتی انجام داد. او برای بررسی نقش نواحی چهره در بروز احساس، در برخی از آزمایش‌ها، بخشی از چهره را پوشاند تا تنها ناحیه مورد نظر مشاهده شود و همچنین برای تعیین نقش نقاط خاص مانند گوشه چشمها، کل صورت را با رنگ تیره پوشاند و در محل این نقاط خالهای سفید نقاشی کرد. بدین ترتیب چندین فیلم ویدیویی از بروز احساس‌های مختلف در شرایط تعیین مختلف، تهیه شد. او سپس طی پرسشنامه‌هایی قضاوت افراد مختلف را درباره احساس بیان شده در این فیلمها جمع‌آوری کرد. او همچنین عکسی از آخرین مرحله بروز هر احساس که اعضای چهره به حداکثر میزان تغییر رسیده بودند تهیه کرده و آنها را نیز به صورت جداگانه به قضاوت افراد گذاشت. در اینجا تنها نگاهی اجمالی به یافته‌های این تحقیق می‌کنیم. برای اطلاع از جزئیات مرجع [6] را ببینید. میانگین دقت شناسایی ۶ احساس مورد بررسی (خوشحالی، غم، ترس، تعجب، خشم و تنفر) با استفاده از فیلم ویدیویی احساس، ۵۹/۷ درصد و با استفاده از تصویر نقطه اوج احساس ۲۹/۴ درصد است. بنابراین فیلم ویدیویی به میزان خیلی زیادی در تشخیص احساس قویتر از تک تصویر است. شکل ۱ دقت تشخیص احساس‌ها را در شرایط مختلف نشان می‌دهد. در این شکل سمت چپ نمودار دقت در حالتی است که صورت عادی و بدون پوشش رنگی مشاهده شده و در سمت راست نمودار دقت شناسایی تنها با استفاده از خالهای سفید است. نگاهی اجمالی به این شکل نشان می‌دهد که بطور کلی تشخیص احساس با استفاده از اطلاعات همه نقاط یک ناحیه بیشتر از شناسایی با استفاده از خالهای سفید مشخص شده در نقاط خاص است. همچنین نقش نواحی مختلف چهره در شناسایی هر احساس کاملاً مشخص است. مثلاً در شکل ۱ (سمت راست) می‌توان دید که قسمت بالای چهره (ناحیه چشمها) نقش بیشتری نسبت به ناحیه پایین در ظهور "ترس" دارد.



شکل ۱. دقت شناسایی احساس در آزمایش Bassili تحت شرایط مختلف [6].

وجود چنین ارتباطی بین نحوه تغییرات صورت و بروز احساس، مشوق تلاشهای زیادی است که از سال ۱۹۹۰ به بعد برای شناسایی خودکار احساس از روی تصویر چهره انجام یافته‌اند [7، 9].

تحقیقات زیادی که تا حالا انجام گرفته، روی پایگاههای متفاوتی آزمایش شده‌اند و بدین سبب از لحاظ دقت، مقایسه آنها با هم ممکن نیست. در اینجا تنها از جنبه بکارگیری ابزارها و روشهای شناسایی، چند نمونه از آنها را مرور می‌کنیم.

اولیه، توابع مشخصه‌ای تعریف شده است. سیستم آنها حرکات صورت را در قالب مجموعه‌ای از کدهای FACS می‌شناسد و دقت کلی آن برابر ۹۲ درصد است. Cohen و همکارانش در [14] با استفاده از توری<sup>۷</sup> تغییرات مکانی-زمانی صورت را استخراج می‌کنند. با بررسی ظهور تغییرات نواحی مختلف چهره و ارتباط آنها با هم، یک شبکه بیز<sup>۸</sup> ساخته می‌شود که در آن ویژگیهای مستقل از هم مشخص شده اند. برای تسریع الگوریتمهای شناسایی، به هر شبکه بیز تعریف شده، درختی برازش می‌شود. اطلاعات زمانی نیز بوسیله یک HMM<sup>۹</sup> مدل می‌شود. Jen-Jier و همکارانش در [15] با استفاده از شار نوری تغییرات و حرکات نواحی خاصی از صورت را در قالب واحدهای حرکت FACS شناسایی می‌کنند. دقت شناسایی حدود ۸۱ درصد بوده است.

Lucey و همکارانش در [16] با استفاده از مدلسازی صورت به روش AAM<sup>۱۰</sup> [17] حرکات و تغییرات چهره را در قالب پارامترهای مدل استخراج و تفسیر می‌کنند. سیستم آنها برخی از واحدهای حرکت را شناسایی می‌کند و دقت آن حدود ۷۰ درصد است.

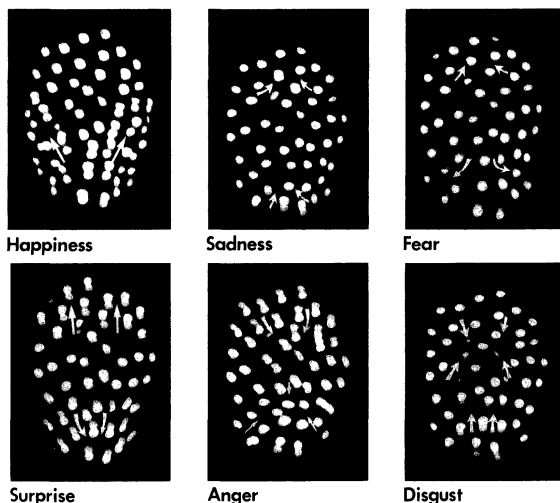
در دو نمونه از کارهای انجام شده از سیستمهای مبتنی بر قانون برای شناسایی حالات چهره استفاده شده که یکی از آنها از یک تصویر و دیگری از دنباله‌ای از تصاویر برای استخراج تغییرات صورت استفاده می‌کند.

همچنانکه در بالا اشاره کردیم، Pantic [18] ابتدا با استفاده از نماهای تمام رخ و نیم رخ، محل اعضای صورت را تشخیص داده و سپس با مقایسه آنها با حالت عادی شخص، تغییرات ایجاد شده را در قالب کدهای FACS شناسایی می‌کند. مثلا، احساس تنفر در نتیجه ظهور واحدهای حرکت (۱۰ و ۱۷ و ۶) یا (۹ و ۱۷ و ۲۶) گزارش می‌شود (جدول ۱). این سیستم واحدهای حرکت را دارای سهم مساوی در تولید احساس فرض می‌کند. مثلا، اگر در تصویر چهره‌ای، کدهای ۹ و ۲۶ شناسایی شود آنگاه شدت ظهور "تنفر" در این نمونه براساس ترکیب اول برابر ۱/۳ یا ۳۳٪ و بر اساس ترکیب دوم برابر ۲/۳ یا ۶۶٪ است. سیستم در نهایت نتیجه بزرگتر را بعنوان میزان ظهور احساس تنفر گزارش می‌کند. بدین ترتیب، خروجی سیستم تعدادی احساس است که واحد‌های حرکت شناسایی شده در تولید آنها نقش دارند. Pantic درباره رفع تناقضهای احتمالی بین احساس‌های شناخته شده، توضیحی نداده است.

Black و Yacoob [7] برای شناسایی احساس از تصاویر ویدیویی، حرکت چهره را در نواحی مختلف به صورت محلی مدل می‌کنند. مثلا با استفاده از مدل نواری شکلی که تنها می‌تواند انحنای عمودی داشته باشد، حرکت لبها را مدل می‌کنند. برای شناسایی احساس، از مدل‌های تغییر مناسب برای نمایش تغییرات چهره استفاده می‌کنند. هرگونه حرکت و تغییر حالات چهره یا سر در هر فریم در قالب پارامترهای مدل ظاهر می‌شود. پس از استخراج تغییرات پارامترها در هر فریم، بر اساس حدود آستانه‌ای که از چند نمونه آزمایشی تهیه شده، تغییرات ناچیز و قابل اغماض حذف می‌شوند. تغییرات باقیمانده در قالب تغییرات میانی، شروع، ادامه یا خاتمه یک احساس شناخته می‌شوند. آرایش خاص این تغییرات برای شناسایی احساس استفاده می‌شود. مثلا برای شناسایی احساس "شادی" باید "شروع خنده"، "ادامه" و "پایان خنده" در فریمهای متوالی صعودی، شناسایی شوند. برای تشخیص هر کدام از این مراحل، حرکات تشکیل دهنده آن مرحله باید همپوشانی<sup>۱۱</sup> قابل ملاحظه‌ای داشته باشند. خروجی سیستم آنها یکی از احساس‌های پایه ودقت شناسایی درست آن ۷۰٪ است.

## ۱-۲ سیستمهای خبره

هوش مصنوعی، ابزارهای طراحی و برنامه نویسی متعددی به جامعه کامپیوتر عرضه کرده است که از این میان می‌توان الگوریتمهای یادگیری ماشین، رباتیک و سیستمهای کنترل را نام برد. سیستم‌های خبره<sup>۱۲</sup> یکی دیگر از دستاوردهای این



شکل ۲. پیکانها، جهت تغییر مکان نقاط خاص روی چهره را هنگام بروز احساس نشان می‌دهند [6]

هرچند ثابت شده که حرکت چهره نقش مهمی در شناسایی احساس دارد، تلاشهایی متعددی نیز صورت گرفته تا این کار را با استفاده از تصویر نقطه اوج احساس و تصویر حالت عادی چهره انجام دهند. Dailey [10] با استفاده از شبکه های عصبی پیشرو، احساس بروز یافته در هر عکس را شناسایی کرده است. نتایج بدست آمده روی پایگاه POFA [2] ۹۰ درصد دقت داشته است. همچنین Pantic [11] از نماهای تمام رخ و نیم رخ در حالت های عادی و نقطه اوج احساس برای شناسایی احساس استفاده کرده است.

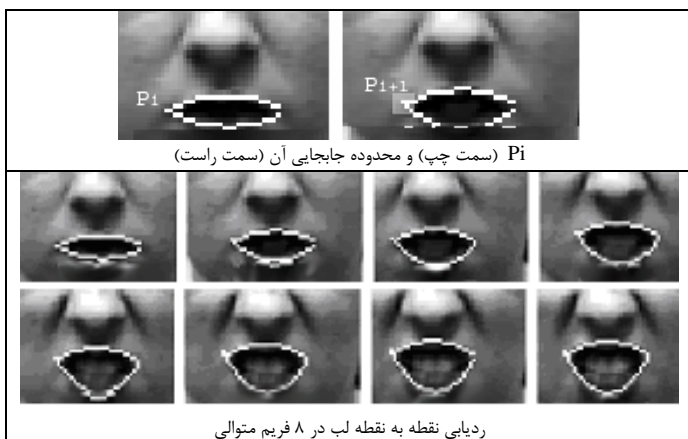
در شناسایی خودکار احساس از تصویر ویدیویی چهره با مسایل متعددی مانند یافتن چهره در تصویر و تعیین محل و ابعاد چشمها، بینی و دهان، آشکار سازی تغییرات صورت و اجزای آن طی یک بازه زمانی خاص و همچنین ایجاد ارتباط این تغییرات با حالت یا احساس شخص مواجه هستیم. هر کدام از این مسایل بسته به شرایط محیطی و شخصی، تنوع و دشواریهای خاص خود را به همراه دارد. مثلا هنگام تشخیص صورت و یافتن محل دقیق اجزای آن، ترکیب و آرایش ظاهر چهره مانند داشتن مو و عینک مشکلات متعددی را پدید می‌آورد که هر کدام موضوع تحقیقات مستقل و وسیعی است.

تحقیقاتی که برای شناسایی احساس یا حالات چهره انجام می‌گیرند، معمولا روی بخش محدودی از این مسایل تمرکز می‌کنند و با ارائه راه‌لهای جدیدتر و بهتر برای مشکلات بخش مذکور، تاثیر آن را بر نتیجه نهایی شناسایی از نظر دقت و کیفیت بررسی می‌کنند. بررسی تمام فعالیت‌های انجام شده در همه مراحل چنین سیستمی، فرصتی بیشتر از این مقاله می‌طلبد و علاقمندان می‌توانند مرور مناسبی بر آنها را در مراجع [۱، ۷-9] ببینند. همچنین، به دلیل اینکه این تحقیقات روی پایگاه‌های متفاوتی آزمایش شده اند، مقایسه آنها از نظر دقت، منطقی و درست نیست.

برای تشخیص و دسته بندی تغییرات چهره، تلاشهای زیادی صورت گرفته است. Essa و همکارانش در [12] با استفاده از نمونه های آموزشی، انرژی حرکت هر کدام از احساس‌های مورد نظر را محاسبه می‌کنند و برای شناسایی نمونه جدید، بعد از محاسبه انرژی حرکت<sup>۱۳</sup>، با استفاده از فاصله اقلیدسی، شبیه‌ترین احساس را انتخاب می‌کنند. با استفاده از ۵۲ نمونه آزمایشی، دقت کلی شناسایی شش احساس تعریف شده در سیستم، ۹۸٪ بوده است. این سیستم قادر به شناسایی احساس‌های مرکب یا تعیین شدت ظهور هر احساس نیست. Cohn و همکارانش در [13] نقاط معینی در نواحی اطراف چشمها، بینی و لبها تعیین کرده و سپس در طی فریمهای متوالی با استفاده از شار نوری<sup>۱۴</sup> این نقاط را دنبال کرده و جایابی آنها را پیدا می‌کنند. برای ارتباط تغییرات هر ناحیه با حرکات

در شکل ۴ فرض کنید  $P_i$  نقطه‌ای در فریم  $I$  بوده و  $P_{i+1}$  محل جدید این نقطه در فریم  $I+1$  باشد. با استفاده از معیار ضریب همبستگی<sup>۱۴</sup>، شبیه ترین نقطه در ناحیه جستجوی محدود در فریم  $I+1$  پیدا می‌شود. ناحیه جستجو در فریم مقصد و اندازه ناحیه‌ای که پیکسل‌های آن گروهی حرکت می‌کنند، بسته به شرایط مساله از جمله نرخ فریم در ثانیه<sup>۱۵</sup> تعیین می‌شوند. از آنجاییکه نقاط مختلف اعضای صورت مستقلا ردیابی می‌شوند، می‌توان آنها را هم زمان ردیابی کرده و زمان اجرا را کاهش داد. در ردیابی یک عضو صورت، ممکن است برخی از نقاط درست ردیابی نشوند. برای کم کردن آثار این نوع خطا و همچنین استخراج راحتتر ویژگیهای طولی، بر مجموعه نقاط ردیابی شده لبها، ابروها، و بالا و پایین چشمها، یک سهمی برازش می‌شود. نمونه‌ای از ردیابی لب به این روش در شکل ۴ نشان داده شده‌است.

برای استخراج تغییرات صورت، ابتدا موقعیت ابروها، چشمها، بینی و لبها در فریم اول به صورت دستی تعیین می‌شود و سپس با استفاده از این الگوریتم، نقاط مربوط به هر عضو ردیابی و تغییرات آن عضو استخراج می‌شود. روش تعیین این تغییرات برای هر عضو متفاوت است. مثلاً برای ابرو، فاصله گوشه‌های ابرو از گوشه‌های نظیر چشم و برای لب، طول و عرض آنها محاسبه می‌شود



شکل ۴. ردیابی یک نقطه (بالا) و ردیابی لب (پایین)

### ۲-۳ استخراج ویژگیهای مورد نیاز

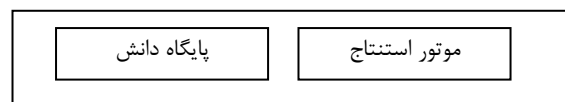
برای استخراج واحدهای حرکت از سه دسته ویژگی استفاده شده است. دسته اول شکل ۵-الف ویژگیهایی از جنس طول هستند. دسته دوم شکل ۵-ب ویژگیهایی از جنس فراوانی و بیانگر تعداد نقاط لبه در بخشهای خاصی از صورت هستند و دسته سوم شکل ۵-ج ویژگیهایی از جنس روشنایی هستند. برای استخراج فراوانی نقاط، با استفاده از الگوریتم Canny [۱۶] لبه های نواحی شکل ۵-ب را پیدا می‌کنیم. این الگوریتم، نقاط لبه را با توجه به توزیع روشنایی محلی شناسایی می‌کند. لبه افقی، عمودی یا مورب را با استفاده از فیلترهای آشکار ساز لبه مانند Sobel جدا می‌کنیم. چینهای نامنظم چانه و ناحیه بین دو چشم را مستقیماً از لبه ها استنتاج می‌کنیم. جدولهای ۲، ۳ و ۴ این مشخصات را تعریف می‌کنند. استخراج مشخصات طولی نسبتاً آسان است. اما در شرایطی بدلیل برخی ویژگیهای طبیعی انسان، تشخیص آنها به پردازشهای مضاعفی نیاز دارد. مثلاً افرادی هستند که هنگام بالا رفتن ابروهایشان موقعیتهای انتهایی دو ابرو تغییر چندانی نمی‌کند و تنها انحناي آن زیاد می‌شود. در این حالت، طول پاره خطهای AE و BD در ۵-الف تقریباً ثابت می‌ماند. یکی از دلایل استفاده از مدل سهمی هم تشخیص اینگونه مسایل است.

علم است که روشهایی را برای "برنامه نویسی مبتنی بر دانش" و کاربردهای آن ارائه می‌کند. سیستم خبره چنین تعریف می‌شود [8].  
سیستم خبره، برنامه‌ای کامپیوتری است که برای مدل سازی توانایی حل مساله انسان طراحی شده است.

هر سیستم خبره دارای دو بخش پایگاه دانش و موتور استنتاج است. نمودار بلوکی یک سیستم خبره در شکل ۳ نشان داده شده است. پایگاه دانش، مجموعه غنی از دانسته های افراد خبره در زمینه مساله مورد بحث است و شامل حقایق قوانین و روابط<sup>۱۳</sup>، بین آنهاست. همانند شیوه استدلال افراد خبره در زمینه مساله مورد نظر، موتور استنتاج با استفاده از دانش موجود در پایگاه دانش، حقایق جدیدی را استنتاج می‌کند و پیشنهادها و راهکارهایی را برای حل مساله مطرح می‌کند. این موتور در واقع مجموعه‌ای از الگوریتم های جستجو و استنتاج همه منظوره است.

همانطور که در شکل ۳ دیده می‌شود، پایگاه دانش از موتور استنتاج جداست و این یکی از مهمترین نقاط قوت سیستم‌های خبره است. زیرا این پایگاه می‌تواند بسته به زمان، مکان و خواسته های کاربر سیستم تغییر کند و حتی سیستم خبره طراحی شده برای یک مساله خاص، با تغییر دانش برای مسایل دیگر بکار گرفته شود. همچنین پایگاه دانش سیستم می‌تواند به صورت افزایشی بهبود یابد. یعنی به مرور زمان، حقایق و قوانین جدیدی به آن اضافه شود و دانش قبلی دستخوش حذف یا تغییر گردد. امروزه روشهای خودکاری برای افزایش و پالایش دانش سیستمهای خبره پیشنهاد شده است [9].

سیستم خبره



شکل ۳. نمودار بلوکی یک سیستم خبره

### ۳- سیستم پیشنهادی

سیستم پیشنهادی در این مقاله، روشی مشابه کار Pantic [11] را برای شناسایی احساس از روی تصویر ویدیویی بکار می‌گیرد. در این سیستم، ابتدا تغییرات صورت در فریمهای متوالی استخراج شده و سپس این تغییرات در مرحله اول به حرکات اولیه و در مرحله بعد به احساس تفسیر می‌شوند.

### ۳-۱ استخراج تغییرات صورت

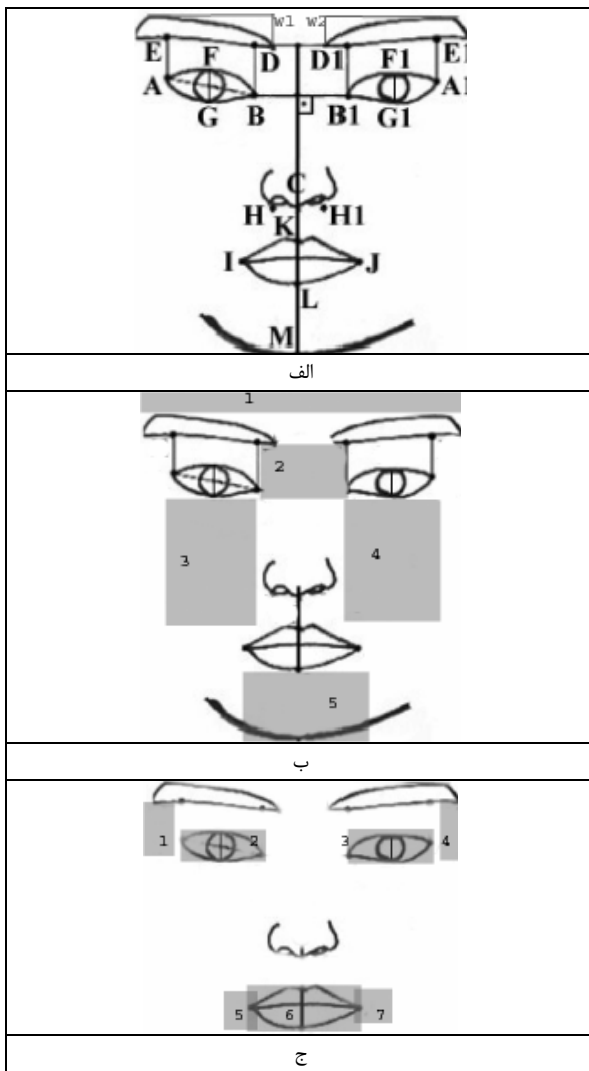
تغییرات صورت ناشی از تغییر شکل اعضای آن هنگام بروز یک حالت یا احساس هستند.

بنابراین باید تغییر شکل و حرکت اعضا را در فریمهای متوالی شناسایی کنیم. تشخیص این تغییرات و بطور کلی تشخیص حرکت، موضوع تخصصی مجزایی است که بر اساس اهمیت و کاربرد آن در بسیاری از زمینه های مرتبط با بینایی ماشین، بطور جداگانه مطالعه می‌شود. نحوه تغییر صورت و اعضای آن ممکن است بسیار پیچیده باشد و بررسی چنین شرایطی موضوع بحث این مقاله نیست. بنابراین تنها حالت بسیار خاص و کنترل شده‌ای در مورد حرکات صورت فرض شده و استخراج تغییرات صورت نیز براساس این فرضها با الگوریتمهای معمول صورت می‌گیرد. فرضهای لحاظ شده برای دنباله تصاویر، چنین هستند:

۱- روشنایی فریمها یکنواخت است و سر حرکت محسوسی ندارد.

۲- تغییر مکان نقطه در دو فریم متوالی اندک است.

۳- تغییر مکان نقاط در یک ناحیه کوچک، گروهی است.



شکل ۵- مشخصات مختلف استفاده شده: الف - مشخصات طولی ب- فراوانی نقاط لبه و ج- روشنایی نواحی [10]

یکی از مشکلات موجود در شناسایی چینه‌ها، وجود چینه‌های طبیعی در صورت اشخاص به ویژه در افراد بزرگسال است. برخی افراد بطور طبیعی در نواحی اطراف چشم و پیشانی چین و چروک‌هایی دارند. در اینجا استفاده از فراوانی نقاط بدلیل حساسیت به پهنای و طول چینه‌ها تاحدی این مشکل را حل کرده است.

همچنین مشخصات مبتنی بر میانگین روشنایی نیز گاهی متأثر از عدم یکنواختی رنگ چهره هستند. همچنین بسته به جهت تابش نور، ممکن است سایه برخی نواحی صورت روی نواحی دیگر بیفتد. در اینگونه حالات، تا جایی که این سایه‌ها در فریم‌های متوالی یکسان باشند، مشکلی پیش نمی‌آید. اما در صورت حرکت سر و جابجایی یا کم و زیاد شدن سایه‌ها، مشخصات مرتبط با روشنایی دقت لازم را نخواهند داشت. در مورد برخی از مشخصات مثل بر آمدن گونه‌ها، این مشکل را می‌توان با استفاده از تصویر نیم رخ چهره حل کرد. همچنین با تصویر سه بعدی چهره می‌توان بدون نیاز به این دسته از ویژگی‌ها، تغییرات مورد نظر را استخراج کرد [11].

بر اساس ویژگی‌های ذکر شده در جدول های ۲، ۳ و ۴ و همچنین توضیحات FACS، واحدهای حرکت را می‌توان شناسایی کرد. مثلاً واحد ۱ در نتیجه بزرگتر شدن ویژگی‌های f7 و f8 رخ می‌دهد. در جدول ۶ فهرست تعدادی از واحدهای بررسی شده در این مقاله و ارتباط آنها با این ویژگی‌های تعریف شده، آمده است. باید توجه داشت که در FACS هیچگونه اشاره‌ای به مشخصات استفاده شده در این تحقیق نشده است. در FACS هر واحد حرکت با یک شماره منحصر بفرده، یک عنوان و توضیحی در باره نحوه بروز واحد حرکت و ماهیچه‌های درگیر در آن مشخص شده‌اند. تبدیل توضیحات ارائه شده در FACS با استفاده از تصاویر نمونه همراه آن انجام شده است. در نسخه اخیر FACS برخی از واحدهای حرکت به دو واحد چپ و راست تقسیم شده‌اند مثلاً واحد حرکت ۱ به دو واحد AU1L (چپ) و AU1R (راست) تقسیم شده است. این تجدید نظر از آنجا ناشی می‌شود که گاهی واحد حرکت ۱ تنها در یکی از ابروها مشاهده می‌شود و احساس بیان شده هم مفهوم متفاوتی از حرکت هر دو ابرو دارد. برخی از واحدهای حرکت مثل AU26 برحسب ارتباط احساس بیان شده با میزان ظهور حرکت به چند واحد مجزا مثل AU26a و AU26b تقسیم می‌شوند. تعریف واحدهای جدید هنوز کاملاً دقیق نیست، از این رو در این تحقیق مورد توجه نبوده‌اند.

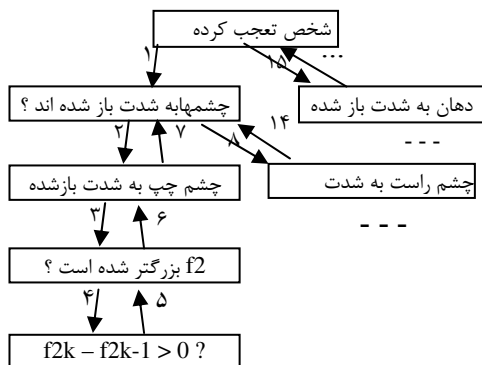
جدول ۲. مشخصات طولی مورد استفاده شکل ۴- الف

کد	مشخصه	توضیح
f1	طول پاره خط AB	طول چشم چپ (از دید ناظر)
f2	طول پاره خط FG	عرض چشم چپ
f3	طول پاره خط A1B1	طول چشم راست (از دید ناظر)
f4	طول پاره خط FIG1	عرض چشم راست
f5	طول پاره خط AE	ارتفاع گوشه خارجی ابروی چپ
f6	طول پاره خط A1E1	ارتفاع گوشه خارجی ابروی راست
f7	طول پاره خط DB	ارتفاع گوشه داخلی ابروی چپ
f8	طول پاره خط D1B1	ارتفاع گوشه داخلی ابروی راست
f9	W1	ارتفاع ابروی چپ
f10	W2	ارتفاع ابروی راست
f11	طول پاره خط KL	عرض دهان
f12	طول پاره خط IJ	طول دهان
f13	طول پاره خط HH1	پهنای بینی
f14	طول پاره خط KC	پهنای لب بالا (ناحیه بین لب و بینی)
f15	فاصله K از IJ	ارتفاع لب بالا
f16	فاصله L از IJ	ارتفاع لب پایین
f17	طول پاره خط LM	ارتفاع چانه

### ۳-۳ شناسایی واحدهای حرکت

پس از استخراج ویژگی‌های مختلف مورد نیاز، آنها را به صورت حقایق اولیه بازنویسی می‌کنیم. قوانینی که بعنوان دانش سیستم تعریف شده‌اند، حقایق جدیدی را از این حقایق اولیه استخراج می‌کنند و این فرایند به صورت افزایشی ادامه پیدا می‌کند. مثلاً این حقیقت که  $f7i+1$  بزرگتر از  $f7i$  است، با استفاده از قانونی بدین شکل بدست می‌آید:  $f7i+1$  بزرگتر از  $f7$  است اگر  $f7i+1 - f7i > 0$ . بنابراین تنها مقادیر ویژگی‌های مورد نظر در دو فریم متوالی به سیستم خبره داده می‌شود و نوع تغییرات آنها با استفاده از قوانینی که تعریف شده‌اند استخراج می‌شود. فرایند انتخاب شده برای روش استنتاج، استدلال پس رونده<sup>۱۶</sup> است. در این روش سیستم خبره از قوانین کلی تر شروع کرده و مسیر استنتاج را به سمت قوانین جزئی و حقایق اولیه شکل می‌دهد. مزیت این روش، فرایند جستجوی متناسب با هدف<sup>۱۷</sup> است. یعنی تنها حقایق و قوانینی بررسی می‌شوند که در استنتاج منتهی به هدف نقش داشته باشند. نمونه چنین استنتاجی در شکل ۶ آمده است.

می‌کنند. همچنین در ادغام شواهد متعدد ظهور یک احساس، از این معیار استفاده می‌شود.



شکل ۶. نمونه‌ای از ترتیب استنتاج پس رونده در سیستم‌های خبره

برای اینکه بتوانیم از این شدت ظهور در مراحل بعدی و تولید قوانین استفاده مناسبی داشته باشیم، شدت ظهور هر واحد حرکت را مطابق رابطه (۱) تعریف می‌کنیم که در آن I شدت ظهور یک واحد، V1 مقدار پیشین و V2 مقدار جدید ویژگی تولید کننده واحد مورد نظر است. برخی از واحدهای حرکت ممکن است از دو منبع مختلف ایجاد شوند. مثلا واحد حرکت ۱ می‌تواند در نتیجه حرکت ابروی چپ یا راست باشد. در اینجا با استفاده از تئوری استنتاج احتمالی [4]، کد شناخته شده را تقویت می‌کنیم. یعنی اگر واحد حرکت ۱ در نتیجه حرکت ابروی چپ با شدت I1 و در نتیجه حرکت ابروی راست با شدت I2 ظاهر شده باشد، شدت کلی واحد ۱ (I) را از رابطه ۲ محاسبه کرده و I1 و I2 را حذف می‌کنیم:

$$I = \text{Min}(1, \text{abs}(V2 - V1) / V1) \quad (1)$$

$$I = I1 + I2 - I1 \times I2 \quad (2)$$

در مواردی که منابع یاد شده بیش از دو مورد باشند، روابط بالا ابتدا برای دو منبع اعمال می‌شوند و سپس نتیجه حاصل با منبع سوم ادغام می‌گردد. با تکرار این مرحله برآیند تمام منابع محاسبه می‌شود. ممکن است واحدی با شدت خیلی کم ظاهر شده باشد. در این صورت، ابتدا واحدهای تکراری را که از منابع مختلف ظاهر شده اند، با استفاده از روشی که در بالا توضیح دادیم، در هم ادغام می‌کنیم. آنگاه واحدهایی را که شدت آنها کمتر از یک حد خاص (مثلا ۰.۱) باشد، حذف می‌کنیم. بعنوان نمونه برای شناسایی واحد حرکت ۱ که بیانگر حرکت رو به بالای گوشه داخلی ابروهاست، از چنین قانونی استفاده می‌کنیم:

$f_{i+1}$  بزرگتر از  $f_i$  است.  
در این صورت واحد حرکت ۱ با شدت  $\min(1, (f_{i+1} - f_i) / f_i)$  ظاهر شده است.

در این قانون  $f_{i+1}$  و  $f_i$ ، ارتفاعهای گوشه داخلی ابروی چپ بترتیب در فریم‌های  $i$  و  $i+1$  هستند. واحدهای دیگر حرکت نیز با قوانین مشابهی شناخته می‌شوند.

### ۴-۳ شناسایی احساس

در این مرحله، واحدهای شناسایی شده در قسمت قبل را به زیر مجموعه‌ای از احساس‌های اولیه (خوشحالی، تعجب، بی‌اعتنائی، تفرق، غم، خشم و ترس) تفسیر

جدول ۳. فراوانی نقاط لبه در نواحی شکل ۴-ب

کد	مشخصه	توضیح
f18	میانگین نقاط لبه ستونهای ناحیه ۱	چینه‌های افقی پیشانی
f19	تعداد نقاط لبه در ناحیه ۲	چینه‌های نامنظم بین چشمها
f20	تعداد نقاط لبه در ناحیه ۳	چین مورب در گونه چپ
f21	تعداد نقاط لبه در ناحیه ۴	چین مورب در گونه راست
f22	تعداد نقاط لبه در ناحیه ۵	چینه‌های نامنظم چانه
f23	میانگین نقاط لبه ستونهای ناحیه ۵	چینه‌های افقی چانه

جدول ۴. میانگین روشنایی نواحی شکل ۴-ج

کد	مشخصه	توضیح
f24	ناحیه ۱	برآمدن گونه چپ
f25	ناحیه ۲	باز یا بسته بودن چشم چپ
f26	ناحیه ۳	باز یا بسته بودن چشم راست
f27	ناحیه ۴	برآمدن گونه راست
f28	ناحیه ۵	باز یا بسته شدن گوشه چپ لب
f29	ناحیه ۶	باز یا بسته شدن گوشه راست لب
f30	ناحیه ۷	باز یا بسته شدن دهان

جدول ۵. نحوه تغییر واحدها با ویژگیها

کد	تغییر ویژگیها
۱	f7 و f8 بزرگتر می‌شوند.
۲	f5 و f6 بزرگتر می‌شوند.
۴	(f5, f7) یا (f6, f8) کوچکتر شده است.
۵	(f2, f5, f7) یا (f4, f6, f8) بزرگتر شده اند.
۶	f24 یا f27 کوچکتر شده است.
۷	f2 و f4 کوچکتر شده اند.
۹	f19 بزرگتر شده است.
۱۰	f14 کوچکتر شده است.
۱۱	f13 کوچکتر شده است.
۱۲	f12 بزرگتر شده است و f11 کوچکتر شده است.
۱۶	f17 کوچکتر شده است.
۱۷	f17 کوچکتر نشده است. f22 بزرگتر شده است.
۱۸	f11 افزایش نیافته و f12 کاهش یافته است.
۲۰	f12 بزرگتر شده است.
۲۳	f12 کوچکتر شده و f15 و f16 بزرگتر شده اند.
۲۴	f15 و f16 کوچکتر شده اند. f17 کوچکتر نشده است.
۲۵	f11 بزرگتر از حد آستانه t0 و کوچکتر از t1 شده است. <sup>۱۸</sup>
۲۶	f11 بزرگتر از حد آستانه t1 و کوچکتر از t2 شده است.
۲۷	f11 بزرگتر از حد آستانه t2 شده است.
۲۸	f11 کوچکتر از حد t0 است. f12 کوچکتر نشده است.
۴۱	f9 یا f10 بزرگتر شده اند. <sup>۱۹</sup>
۴۲	f9 یا f10 بزرگتر شده اند. f2 > 0 و f4 > 0.
۴۳	f2=0 و f4=0

در اینجا برای هر واحد حرکت، شدتی تعریف می‌شود که حاکی از میزان وقوع یا ظهور آن واحد است. برخی از واحدهای حرکت نظیر ۲۵ و ۲۶ و ۲۷ (جدول ۱) برای تفکیک ازم به چنین معیاری نیاز دارند و واحدهای دیگر نیز حداقل برای چشم پوشی از حرکات اندک، از این معیار بعنوان حد آستانه‌ای مناسب استفاده

پایگاه دربردارنده اشخاص با ملیت، نژاد، جنس و سنهای مختلف است. شکل ۷ تعدادی از تصاویر این پایگاه را نشان می‌دهد. هر دنباله تصویر این پایگاه بیانگر یکی از احساس‌های جدول ۶ مربوط به یک نفر است. سر شخص حرکت چندانی ندارد.



شکل ۷. نمونه‌هایی از تصاویر پایگاه Cohn-Kanade [۲۲]

عمده‌ترین بخش محاسبات در ردیابی نقاط به محاسبه ضریب همبستگی اختصاص دارد. برای بلوکی به ابعاد  $n \times n$  در فضای جستجوی  $R \times R$ ، این کمیت  $R^2$  بار محاسبه می‌شود. ضریب همبستگی مطلق دوناچی  $X \times n$  و  $Y \times n$  چنین محاسبه می‌شود که در آن به  $4n^2$  عمل ضرب نیاز است:

$$r = (\|X\| \times \|Y\|) / \|X \cdot Y\| \quad (۳)$$

در این رابطه،  $\| \cdot \|$  نشانگر اندازه اقلیدسی بلوک است. بنابراین تعداد کل عمل ضرب در ردیابی نقاط انتخاب شده بین دو فریم برابر است با:

$$N = 4 K \times R^2 \times n^2 \quad (۴)$$

که در آن  $K$  تعداد نقاط انتخاب شده است.

برای پیاده‌سازی ردیابی و استخراج ویژگی از نرم افزار MATLAB6.5 و برای پیاده‌سازی سیستم خیره از ابزار [23]CLIPS استفاده کردیم. برای ردیابی اعضای چهره در فریمهای متوالی ۵۶ نقطه را انتخاب کردیم. اندازه بلوک اطراف هر نقطه را  $5 \times 5$  و ناحیه جستجو را  $13 \times 13$  انتخاب کردیم.

در مرحله آموزش، با استفاده از چند تصویر این بانک حد آستانه ۱۰٪ برای ظهور هر واحد حرکت تعیین شد. آنگاه ۱۰۵ دنباله غیر از مجموعه آموزش برای آزمایش انتخاب شد. جدول ۸ ماتریس سردرگمی<sup>۱۸</sup> شناسایی واحدهای حرکت را نشان می‌دهد. سطر و ستون اول این ماتریس، کد واحدهای حرکت را نشان می‌دهند و درآیه سطر  $m$  و ستون  $n$  نشان دهنده تعداد واحدهای حرکت با کد  $m$  است که سیستم ما آنها را واحد حرکت  $n$  شناخته است. مثلاً عدد ۳ در درآیه  $(۱, ۲)$  بیانگر این است که سیستم ۳ مورد از واحدهای حرکت ۲ را به اشتباه واحد ۱ شناسایی کرده است. قطر اصلی این جدول نشان دهنده تعداد واحدهای حرکت  $m$  است که سیستم نیز آنها را  $m$  شناخته است و بنابراین تعداد نمونه‌های شناخته شده درست را در بردارد. ستون  $X$  تعداد نمونه‌هایی از واحد  $h$  را نشان می‌دهد که سیستم آنها را  $z$  زده است. ستون  $O$  تعداد واحدهایی را نشان می‌دهد که بدلیل داشتن شدت ظهوری کمتر از حد آستانه تعیین شده، حذف شده‌اند. ستون آخر دقت شناسایی درست را برحسب درصد بیان می‌کند. دقت<sup>۱۹</sup> شناسایی از تقسیم مقادیر قطر اصلی (شناسایی درست) بر مجموع سطرها (تعداد نمونه‌های موجود هر واحد حرکت) بدست آمده است. سطر آخر جدول بیانگر میزان صحت (اعتبار)<sup>۲۰</sup> اشاره کند نتایج بدست آمده است. این کمیت از تقسیم مقدار قطر اصلی بر مجموع مقادیر ستون محاسبه شده است.

می‌کنیم. در اینجا می‌توان از برداشت اشخاص مختلف استفاده کرد. همچنین می‌توان با تکیه بر تشابه احساس‌های پایه بین افراد مختلف، از تحقیقات انجام یافته برای تطابق مجموعه واحدهای حرکت و احساس بهره برد. با استفاده از هریک از این روشها یا تلفیق آنها می‌توان جدولی مانند جدول ۶ را برای شناسایی احساس تدوین کرد:

جدول ۶. رابطه احساس‌ها و واحدهای حرکت

احساس	نمونه مجموعه واحدهای حرکت تشکیل دهنده
شادی	۱۲،۶،۲۵
تعجب	۲۷،۲،۱
غم	۴،۷،۱۷
بی‌اعتنایی	۱،۲،۲۷
خشم	۴،۵،۳۸
تفر	۴،۷،۹
ترس	۲،۲۰،۲۵

ما از تلفیق دو روش استفاده کردیم. بدین معنی که ابتدا مجموعه‌ای از تصاویر را به چند نفر نشان داده و احساسی را که از تصویر برداشت می‌کردند، جمع‌آوری کردیم. در مواردی که عقاید افراد مختلف در مورد احساس ظاهر شده در تصویر متفاوت بود، از نتایج کار Bassili استفاده کردیم [6]. احساس‌های مختلف موجود در یک فریم بر اساس واحدهای حرکت موجود در آن فریم با قوانینی به شکل زیر شناسایی می‌شوند:

- واحد حرکت ۱۲ با شدت  $I_1$  ظاهر شده است.
- واحد حرکت ۲۵ با شدت  $I_2$  ظاهر شده است.
- آنگاه:
- "شادی" با شدت  $(I_1 + I_2) / 2$  ظاهر شده است.

از آنجا که هر احساس می‌تواند در نتیجه ظهور بیش از یک دسته از ویژگیها رخ داده باشد، مشابه ادغام واحدهای حرکت، احساس‌های تکراری در هم ادغام شده و در قالب یک احساس با شدت قویتر ظاهر می‌شوند. همچنین در شرایط خیلی معدودی ممکن است احساس‌های متناقض در یک فریم ظاهر شوند. قوانینی تعریف شده است که در چنین حالاتی، احساس با شدت کمتر را حذف کرده و تناقض را از بین می‌برد. بعنوان دانش (حقایق) اولیه، احساس‌های متناقض به شکل زیر تعریف شده‌اند:

#### Contradict (Happy, Sad)

این گزاره بدین معنی هست که حالات "خوشحال" و "غمگین" با هم ناسازگارند. آنگاه، قانونی به شکل زیر تناقضهای احتمالی را حذف می‌کند.

- در فریم  $n$  احساس  $e_1$  با شدت  $I_1$  ظاهر شده است.
- در فریم  $n$  احساس  $e_2$  با شدت  $I_2$  ظاهر شده است.
- $e_1$  و  $e_2$  متناقض هستند.
- $I_1 > I_2$
- آنگاه:  $e_2$  را حذف کن.

#### ۴- نتایج آزمایش

برای آزمون سیستم از مجموعه تصاویر ویدیویی پایگاه Cohn-Kanade استفاده شده است [۱]. این پایگاه شامل بیش از ۵۰۰ دنباله تصویر خاکستری با اندازه  $640 \times 490$  پیکسل است که به صورت دستی واحدهای آنها استخراج شده است.



## ۵- نتیجه گیری

تغییرات چهره بیان می‌شوند [۱]. همانطور که در بخش اول هم اشاره شد، واحدهای حرکت همه انواع تغییرات غیر قابل تقسیم صورت را در بر می‌گیرد. بدین ترتیب می‌توان هر نوع حالت چهره را که مرکب از حرکات اولیه است، تعریف کرده و با توسعه دانش سیستم آن را شناسایی کرد. اهمیت این کار زمانی آشکارتر می‌شود که بخواهیم حالتها و احساس‌هایی غیر از موارد کلیشه‌ای مطرح شده در اینجا را بشناسیم. مزیت دیگر این گام اولیه، خصوصی سازی سیستم برای افراد و جوامع خاص با ویژگیهای فرهنگی خودشان است. مثلا ابراز مخالفت در برخی جوامع با حرکت سر و در برخی دیگر با حرکت لبها و چانه بیان می‌شود. یکی دیگر از زمینه‌های فعالیتهای آتی، افزودن امکان خصوصی سازی خودکار به سیستم است.

مبنای دانش بکار رفته در سیستم، FACS 0/1 است که مرتبا در حال بازنگری، تصحیح و تکامل است. اعمال آخرین تغییرات FACS برای شناسایی واحدهای حرکت، یکی دیگر از کارهایی است که باید انجام پذیرد.

## سیاسگزاری

از آقای پروفیسور Zara Ambadar از دانشگاه پیتسبورگ که پایگاه Cohn-Kanade [23] را در اختیار ما قرار دادند و پروفیسور Scherer از دانشگاه ژنو که مجموعه مقالات و گزارشهای فنی مرتبط با نتیجه مطالعات روانشناختی خود را در زمینه احساس‌ها در اختیار ما قرار دادند، صمیمانه تشکر می‌کنیم.

## مراجع

[۱] م. منصوریزاده، شناسایی احساس با استفاده از تصویر ویدیویی چهره، پایان‌نامه کارشناسی ارشد، بخش مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، ۱۳۸۲.

[2] P. Ekman, W. V. Friesen, and J. C. Hager, *Facial Action Coding System (FACS), The Manual*. Salt Lake City: A Human Face, 2002.

[3] P. Ekman, "Facial Expressions," in *Handbook of Cognition and Emotion*, T. Dalgleish and M. Power, Eds. New York: John Wiley & Sons Ltd, 1999.

[4] D. Galati, K.R. Scherer and P.R. Bitti, "Voluntary facial expression of emotion: Comparing congenitally blind to normal sighted encoders," *Journal of Personality and Social Psychology*, Vol. 73, No. 6, pp. 1363-79, 1997.

[5] T. Striano and U. Liskowski, "Sensitivity to the Context of Facial Expression in the Still Face at 3-, 6-, and 9-Months of Age," *Infant Behavior & Development*, Vol. 28, No.1, pp. 10-19, 2005.

[6] J. N. Bassili, "Emotion Recognition: The Role of Facial Movement and the Relative Importance of Upper and Lower Areas of the Face," *Journal of Personality and Social Psychology*, Vol. 37, No. 11, 1979.

[7] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying Facial Actions," *IEEE TPAMI*, Vol. 21, No.10, pp. 974-989, 1999.

[8] B. Fasel and J. Luetten, "Automatic Facial Expression Analysis: A Survey," *Pattern Recognition*, Vol. 36, No. 1, pp. 259-275, 2003.

[9] M. Pantic and L. J. M. Rothkrantz, "Automatic Analysis of Facial Expressions: The State of The Art," *IEEE TPAMI*, Vol. 22, No.12, pp. 1424-1445, 2000.

[10] M. N. Dailey, G. W. Cottrell, C. Padgett, and R. Adolphs, "EMPATH: A Neural Network that Categorizes Facial

در این مقاله سیستمی برای شناسایی احساس از روی تصاویر ویدیویی چهره، با استفاده از سیستمهای خبره ارائه کردیم. ارائه روش یکسانی برای محاسبه شدت ظهور واحدهای حرکت، تقویت نتایجی که از منابع مختلف بدست می‌آیند و رفع خودکار تناقض در شناسایی احساس، از نقاط قوت این سیستم است. با توجه به اینکه اغلب کارهای شناسایی احساس از روی تصاویر ویدیویی قادر به شناسایی احساس‌های مرکب نیستند، شناسایی مجموعه‌ای از احساس‌ها از نو آوری‌های سیستم پیشنهادی است. حذف واحد حرکت‌هایی که با شدت ظهور پایینتر از حد آستانه معینی ظاهر شده اند، سبب می‌شود شدت ظهور احساس‌های مرتبط از این امر تاثیر نپذیرند. بهره‌گیری از سیستم خبره برای شناسایی واحدها و احساس‌ها، انعطاف زیادی به دانش سیستم می‌دهد. بنحوی که آموزش سیستم و حتی پالایش و اصلاح دانش موجود به صورت خودکار و با استفاده از قوانینی مشابه قوانین بکار رفته برای شناسایی قابل انجام است.

احساس‌های تعریف شده شکل بروز مشابهی در افراد جوامع مختلف دارند. با این وجود، شدت ظهور آن در افراد مختلف ممکن است متفاوت باشد. برای خصوصی و سفارشی کردن این سیستم برای افراد خاص لازم است سیستم با اطلاعات آن فرد آموزش داده شود. خودکار بودن فرایند پالایش دانش در سیستمهای خبره به ما کمک خواهد کرد تا بتوانیم فرایند خصوصی سازی را نیز به صورت خودکار انجام دهیم.

دقت شناسایی احساس در حد ۸۰ درصد با سیستمهای معرفی شده در بخش ۲ قابل مقایسه است. عمده خطای سیستم (بیش از ۵۰ درصد) در شناسایی واحدهای حرکت، بخاطر ردیابی اشتباه برخی نقاط اعضای چهره پیش می‌آید و بکارگیری روشی قویتر برای ردیابی، می‌تواند مفید و موثر باشد. یکی از مهمترین راههای ابراز احساس، حرکت سر در جهت‌های مختلف است. همچنین گاهی بروز احساس در حالت‌های مختلف سر معنی متفاوتی از احساس واقعی دارد. مثلا کج کردن سر به همراه خنده، بیانگر تمسخر است. یکی از مهمترین بخش تحقیقات آینده، ردیابی و استخراج حرکات سر است.

یکی از مشکلات شناسایی واحدها، از دست رفتن اطلاعات سه بعدی، هنگام استفاده از تصویر دو بعدی است. سیستم فعلی برای استفاده از تصاویر سه بعدی، نیاز به تعریف مجدد تعدادی از ویژگیها دارد. از آن جمله باید نقاط جدیدی مانند بالاترین نقطه گونه‌ها و چانه را نام برد. با وجود حرکت سر که می‌تواند جایجایی یا دورانی باشد، ممکن است برخی از اعضای صورت در فریمی ناپدید شوند. مثلا وقتی سر چرخشی افقی کرده باشد ممکن است اعضای یک نیمه صورت ناپدید شده باشند. پیدا کردن مجدد آنها در یک فریم و ادامه کار ردیابی یکی دیگر از فعالیتهای آینده است.

AAM [17] روشی برای مدل کردن صورت است که هم از مشخصات هندسی اجزا و هم از ترکیب بافت و روشنایی برای نمایش صورت استفاده می‌کند. این نوع مدلسازی ضمن استفاده از نحوه ترکیب اجزای صورت، کاهش ابعاد قابل ملاحظه‌ای نیز به ارمغان می‌آورد. در آینده سعی خواهیم کرد با استفاده از این مدلسازی، سیستم را بازنویسی کنیم.

از آنجاییکه در محیط واقعی و طبیعی نور و پس زمینه بسیار متنوع و معمولا خارج از کنترل است، سیستم باید خود را به صورت پویا با این شرایط تطبیق دهد. در این سیستم ابتدا واحد حرکات را شناسایی کردیم و سپس با استفاده از نتایج این بخش، به شناسایی احساس پرداختیم. در اینجا، تنها احساس‌های پایه را که بر اساس مطالعات روانشناسان شکل نمایش مشابهی بین جوامع مختلف دارند، مورد توجه قرار دادیم و بدین وسیله توانستیم بنیه نظری این کار را بر تحقیقات انجام شده استوار کنیم. اما بدیهی است که احساس‌های آدمها منحصر به این چند مورد نیست و در برخی موارد صدها احساس برشمرده شده که همگی بوسیله حالات و



**نصرا... مقدم چرکری** در سال ۱۳۶۵ کارشناسی خود را در مهندسی کامپیوتر از دانشگاه شهید بهشتی و کارشناسی ارشد مهندسی کامپیوتر و دکتری مهندسی سیستم‌های اطلاعات خود را از دانشگاه یاماناشی ژاپن به ترتیب در سالهای ۱۳۷۱ و ۱۳۷۴ دریافت کرد. او اکنون استادیار بخش مهندسی برق و کامپیوتر دانشگاه تربیت مدرس است. زمینه های پژوهشی مورد علاقه او بینایی ماشین، پردازش تصویر و الگوریتم‌های موازی است. آدرس پست الکترونیک: [charkari@modares.ac.ir](mailto:charkari@modares.ac.ir)



**احسان‌اله کبیر** کارشناسی ارشد پیوسته خود را در مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستم‌های الکترونیک از دانشگاه اسکس در انگلستان، به ترتیب در سالهای ۱۳۶۴ و ۱۳۶۹ دریافت کرد. او اکنون دانشیار بخش مهندسی برق دانشگاه تربیت مدرس است. زمینه های پژوهشی مورد علاقه او بازشناسی الگو، به ویژه بازشناسی متون چاپی و دستنویس و بینایی کامپیوتری است. آدرس پست الکترونیک: [kabir@modares.ac.ir](mailto:kabir@modares.ac.ir)

- <sup>1</sup> Action Unit
- <sup>2</sup> Facial Action Coding System
- <sup>3</sup> Intensity
- <sup>4</sup> Expert Systems
- <sup>5</sup> Motion Energy
- <sup>6</sup> Optical Flow
- <sup>7</sup> Wire Frame
- <sup>8</sup> Bayesian Network
- <sup>9</sup> Hidden Markov Model
- <sup>10</sup> Active Appearance Models
- <sup>11</sup> Overlap
- <sup>12</sup> Expert Systems
- <sup>13</sup> Facts, Rules, Relationships
- <sup>14</sup> Correlation Coefficient
- <sup>15</sup> Frame rate
- <sup>16</sup> Backward Reasoning
- <sup>17</sup> Goal Oriented
- <sup>18</sup> Confusion Matrix
- <sup>19</sup> Accuracy
- <sup>20</sup> Confidence

Expressions," *Journal of Cognitive Neuroscience*, Vol. 14, No. 8, pp. 1158-1173, 2002.

[11] M. P. a. L. J. M. Rothkrantz, "An Expert System for Automatic Analysis of Facial Expression," *Image and Vision Computing Journal*, Vol. 18, No. 11, pp. 881-905, 2000.

[12] I. A. Essa and A. P. Pentland, "Coding, Analysis, Interpretation, and Recognition of Facial Expressions," *IEEE TPAMI*, Vol. 19, No. 7, pp. 757-763, 1997.

[13] J. F. Cohn, A. J. Zlochower, J. J. Lien, and T. Kanade, "Feature-Point Tracking by Optical Flow Discriminates Subtle Differences in Facial Expression," in *3rd IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, pp. 396-401, 1998.

[14] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial Expression Recognition from Video Sequences: Temporal and Static Modeling," *Computer Vision and Image Understanding*, Vol. 91, No.1-2, pp. 160-187, 2003.

[15] J. J. J. Lien, T. Kanade, J. F. Cohn, and C. C. Li, "Detection, Tracking, and Classification of Action Units in Facial Expression," *Journal of Robotics and Autonomous Systems*, Vol. 31, No.3, pp. 131-146, 2000.

[16] S. Lucey, I. Matthews, C. Hu, Z. Ambadar, F. De la Torre Frade, and J. Cohn, "AAM Derived Face Representations for Robust Facial Action Recognition," in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pp. 155-162, 2006.

[17] T. F. Cootes, G. J. Edwards and C. J. Taylor, "Active Appearance Models," *Lecture Notes in Computer Science*, Vol. 1407, pp. 484, 1998.

[18] M.J.Black and Y. Yaccob, "Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion," *Intl. Journal of Computer Vision*, Vol. 25, No. 1, pp. 23-48, 1997.

[19] J. Durkin, *Expert Systems, Design and Development*: Macmillan Publishing Company, 1994.

[20] S. C. Park, S. Piramuthu, and M. J. Shaw, "Dynamic Rule Refinement in Knowledge-Based Data Mining Systems," *Decision Support Systems*, Vol. 31, No. 2, pp. 205-222, 2001.

[21] J. Canny, "A Computational Approach to Edge Detection," *IEEE TPAMI*, Vol. 8, No. 6, pp. 679-698, 1986.

[22] T. Kanade, J. F. Cohn, and Y. L. Tian, "Comprehensive Database for Facial Expression Analysis," in *4th IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, pp. 46-53, 2000.

[23] G. Riley, "CLIPS, A Tool for Building Expert Systems," available at <http://www.ghg.net/clips/CLIPS.html>, 2003.



**محرم منصوری زاده** کارشناسی مهندسی کامپیوتر خود را از دانشگاه اصفهان و کارشناسی ارشد خود را در همین رشته از دانشگاه تربیت مدرس به ترتیب در سالهای ۱۳۸۰ و ۱۳۸۲ دریافت کرد.

او اکنون دانشجوی دوره دکتری مهندسی کامپیوتر در دانشگاه تربیت مدرس است. زمینه های پژوهشی مورد علاقه

او بازشناسی الگو، سیستم‌های مبتنی بر قانون و ارتباط انسان و ماشین است.

آدرس پست الکترونیک: [mansoor@modares.ac.ir](mailto:mansoor@modares.ac.ir)