

شناسایی انسان از روی نحوه راه رفتن با کمک توصیف‌گر بافت پویا

بهناز عبدالمهی^۱ و نیلوفر قیصری^۲

۱- کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر- دانشگاه صنعتی اصفهان- اصفهان- ایران

b.abdolahi@ec.iut.ac.ir

۲- استادیار، دانشکده مهندسی برق و کامپیوتر- دانشگاه صنعتی اصفهان- اصفهان- ایران

n.gheissari@cc.iut.ac.ir

چکیده: امروزه بررسی حرکات انسان بخش مهمی از تحقیقات در زمینه بینایی ماشین را به خود اختصاص داده است. مطالعات نشان داده‌اند که انسان‌ها توانایی زیادی در شناسایی آشنایان خود از روی برخی رفتارها، از جمله شیوه راه رفتن آن‌ها دارند. از این‌رو، محققان بر آن شدند تا از این قابلیت انسان‌ها در ساخت سیستم‌های شناسایی استفاده نمایند. این روش با امکان شناسایی در فواصل دور، عدم نیاز به اطلاع و یا همکاری فرد و همچنین، عدم امکان پنهان نمودن، کارایی بالایی به ویژه در محیط‌های عمومی ارائه می‌دهد. در این مقاله روش شناسایی جدیدی مبتنی بر توصیف‌گرهای بافت پویا ارائه می‌گردد. بافت‌های پویا شامل ترکیبی از اطلاعات ظاهری و حرکتی هستند که با بررسی دو بعد مکان و زمان به دست می‌آیند. از آنجایی که راه رفتن انسان شامل اطلاعات مهمی در دو بعد زمان و مکان است، می‌تواند با کمک ویژگی‌های بافت پویا توصیف گردد. برای به دست آوردن این ویژگی‌ها، نقاط کلیدی مکان-زمان استخراج و با کمک الگوی دودویی محلی با سه صفحه متعامد که نمونه‌ای از توصیف‌گرهای بافت پویاست، توصیف می‌گردند. سپس با استفاده از فرهنگ لغت به دست آمده از دسته بندی ویژگی‌ها، مدل آموزشی از شیوه راه رفتن هر فرد مشخص می‌شود. کارایی روش پیشنهادی بر روی دو پایگاه داده KTH و IXMAS بررسی شده است.

واژه های کلیدی: آنالیز حرکات انسان، الگوی دودویی محلی با سه صفحه متعامد، بافت پویا، شناسایی از روی نحوه راه رفتن، فرهنگ لغت، نقاط کلیدی مکان-زمان.

۱- مقدمه

هوشمند و سیستم‌های تشخیص هویت اشاره نمود.

شیوه راه رفتن هر شخص در مقایسه با افراد دیگر یکتاست. گام برداشتن انسان، فعالیتی پیچیده است که شامل حرکات همزمان اعضای بدن و تعامل بین آنهاست. از این‌رو، می‌توان از تفاوت بین حرکات و ساختار اعضای بدن در افراد مختلف، به عنوان ویژگی‌های خاص برای شناسایی استفاده نمود. از مزایای روش شناسایی از روی نحوه راه رفتن انسان می‌توان به امکان شناسایی فرد در فواصل دور، عدم نیاز به اطلاع و یا همکاری فرد و همچنین عدم امکان پنهان نمودن آن از دید ناظر و سیستم مراقبتی اشاره نمود. این روش می‌تواند در محیط‌های عمومی و با یک دنباله تصاویر با وضوح پایین از حرکات فرد، کارایی بالایی نسبت

امروزه بررسی رفتار و حرکات انسان از جمله راه رفتن، توجه بسیاری از محققان را به خود جلب نموده و به علت دامنه کاربرد وسیع از اهمیت بالایی برخوردار است. از جمله کاربردهای آن می‌توان در شاخص‌گذاری ویدئو، سیستم‌های نظارتی، تحلیل فیلم‌های ورزشی، محیط‌های

تاریخ ارسال مقاله : ۱۳۹۱/۱۲/۱۲

تاریخ پذیرش مقاله : ۱۳۹۲/۰۵/۲۰

نام نویسنده مسؤول : بهناز عبدالمهی

نشانی نویسنده مسؤول : ایران - کرمان - انتهای اتوبان هفت باغ-

دانشگاه تحصیلات تکمیلی و فناوری پیشرفته

همچنین، به علت محاسبه ویژگی‌های بافت پویا بر روی تصاویر ویدئویی، اطلاعات مربوط به شکل ظاهری (موقعیت، جهت و زاویه دید) هم وارد مسأله می‌گردد. از آنجایی که حرکات انسان از جمله راه رفتن نیز شامل اطلاعاتی در دو بعد مکان و زمان هستند، از این رو می‌توان از توصیف‌گرهای بافت پویا برای تحلیل استفاده نمود. الگوی دودویی محلی با سه صفحه متعامد^۵ (LBP-TOP) نمونه‌ای از توصیف‌گرهای بافت پویاست [۱۵] که تاکنون در حوزه‌های مختلف مانند توصیف صورت [۱۶، ۱۷]، گفتار [۱۸، ۱۹]، حالت دست [۲۰]، تشخیص حرکات [۷، ۲۱] و شناسایی انسان از روی نحوه راه رفتن [۲۲] به‌طور موفقیت‌آمیزی به‌کار گرفته شده است.

استخراج ویژگی‌های مناسب و متمایز یک پیش‌نیاز اساسی برای مرحله توصیف است که به دو روش سراسری و محلی انجام می‌گیرد. در روش سراسری از تمام اطلاعات دنباله ویدئویی استفاده می‌شود. در این روش طرحواره^۶ با کمک الگوریتم‌های جداسازی پس‌زمینه استخراج می‌شود و سپس کل طرحواره به عنوان ناحیه کلیدی در نظر گرفته شده و به ورودی توصیف‌گر داده می‌شود. در این روش همه نقاط از اهمیت یکسانی برخوردارند، به همین علت، این روش توانایی تعمیم‌دهی بالایی دارد، اما این روش نسبت به نویز، زاویه دید و همپوشانی حساس است. برای نمونه، بلنک و همکارانش [۸] با در نظر گرفتن نقاط متناظر در هر فریم و جداسازی پس‌زمینه، طرحواره انسان را در هر فریم استخراج نموده، با کنار هم قرار دادن این طرحواره‌ها یک حجم سه بعدی که اطلاعات مهم مکانی و زمانی را در بردارد، تشکیل می‌دهند. سپس حرکت مورد نظر با کمک یک توصیف‌گر مناسب تحلیل می‌شود. وجود نویز در کارایی این روش بسیار مؤثر است، زیرا وجود نویز در کناره حجم مکان-زمان در به‌دست آوردن نقاط برجسته اختلال ایجاد می‌کند.

در روش محلی اطلاعات نواحی خاصی از دنباله ویدئویی که شامل تغییرات ناگهانی در دو بعد مکان و زمان هستند، در نظر گرفته می‌شود. هر ناحیه به طور جداگانه با کمک یک توصیف‌گر مناسب توصیف گردیده، سپس با ترکیب اطلاعات نواحی مختلف، یک توصیف کلی به وجود

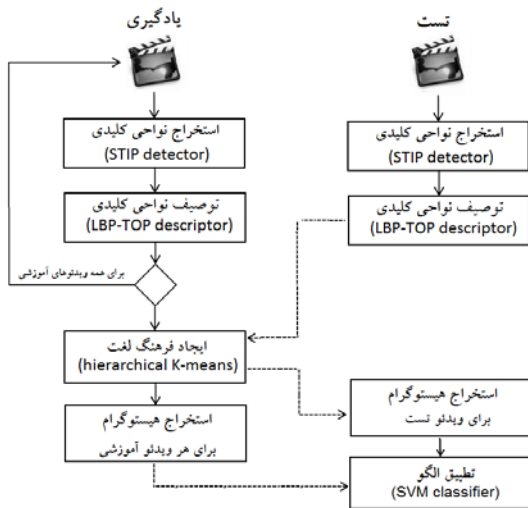
به سایر روش‌های شناسایی ارائه دهد. برای این دسته از سیستم‌ها کاربردهای فراوانی وجود دارد که از آن جمله می‌توان به نظارت، کنترل و تحلیل فعالیت‌های انسان اشاره نمود.

هدف این مقاله بررسی تحقیقات پژوهشی در زمینه تحلیل حرکات و ارائه روشی برای شناسایی از روی نحوه راه رفتن افراد است. تاکنون رهیافت‌های متعددی در این زمینه ارائه شده است که در اکثر آنها از اطلاعات حرکتی [۱-۳] یا شکل ظاهری [۴-۶] بهره گرفته شده است اما بررسی‌ها نشان داده‌اند که استناد به یک بعد به تنهایی نمی‌تواند تمامی اطلاعات مورد نیاز را فراهم کند. روش‌های مبتنی بر شکل به طور طبیعی نسبت به تغییرات ظاهری مانند حمل کوله‌پشتی حساس هستند. این روش‌ها نسبت به نویز، زمینه نویزی و یا حتی زمینه متحرک حساسیت کمتری نشان می‌دهند، اما زمان بر بوده، هزینه محاسباتی زیادی برای تطبیق‌های پیچیده به سیستم تحمیل می‌کنند. روش‌های مبتنی بر حرکت هیچ مدل خاصی برای بدن انسان در نظر نمی‌گیرند و از تغییرات وابسته به زمان، مانند تغییر در سرعت راه رفتن به عنوان اطلاعات استفاده می‌کنند.

تحلیل اطلاعات مکان-زمان شامل اطلاعات مکانی به معنای حالت ظاهری فرد در هر زمان و نیز اطلاعات زمانی به معنای موقعیت حرکتی بدن فرد در هر مکان است. امروزه تحلیل و بررسی حرکات انسان در فضای مکان-زمان بسیار مورد توجه قرار گرفته است [۷-۹]. از این رو، در تحقیقات اخیر برای شناسایی از روی نحوه راه رفتن نیز از اطلاعات در دو بعد مکان و زمان بهره گرفته شده است. برای نمونه، بن‌عبدالقادر و همکارانش [۱۰] از نمودار خود متشابه^۱ برای شناسایی استفاده نمودند. لیو و همکارانش [۱۱] الگوی حاشیه که الگوی دیگری برای شناسایی است، معرفی کردند که در [۱۲] به یک الگوی حاشیه برپایه واریانس شکل^۲، توسعه می‌یابد. هنگ و ونگ [۱۳] نیز شناسایی را با کمک تصاویر انرژی صفحه Xt به‌دست آمده از فضای xyt انجام دادند و کویبا و اتسو [۱۴] از خودهمبستگی محلی سطح بالای مکعبی^۳ برای شناسایی حرکات بهره گرفتند.

بافت‌های پویا^۴ در طی زمان تغییرات آماری دارند. از این رو، به طور ذاتی شامل اطلاعات حرکتی هستند،

می‌گردند. در این مرحله به ازای هر یک از نقاط کلیدی یک بردار ویژگی به‌دست می‌آید. سپس با استفاده از یک الگوریتم دسته‌بندی^۱ مناسب، بردارهای ویژگی به‌دست آمده به کلمات یک فرهنگ لغت نگاشت می‌گردند. سرانجام هر کلاس آموزشی با هیستوگرامی از تعداد تکرارهای کلمات این فرهنگ لغت مدل می‌شود. از این‌رو، برای هر کلاس آموزشی یک مدل یکتا به‌دست می‌آید. برای یک دنباله ویدئویی تست، پس از به‌دست آوردن بردار ویژگی و مقایسه آن با فرهنگ لغت موجود، هیستوگرام نمونه تست ایجاد می‌گردد. سپس با کمک یک الگوریتم طبقه‌بندی مناسب، نمونه تست به مدل آموزشی مناسب نسبت داده می‌شود.



شکل (۱): مراحل روش پیشنهادی

۳- استخراج نواحی کلیدی

حرکات انسان در یک دنباله تصویر، دارای تغییرات آماری در دو بعد مکان و زمان هستند. نقاط نشان دهنده این تغییرات، به‌علت دارا بودن اطلاعات کلیدی به عنوان نقاط برجسته برای تحلیل و بررسی حرکات محسوب می‌شوند [۲۶، ۲۷]. در این مرحله نواحی کلیدی با استفاده از آشکار ساز نقاط کلیدی مکان-زمان که نشان دهنده تغییرات ناگهانی در دو بعد مکان و زمان است، مشخص می‌شوند. در این روش، توصیفی از صحنه به‌صورت مجموعه‌ای از نواحی کلیدی مستقل در اطراف هر یک از این نقاط بیان می‌شوند (شکل ۲). سپس از ترکیب این نواحی، یک نمایش

می‌آید. در روش محلی نیازی به جداسازی پس زمینه وجود ندارد. همچنین، این روش نسبت به همپوشانی جزئی، تغییر در زاویه دید و شکل ظاهری نیز مقاوم است. در این روش به‌علت عدم استفاده از تمام نقاط دنباله ویدئویی ممکن است برخی اطلاعات از دست برود، اما با استخراج نقاط کلیدی مناسب و متمایز، اطلاعات اساسی دنباله ویدئویی استخراج می‌شود. از این‌رو، این روش عملکرد مناسبی خواهد داشت. آشکارساز دو بعدی هریس یک آشکار ساز معروف در بعد مکان است [۲۳]، اما از آنجایی که نقاط شناسایی شده در یک بعد مکان و یا زمان تمامی اطلاعات مورد نیاز را در اختیار قرار نمی‌دهند، از این‌رو، لپتو و لیندبرگ آشکارساز دو بعدی هریس را به حالت سه بعدی توسعه دادند [۲۴] و سپس برای مقاوم نمودن آن نسبت به مقیاس، از عملگر نرمال دو بعدی لاپلاسیین استفاده نمودند [۲۵]. در این روش، رویدادهای مکان-زمانی با به‌دست آوردن بیشینه عملگر نرمال دو بعدی لاپلاسیین تحت مقیاس‌های مکان و زمان، شناسایی می‌شوند.

در این مقاله روشی برپایه فرهنگ لغت^۷ برای شناسایی از روی نحوه راه رفتن با استفاده از توصیف‌گرهای بافت پویا ارائه شده است. در این روش، به‌جای توصیف کل تصویر [۲۲]، از مزایای روش محلی برای استخراج ویژگی‌های کلیدی استفاده شده و سپس هر ویژگی، پس از توصیف، به لغتی در فرهنگ لغت نگاشته می‌شود. توصیف‌گر LBP-TOP نیز به توصیف‌گری مقاوم به چرخش توسعه داده شده است.

در ادامه، مراحل روش ارائه شده برای شناسایی به اختصار معرفی و سپس هریک از مراحل به طور کامل شرح داده می‌شوند. پس از آن نتایج تجربی به‌دست آمده بر روی دو پایگاه داده مطرح می‌گردند.

۲- راه حل پیشنهادی

مراحل اصلی روش ارائه شده در شکل (۱) بیان شده‌اند. در این روش، ابتدا نواحی مهم و کلیدی توسط یک آشکارساز مکان-زمان که یک روش محلی است، از دنباله‌های ویدئویی آموزشی استخراج می‌شوند و با استفاده از الگوی دودویی محلی با سه صفحه متعامد، توصیف

شناسایی انسان از روی نحوه راه رفتن با کمک توصیف‌گر بافت پویا

این ماتریس 3×3 و شامل مشتقات مرتبه اول در دو

بعد مکان و زمان است:

$$\mu = g(\cdot; \sigma_i^2, \tau_i^2) * \begin{pmatrix} L_x^2 & L_x L_y & L_x L_t \\ L_x L_y & L_y^2 & L_y L_t \\ L_x L_t & L_y L_t & L_t^2 \end{pmatrix} \quad (3)$$

روابط بین مقیاس‌های یکپارچه σ_i^2, τ_i^2 و مقیاس‌های محلی σ_i^2, τ_i^2 به صورت $\sigma_i^2 = s\sigma_i^2$ و $\tau_i^2 = s\tau_i^2$ بیان می‌شود. مشتقات مرتبه اول بدین صورت تعریف می‌شود:

$$L_x(\cdot; \sigma_i^2, \tau_i^2) = \partial_x(g * f) \quad (4)$$

$$L_y(\cdot; \sigma_i^2, \tau_i^2) = \partial_y(g * f) \quad (5)$$

$$L_t(\cdot; \sigma_i^2, \tau_i^2) = \partial_t(g * f) \quad (6)$$

برای شناسایی نقاط کلیدی، نواحی از f که در آن مقادیر ویژه $\lambda_1, \lambda_2, \lambda_3$ از تابع μ دارای مقادیر بالایی هستند، جستجو می‌گردند.

$$H = \det(\mu) - k \text{trace}^3(\mu) \\ = \lambda_1 \lambda_2 \lambda_3 - k(\lambda_1 + \lambda_2 + \lambda_3)^3 \quad (7)$$

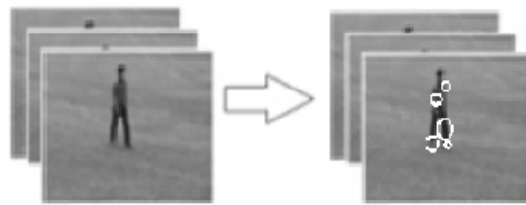
با معرفی نسبت‌های $\alpha = \lambda_2 / \lambda_1$ و $\beta = \lambda_3 / \lambda_1$ می‌توان تابع H را به صورت زیر بیان نمود:

$$H = \lambda_1^3 (\alpha\beta - k(1 + \alpha + \beta)^3) \quad (8)$$

از آنجایی که $H \geq 0$ است، فرض $k \leq \alpha\beta / (1 + \alpha + \beta)^3$ را در نظر می‌گیریم. مقدار بیشینه k برابر با $k = 1/27$ بوده، با در نظر گرفتن $\alpha = \beta = 1$ به دست می‌آید.

با فرض مقادیر بالای k ، نقاط بیشینه H نشان دهنده نقاطی با تغییرات قابل توجهی در هر دو جهت مکان و زمان است. با در نظر گرفتن مقدار بیشینه ۲۳ برای α و β ، مقدار k مورد نیاز برای تابع H ، $k \approx 0.005$ به دست می‌آید. از این‌رو، شناسایی نقاط کلیدی مکان-زمان تصویر f با استفاده از نقاط بیشینه H در دو حوزه مکان و زمان انجام می‌گیرد.

کلید به وجود می‌آید.



شکل (۲): نواحی کلیدی به دست آمده از آشکار ساز نقاط کلیدی مکان-زمان

۳-۱- آشکار ساز نقاط کلیدی مکان-زمان

ویژگی‌های محلی تصویر یک نمایش انتزاعی از الگوهای تصویر و یا تفسیری از داده‌های ویدئویی را بیان می‌کنند. لپتو و لیندبرگ [۲۵] برای کشف رویدادهای مکانی و زمانی، آشکار ساز دو بعدی هریس و فورسترن [۲۳] را که نقاط کلیدی محلی در بعد مکان را شناسایی می‌کند، توسعه داده‌اند. در این روش نواحی‌ای که دارای تغییرات محلی در دو بعد مکان و زمان هستند، شناسایی شده و سپس با بیشینه کردن عملگر نرمال دو بعدی لاپلاسیان تحت مقیاس‌های مکان و زمان، وسعت نواحی کشف شده تخمین زده می‌شوند.

برای مدل کردن دنباله مکانی و زمانی تصویر از تابع $f: R^2 \times R \rightarrow R$ استفاده می‌شود و نمایش خطی آن $L: R^2 \times R \times R_+^2 \rightarrow R$ با کانولوشن f و یک هسته گاوسی ناهمسان با متغیر مکانی σ_i^2 و متغیر زمانی τ_i^2 نشان داده می‌شود:

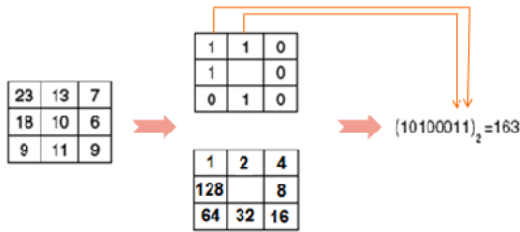
$$L^{sp}(\cdot; \sigma_i^2, \tau_i^2) = g(\cdot; \sigma_i^2, \tau_i^2) * f(\cdot) \quad (1)$$

در حالی که هسته گاوسی جداپذیر در دو بعد مکان و زمان بدین صورت تعریف می‌شود:

$$g^{sp}(x, y, t; \sigma_i^2, \tau_i^2) = \frac{1}{\sqrt{(2\pi)^3 \sigma_i^4 \tau_i^2}} \exp(-(x^2 + y^2)/2\sigma_i^2 - t^2/2\tau_i^2) \quad (2)$$

استفاده از یک پارامتر مقیاس جداگانه برای حوزه زمان ضروری است، زیرا حوزه‌های مکان و زمان برای هر رویداد کاملاً مستقل هستند. از این‌رو، یک ماتریس گشتاور مرتبه ۲ مکان-زمان معرفی می‌گردد.

مقدار الگو با مقایسه روشنایی نقطه مرکزی با نقاط دیگر در این همسایگی به دست می‌آید. اگر مقدار روشنایی آن‌ها بیشتر یا مساوی مقدار نقطه مرکزی باشد، یک و در غیر این صورت مقدار صفر جایگزین این نقاط خواهد شد. این مقدار برای نقاط قطری با درون‌یابی تخمین زده می‌شود (شکل ۴).



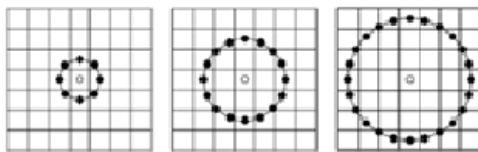
شکل (۴): نمایش عملگر پایه LBP

در این مرحله الگوی دودویی از تفاوت روشنایی نقطه مرکزی و نقاط مجاور در همسایگی مذکور به دست می‌آید:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (9)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

در این فرمول s تابع علامت، g_c مقدار روشنایی نقطه مرکزی و g_p مقدار روشنایی نقاط همسایگی هستند. متغیرهای P و R نیز به ترتیب تعداد نقاط و شعاع همسایگی را مشخص می‌کنند. شکل (۵) نمونه‌هایی از مجموعه‌های همسایگی متقارن برای مقادیر مختلف (P,R) را نشان می‌دهد.



شکل (۵): مجموعه‌های همسایگی $(۸,۱)$ ، $(۱۶,۲)$ ، $(۲۴,۳)$ [۳۰]

با فرض داشتن تصویری با مختصات $N \times M$ ، پس از محاسبه مقدار الگو برای هر نقطه، بافت تصویر با یک هیستوگرام بیان می‌شود:

$$H(k) = \sum_{i=1}^N \sum_{j=1}^M f(LBP_{P,R}(i,j), k) \quad (10)$$

$$k \in [0, K], \quad f(x, y) = \begin{cases} 1 & x=y \\ 0 & otherwise \end{cases}$$

۴- توصیف نواحی کلیدی

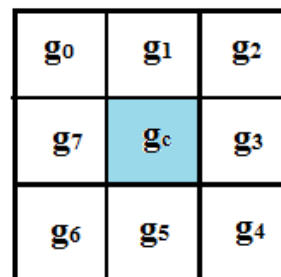
بافت پویا دنباله‌ای از تصاویر متحرک با تغییرات آماری در طی زمان است. از این رو، می‌توان از توصیف‌گرهای بافت پویا برای توصیف راه رفتن افراد در فضای مکان - زمان استفاده نمود. بنابراین، در این مرحله از الگوی دودویی محلی با سه صفحه متعامد (LBP-TOP) برای توصیف نواحی کلیدی به دست آمده در مرحله قبل بهره گرفته می‌شود.

این الگو تعمیمی از الگوی دودویی محلی (LBP) است. الگو LBP یک کد دودویی تولید می‌کند که از مقایسه یک همسایگی از نقاط با نقطه مرکزی به دست می‌آید و توصیف کننده الگوی بافت محلی است. با استفاده از LBP، بافت تصویر با کمک هیستوگرام کدهای دودویی برای هر نقطه در صفحه مکانی توصیف می‌گردد. این در صورتی است که در هنگام استفاده از LBP-TOP، اطلاعات دنباله ویدئویی در دو بعد مکان و زمان در نظر گرفته می‌شود.

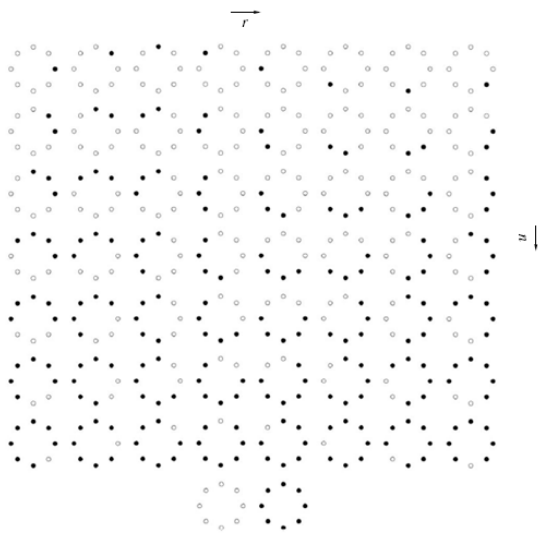
۴-۱- الگوی دودویی محلی

الگوی دودویی محلی (LBP) یک عملگر برای توصیف ساختار مکانی بافت محلی تصویر است. اجالا و همکارانش در توسعه کارهای گذشته خود [۲۸، ۲۹]، روشی برپایه الگوهای دودویی محلی ارائه داده‌اند که از نظر علمی و محاسباتی ساده بوده، نسبت به تغییرات مقیاس و چرخش مقاوم است [۳۰].

در این الگو، ابتدا بافت تصویر به صورت یک همسایگی محلی 3×3 تعریف می‌شود که $g_i (i=c, 0, \dots, 7)$ برابر با مقدار روشنایی نقاط این همسایگی است (شکل ۳).



شکل (۳): همسایگی یکنواخت مدور 3×3 [۳۰]



شکل (۶): الگوی دودویی یکنواخت به ازای $P=8$ [۳۱]

۴-۲- الگوی دودویی محلی هیستوگرام فوریه

الگوی دودویی محلی هیستوگرام فوریه (LBP-HF)؛ یک توصیف‌گر مقاوم به چرخش مبتنی بر الگوی دودویی یکنواخت (LBP^{u2}) است که با محاسبه تبدیل فوریه گسسته بر روی هیستوگرام‌های LBP^{u2} به دست می‌آید. برخلاف اکثر توصیف‌گرهای بافت که به صورت محلی مقاوم به چرخش هستند، این توصیف‌گر به صورت کلی برای سراسر ناحیه قابل توصیف نسبت به چرخش مقاوم است [۳۲].

با فرض هیستوگرام LBP^{u2} به صورت $h_I(U_P(n, r))$ ، دوران تصویب ورودی I به اندازه α درجه، باعث یک انتقال مدور در هیستوگرام در امتداد هر سطر است.

$$h_I^\alpha(U_P(n, r+a)) = h_I(U_P(n, r)) \quad (14)$$

$$(\alpha = a \frac{360}{P}, a = 0, 1, \dots, P-1)$$

در اینجا، از تبدیل فوریه گسسته^۱ برای حذف تأثیر چرخش بر روی سطرهای هیستوگرام استفاده می‌شود.

$$H(n, u) = \sum_{r=0}^{P-1} h_I(U_P(n, r)) e^{-i2\pi ur/P} \quad (15)$$

که $H(n, u)$ مقدار تبدیل فوریه مربوط به n امین سطر

که مقدار بیشینه الگوی دودویی محلی است.

با چرخش تصویر، مقادیر g_i می‌توانند بر روی محیط دایره و گرداگرد g_c حرکت نمایند. از این رو، چرخش یک الگوی دودویی خاص به طور طبیعی، یک مقدار جدید برای الگوی LBP به دست می‌دهد. برای حذف تأثیر چرخش و دادن یک برچسب یکتا به هر الگوی دودویی محلی مقاوم به چرخش، عملگر LBP بدین فرم معرفی می‌شود:

$$LBP_{P,R}^{ri} = \min \left\{ ROR(LBP_{P,R}, i) \mid i = 0, 1, \dots, P-1 \right\} \quad (11)$$

که $ROR(x, i)$ یک چرخش مدور ساعتگرد (شیفت به راست) به تعداد i بر روی x اعمال می‌کند و این چرخش تا رسیدن هر الگو به مقدار کمینه خود ادامه می‌یابد.

کارایی الگوهای متفاوت به ساختار مکانی الگوها نسبت داده می‌شود. برای این منظور یک معیار همسان (U) تعریف می‌شود.

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (12)$$

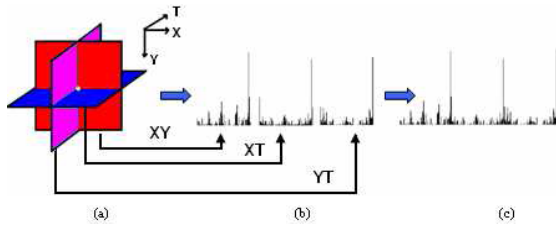
مقدار U در یک الگوی دودویی محلی، نشان دهنده تعداد تغییرات مکانی (انتقال بین ۰ و ۱) است. برای مثال، مقدار U در الگوی ۰۰۰۰۰۰۰۰ برابر با ۰ و در الگوی ۱۰۰۰۰۰۰۰ برابر با ۲ است (شکل ۶).

الگوی دودویی محلی یکنواخت مقاوم به چرخش به الگوی ظاهری یکسان رجوع می‌کند که انتقال و یا انفصال‌های محدودی در یک نمایش دودویی مدور دارد.

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (13)$$

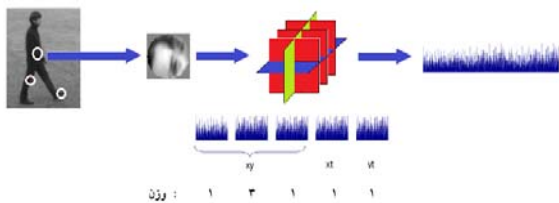
الگوهای دودویی یکنواخت ($U \leq 2$) که بیانگر ویژگی‌های اصلی بافت هستند، اکثریت الگوها را تشکیل می‌دهند. از این رو، به تمام الگوهای دودویی غیریکنواخت یک مقدار یکسان (P+1) تخصیص داده می‌شود.

یک از صفحات yt, xt, xy به‌طور جداگانه محاسبه می‌گردد و هیستوگرام مربوط به هر صفحه مشخص می‌شود. سپس هیستوگرام نهایی با الحاق این هیستوگرام‌ها به‌دست می‌آید.



شکل (۸): تشکیل هیستوگرام ویژگی [۱۶]

از آنجایی که تغییرات ظاهری که در بعد مکان رخ می‌دهند، بیشتر از تغییرات حرکتی در بعد زمان هستند، برای تغییرات در بعد مکان وزن بیشتری در نظر گرفته می‌شود. برای این هدف، پس از استخراج نقاط کلیدی دنباله ویدئویی، اطراف هر نقطه کلیدی سه صفحه در فضای مکانی xy (در $\frac{1}{2}, \frac{1}{4}$ و $\frac{3}{4}$ فضای مکانی) و یک صفحه در مرکز هر فضای زمانی (xt و yt) با ابعاد مشخص (25×25) در نظر گرفته می‌شود. به صفحه مرکزی مکانی نیز وزنی سه برابر وزن صفحات دیگر اعمال می‌شود. از این‌رو نسبت وزن بعد مکان به زمان پنج به دو خواهد بود (شکل ۹).



شکل (۹): توصیف ناحیه کلیدی با کمک عملگر LBP-OP

TOP

در بسیاری از کاربردهای پیشین توصیف‌گر LBP-OP، از عملگر LBP ساده برای توصیف صفحات متعامد استفاده شده است [۳۳]. اما در این مقاله، فضای مکانی (صفحه xy) با استفاده از عملگر LBP-HF و فضای زمانی (صفحات xt و yt) با کمک عملگر LBP مقاوم به چرخش (LBP^{riu2}) توصیف می‌گردند. از این‌رو توصیف‌گر LBP_TOP توسعه یافته نسبت به چرخش و

هیستوگرام $h_I(U_P(n, r))$ است.

از آنجایی که یک انتقال مدور باعث یک تغییر زاویه فاز در ضرایب تبدیل فوریه می‌گردد، لذا با فرض در معادله زیر به‌دست می‌آید:

$$H'(n, u) = H(n, u) e^{-i2\pi u a / P} \quad (16)$$

بنابراین، برای هر $1 \leq n_1$ و $n_2 \leq P-1$:

$$\begin{aligned} H'(n_1, u) \overline{H'(n_2, u)} &= H(n_1, u) e^{-i2\pi u a / P} \overline{H(n_2, u) e^{i2\pi u a / P}} \\ &= H(n_1, u) \overline{H(n_2, u)} \end{aligned} \quad (17)$$

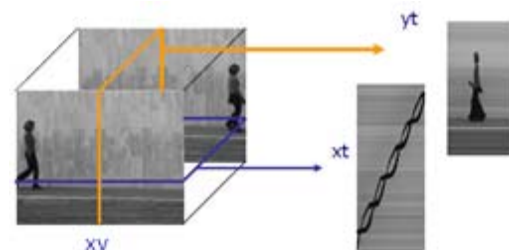
که $\overline{H(n_2, u)}$ نشان دهنده مزدوج $H(n_2, u)$ خواهد بود. با در نظر گرفتن $1 \leq n_1$ ، $n_2 \leq P-1$ و ویژگی‌های $0 \leq u \leq P-1$:

$$LBP^{u2} - HF(n_1, n_2, u) = H(n_1, u) \overline{H(n_2, u)} \quad (18)$$

نسبت به انتقالات مدور سطرهای هیستوگرام $h_I(U_P(n, r))$ مقاوم بوده و در نتیجه به چرخش تصویر ورودی $I(x, y)$ نیز مقاوم است.

۳-۴ LBP-TOP توسعه یافته

عملگر LBP توصیف‌کننده الگوی بافت تصویر در بعد مکان است؛ در صورتی که توصیف‌گر LBP-TOP اطلاعات دنباله ویدئویی را در دو بعد مکان (صفحه xy) و زمان (صفحات xt و yt) در نظر می‌گیرد (شکل ۷).



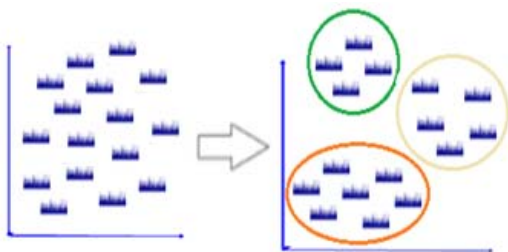
شکل (۷): نمایش صفحات xy, xt, yt راه رفتن [۲۲].

همانطور که در شکل (۸) دیده می‌شود، ویژگی‌های هر

دنباله ویدئویی است که دارای تغییراتی در بعد مکان و زمان است، مانند زانو، قوزک پا و آرنج.

تغییرات مقیاس مقاوم است.

۴-۴- تحلیل چند نمایشی



شکل (۱۱): ایجاد فرهنگ لغت از مجموعه بردارهای ویژگی سپس برای بهبود کارایی، درصدی از دسته‌هایی با توزیع یکنواخت، به علت دارا بودن سهم کمتر در نتیجه نهایی، از فرهنگ لغت حذف می‌شوند.

برای افزایش دقت مرحله توصیف، از روش چندنمایشی استفاده می‌شود. هدف این روش ترکیب چندین هسته از عملگرهای برپایه LBP، با پارامترهای P و R متفاوت است و هیستوگرام نهایی از اتصال هیستوگرام‌های حاصل از هر هسته به دست می‌آید (شکل ۱۰). در اینجا برای توصیف هر صفحه از توصیف‌گر LBP-TOP با دو هسته ۸ و ۱۶ نقطه همسایگی و شعاع‌های یک و دو استفاده می‌شود.



شکل (۱۰): تحلیل روش چند نمایشی

۵-۱- K-means سلسله مراتبی

اجرای الگوریتم K-means سلسله مراتبی یک درخت K-d می‌سازد که K تعداد دسته در هر اجرا و d تعداد لایه‌های درخت هستند. در این روش، ابتدا الگوریتم K-means اصلی بر روی داده‌ها اعمال می‌گردد و K مرکز دسته به دست می‌آید. سپس داده‌ها به K گروه تقسیم می‌شوند؛ به طوری که هر گروه شامل نزدیکترین داده‌ها به مرکز یک دسته خاص است. پس از آن، الگوریتم K-means اصلی، جداگانه بر روی هر گروه اعمال می‌گردد و این کار به صورت بازگشتی تا d لایه انجام می‌گیرد. پس از اجرای کامل الگوریتم K-means سلسله مراتبی، K^d دسته (برگ‌های درخت) به دست می‌آید. در این مقاله از فاصله اقلیدسی به عنوان معیار فاصله برای تعیین فاصله داده‌ها و مراکز دسته‌ها استفاده شده است.

در پایان این مرحله هر ناحیه کلیدی با یک هیستوگرام و یا بردار ویژگی، و هر دنباله ویدئویی با مجموعه‌ای از بردارهای ویژگی مشخص می‌شود.

۵- ایجاد فرهنگ لغت

تعداد نواحی کلیدی به دست آمده در مرحله اول برای هر دنباله ویدئویی متفاوت خواهد بود. از این رو، پس از اعمال توصیف‌گر بافت پویا، تعداد بردارهای ویژگی به دست آمده برای هر دنباله نیز متفاوت خواهد بود و این مسأله، مقایسه دنباله‌ها را با مشکل مواجه می‌کند. با ایجاد فرهنگ لغت می‌توان بر این مشکل غلبه نمود.

برای ایجاد فرهنگ لغت، با کمک الگوریتم K-means سلسله مراتبی، بردارهای ویژگی مجموعه آموزشی دسته‌بندی می‌شوند، هر دسته نشان دهنده یک کلمه است و این کلمات در کنار یکدیگر یک فرهنگ لغت تشکیل می‌دهند (شکل ۱۱). هر کلمه در اینجا نمایانگر ناحیه‌ای از

۶- استخراج هیستوگرام

در این مرحله، هر دنباله ویدئویی آموزشی با هیستوگرامی از تعداد تکرارهای کلمات مربوط به دنباله مورد نظر در فرهنگ لغت تعریف می‌شود (شکل ۱۲).

الگوریتم به‌علت تعیین دسته‌بندی اولیه تصادفی، نتایج متفاوتی در هر اجرا ارائه می‌دهد. از این‌رو، نتایج بیان شده پس از بررسی ۱۰ بار اجرای الگوریتم تعیین گردیده‌اند.

۸-۱- پایگاه داده KTH

پایگاه داده KTH یکی از مشهورترین و پیچیده‌ترین پایگاه داده‌های تحلیل حرکات انسان است. این پایگاه داده شامل ۲۵ نفر است که هر یک شش عمل انجام می‌دهند. در این مقاله روش پیشنهادی بر روی عمل راه رفتن افراد این پایگاه داده اعمال می‌شود. دنباله‌های تصویر افراد در شرایط متفاوتی تهیه شده است که در این‌جا سه حالت آن در نظر گرفته می‌شود: دنباله راه رفتن نرمال، دنباله راه رفتن با تغییر در مقیاس و دنباله راه رفتن با تغییر در پوشش افراد. نمونه‌هایی از این دنباله‌ها در شکل (۱۳) قابل مشاهده است. زاویه دید در دنباله‌های اول و سوم از نیم‌رخ و دنباله دوم به‌صورت قطری است.



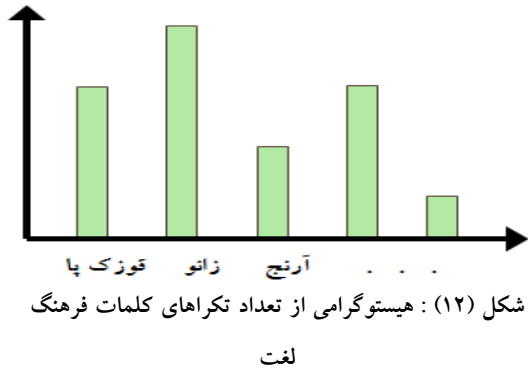
شکل (۱۳): نمونه‌هایی از پایگاه داده KTH

از چپ به راست به ترتیب دنباله‌های راه رفتن نرمال، راه رفتن با تغییر در مقیاس و راه رفتن با تغییر در پوشش افراد نمایش داده شده است.

قابل توجه است که پس‌زمینه در پایگاه داده KTH ثابت است و از آنجایی که روش پیشنهادی بر روی نقاط کلیدی که نقاطی با تغییرات در مکان و زمان هستند، انجام می‌گیرد، نیازی به حذف پس‌زمینه وجود ندارد.

در فرایند یادگیری، فرهنگ لغت با استفاده از الگوریتم K-means سلسله‌مراتبی و با در نظر گرفتن $K=2$ ، $level=10$ با (۲۸۱۰) ۱۰۲۴ کلمه ایجاد می‌شود و سپس با حذف دسته‌های یکنواخت ۷۰۰ کلمه باقی می‌ماند. در پایان، برای هر فرد یک مدل آموزشی و به ازای ۲۵ فرد موجود در پایگاه داده، ۲۵ مدل آموزشی خواهیم داشت.

هیستوگرام به‌دست آمده به عنوان مدل آموزشی برای دنباله ویدئویی ذخیره می‌گردد.



شکل (۱۲): هیستوگرامی از تعداد تکرارهای کلمات فرهنگ لغت

۷- تطبیق الگو

در این مرحله، نواحی کلیدی دنباله ویدئویی تست استخراج و توصیف می‌گردند. سپس مجموعه بردارهای ویژگی به‌دست آمده، با کلمات فرهنگ لغت ایجاد شده در مرحله یادگیری مقایسه می‌شوند و هر بردار در دسته متناسب با خود قرار می‌گیرد. سپس هیستوگرام نهایی دنباله ویدئویی تست به‌دست می‌آید. در این مرحله، با استفاده از یک طبقه‌بند مناسب هیستوگرام دنباله تست با مدل‌های آموزشی به‌دست آمده در فرایند یادگیری مقایسه می‌شود و نمونه تست به کلاس مشابه منطبق می‌گردد.

در این مقاله برای طبقه‌بندی از ماشین بردار پشتیبان (SVM) استفاده شده است. این طبقه‌بند شامل توابع هسته‌ای مختلفی است. از این‌رو، نتایج با برخی توابع از قبیل "چند جمله‌ای"، "گوسی" و "تابع پایه شعاعی" بررسی شد و سرانجام به‌علت نتایج بهتری که با تابع هسته‌ای "چندجمله‌ای" به‌دست آمد، این تابع به عنوان هسته SVM انتخاب گردید.

۸- نتایج تجربی

در این بخش، کارایی روش پیشنهادی بر روی دو پایگاه داده KTH [۳۳] و IXMAS [۳۴] بررسی می‌گردد.

از آنجایی که در روش ارائه شده، برای ایجاد فرهنگ لغت از الگوریتم K-means استفاده شده است و این

تست و با برعکس، به علت عدم تناسب در مقیاس، زاویه دید و پوشش، پایین‌ترین نتایج را ارائه می‌دهند.

در جدول (۱-ب) نیز شرایط یکسانی صدق می‌کند. استفاده از دنباله‌های "راه رفتن با تغییر مقیاس" و "تغییر پوشش" برای مرحله یادگیری، به علت دربرداشتن اطلاعات مقیاس، زاویه دید و پوشش در شرایط متفاوت، بهترین نتیجه را ارائه می‌دهد. در صورتی که با استفاده از دنباله‌های "راه رفتن نرمال" و "تغییر پوشش" برای مرحله یادگیری، اطلاعات مقیاس و زاویه دید نادیده گرفته شده و با استفاده از دنباله‌های "راه رفتن نرمال" و "تغییر مقیاس" برای مرحله یادگیری، اطلاعات پوشش نادیده گرفته می‌شوند و از این‌رو نتایج پایین‌تری به دست می‌آید.

۸-۲- پایگاه داده IXMAS

پایگاه داده IXMAS یکی از مشهورترین پایگاه داده‌های چندنمایی^{۱۲} تحلیل حرکات انسان است. این پایگاه داده شامل ۱۲ نفر است که هر یک ۱۴ عمل انجام می‌دهند. در این مقاله روش پیشنهادی بر روی عمل راه رفتن ۱۰ نفر اعمال می‌شود. دنباله‌های تصویر در این پایگاه داده تحت پنج نما گرفته شده‌اند که در این‌جا چهار نما مد نظر است و نمای پنجم که از بالای سر فرد گرفته شده است، به علت دارا نبودن اطلاعات کافی حذف می‌شود (شکل ۱۴).



شکل (۱۴): نمونه‌هایی از پایگاه داده IXMAS

نمونه‌ها از چپ به راست به ترتیب با دوربین‌های شماره ۱، ۲، ۳ و ۴ گرفته شده‌اند.

در این پایگاه داده، همان‌طور که در شکل (۱۵) دیده می‌شود، راه رفتن هر فرد به صورت دایره‌ای انجام می‌گیرد.



شکل (۱۵): نمونه‌ای از حرکت فردی تحت دوربین شماره ۱.

در جدول (۱) نتایج روش ارائه شده بر روی پایگاه داده KTH نشان داده شده است.

جدول (۱): نتایج روش پیشنهادی بر روی پایگاه داده KTH (الف)

مجموعه آموزش	نرمال	تغییر مقیاس	تغییر مقیاس	تغییر پوشش	تغییر پوشش
مجموعه تست	تغییر مقیاس	تغییر پوشش	نرمال	نرمال	تغییر مقیاس
نتایج	٪۷۳٫۱	٪۷۷٫۳	٪۷۲	٪۶۳	٪۷۵٫۳

(ب)

مجموعه آموزش	تغییر مقیاس و تغییر پوشش	نرمال و تغییر پوشش	نرمال و تغییر مقیاس
مجموعه تست	نرمال	تغییر مقیاس	تغییر پوشش
نتایج	٪۸۲٫۴	٪۷۴٫۷	٪۷۸٫۴

در جدول (۱-الف) از سه حالت موجود، دنباله‌های ویدئویی در یک حالت برای یادگیری و در حالتی دیگر برای تست به کار می‌رود. در جدول (۱-ب) از سه حالت موجود، دو حالت برای یادگیری و یک حالت برای تست استفاده می‌شود. میانگین نتایج جدول الف، ۷۰٫۷۵ درصد و میانگین نتایج جدول ب، ۷۸٫۵ درصد است. همان‌طور که از نتایج بر می‌آید، در روش دوم به علت وجود اطلاعات جزئیات بیشتر در مرحله یادگیری، نتایج بهتری به دست آمده است.

با بررسی جدول (۱-الف) در می‌یابیم دنباله "راه رفتن نرمال"، شامل اطلاعاتی از هر دو دنباله دیگر است. این دنباله با دنباله "راه رفتن با تغییر مقیاس" از نظر پوشش و بافت تصویر، و با دنباله "راه رفتن با تغییر پوشش" از لحاظ زاویه دید و مقیاس مشابه است. از این‌رو، استفاده از دنباله "راه رفتن نرمال" برای مرحله یادگیری، اطلاعات مناسب و نتایج قابل قبولی را ارائه می‌دهد. دنباله‌های "راه رفتن نرمال" و "راه رفتن با تغییر پوشش" تناسب بیشتری با یکدیگر دارند و از این‌رو، اطلاعات مناسبتر و نتایج بهتری را ارائه می‌دهند. استفاده از دنباله "راه رفتن با تغییر مقیاس" برای یادگیری و دنباله "راه رفتن با تغییر پوشش" برای

در زندگی عادی عمل راه رفتن در شرایط مختلفی از زاویه دید، چرخش، سرعت، مقیاس و تغییرات ظاهری انجام می‌گیرد و یک روش مطلوب برای شناسایی باید همه این شرایط را در نظر گیرد. هر یک از پایگاه داده‌های ارائه شده برای راه رفتن، برخی شرایط را در هنگام راه رفتن اعمال کرده، برخی دیگر را، با توجه به روش مورد استفاده، ثابت در نظر می‌گیرند. برای مثال، در برخی پایگاه داده‌ها، تغییرات ظاهری نادیده گرفته شده و راه رفتن افراد با پوشش و شکل ظاهری یکسان انجام می‌گیرد، در برخی نیز، در زمان راه رفتن، مقیاس و یا زاویه دید ثابت در نظر گرفته می‌شود، در صورتی که در این مقاله با استفاده از پایگاه داده‌های مذکور، عمل راه رفتن در شرایط مختلف بررسی شده است. روش ارائه شده در این مقاله، با در نظر گرفتن اطلاعات در بعد زمان، نسبت به تغییرات سرعت مقاوم است. همچنین، به علت استفاده از ویژگی‌های محلی نسبت به زاویه دید و تغییرات ظاهری مقاوم بوده، با توسعه توصیفگر LBP-TOP، نسبت به چرخش و مقیاس نیز مقاوم است، از این رو، نتایج مناسبی به دست آمده است.

همان‌گونه که ذکر شد، از آنجایی که تاکنون از این پایگاه داده‌ها برای هدف "شناسایی از روی نحوه راه رفتن" استفاده نشده است، از این رو امکان مقایسه دقیق با کارهای گذشته وجود ندارد، اما در مقاله [۲۲]، که منبعی برای این تحقیق بوده است، روشی مبتنی بر بافت پویا برای شناسایی از روی نحوه راه رفتن بر روی پایگاه داده CMU، با میانگین نتایج ۸۲٪، ارائه شده است. در مقاله مذکور فقط تغییرات ظاهری و سرعت در نظر گرفته شده است. همچنین، نیاز به حذف پس زمینه یکی از محدودیت‌های روش ارائه شده در این مقاله است. جداسازی پس زمینه باعث بهبود نتایج می‌گردد، اما از نظر محاسباتی بسیار پیچیده و زمان‌بر است. از این رو، سعی می‌شود در صورت امکان از این مرحله اجتناب کرده، بر روی داده‌های اصلی کار شود. در روش ارائه شده در این تحقیق، بدلیل استفاده از اطلاعات محلی نیازی به حذف پس زمینه نبوده و همچنین در پایگاه داده مورد استفاده نیز شرایط بیشتری در نظر گرفته شده است. لذا با توجه به بهبودهای انجام شده و مقایسه نتایج با مقاله [۲۲]، پیش‌بینی می‌شود که روش ارائه

در فرایند یادگیری، فرهنگ لغت با استفاده از الگوریتم K-means سلسله مراتبی و با در نظر گرفتن $k=2$ ، level=11 با (۲۸۱۱) ۲۰۴۸ کلمه ایجاد می‌شود و سپس با حذف دسته‌های یکنواخت ۱۷۰۰ کلمه باقی می‌ماند. در این پایگاه داده، برای هر فرد چهار مدل آموزشی و هر مدل تحت یک نما به دست می‌آید. از این رو در پایان فرایند یادگیری، به ازای ۱۰ فرد موجود در پایگاه داده، ۴۰ مدل آموزشی خواهیم داشت.

در این پایگاه داده هر فرد عمل راه رفتن را سه بار و در شرایط یکسان انجام می‌دهد. از این رو تحت هر نما، برای هر فرد سه دنباله ویدئویی راه رفتن وجود دارد. برای آزمایش روش پیشنهادی، هر بار دو دنباله برای مرحله یادگیری و دنباله دیگر برای تست به کار برده می‌شود. در جدول (۲) نتایج روش پیشنهادی بر روی پایگاه داده IXMAS نشان داده شده است. همان‌گونه که در این جدول قابل مشاهده است، میانگین نتایج ۷۳٫۸ درصد است.

جدول (۲): نتایج روش پیشنهادی بر روی پایگاه داده

IXMAS				
نتایج	دوربین ۱	دوربین ۲	دوربین ۳	دوربین ۴
نتایج	۶۶٫۵٪	۷۲٫۲٪	۸۱٪	۷۵٫۵٪

نتایج برای هر دوربین به طور جداگانه بیان شده است. همان‌گونه که در این جدول دیده می‌شود، دنباله تصویر به دست آمده از دوربین شماره ۳ اطلاعات بیشتری در برداشته و از این رو بهترین نتایج را ارائه می‌دهد.

۸-۳- تحلیل و بررسی

در مقاله [۳۵] روش ارائه شده بر روی پایگاه داده KTH برای شناسایی حرکات انسان اعمال گردید و نتایج مطلوبی به دست آمد. سپس سعی شد تا از روشی مشابه بر روی داده‌های عمل راه رفتن در این پایگاه داده، برای شناسایی از روی نحوه راه رفتن نیز استفاده گردد. از این رو در این مقاله به جای استفاده از پایگاه داده‌های ارائه شده برای راه رفتن، از پایگاه داده‌های KTH و IXMAS، که در زمینه تحلیل حرکات انسان هستند، استفاده شده است.

شده در این تحقیق نتایج مطلوبی ارائه داده است.

۹- نتیجه‌گیری

در این مقاله روشی بر پایه فرهنگ لغت برای شناسایی انسان از روی نحوه راه رفتن ارائه شده است. راه رفتن انسان دارای اطلاعات مهمی در هر دو بعد مکان و زمان است از این رو، در این روش از توصیف‌گرهای بافت پویا بهره گرفته شده است. در این مقاله توصیف‌گر LBP-TOP به عنوان یک توصیف‌گر مقاوم به چرخش توسعه یافته و از آن برای توصیف ویژگی‌های محلی استفاده می‌شود. سپس با کمک الگوریتم K-means سلسله مراتبی و ایجاد فرهنگ لغتی از بردارهای ویژگی، هر دنباله ویدئویی از راه رفتن افراد با هیستوگرامی از تکرار کلمات این فرهنگ لغت نمایش می‌یابد. کارایی روش پیشنهادی بر روی دو پایگاه داده KTH و multi view IXMAS بررسی گردیده و نتایج مطلوبی به دست آمده است.

مراجع

- [8] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes", ICCV, pp. 1395-1402, 2005.
- [9] R. Mattivi and L. Shao, "Human Action Recognition Using LBP-TOP as Sparse Spatio-Temporal Feature Descriptor", CAIP, pp. 740-747, 2009.
- [10] C. BenAbdelkader, R. G. Cutler, and L. S. Davis, "Gait recognition using image self-similarity", EURASIP Journal on Applied Signal Processing, Vol. 24, No. 4, pp. 572-585, 2004.
- [11] Y. Liu, R. Collins, and Y. Tsin, "Gait sequence analysis using frieze patterns", Computer Vision—ECCV, pp. 733-736, 2002.
- [12] S. Lee, Y. Liu, and R. Collins, "Shape variation-based frieze pattern for robust gait recognition", IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.
- [13] G. C. Huang and Y. H. Wang, "Human gait recognition based on XT plane energy images", International Conference on Wavelet Analysis and Pattern Recognition, ICWAPR, pp. 1134-1138, 2007.
- [14] T. Kobayashi and N. Otsu, "Action and simultaneous multiple-person identification using cubic higher-order local auto-correlation", Pattern Recognition, pp. 741-744, 2004.
- [15] J. Ren, X. Jiang, and J. Yuan, "Dynamic texture recognition using enhanced lbp features", IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP), 2013.
- [16] Z. Guoying, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 6, pp. 915-928, 2007.
- [17] J. Komulainen, A. Hadid, and M. Pietikäinen, "Face Spoofing Detection using Dynamic Texture", International Workshop on Computer Vision With Local Binary Pattern Variants - ACCV, 2012.
- [18] G. Zhao, M. Pietikäinen and A. Hadid, "Local spatiotemporal descriptors for visual recognition of spoken phrases", International Workshop on Human-centered Multimedia, pp. 57-66, 2007.
- [19] Q. Zhu, N. Chatlani, and J. J. Soraghan, "1-D Local binary patterns based VAD used INHMM-based improved speech recognition", the 20th European Signal
- [1] C. B. Abdelkader, L. Davis, and R. Cutler, "Motion-based recognition of people in eigengait space", pp. 267-272, 2002.
- [2] J. Han and B. Bhanu, "Individual recognition using gait energy image", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 2, pp. 316-322, 2006.
- [3] S. Fazekas, T. Amiaz, D. Chetverikov, and N. Kiryati, "Dynamic texture detection based on motion analysis", International Journal of Computer Vision, Vol. 82, No. 1, pp. 48-63, 2009.
- [4] F. Tafazzoli and R. Safabakhsh, "Model-based human gait recognition using leg and arm movements", Engineering Applications of Artificial Intelligence, Vol. 23, No. 8, pp. 1237-1246, 2010.
- [5] G. V. Veres, L. Gordon, J. N. Carter, and M. S. Nixon, "What image information is important in silhouette-based gait recognition?", CVPR, pp. 776-782, 2004.
- [6] G. Zhao, L. Cui, and H. Li, "Gait recognition using fractal scale", Pattern Analysis & Applications, Vol. 10, No. 3, pp. 235-246, 2007.
- [7] V. Kellokumpu, G. Zhao, and M. Pietikäinen,

- Pietikäinen, "Rotation invariant image description with local binary pattern histogram fourier features", *Image Analysis*, pp. 61-70, 2009.
- [33] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Dynamic textures for human movement recognition", pp. 470-476, 2010.
- [34] J. Liu and M. Shah, "Learning human actions via information maximization", pp. 1-8, 2008.
- [35] B. Abdolahi, S. Ghasemi, and N. Gheissari, "Human Motion Analysis using Dynamic Textures", in 16th International Iran Computer Society conference, Shiraz, Iran, 2012.
- Processing Conference (EUSIPCO), pp. 1633-1637, 2012.
- [20] B. Kim, S.-H. Lee, M.-K. Sohn, D.-J. Kim, and H. Kim, "Hand Pose Recognition Using Local Binary Patterns and Random Forests Classifier", *Bulletin of Networking, Computing, Systems, and Software*, Vol. 2, No. 1, pp. 64-65, 2013.
- [21] V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Recognition of human actions using texture descriptors", *Machine Vision and Applications*, Vol. 22, No. 5, pp. 767-780, 2011.
- [22] M. Tistarelli, M. Nixon, V. Kellokumpu, G. Zhao, S. Li, and M. Pietikäinen, "Dynamic Texture Based Gait Recognition", *ICB, LNCS*. 5558, pp. 1000-1009, 2009.
- [23] C. Harris and M. Stephens, "A combined corner and edge detector", *Alvey Vision Conference*, pp. 147-152, 1988.
- [24] I. Laptev and T. Lindeberg, "Space-time interest points", *International Conference on Computer Vision*, pp. 432-439, 2003.
- [25] I. Laptev, "On space-time interest points", *International Journal of Computer Vision*, Vol. 64, No. 2/3, pp. 107-123, 2005.
- [26] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-8, 2008.
- [27] A. Kovashka and K. Grauman, "Learning a hierarchy of discriminative space-time neighborhood features for human action recognition", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2046-2053, 2010.
- [28] T. Ojala, K. Valkealahti, E. Oja, and M. Pietikäinen, "Texture discrimination with multidimensional distributions of signed gray-level differences", *Pattern Recognition*, Vol. 34, No. 3, pp. 727-739, 2001.
- [29] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions", *Pattern Recognition*, Vol. 33, No. 1, pp. 43-52, 2000.
- [30] T. Ojala, M. Pietikäinen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp. 971-987, 2002.
- [31] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching", *Pattern Recognition*, Vol. 43, No.3, pp. 706-719, 2010.
- [32] T. Ahonen, J. Matas, C. He, and M.

¹ Self-Similarity Plots

² Shape Variance-Based Frieze Pattern

³ Cubic Higher order Local Autocorrelation

⁴ Dynamic Texture

⁵ Local Binary Pattern-Three Orthogonal Planes

⁶ Silhouette

⁷ Virtual Dictionary

⁸ Clustering

⁹ Local Binary Pattern Fourier Histogram

¹⁰ Discrete Fourier Transform

¹¹ Radial Basis Function

¹² Multiview

