

Social Groups Detection by Using Support Vector Machine in Video

Ali Akbari¹, Hassan Farsi², Sajad Mohamadzadeh^{3*}

1- Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

2- Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

*3- Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

¹ali.akbari@birjand.ac.ir, ²hfarsi@birjand.ac.ir, and ^{3*}s.mohamadzadeh@birjand.ac.ir

Corresponding author's address: Sajad Mohamadzadeh, Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

Abstract- Detecting social groups is one of important and complex problems which has been concerned recently. Detecting social groups and relation between group members will be necessary for human robots in near future. Databases have some information including trajectories and also labels of members. The target is to detect social groups that contains at least two people or detecting individual motion of the people. In the proposed method, for detecting social groups, physical distance, temporal causality and shape similarity features are used. The required time to extract these features is lower than the other suggested features. In addition to precision and recall, the effectiveness of the proposed method in terms of required time for training and testing data is also examined. Lower required time provides greater ability to implement for human robots. The proposed method provides acceptable results in valid databases and is compared to existing methods in terms of statistical results and the required time.

Keywords- Detecting Social Groups, Physical Distance, Temporal Causality, Shape Similarity, Crowd Place.



تشخیص گروه‌های اجتماعی با استفاده از ماشین بردار پشتیبان در ویدئو

علی اکبری^۱، حسن فرسی^۲، سجاد محمدزاده^{۳*}

۱- دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

۲- دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

۳* - دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

¹ali.akbari@birjand.ac.ir, ²hfarsi@birjand.ac.ir, and ^{3*}s.mohamadzadeh@birjand.ac.ir

* نشانی نویسنده مسئول: خراسان جنوبی، بیرجند، دانشگاه بیرجند، سجاد محمدزاده.

چکیده - تشخیص گروه‌های اجتماعی یکی از مسائل مهم و پیچیده‌ای است که در چند سال اخیر مورد توجه قرار گرفته است. تشخیص گروه‌های اجتماعی و همچنین نحوه ارتباط با اعضای گروه‌های اجتماعی از جمله مسائلی است که ربات‌های انسان‌نما در آینده نزدیک به آن نیاز دارند. پایگاه داده‌ها حاوی اطلاعات مسیر افراد و همچنین شماره افراد می‌باشند. هدف، یافتن گروه‌های اجتماعی با حداقل دو نفر و یا تشخیص انفرادی بودن حرکت افراد است. در روش پیشنهادی برای تشخیص گروه‌های اجتماعی از ویژگی‌های فاصله افراد، شباهت مسیر طی شده و شباهت دو دنباله حرکت بین هر دو نفر استفاده شده است. استخراج این ویژگی‌ها نسبت به ویژگی‌های بکار رفته تاکنون نیاز به زمان کمتری دارد. علاوه بر دقت و فراخوانی، کارایی روش پیشنهادی از نظر زمان مورد نیاز برای آموزش و آزمایش داده نیز بررسی شده است. هر چه زمان مورد نیاز کمتر باشد، قابلیت پیاده سازی برای ربات‌های انسان‌نما بیشتر خواهد شد. روش پیشنهادی در پایگاه داده‌های معتبر نتایج قابل قبولی داشته و در ادامه با روش‌های موجود از نظر نتایج آماری و زمان مورد نیاز، مقایسه شده است.

واژه‌های کلیدی: تشخیص گروه‌های اجتماعی، ویژگی فاصله، ویژگی تشابه تغییرات حرکت افراد، ویژگی تشابه شکل مسیر طی شده، مکان شلوغ.

۱- مقدمه

بر اساس مطالعات، بیشتر گروه‌های اجتماعی دو نفری هستند و گروه‌های اجتماعی سه نفری و چهار نفری از نظر تمایل افراد برای حضور در این گروه‌ها و حرکت در شلوغی در رتبه بعد قرار دارند [۳]. بر اساس پژوهشی دیگر برای تعریف گروه، از تاثیر پذیری شخصی از شخص یا اشخاص دیگر در حرکت کردن میان شلوغی کمک گرفته شده است [۴]. شخصی که وارد گروه می‌شود تحت تاثیر گروه، سرعت و جهت آن شخص تغییر کرده و مطابق گروه می‌شود.

در ادامه به مرور کارهای انجام شده در بخش دوم اشاره شده است. در بخش سوم، ویژگی‌های استفاده شده در روش پیشنهادی و فرآیند یافتن گروه‌های اجتماعی بیان شده است. در بخش چهارم، روش پیشنهادی با روش‌های موجود مقایسه شده است و در پایان به نتایج پژوهش اشاره شده است.

با گذر زمان و گسترش استفاده از دوربین برای حفظ امنیت و تشخیص ناهنجاری‌های اجتماعی، اهمیت پردازش داده‌های ویدئویی افزایش یافته است. تشخیص گروه‌های اجتماعی برای تشخیص ناهنجاری‌های اجتماعی و گروه‌های خطرناک از سوی دولت‌ها مورد توجه قرار گرفته است [۱]. تشخیص گروه‌های سرقت و گروه‌های عملیات تروریستی از اهمیت بالایی برخوردار است. برای تشخیص گروه‌هایی با این اهداف، تشخیص گروه‌های اجتماعی یک پیش نیاز محسوب می‌شود.

بر اساس تحقیقات اخیر مردم تمایل دارند در گروه‌ها و دسته‌های اجتماعی و تحت تاثیر آن‌ها در شلوغی حرکت کنند [۲]. حرکت کردن در گروه به معنی زیر مجموعه بودن در گروه نیز می‌باشد، به عبارت دیگر یک گروه شامل دو یا چند نفر را می‌توان از نظر مسیر حرکت، سرعت حرکت و جهت حرکت تقریباً یکسان در نظر گرفت.

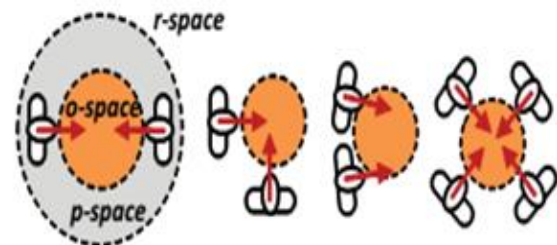
۲- مروری بر کارهای انجام شده

در مرجع [۵] تشخیص گروه‌های اجتماعی برای هدایت رفتار ربات در کارهای تیمی همراه با انسان مورد بررسی قرار گرفته است. در مرجع [۶] از گروه‌های اجتماعی برای جابجایی ربات در بین گروه‌های انسانی کمک گرفته شده است. تشخیص گروه‌های مکالمه و اجتماعی برای ربات در روش مرجع [۷]، با استفاده از جهت پایین تنه افراد و فاصله افراد انجام شده است. با گذر زمان، بر اهمیت تشخیص گروه‌های اجتماعی، در رابطه با ربات‌های انسان‌نما افزوده می‌شود.

روش‌های تشخیص گروه‌های اجتماعی که در آن‌ها از مسیر حرکت افراد استفاده می‌شود، به سه بخش بر پایه گروه^۱، بر پایه ترکیب گروه‌های منحصر به فرد^۲ و بر پایه فردی^۳ تقسیم می‌شوند [۸]. در بخش مبتنی بر پایه گروه، گروه‌های جزئی بررسی می‌شوند. در روش مرجع [۹]، که مبتنی بر پایه گروه می‌باشد، با استفاده از مشتق دوم ماتریس اطلاعات مسیر حرکت افراد، گروه‌های اجتماعی تشخیص داده شده است. در بخش ترکیب گروه‌های منحصر به فرد از اطلاعات با اهمیت‌تری همچون تعقیب گروه‌ها استفاده شده است [۱۰].

در بخش تشخیص گروه‌های اجتماعی بر پایه فردی از اطلاعات مسیر تک تک افراد استفاده می‌شود. در روش مرجع [۱۱]، ابتدا به تشخیص مقدماتی سر افراد پرداخته و سپس با کمک چگالی احتمال شلوعی در منطقه شلوع، با دقت بیشتری سر افراد را بازشناسی کرده است.

در روش [۱۲]، بر پایه موقعیت افراد، فاصله افراد و جهت سر افراد، گروه‌های اجتماعی تشخیص داده شده است. در این روش، گروه‌های ثابت و بدون حرکت تشخیص داده شده اند. در شکل ۱، حالت‌های مختلف در نظر گرفته شده در مقاله [۱۲] نشان داده شده است.



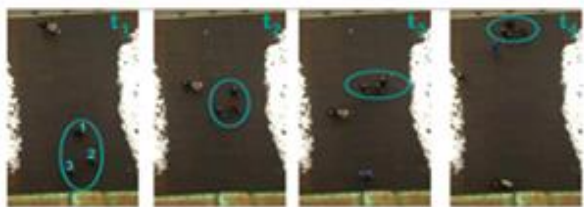
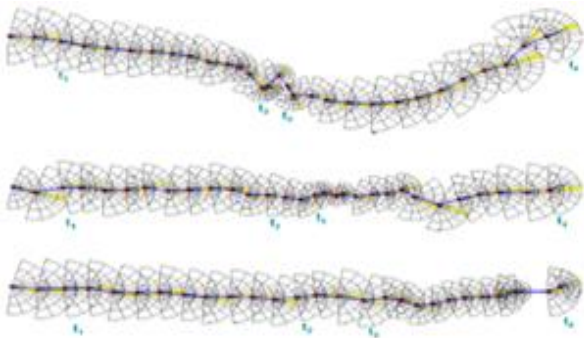
شکل ۱: حالت‌های مختلف گروه‌های اجتماعی در مقاله [۱۲].

در مرجع [۱۳]، با استفاده از ویژگی‌های فاصله، تفاوت سرعت و اطلاعات همپوشانی زمانی^۴ با SVM^۵، گروه‌های اجتماعی تشخیص داده شده است. در مرجع [۱۴]، از روش‌های تشابه احتمالاتی

تغییرات دو نفر در چند فریم و ناحیه‌بندی نرم^۶ برای تشخیص گروه‌های اجتماعی استفاده شده است.

در مرجع [۸]، از ویژگی‌های فاصله، تشابه تغییرات حرکت افراد در فریم‌های متوالی^۷، تشابه شکل مسیر طی شده توسط افراد^۸ و تشابه نقشه گرمایی افراد در فریم‌های قبلی استفاده شده است. در این روش، با محاسبه تشابه نقشه گرمایی افراد در فریم‌های قبلی، هزینه محاسباتی زیادی به روش اعمال می‌شود.

در روشی دیگر از ویژگی‌های مجاورت^۹ و سرعت^{۱۰} برای تشخیص گروه‌های اجتماعی استفاده شده است [۱۵]. در این روش در حالت‌های خیلی شلوع، از روش‌های مارکوف، از ویژگی‌های مجاورت و سرعت کمک گرفته شده و در حالت‌هایی که خیلی شلوع نیست از دنبال کردن افراد در فریم‌های متوالی، گروه‌های اجتماعی تشخیص داده شده است. در شکل ۲ بررسی موقعیت افراد در فریم‌های متوالی، نشان داده شده است.



شکل ۲: دنبال کردن افراد در فریم‌های متوالی در شلوعی کم برای تشخیص گروه‌های اجتماعی، از بالا به پایین: دنبال کردن شخص شماره ۱، دنبال کردن شخص شماره ۲، دنبال کردن شخص شماره ۳. [۱۵]

انواع مختلف گروه‌های اجتماعی که شامل گروه‌های انتقال، گروه‌های با فاصله کم در محیط شلوع و گروه‌های سلسله مراتبی هستند در شکل ۳ نشان داده شده است [۸].

گروه‌های انتقال^{۱۱}، برای حالت صف ایستادن یا انتقال از جایی که حالت ورودی یک نفره است تعریف شده است. گروه‌های با فاصله کم در محیط شلوع^{۱۲} با گروه‌های اجتماعی در محیط غیر شلوع تفاوت دارد و این تفاوت مربوط به تراکم بالاتر یا فاصله کم افراد گروه در محیط شلوع است. گروه‌های اجتماعی سلسله مراتبی^{۱۳} به گروه‌هایی گفته می‌شوند که شامل زیر گروه‌هایی است که با

مرجع [۸] می‌باشد، ویژگی‌های استفاده شده در مرجع [۸]، بیان می‌شوند. در ادامه، با استفاده از روش‌های انتخاب ویژگی، ویژگی‌های بهینه انتخاب شده و نتایج براساس این ویژگی‌ها گزارش می‌شوند.

۳-۱- تعریف مساله

برای بررسی هم گروه بودن افراد در گروه‌های اجتماعی، در آموزش و آزمایش، شناخت پارامترهای مساله و تابعی برای نشان دادن هم گروه بودن دو نفر یا چند نفر در گروه‌های اجتماعی، الزامی است. بنابراین ابتدا تابع مساله براساس رابطه (۱) معرفی می‌شود [۸].

$$cc = \max_{y \in z(m)} \sum_{x \in y} \sum_{a \neq b \in x} W_d^{ab} \quad (1)$$

در این رابطه، m همه افراد درون فریم ویدئو هستند که روش پیشنهادی باید گروه‌های اجتماعی بین آن‌ها را تشخیص دهد. تمام حالت‌های گروه‌بندی افراد و y حالتی از گروه‌بندی است که افراد a و b در یک گروه x قرار دارند. در رابطه (۱)، x گروهی شامل افراد a و b است و d ماتریس ویژگی بین همه افراد درون ویدئو است. فردی که جداگانه حرکت می‌کند، در گروهی با تعداد یک فرد قرار داده می‌شود. در آموزش و آزمایش هدف ماکزیمم کردن cc است. در نتیجه پارامترهایی در رابطه (۲) معرفی می‌شوند که اگر دو نفر در یک گروه باشند، W_d^{ab} عددی مثبت باشد و برعکس اگر دو نفر در یک گروه نباشند W_d^{ab} عددی منفی و نزدیک صفر باشد [۸]. زمان زیادی برای یافتن ماکزیمم cc بین همه حالت‌های موجود بین افراد، لازم است.

$$W_d^{ab} = \alpha(1 - d(a, b)) - \beta(d(a, b)) \quad (2)$$

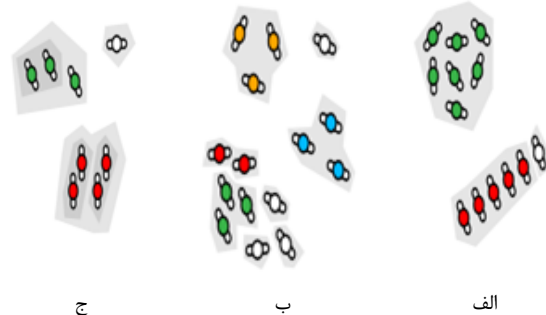
در رابطه (۲)، اگر افراد a و b در یک گروه قرار دارند، لازم است W_d^{ab} عددی مثبت شود و وظیفه ماشین بردار پشتیبان در مرحله آموزش، در حالت هم گروه بودن افراد a و b ، یافتن پارامترهای α و β است به گونه‌ای که W_d^{ab} عددی مثبت شود. بنابراین در رابطه (۱)، اگر گروه‌های اجتماعی درست تشخیص داده شوند، W_d^{ab} عددی مثبت می‌شود و در نتیجه cc ماکزیمم می‌شود.

در روش پیشنهادی، یک سطر از ماتریس ویژگی مربوط به افراد a و b ، از رابطه (۳) حاصل می‌شود.

$$d(a, b) = [d_{ph}, d_{sh}, d_{ca}]_{a,b} \quad (3)$$

بین هر دو نفر از افراد حاضر در ویدئو، ویژگی‌های فاصله، تشابه تغییرات حرکت افراد و تشابه شکل مسیر طی شده محاسبه می‌شود و ماتریس ویژگی از ترکیب این سه ویژگی بین هر دو نفر ساخته می‌شود. در رابطه ۳، d_{ph} ویژگی فاصله فیزیکی افراد، d_{sh} ویژگی تشابه شکل مسیر طی شده توسط افراد و d_{ca} ویژگی تشابه تغییرات حرکت افراد است. هر سه ویژگی بین هر دو نفر حاضر در

ترکیب یکدیگر گروه بزرگتری را تشکیل می‌دهند. برای مثال در شکل ۳، گروه سلسله مراتبی ۴ نفره شامل دو گروه ۲ نفره است. با وجود حالت‌های مختلف گروه‌های اجتماعی نشان داده شده در شکل ۳، تشخیص انواع لین گروه‌ها، از اهمیت بالایی برخوردار می‌باشد.



شکل ۳: انواع مختلف گروه‌های اجتماعی [۸]، الف: گروه‌های انتقال، ب: گروه‌های با فاصله کم در محیط شلوغ، ج: گروه‌های سلسله مراتبی.

در مرجع [۱۶]، یک روش بدون آموزش معرفی شده است که در آن از اطلاعات گروه‌ها در فریم قبل استفاده می‌شود. در این روش از یک مدل مخلوط نامعین^{۱۴} احتمالات گروه بودن با ویژگی میانگین فاصله در فریم‌های متوالی بررسی شده است. در این روش، فقط از ویژگی فاصله استفاده شده و با در نظر گرفتن ویژگی‌های بیشتر، تشخیص گروه‌های اجتماعی با دقت بالاتر دست یافتنی خواهد شد. در مرجع [۱۷]، با استفاده از اطلاعات مسیر افراد به تشخیص گروه‌های اجتماعی پرداخته شده است. برای یافتن گروه‌های اجتماعی چندین حالت مختلف بررسی شده است: گروه‌هایی که افراد به یک نقطه نزدیک می‌شوند، گروه‌های ثابت و گروه‌هایی که افراد از یک نقطه دور می‌شوند. تمام حالت‌ها بین همه افراد بررسی می‌شود. در مطالعه‌ای دیگر، براساس ناحیه‌بندی جهت‌های افراد به تشخیص گروه‌های اجتماعی پرداخته شده است [۱۸]. در مرجع [۱۹]، برای تشخیص گروه‌های اجتماعی از جهت‌یابی افراد در موقعیت‌های مختلف، استفاده شده است. در مرجع [۲۰]، نویسندگان با استفاده از همبستگی مکانی و زمانی بین مسیر طی شده توسط افراد به تشخیص گروه‌های اجتماعی پرداخته‌اند.

۳- روش پیشنهادی

آموزش و آزمایش در تشخیص گروه‌های اجتماعی به استخراج ویژگی و همچنین روابطی که نشان دهد دو نفر یا چند نفر در یک گروه هستند، نیاز دارند. در این بخش، ابتدا تعریف مساله بیان می‌شود که نشان می‌دهد دو نفر یا چند نفر چگونه در یک گروه قرار می‌گیرند. سپس با توجه به اینکه روش پیشنهادی بر مبنای

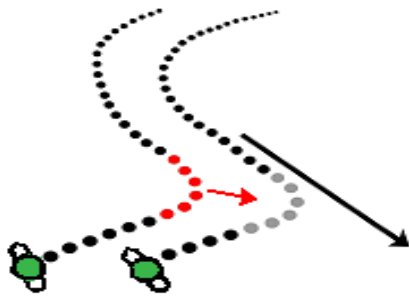
در این رابطه، مدل مخلوط گاوسی بین فرد a و فرد b در لحظه t بدست می‌آید. در رابطه (۴)، P_a^t موقعیت ذره a در لحظه t می‌باشد. در این مدل، N مدل گاوسی ساده با میانگین صفر بوده و واریانس‌ها همان مرزهای جدول تئوری مجاورت، یعنی $0/5$ و $1/2$ و $3/7$ و $7/6$ می‌باشند. در نتیجه بین فرد a و فرد b در لحظه t با هر کدام از چهار واریانس یک خروجی محاسبه می‌شود و خروجی نهایی میانگین چهار گاوسی با واریانس‌های مختلف است. در ادامه برای کل فریم متوالی در ویدئو بین هر دو نفر، لازم است میانگین مدل مخلوط گاوسی بر اساس رابطه (۵) محاسبه شود.

$$d_{ph} = \frac{1}{T} \sum_{t \in T} GMM(P_a^t - P_b^t) \quad (5)$$

توسط این رابطه ویژگی فاصله فیزیکی برای تشخیص گروه‌های اجتماعی استخراج می‌شود. T زمان کل ویدئو است. خروجی ویژگی فاصله فیزیکی بین هر دو نفر یک عدد خواهد شد.

۳-۳- ویژگی تشابه تغییرات حرکت افراد در فریم‌های متوالی

در همه پایگاه داده‌ها مکان و سرعت افراد در جهت‌های مکانی در هر فریم داده شده است. یعنی در فریم، شماره افراد با موقعیت و سرعت در جهت x و y داده شده است. براساس این اطلاعات، ابتدا احتمال شرطی شباهت مسیر فرد a به مسیر فرد a در n فریم قبلی محاسبه شده و در ادامه احتمال شرطی شباهت مسیر فرد a به مسیر فرد b در n فریم قبلی محاسبه می‌شود. اگر احتمال شرطی دوم نزدیک به احتمال شرطی اول شود، تشابه تغییرات حرکت فرد a و فرد b در فریم‌های متوالی وجود دارد. در شکل ۵ تشابه تغییرات حرکت افراد در فریم‌های متوالی نشان داده شده است.



شکل ۵: ویژگی تشابه تغییرات حرکت افراد در فریم‌های متوالی [۸].

این مساله به صورت ذیل فرمول بندی می‌شود [۸].

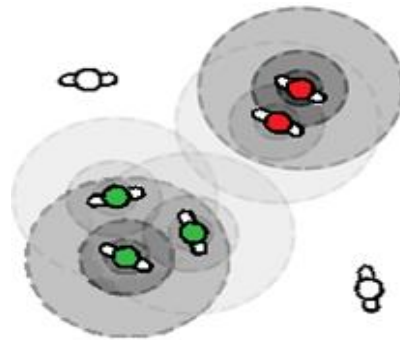
$$RSS_c = \sum_{t=1}^T (V_a(t) - P(V_a(t)|\bar{V}_a(n)))^2 \quad (6)$$

در رابطه (۶)، $\bar{V}_a(n)$ اطلاعات ذخیره شده مربوط به مسیر طی شده توسط فرد a از $t-n$ الی $t-1$ است. خروجی رابطه مربع اختلاف مسیر فرد a و احتمال شرطی شباهت مسیر فرد a ، به مسیر فرد a در n فریم قبلی از حال حاضر را برای همه زمان‌های ویدئو نشان

ویدئو محاسبه می‌شود. در ماتریس ویژگی ستون اول و دوم هر سطر شماره افراد و ستون سوم تا پنجم ماتریس، ویژگی‌های محاسبه شده است. بنابراین ماتریس ویژگی نهایی شامل شماره افراد و ویژگی بین هر دو نفر در همان سطح افراد است. با توجه به تعداد زیاد افراد در ویدئو، یافتن سه ویژگی بین هر دو نفر، هزینه محاسباتی زیادی خواهد داشت. به عنوان مثال، ماتریس ویژگی در فیلمی با حضور چهل نفر، بیش از هزار سطر دارد. در ادامه به معرفی ویژگی‌های استفاده شده در روش پیشنهادی خواهیم پرداخت.

۳-۲- ویژگی فاصله فیزیکی

یکی از ویژگی‌های مهم در تشخیص گروه‌های اجتماعی بررسی فاصله فیزیکی افراد در گروه است. در شکل ۴ اهمیت ویژگی فاصله فیزیکی بهتر نشان داده شده است.



شکل ۴: ویژگی فاصله فیزیکی در تشخیص گروه‌های اجتماعی [۸].

در مرجع [۲۱] به بیان تئوری مجاورت براساس ویژگی فاصله فیزیکی پرداخته شده است که در جدول ۱ نشان داده شده است.

جدول ۱: تئوری مجاورت [۲۱]

عنوان	فاصله بر حسب متر	توضیحات
صمیمی	۰-۰/۵	امکان درگیری بین افراد وجود دارد.
شخصی	۰/۱-۵/۲	روابط آشنا
اجتماعی	۱/۳-۲/۷	روابط رسمی
عمومی	۳/۷-۷/۶	کاملاً غیر شخصی

براساس تئوری مجاورت، چهار کلاس برای تفکیک نوع رابطه بین افراد براساس فاصله دو فرد، وجود دارد. هر کلاس با کلاس دیگر با مرزی که در جدول نشان داده شده، تفکیک شده است. در روش پیشنهادی از مدل مخلوط گاوسی که الهام گرفته از تئوری مجاورت می‌باشد، استفاده شده است. در ادامه به معرفی مدل مخلوط گاوسی پرداخته می‌شود که توسط رابطه (۴) بیان می‌گردد.

$$GMM(P_a^t - P_b^t) = \frac{1}{4} \sum_{i=1}^4 N(P_a^t - P_b^t | 0, \sigma_i) \quad (4)$$

توسط فرد a و فرد b می‌باشند. تابع هزینه انباشته به صورت زیر تعریف می‌شود [۸].

$$\gamma_{ab}(i, j) = D_{ab}^{ij} + \min\{\gamma_{ab}(i-1, j) + \gamma_{ab}(i-1, j-1) + \gamma_{ab}(i, j-1)\} \quad (10)$$

برای محاسبه تابع هزینه انباشته $\gamma_{ab}(i, j)$ ، براساس رابطه (۱۰)، باید ابتدا D_{ab}^{11} ، که برابر $\gamma_{ab}(1, 1)$ می‌باشد، محاسبه شود. در ادامه $\gamma_{ab}(i, j)$ براساس موقعیت مکانی دو فرد در دنباله زمانی حاصل می‌شود. در نهایت ویژگی تشابه شکل مسیر طی شده توسط افراد از رابطه (۱۱)، حاصل می‌شود [۸].

$$d_{sh}(a, b) = \gamma_{ab}(f, g) / \max(f, g) \quad (11)$$

در رابطه (۱۱)، f و g به ترتیب نشان دهنده مسیر طی شده توسط فرد a و فرد b می‌باشند. چون دو عدد f و g می‌توانند متفاوت باشند، بر ماکزیمم این دو عدد تقسیم شده است.

۵-۳- ویژگی تشابه نقشه گرمایی

ابتدا پارامتر نقشه فعالیت فرد a با استفاده از رابطه (۱۲) براساس نقشه گرمایی متوسط شخص در یک قطعه مستطیلی کوچک p و q حاصل می‌شود. قطعه مستطیلی R و C از اطلاعات مسیر طی شده توسط افراد a و b حاصل می‌شود و برابر مستطیلی است که دو نفر در طول و عرض آن در پنجره زمانی حرکت داشته‌اند.

$$H_a(i, j) = \sum_{p=1}^R \sum_{q=1}^C \bar{E}(p, q) e^{-\tau \| (p-i, q-j) \|} \quad (12)$$

در رابطه (۱۲)، نقشه فعالیت فرد a محاسبه می‌شود. در این رابطه، τ پنجره زمانی از ویدئو است. در ادامه ویژگی نقشه گرمایی بین افراد a و b از رابطه (۱۳) محاسبه می‌شود.

$$d_{ne}(a, b) = \sum_{i=1}^R \sum_{j=1}^C H_a(i, j) H_b(i, j) \quad (13)$$

خروجی رابطه (۱۳)، ویژگی تشابه نقشه گرمایی بین افراد a و b می‌باشد که نشان دهنده اشتراک نقشه فعالیت افراد a و b در قطعه مستطیلی R و C می‌باشد.

در بخش بعد، ابتدا با استفاده از روش‌های انتخاب ویژگی، بهترین ویژگی‌ها انتخاب می‌شوند و سپس با استفاده از نتایج بدست آمده، سه ویژگی برای روش پیشنهادی انتخاب می‌شود. در انتها با مقایسه روش پیشنهادی با روش‌های موجود، کارایی روش پیشنهادی نشان داده شده است.

۴- نتایج روش پیشنهادی و مقایسه با روش‌های موجود

۴-۱- پایگاه داده‌ها

با توجه به تعداد زیاد افراد در ویدئو و همچنین عبور افراد از مسیرهای مختلف، تشخیص گروه‌ها با مشکل مواجه می‌شود. عبور گروه‌ها از کنار هم نیز باعث پیچیدگی مساله می‌شود. در ادامه

می‌دهد. در این رابطه تشابه تغییرات حرکت فرد به تغییرات حرکت همان شخص در فریم‌های قبل محاسبه می‌شود و در رابطه (۷)، تشابه تغییرات حرکت فرد به تغییرات حرکت شخصی دیگر در فریم‌های قبل محاسبه می‌شود [۸]. به عبارت دیگر، تشابه تغییرات حرکت فرد a به تغییرات حرکت فرد b در فریم‌های قبل محاسبه می‌گردد. در روابط (۶) و (۷)، P احتمال شرطی است.

$$RSS_u = \sum_{t=1}^T (V_a(t) - P(V_a(t)|\bar{V}_b(n)))^2 \quad (7)$$

رابطه (۸)، تشابه تغییرات حرکت افراد a و b در فریم‌های متوالی است [۸].

$$S_{b \rightarrow a} = \frac{(RSS_c - RSS_u)/n}{RSS_u / (T - 2n - 1)} \quad (8)$$

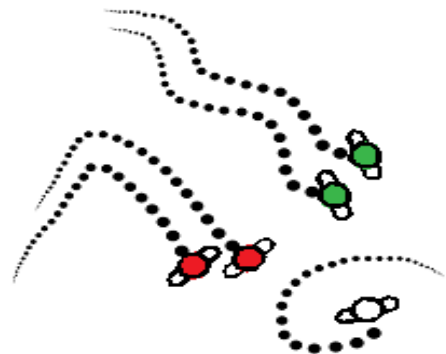
حال اگر تشابه تغییرات حرکت فرد به تغییرات حرکت خود فرد و تغییرات حرکت شخص دیگر عددی نزدیک باشد، ویژگی تشابه تغییرات حرکت افراد a و b به سمت صفر میل می‌کند. در رابطه زیر از تابع احتمال فیشر^{۱۵} استفاده شده است.

$$d_{ca}^k(a, b) = \max_{S \in \{S_{b \rightarrow a}, S_{a \rightarrow b}\}} \int_0^S F(x \setminus n, k - 2n - 1) dx \quad (9)$$

خروجی رابطه (۹)، ویژگی تشابه تغییرات حرکت افراد a و b در فریم‌های متوالی است.

۴-۳- ویژگی تشابه شکل مسیر طی شده توسط افراد

در این پژوهش از الگوریتم اندازه‌گیری شباهت بین دو دنباله زمانی^{۱۶} استفاده شده است. در شکل ۶ اهمیت ویژگی تشابه شکل مسیر طی شده توسط افراد در تشخیص گروه‌های اجتماعی نشان داده شده است.



شکل ۶: ویژگی تشابه شکل مسیر طی شده توسط افراد [۸].

برای محاسبه تابع هزینه انباشته $\gamma_{ab}(i, j)$ نیاز به محاسبه ماتریس فاصله اقلیدسی مسیر فرد a در لحظه اول با مسیر فرد b در لحظه اول یعنی D_{ab}^{11} است. D_{ab}^{ij} ماتریس فاصله اقلیدسی مسیر فرد a در لحظه i با مسیر فرد b در لحظه j است. ابعاد ماتریس فاصله اقلیدسی $f \times g$ است که f و g به ترتیب نشان دهنده مسیر طی شده

Student003 به همگرایی رسید. این عملیات ماکزیمم‌سازی فراخوانی با استفاده از PSO پانزده بار اجرا شده است و میانگین وزن‌های نهایی به صورت زیر بدست آمد. وزن نهایی ویژگی فاصله افراد برابر ۰/۴۵، وزن نهایی ویژگی شباهت مسیر طی شده برابر ۰/۸۰، وزن نهایی ویژگی شباهت دو دنباله حرکت برابر ۰/۶۳ و وزن نهایی ویژگی نقشه گرمایی برابر ۰/۱۱ می‌باشد.

بنابراین با توجه به اعدادی که در مرحله انتخاب ویژگی بدست آمد و کم بودن ضریب ویژگی نقشه گرمایی، روش پیشنهادی براساس حذف ویژگی نقشه گرمایی ارائه شده است. یعنی در ادامه تشخیص گروه‌های اجتماعی با استفاده از SVM با سه ویژگی فاصله افراد، شباهت مسیر طی شده و شباهت دو دنباله حرکت انجام شده است.

برای آموزش و آزمایش در پردازش ویدئو، از نرم افزار متلب و سروری با ۶۴ هسته موازی استفاده شده است. نتایج مقایسه روش پیشنهادی با سایر روش‌ها در جداول ۳، ۴ و ۵ آورده شده است. سایر نتایج در سیستم با هفت هسته موازی انجام شده است. در جدول ۳، مقایسه نتایج در پایگاه داده Student003 و در جدول ۴، مقایسه نتایج در پایگاه داده GVEII نشان شده است.

جدول ۳: مقایسه نتایج در پایگاه داده Student003

دقت	فراخوانی	
۸۰/۰±۳۱/۴۱	۸۰/۰±۲۷/۲۸	روش پیشنهادی
۸۲/۰±۳۶/۳۱	۷۴/۰±۱۶/۲۱	روش [۸]
۷۷/۲۸	۷۳/۶۹	روش [۱۵]
۵۶/۷۶	۷۶/۰۲	روش [۱۳]
۷۱/۱۲	۷۸/۷۶	روش [۱۶]
۴۰/۴۸	۴۸/۶۳	روش [۹]

جدول ۴: مقایسه نتایج در پایگاه داده GVEII

دقت	فراخوانی	
۸۲/۰±۹۱/۲۲	۸۱/۰±۳۲/۰۹	روش پیشنهادی
۸۳/۰±۹۶/۲۶	۸۰/۰±۷۶/۱۵	روش [۸]
۷۹/۴۵	۷۶/۳۳	روش [۱۷]
۷۶/۴۲	۷۸/۹۳	روش [۲۰]
۶۹/۷۳	۶۲/۴۸	روش [۱۹]
۷۴/۸۱	۵۷/۲۷	روش [۱۸]

نتایج در جدول ۳ و ۴ برای روش پیشنهادی و مرجع [۸] از بین پنج مرحله اجرا میانگین گرفته شده و واریانس کنار میانگین نوشته شده است. شرایط در نظر گرفته شده مانند سایر مراجع به خصوص مرجع [۸] می‌باشد یعنی از هر فیلم ۱۰۰ ثانیه به صورت تصادفی انتخاب شده است و بعد از ۵ بار اجرا میانگین معیارها بیان شده است. در جدول ۵، زمان مورد نیاز در آموزش و آزمایش روش پیشنهادی و مرجع [۸] که از درصد دقت و فراخوانی بالاتری برخوردار بودند، در شرایط یکسان با هم مقایسه شده‌اند.

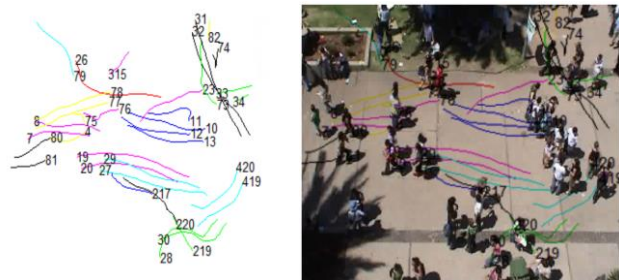
جدول مربوط به تعداد افراد در دو پایگاه داده آورده شده است. فاصله متوسط افراد در گروه‌ها در دو پایگاه داده متفاوت است و این نیاز به آموزش را نشان می‌دهد. بنابراین به منظور بررسی عملکرد روش پیشنهادی از پایگاه داده‌های Student003 [۸] و GVEII [۲۲] استفاده شده است. در جدول ۲ ویژگی‌های این دو پایگاه داده آورده شده‌اند.

جدول ۲: مقایسه دو پایگاه داده

v	p	G	d ₁	d ₂	
۲۰	۴۰۶	۱۰۸	۰/۴۱	۰/۷۰	Student 003
۳۰	۶۳۰	۲۰۷	۰/۷۷	۱/۶۶	GVEII

در این جدول متغیرهای v و p و g و d₁ و d₂ به ترتیب نشان دهنده تعداد ویدئو، تعداد افراد، تعداد گروه‌ها، فاصله حداقل افراد در گروه (برحسب متر)، فاصله حداکثر افراد در گروه (بر حسب متر) می‌باشند.

این پایگاه داده‌ها شامل داده‌های مسیر افراد و شماره افراد می‌باشند. در شکل ۷، مسیر و شماره افراد در دنباله‌ای از فریم‌ها نشان داده شده است.



شکل ۷: مسیر حرکت افراد و شماره افراد در دنباله‌ای از فریم‌ها.

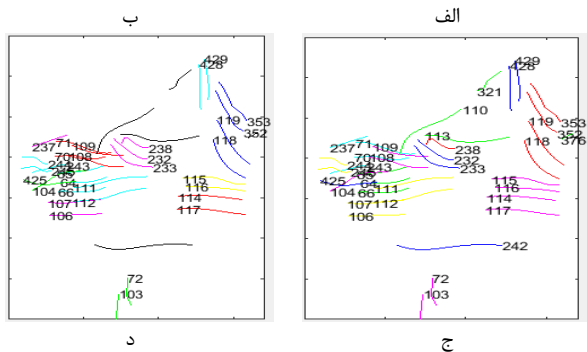
برای مقایسه بهتر نتایج، از پارامترهای دقت^{۱۷} و فراخوانی^{۱۸} استفاده شده است. دقت حاصل تقسیم تعداد گروه‌هایی که به درستی تشخیص داده شده به تعداد همه گروه‌های تشخیص داده شده است و فراخوانی حاصل تقسیم تعداد گروه‌هایی که به درستی تشخیص داده شده به تعداد گروه‌های درست پایگاه داده است.

۲-۴- انتخاب ویژگی با استفاده از الگوریتم بهینه‌سازی اجتماع ذرات^{۱۹} (PSO)

با استفاده از الگوریتم هوشمند PSO، به هر کدام از چهار ویژگی فاصله افراد، شباهت مسیر طی شده، شباهت دو دنباله حرکت و نقشه گرمایی، یک وزن اولیه بین صفر و یک به صورت تصادفی داده شده است. تابع هزینه یا هدف، ماکزیمم کردن فراخوانی در نظر گرفته شده است. با کمک PSO و با جمعیت ده ذره و در هر ذره به SVM اجازه داده شده که ۱۰۰ مرحله وزن‌های خود را بروزسانی کند. در نهایت PSO بعد از بیست تکرار برای پایگاه داده

دقت مرجع [۸] به این معنی است که تعداد گروه‌های کمتری تشخیص داده شده است اما فراخوانی این تعداد بالاتر از مرجع [۸] می‌باشد. دلیل این فرایند حذف ویژگی نقشه گرمایی است. به عبارت دیگر، با حذف شدن ویژگی نقشه گرمایی، هم سرعت روش پیشنهادی افزایش و هم درستی گروه‌های پیدا شده نیز افزایش یافته است. به عنوان مثال از بیست گروه در یک ویدئو، مرجع [۸] هفده گروه و روش پیشنهادی شانزده گروه تشخیص داده‌اند که قابلیت اطمینان و درستی شانزده گروه بالاتر است.

در شکل‌های ۸ و ۹، نتایج روش پیشنهادی در دو ویدئو از بیست ویدئوی پایگاه داده Student003 و شکل‌های ۱۰ و ۱۱ در دو ویدئو از سی ویدئوی پایگاه داده GVEII نشان داده شده است. در این شکل‌ها، شکل (الف) فریم ویدئو ورودی، شکل (ب) نتایج تشخیص گروه‌های اجتماعی در این فریم براساس ویژگی‌های وابسته به فریم‌های قبلی، شکل (ج) نتایج تشخیص گروه‌های اجتماعی در این فریم براساس ویژگی‌های وابسته به فریم‌های قبلی می‌باشد که دو شخص با یک رنگ به معنی هم گروه بودن دو شخص است، شکل (د) گروه‌های اجتماعی که در پایگاه داده مربوطه، داده شده و دو شخص با یک رنگ به معنی هم گروه بودن آن دو شخص است.



شکل ۸: الف: فریم ویدئو ورودی، ب: فریم ویدئو خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی، ج: گروه‌های اجتماعی که در پایگاه داده Student003 داده شده، د: خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی.

جدول ۵: مقایسه زمان مورد نیاز برای تمام ویدئوها در آموزش و آزمایش برای دو پایگاه داده

Student003	GVEII	
۱۲ ساعت	۱۶ ساعت	زمان مورد نیاز در روش پیشنهادی
۶۴ ساعت	۷۸ ساعت	زمان مورد نیاز در روش [۸]

همانطور که در جدول ۵، مشاهده می‌شود، روش پیشنهادی به زمان بسیار کمتری نسبت به مرجع [۸] برای آموزش و آزمایش نیاز دارد. بنابراین بازدهی روش پیشنهادی با در نظر گرفتن معیارهای دقت، فراخوانی و زمان از سایر روش‌های ذکر شده بهتر می‌باشد. برای بررسی بیشتر روش پیشنهادی در جدول ۶ درصد بخش‌های آموزش و آزمایش تغییر داده شده است و عملکرد روش پیشنهادی با استفاده از معیارهای دقت و فراخوانی ارزیابی شده است.

جدول ۶: مقایسه عملکرد SVM در درصد‌های مختلف آموزش و آزمایش برای دو پایگاه داده GVEII و Student003

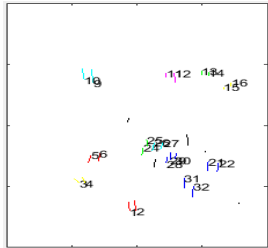
Student003	GVEII	پایگاه داده	درصد آموزش و آزمایش
دقت فراخوانی	دقت فراخوانی	دقت فراخوانی	دقت فراخوانی
۷۴/۶۳	۷۵/۳۶	۷۴/۷۴	۷۶/۵۵
±۰/۳۲	±۰/۲۴	±۰/۳۳	±۰/۲۸
۷۰/۵۱	۷۲/۱۲	۷۱/۱۸	۷۲/۸۵
±۰/۳۹	±۰/۴۱	±۰/۲۹	±۰/۳۱
۶۶/۲۳	۶۷/۷۸	۶۶/۸۹	۶۸/۵۷
±۰/۴۲	±۰/۳۷	±۰/۴۵	±۰/۴۷

در جدول ۶، ۸۰ درصد آموزش از ابتدای فیلم‌ها و ۲۰ درصد انتهایی فیلم برای آزمایش انتخاب شده است. در حالت سوم، ۶۰ درصد آموزش از ابتدای فیلم‌ها و ۴۰ درصد انتهایی فیلم برای آزمایش انتخاب شده است. نتایج از بین پنج مرحله اجرا میانگین گرفته شده و واریانس پایین میانگین نوشته شده است.

همانطور که در بخش‌های قبل ذکر شد، در مرجع [۸] از ویژگی‌های فاصله افراد، شباهت مسیر طی شده و شباهت دو دنباله حرکت و نقشه گرمایی بین هر دو نفر استفاده شده و محاسبه ویژگی نقشه گرمایی بین دو نفر در مرحله استخراج ویژگی زمان زیادی نیاز دارد. در روش پیشنهادی ویژگی نقشه گرمایی حذف شده است که هم باعث کاهش زمان محاسباتی برای تشخیص گروه‌های اجتماعی شده است و هم عملکرد روش پیشنهادی بهبود یافته است. کمتر بودن دقت روش پیشنهادی از



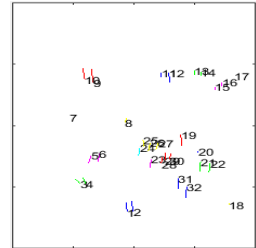
ب



د



الف

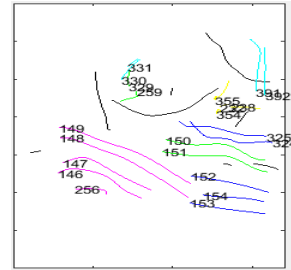


ج

شکل ۱۱: الف: فریم ویدئو ورودی، ب: فریم ویدئو خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی، ج: گروه‌های اجتماعی که در پایگاه داده GVEII داده شده، د: خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی.



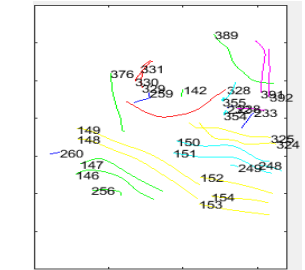
ب



د



الف



ج

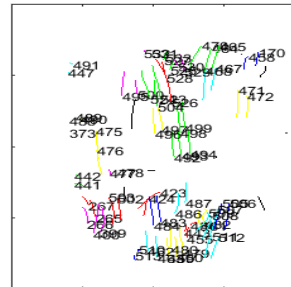
شکل ۹: الف: فریم ویدئو ورودی، ب: فریم ویدئو خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی، ج: گروه‌های اجتماعی که در پایگاه داده Student003 داده شده، د: خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی.

۵- نتیجه‌گیری

در این پژوهش با استفاده از روش انتخاب ویژگی PSO ویژگی‌های فاصله، تشابه تغییرات حرکت افراد در فریم‌های متوالی و تشابه شکل مسیر طی شده توسط افراد انتخاب و استفاده شدند و روش پیشنهادی با استفاده از SVM به نتایج قابل قبولی در مقایسه با کارهای گذشته دست یافته است. برای سیستم‌هایی مانند ربات‌های انسان‌نما، قابلیت پیاده‌سازی و نیاز به زمان کمتر، اهمیت بالایی دارد. با حذف ویژگی نقشه گرمایی از مجموع ویژگی‌های مرجع [۸] و کاهش اندک دقت، فراخوانی در تشخیص افزایش یافته است. زمان مورد نیاز و در نتیجه پیچیدگی سیستم به شدت کاهش یافته است. برای بهبود تشخیص گروه‌های اجتماعی باید از ویژگی‌هایی جدیدی استفاده شود که دقت و فراخوانی روش را افزایش داده و همچنین محاسبات سنگین به روش تحمیل نشود.



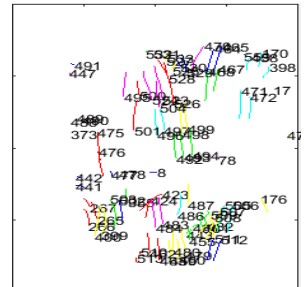
ب



د



الف



ج

شکل ۱۰: الف: فریم ویدئو ورودی، ب: فریم ویدئو خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی، ج: گروه‌های اجتماعی که در پایگاه داده GVEII داده شده، د: خروجی تشخیص گروه‌های اجتماعی در روش پیشنهادی، در ج و د.

مراجع

- [۱] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 935-942, 2009.
- [۲] M. Moussaïd, N. Perozo, S. Garnier, D. Helbing, and G. Theraulaz, "The walking behaviour of pedestrian social groups and its impact on crowd dynamics," *PLoS one*, vol. 5, p. e10047, 2010.
- [۳] M. Van Zomeren, T. Postmes, and R. Spears, "Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives," *Psychological bulletin*, vol. 134, p. 504, 2008.

and first results," *Pattern Recognition Letters*, vol. 44, pp. 16-29, 2014.

پاورقی‌ها:

- ¹ Group-based
- ² Individual-group joint
- ³ Individual-based
- ⁴ Time overlap information
- ⁵ Support Vector Machine (SVM)
- ⁶ Soft segmentation
- ⁷ Motion causality
- ⁸ Trajectory shape
- ⁹ Proximity
- ¹⁰ Velocity
- ¹¹ Transitivity
- ¹² Density invariance
- ¹³ Hierarchical coherence
- ¹⁴ Infinite mixture model
- ¹⁵ Fisher-Snedecor
- ¹⁶ Dynamic Time Warping (DTW)
- ¹⁷ Precision
- ¹⁸ Recall
- ¹⁹ Particle Swarm Optimization

- [۴] M. Luber, J. A. Stork, G. D. Tipaldi, and K. O. Arras, "People tracking with human motion predictions from social forces," in *Robotics and Automation (ICRA), IEEE International Conference on*, pp. 464-469, 2010.
- [۵] T. Iqbal, M. Moosaei, and L. D. Riek, "Tempo adaptation and anticipation methods for human-robot teams," in *RSS, Planning HRI: Shared Autonomy Collab. Robot. Workshop*, 2016.
- [۶] T. Iqbal, S. Rack, and L. D. Riek, "Movement coordination in human-robot teams: a dynamical systems approach," *IEEE Transactions on Robotics*, vol. 32, pp. 909-919, 2016.
- [۷] M. Vázquez, A. Steinfeld, and S. E. Hudson, "Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pp. 3010-3017, 2015.
- [۸] F. Solera, S. Calderara, and R. Cucchiara, "Socially constrained structural learning for groups detection in crowd," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, pp. 995-1008, 2016.
- [۹] J. Shao, C. Change Loy, and X. Wang, "Scene-independent group profiling in crowd," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2219-2226, 2014.
- [۱۰] S. K. Pang, J. Li, and S. J. Godsill, "Detection and tracking of coordinated groups," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 472-502, 2011.
- [۱۱] M. Rodriguez, I. Laptev, J. Sivic, and J.-Y. Audibert, "Density-aware person detection and tracking in crowds," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 2423-2430, 2011.
- [۱۲] M. Cristani, L. Bazzani, G. Paggetti, A. Fossati, D. Tosato, A. Del Bue, *et al.*, "Social interaction discovery by statistical analysis of F-formations," in *BMVC*, p. 4, 2011.
- [۱۳] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and where are you going?," in *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 1345-1352, 2011.
- [۱۴] M.-C. Chang, N. Krahnstoever, and W. Ge, "Probabilistic group-level motion analysis and scenario recognition," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 747-75, 2011.
- [۱۵] W. Ge, R. T. Collins, and R. B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, pp. 1003-1016, 2012.
- [۱۶] M. Zanotto, L. Bazzani, M. Cristani, and V. Murino, "Online bayesian nonparametrics for group detection," in *Proc. of BMVC*, 2012.
- [۱۷] J. Shao, C. C. Loy, and X. Wang, "Learning scene-independent group descriptors for crowd understanding," *IEEE transactions on circuits and systems for video technology*, vol. 27, pp. 1290-1303, 2017.
- [۱۸] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1-6, 2007.
- [۱۹] R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," in *European conference on computer vision*, pp. 439-452, 2010.
- [۲۰] B. Zhou, X. Tang, and X. Wang, "Coherent filtering: detecting coherent motions from crowd clutters," in *Computer Vision-ECCV 2012*, ed: Springer, pp. 857-871, 2012.
- [۲۱] S. D. Reicher, R. Spears, and T. Postmes, "A social identity model of deindividuation phenomena," *European review of social psychology*, vol. 6, pp. 161-198, 1995.
- [۲۲] S. Bandini, A. Gorrini, and G. Vizzari, "Towards an integrated approach to crowd analysis and crowd synthesis: A case study