

محمد بن محمد  
الشيخ  
الشيخ  
الشيخ  
الشيخ



## سوگند نامه دانش آموختگان دانشگاه حکیم سبزواری

کزین برتر اندیشه بر نگذرد

به نام خداوند جان و خرد

اینک که به خواست آفریدگار پاک، کوشش خویش و بهره‌گیری از دانش استادان و سرمایه‌های مادی و معنوی این مرز و بوم، توشه‌ای از دانش و خرد گردآورده‌ام، در پیشگاه خداوند بزرگ سوگند یاد می‌کنم که در به کارگیری دانش خویش، همواره بر راه راست و درست گام بردارم. خداوند بزرگ، شما شاهدان، دانشجویان و دیگر حاضران را به عنوان داورانی امین گواه می‌گیرم که از همه دانش و توان خود برای گسترش مرزهای دانش بهره‌گیرم و از هیچ کوششی برای تبدیل جهان به جایی بهتر برای زیستن، دریغ نورزم. پیمان می‌بندم که همواره کرامت انسانی را در نظر داشته باشم و هموعان خود را در هر زمان و مکان تا سر حد امکان یاری دهم. سوگند می‌خورم که در به کارگیری دانش خویش به کاری که با راه و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مبادینت دارد دست نیازم. همچنین در سایه اصول جهان‌شمول انسانی و اسلامی، پیمان می‌بندم از هیچ کوششی برای آبادانی و سرافرازی میهن و هم‌میهنانم فروگذاری نکنم و خداوند بزرگ را به یاری طلبم تا همواره در پیشگاه او و در برابر وجدان بیدار خویش و ملت سرافراز، بر این پیمان تا ابد استوار بمانم.

نام و نام خانوادگی و امضای دانشجو

ریحانه رشیدی



دانشگاه حکیم سبزواری

دانشکده برق

پایان نامه جهت دریافت درجه کارشناسی ارشد رشته برق گرایش الکترونیک

شناسایی عمل انسان در تصاویر ویدئویی به کمک ویژگی‌های محلی  
و بسته‌ای از کلمات تصویری

استاد راهنما:

دکتر جواد حدادنیا

نگارش:

ریحانه رشیدی

بهمن ماه ۱۳۹۰

تقدیم به

پدر و مادرم

که از نگاهشان صلابت، از رفتارشان محبت، و از صبرشان ایستادگی را آموختم

و به رضا

اسطوره زندگی ام، پناه خستگی ام و امید بودم

و نیز پدر و مادر، همسرم؛ همیاری و همدلی آنان را از یاد نمی برم

و برای امیر هوشنگ و ملیحه و عبدالرضا که وجودشان برایم شادی بخش است و دگرگرم کننده

تشکر ویژه ای دارم از استاد عزیز و گرامی که افتخارم جناب آقای دکتر حواد حداد دنیا که همواره مرا با راهنمایی های ارزشمندشان یاری نموده اند و با صبر و متانت مثال زدنی در تمام محظات این دوره ارزشمند با من همراهی نموده اند.

از جناب آقای دکتر مجید تقائی نژاد و جناب آقای دکتر علیرضا شانی که برای داوری این پایان نامه قبول زحمت نموده و وقت گرانهای خود را در اختیار اینجانب قرار دادند، و همچنین از همه عزیزانی که در این راه مراباری نموده اند تشکر و قدردانی می نمایم.

ما موجودات فانی، از مواد ابدی و پایدار آفرینش، علی رغم عمر کوتاه خودمان، ارزش هایی می آفرینیم که می توانند عمر جاوید داشته باشند.

آلبرت اینشتین

## چکیده

هدف تحقیق حاضر شناسایی اعمال انسانی در تصاویر ویدئویی است که بدین منظور از بیان محلی تصاویر ویدئویی بهره گرفته‌ایم. یکی از راهکارهای ارائه شده در این پایان‌نامه تلفیق توصیفات ظاهری و بیان‌های حرکتی است که نهایتاً به یک توصیف‌کننده مکان-زمان محلی منجر می‌گردد. تا کنون ویژگی‌های محلی گوناگونی به منظور بیان اطلاعات ظاهری ارائه شده‌اند. ما با به کارگیری توصیف‌کننده جدیدی که حاصل توصیف محلی PHOG بر روی نقاط حاصل از آشکارساز هریس می‌باشد، توانستیم از مزایای توصیف‌کننده‌های دوبعدی به خوبی بهره ببریم. علاوه بر این استفاده از آشکارساز و توصیف‌کننده دوبعدی دیگری با نام سیفت نیز در قالب توصیف‌کننده‌های مکان-زمان محلی جهت مقایسه و ارزیابی بهتر روش پیشنهادی به طور جداگانه بیان گردیده است. برای توصیف زمانی اعمال انسانی نیز از بیان هیستوگرامی بردارهای شار نوری استفاده نمودیم و با تلفیق این دو بیان، به یک توصیف‌کننده واحد دست یافتیم که هم اطلاعات مکانی و هم زمانی مربوطه را شامل می‌شود. در ادامه با انتقال این توصیف‌کننده‌ها به فضای بسته‌ای از کلمات تصویری، به هیستوگرام‌هایی متناظر با هر ویدئو دست می‌یابیم که در واقع ورودی مرحله آموزش و دسته‌بندی خواهند بود. نتایج ارزیابی سیستم شناسایی مذکور بر روی دو پایگاه داده معروف از اعمال انسانی ارائه گردیده است. هر یک از این پایگاه‌های داده تعدادی عمل مختلف را شامل می‌شوند که توسط چندین عامل انسانی و در شرایط متفاوت انجام گرفته‌اند. نتایج به دست آمده نشان‌دهنده برتری روش به کاربرنده ویژگی محلی Harris-PHOG در مقایسه با روش سیفت است. همچنین در مقایسه‌ای که با نتایج دیگر مراجع مرتبط صورت گرفت، راهکار پیشنهادی یعنی به کارگیری توصیف‌کننده‌های مکان-زمان محلی توانست به نتایج برتر و یا قابل قیاس با برترین نتایج موجود دست یابد.

**کلمات کلیدی:** شناسایی عمل انسان، ویژگی‌های محلی، آشکارساز هریس، PHOG، شار نوری،

SIFT، بسته‌ای از کلمات تصویری.

## فهرست مطالب

فصل اول: مقدمه	۱
۱-۱ تعریف شناسایی عمل انسان	۳
۲-۱ راهکارهای ارائه شده در این پایان نامه	۵
۳-۱ ساختار پایان نامه	۶
فصل دوم: کارهای صورت گرفته در زمینه شناسایی عمل انسان	۷
۱-۲ تقسیم بندی روش های موجود	۹
۲-۲ بیان فضایی عمل	۹
۱-۲-۲ مدل های بدن	۹
۲-۲-۲ مدل های کل نگر	۱۴
۳-۲-۲ ویژگی های محلی	۱۷
۳-۲ بیان زمانی عمل	۲۱
۱-۳-۲ مدل های دینامیک عمل	۲۱
۲-۳-۲ الگوهای عمل	۲۳
۳-۳-۲ آماره های زمانی	۲۵
فصل سوم: ساختار کلی سیستم شناسایی عمل	۲۹
۱-۳ روال کلی حل مسئله	۳۱
۲-۳ بیان اطلاعات ظاهری	۳۴
۱-۲-۳ ویژگی های محلی	۳۴
۲-۲-۳ نگاهی به تاریخچه روش های مبتنی بر ویژگی های محلی	۳۶
۳-۲-۳ خصوصیات ویژگی محلی ایده آل	۳۷
۳-۳ بیان اطلاعات حرکتی	۴۰
۱-۳-۳ تئوری شار نوری	۴۱



## فهرست مطالب

فصل چهارم: توصیف کننده های مکان-زمان محلی در قالب بسته ای از کلمات تصویری	۴۷
۱-۴ راهکارهای پیشنهادی	۴۹
۲-۴ توصیف کننده های محلی در بر گیرنده ریخت و حرکات	۵۲
۳-۴ معرفی ویژگی های محلی به کار رفته به منظور بیان ظاهری اعمال	۵۳
۱-۳-۴ آشکارساز هریس	۵۳
۲-۳-۴ توصیف کننده PHOG	۵۹
۳-۳-۴ آشکار ساز و توصیف کننده سیفت	۶۲
۴-۴ استفاده از شار نوری برای بیان اطلاعات حرکتی	۸۴
۱-۴-۴ گزینش نقاط کلیدی پر معنا توسط بردارهای شار نوری	۸۴
۲-۴-۴ هیستوگرام شار نوری به عنوان توصیف کننده زمانی	۸۶
۵-۴ بسته ای از کلمات تصویری	۹۰
۱-۵-۴ روال کلی الگوریتم	۹۱
۲-۵-۴ آموزش کتاب کد	۹۲
۳-۵-۴ ایجاد بردار متناظر با ویدئو از روی کتاب کد بصری	۹۳
۶-۴ آموزش و دسته بندی	۹۵
۱-۶-۴ دسته بند SVM	۹۶
فصل پنجم: ارزیابی نتایج	۱۰۱
۱-۵ ملاک ارزیابی در شناسایی عمل انسانی	۱۰۳
۱-۱-۵ پایگاه داده KTH	۱۰۳
۲-۱-۵ پایگاه داده ویزمن	۱۰۵
۲-۵ نحوه ارزیابی	۱۰۶
۳-۵ ارزیابی روش مبتنی بر Harris-PHOG	۱۰۹

## فهرست مطالب

۱۱۰	۱-۳-۵ نتایج اعمال روش پیشنهادی Harris-PHOG بر روی پایگاه داده KTH
۱۱۲	۲-۳-۵ نتایج اعمال روش پیشنهادی Harris-PHOG بر روی پایگاه داده ویزمن
۱۱۳	۴-۵ ارزیابی روش مبتنی بر آشکارساز و توصیف کننده سیفت
۱۱۳	۱-۴-۵ نتایج اعمال روش پیشنهادی با سیفت بر روی پایگاه داده KTH
۱۱۴	۲-۴-۵ نتایج اعمال روش پیشنهادی با سیفت بر روی پایگاه داده ویزمن
۱۱۵	۵-۵ مقایسه روش های مبتنی بر سیفت و Harris-PHOG
۱۱۶	۶-۵ مقایسه نتایج روش های پیشنهادی با سایر کارهای صورت گرفته
۱۱۸	۷-۵ پارامترهای تأثیرگذار و میزان اثر آنها
۱۱۸	۱-۷-۵ بررسی تأثیر میزان مشارکت نقاط کلیدی مجموعه آموزش بر نتیجه دسته بندی
۱۲۰	۲-۷-۵ بررسی تأثیر نوع کرنل به کار رفته در دسته بند SVM بر نتیجه دسته بندی
۱۲۰	۳-۷-۵ تعیین سایر پارامترهای مؤثر بر نتیجه دسته بندی
۱۲۳	فصل ششم: نتیجه گیری و پیشنهادها
۱۲۷	مراجع

## فهرست جداول

جدول (۱-۵) مقایسه نتایج حاصل از دو روش پیشنهادی در دو پایگاه داده تحت بررسی ----- ۱۱۶

جدول (۲-۵) مقایسه نتایج روش های پیشنهادی با سایر کارهای صورت گرفته ----- ۱۱۷

جدول (۳-۵) تأثیر نوع کرنل به کار رفته در دسته بند بر روی درصد صحت دسته بندی ----- ۱۲۰

## فهرست اشکال

- شکل (۱-۲) نمایش نشانگرهای نوری متحرک ----- ۱۰
- شکل (۲-۲) بیان های مختلف از مدل های بدن ----- ۱۲
- شکل (۳-۲) بیان های کل نگر ----- ۱۳
- شکل (۴-۲) بیان های وضعیت محلی ----- ۱۸
- شکل (۵-۲) روش های به کار برنده الگوهای عمل ----- ۲۴
- شکل (۱-۳) ویژگی های محلی آشکار شده در یک تصویر نمونه ----- ۳۴
- شکل (۲-۳) مثالی از حجم کلیدی استخراج یافته از ویدئو ----- ۴۰
- شکل (۳-۳) نمایشی از مسئله روزنه ----- ۴۳
- شکل (۱-۴) فلوچارت کلی روش پیشنهادی ----- ۵۱
- شکل (۲-۴) نحوه انتخاب ویژگی های محلی ----- ۵۴
- شکل (۳-۴) استفاده از ماتریس هسین برای استخراج ویژگی های محلی ----- ۵۸
- شکل (۴-۴) روال توصیف کنندگی در PHOG ----- ۶۰
- شکل (۵-۴) نحوه ساخت هرم فضا-مقیاس در آشکارساز سیفت ----- ۶۷
- شکل (۶-۴) نحوه تعیین اکسترمم های محلی در هرم فضا-مقیاس ----- ۶۸
- شکل (۷-۴) نمودارهای تعیین فرکانس نمونه برداری در مقیاس ----- ۷۰
- شکل (۸-۴) نمودار تعیین فرکانس نمونه برداری در مکان ----- ۷۲
- شکل (۹-۴) نقاط کلیدی استخراج شده با در نظر گرفتن آستانه گیری ----- ۷۵
- شکل (۱۰-۴) نمودار پایداری نقطه کلیدی با افزایش نویز ----- ۷۹
- شکل (۱۱-۴) رو به طرف بالا نمودن ناحیه اطراف نقطه کلیدی همراه با جهتش ----- ۸۱
- شکل (۱۲-۴) تعیین هیستوگرام های جهت برای توصیف کننده سیفت ----- ۸۱
- شکل (۱۳-۴) آستانه گیری بر روی دامنه شار نوری و گزینش نقاط کلیدی پرمعنا ----- ۸۵
- شکل (۱۴-۴) تشکیل هیستوگرام شار نوری توسط بردارهای شار نوری ----- ۸۸
- شکل (۱۵-۴) قالب کلی روش بسته ای از کلمات تصویری ----- ۹۱

## فهرست اشکال

- شکل (۴-۱۶) نمایش ساختار شبکه عصبی MLP ----- ۹۶
- شکل (۴-۱۷) ضرورت انتخاب دسته بند مناسب ----- ۹۷
- شکل (۴-۱۸) نمایش بردارهای پشتیبان در SVM خطی نمونه ----- ۹۹
- شکل (۵-۱) فریم های نمونه از پایگاه داده KTH ----- ۱۰۵
- شکل (۵-۲) فریم های نمونه از پایگاه داده ویزمن ----- ۱۰۶
- شکل (۵-۳) ماتریس درهم آمیختگی مربوط به اعمال روش Harris-PHOG بر روی پایگاه داده KTH ----- ۱۱۰
- شکل (۵-۴) ماتریس درهم آمیختگی مربوط به اعمال روش Harris-PHOG بر روی پایگاه داده ویزمن ----- ۱۱۲
- شکل (۵-۵) ماتریس درهم آمیختگی مربوط به اعمال روش سافت بر روی پایگاه داده KTH ----- ۱۱۴
- شکل (۵-۶) ماتریس درهم آمیختگی مربوط به اعمال روش سافت بر روی پایگاه داده ویزمن ----- ۱۱۵
- شکل (۵-۷) نمودار تأثیر میزان مشارکت نقاط کلیدی در ساخت کتاب کد بر نتیجه دسته بندی ----- ۱۱۹

# فصل اول

مقدمه

## ۱-۱ تعریف شناسایی عمل انسان

شناسایی عمل انسان یکی از مباحث به روز در حوزه بینایی ماشین است که کاربردهای گسترده‌ای شامل واسطه‌های انسان-رایانه، شاخص گذاری ویدئو بر اساس محتوا، نظارت ویدئویی، روباتیک و غیره را در بر می‌گیرد. از نظر تاریخی شناسایی بصری عمل انسان به زیرمجموعه‌هایی همچون شناسایی رفتار<sup>۱</sup> برای واسطه‌های انسان-رایانه [۲،۱]، شناسایی حالات صورت<sup>۲</sup> [۳]، و شناسایی وضعیت حرکت<sup>۳</sup> برای نظارت‌های ویدئویی [۴] تقسیم‌بندی می‌شود.

شناسایی عمل انسان عبارت است از فرایند نام‌دهی به اعمال انسان، معمولاً در قالب فعل ساده مربوط به آن عمل، با استفاده از مشاهدات دریافتی. از نظر فنی، به دنباله‌ای از حرکات ایجاد شده توسط یک عامل انسانی در جریان انجام کاری عمل اطلاق می‌گردد. بنابراین یک عمل در واقع مفهومی چهار بعدی است که می‌توان آن را به بخش‌های فضایی<sup>۴</sup> و زمانی<sup>۵</sup> تجزیه نمود. ما برای شناسایی عمل انسان تنها به مشاهدات بصری می‌پردازیم، ولی باید توجه داشت که اعمال از طریق دیگر دریافته‌ها همچون صدا نیز قابل شناسایی می‌باشند. برچسب یک عمل نامی است که یک عامل انسانی عادی قادر به فهمیدن و انجام آن باشد. وظیفه سیستم شناسایی عمل، نام‌دهی به اعمال است یعنی تعیین برچسبی برای عمل که بتواند یک عمل نمونه را حتی زمانی که توسط افراد مختلف و از نقطه‌نظرهای متفاوت انجام گرفته و یا تغییرات زیاد در حالت و سرعت موجود است، به بهترین نحو توصیف کند.

اگرچه برای یک انسان شناسایی هزاران دسته گوناگون از اعمال انسانی امری طبیعی است، ولی این شناسایی برای سیستم‌های بینایی ماشین همچنان با چالش همراه است. این مشکل اساساً به دو جنبه باز می‌گردد: (۱) تغییرات وسیع در اعمال انسانی به دلیل تغییرات ذاتی، حالات مختلف بدن،

---

<sup>1</sup> Gesture recognition

<sup>2</sup> Facial expression recognition

<sup>3</sup> Movement behavior recognition

<sup>4</sup> spatial

<sup>5</sup> temporal

تبدیلات بصری، استتار، نویز و غیره؛ (۲) شباهت‌های بین کلاسی ناشی از الگوهای بصری مشترک میان کلاس‌های متفاوت.

به طور کلی یک سیستم شناسایی عمل انسان دو وظیفه اصلی زیر را شامل می‌شود.

- استخراج ویژگی<sup>۱</sup> که وظیفه اصلی در شناسایی عمل است و استخراج نشانه‌های وضعیتی و حرکتی از ویدئو با تمایزدهی خوب نسبت به اعمال انسانی را در بر می‌گیرد. بیان‌های گوناگونی از استخراج ویژگی قابل استفاده است که در هر یک بایستی مواردی چون موقعیت فرد، پایداری در برابر استتار جزئی، شلوغی پس زمینه، وجود سایه‌ها و روشنایی‌های متفاوت پاسخ داده شوند. گذشته از این، بیان مورد نظر باید نسبت به انواع متفاوت هیكل افراد و طرز پوشش آنها حساسیت زیادی نشان ندهد.

- آموزش عمل و دسته‌بندی<sup>۲</sup> که عبارتند از مراحل آموزش مدل آماری عمل از روی ویژگی‌های استخراج یافته، و به کار بردن آن برای دسته‌بندی مشاهدات جدید. یکی از چالش‌های بزرگ در این مسیر، مقابله با تغییرات زیادی است که یک کلاس عمل خاص می‌تواند از خود بروز دهد، به ویژه اگر عمل توسط افراد مختلف با جنسیت و اندازه‌های متفاوت و با سرعت و شیوه متفاوت انجام پذیرفته باشد. گروه‌های مختلف اعمال که از نظر ما تعریف روشن و مشخصی دارند، همانند لگد زدن، مشت زدن، یا کف زدن، ممکن است هنگام انجام، بسیار متغیر باشند. بنابراین طراحی مدلی که ویژگی‌های مشخصه هر عمل را در بر داشته و در عین حال با تغییرات ممکنه سازگاری مناسبی داشته باشد، مبحث مهمی است.

---

<sup>1</sup> Feature extraction

<sup>2</sup> Action learning and classification



## ۱-۲ راهکارهای ارائه شده در این پایان نامه

با توجه به نتایج برتر ارائه شده توسط روش‌های به کار برنده ویژگی‌های محلی، بر آن شدیم تا از این ویژگی‌ها به منظور توصیف اعمال انسانی بهره بگیریم. روش پیشنهادی ما شامل تعریف یک سری توصیف‌کننده مکان-زمان محلی<sup>۱</sup> می‌گردد که با تلفیق ویژگی‌های محلی توصیف‌کننده ظاهر و حرکت، به بیان خوبی از جنبه‌های مکانی و زمانی اعمال انسان می‌انجامند. اساس محاسبه این توصیف‌کننده‌ها بر نقاط کلیدی حاصل از مرحله آشکارسازی واقع است. در این مرحله با افزودن یک آستانه‌گیری حرکتی، موفق شدیم تا نقاط کم ارزش را که فاقد حرکت کافی در فریم‌های متوالی ویدئو هستند از مجموعه نقاط کلیدی کنار بگذاریم و بدین ترتیب توصیفی مفیدتر و با تمایزدهندگی بالاتر از هر ویدئو ارائه دهیم. در انتخاب ویژگی محلی نیز پیشنهاد ما استفاده از آشکارساز هریس به همراه توصیف‌کننده PHOG<sup>۲</sup> است. به کارگیری این روش در کنار توصیف‌کننده حرکتی حاصل از شار نوری تصاویر، نتایج بسیار خوبی را در پایگاه‌های داده محک اعمال انسانی ارائه نمود. همچنین به منظور انجام مقایسه‌ای ملموس‌تر، ویژگی محلی سift (SIFT)<sup>۳</sup> را که یکی از بهترین توصیف‌کننده‌های دوبعدی محسوب می‌گردد در قالب روش پیشنهادی توصیف‌کننده‌های مکان-زمان به کار گرفتیم و نتایج حاصله را مورد بررسی قرار دادیم.

به منظور بیان هر یک از ویدئوهای نمونه توسط این توصیف‌کننده‌های محلی، به سراغ الگوریتم بسته‌ای از کلمات تصویری (BOVW)<sup>۴</sup> رفتیم. در این ساختار، هر نمونه به صورت هیستوگرامی از تعداد رخداد‌های کلمات تصویری مشخص بیان می‌شود، به گونه‌ای که بردن نمونه‌ها به چنین فضایی مسئله را به یک مسئله آموزش و دسته‌بندی بدل می‌نماید.

---

<sup>1</sup> Local spatio-temporal descriptors

<sup>2</sup> Pyramid Histogram Of Oriented Gradients

<sup>3</sup> Scale Invariant Feature Transform

<sup>4</sup> Bag Of Visual Words

## ۳-۱ ساختار پایان نامه

ساختار این پایان نامه بدین ترتیب است، در فصل دوم پس از تقسیم بندی راهکارهای موجود حل مسئله، به بیان روش های انجام گرفته در زمینه شناسایی اعمال انسانی خواهیم پرداخت. فصل سوم روال کلی حل مسئله در این پایان نامه را بازگو نموده و در ادامه به شرح الگوریتم های به کار رفته به منظور توصیف ظاهری و حرکتی ویدئوها می پردازد. در فصل چهار راهکارهای پیشنهادی این پایان نامه شامل بیان اعمال با توصیف کننده های مکان-زمان محلی و چگونگی ساخت این توصیف کننده ها مورد بررسی قرار گرفته و انتقال به فضای بسته ای از کلمات تصویری به تفصیل بیان می گردد و نیز مرحله نهایی که عبارت است از آموزش دسته بند و دسته بندی اعمال انسانی شرح داده خواهد شد. در فصل پنجم به ارزیابی نتایج به دست آمده توسط روش های پیشنهادی و مقایسه آنها با سایر کارهای صورت گرفته می پردازیم و با ارائه نتیجه گیری و بیان پیشنهادها در فصل ششم بحث را به پایان می بریم.

## فصل دوم

کارهای صورت گرفته در زمینه شناسایی عمل انسان

## ۱-۲ تقسیم بندی روش های موجود

روش های گوناگون شناسایی عمل را می توان از جنبه های مختلفی دسته بندی نمود. بسته به اینکه ساختارهای زمانی و یا فضایی موجود در ویدئوها بررسی شوند، روش های موجود را به دو دسته متفاوت می توان تقسیم نمود. در این فصل ابتدا روش های مختلفی که به بیان ساختار فضایی اعمال می پردازند را از نظر می گذرانیم و در ادامه روش هایی را بررسی می کنیم که از مدل نمودن ساختار زمانی اعمال بهره می گیرند.

## ۲-۲ بیان فضایی عمل

برای ایجاد تمایز بین اعمال از دیدگاه بصری، به بیانی فضایی از هر عمل نیاز داریم. همان طور که ذکر شد یکی از مراحل اولیه در شناسایی عمل، استخراج ویژگی هایی از تصویر است که نسبت به وضعیت و حرکت بدن فرد تمایزدهنده باشند. بدین منظور بیان های متفاوتی پیشنهاد شده اند که بر اساس میزان اطلاعات سطح بالایی که ارائه می دهند و میزان کارایی آنها در استخراج ویژگی سنجیده می شوند. می توان این روش های مختلف بیان را به سه گروه اصلی تقسیم بندی نمود: مدل های بدن<sup>۱</sup>، مدل های تصویر<sup>۲</sup>، و ویژگی های محلی<sup>۳</sup>.

## ۱-۲-۲ مدل های بدن

در روش اول در واقع ساختار فضایی اعمال بر اساس بدن فرد بیان می گردد. در هر یک از فریم های ویدئوی تحت بررسی، حالت بدن فرد از طریق ویژگی های تصویر گوناگونی که موجودند بازیابی شده و سپس شناسایی عمل بر مبنای این تخمین های حالت صورت می گیرد. چنین روشی یک راهکار ذاتی

---

<sup>1</sup> Body models

<sup>2</sup> Image models

<sup>3</sup> Local features