



Human identification and tracking from images

Mohammad Nazim Jafari¹, Mohammad Rahim Tahiri¹

Abstract

The method used to identify and track humans in this research is based on edge detection. After subtracting the foreground from the background, using the edging method, the edges of the foreground areas are obtained; Then, the identification algorithm identifies people from top to bottom and removes them from the image. Finally, by using the Gaussian probability distribution, people are tracked according to their position in the previous form. This algorithm works well in crowded places and does not suffer from overlaps. Considering that stereo images are used in this study, the distance of people to the camera can also be determined.

Keywords: human identification, human tracking, Gaussian distribution, edge detection.



فصلنامه علمی - تحقیقی علوم فنی

پوهنتون خاتم النبیین (ص)

سال دوم ، شماره دوم، بهار ۱۴۰۳

صفحه ۸۷ - ۹۸



شناسایی و ردیابی انسان از روی تصاویر

محمدناظم جعفری^۱، محمدرحیم طاهری^۲

چکیده

روش استفاده شده برای شناسایی و ردیابی انسان در این تحقیق براساس لبه‌گیری عمل می‌کند. پس از تفریق پیش‌زمینه از پس‌زمینه، با استفاده از روش لبه‌گیری، لبه‌های نواحی پیش‌زمینه به دست می‌آیند؛ سپس الگوریتم شناسایی، اشخاص را از بالا به پایین شناسایی و از تصویر حذف می‌کند. در نهایت با استفاده از توزیع احتمال گوسی، افراد با توجه به جایگاه‌شان در فرم قبلی، ردیابی می‌شوند. این الگوریتم در مکان‌های شلوغ به خوبی عمل می‌کند و در مقابل هم‌پوشانی‌ها دچار مشکل نمی‌شود. با توجه به اینکه از تصاویر استریویی در این مطالعه استفاده می‌شود، فاصله افراد نسبت به دوربین را نیز می‌توان تعیین کرد. واژگان کلیدی: شناسایی انسان، ردیابی انسان، توزیع گوسی، لبه‌گیری.

^۱ دانشگاه کاتب، کابل، افغانستان. m.n.jafari@kateb.edu.af

^۲ دانشگاه کاتب، کابل، افغانستان. mohammadrahimtaheri@gmail.com

از دیرباز تاکنون، شناسایی و به‌دنبال آن ردیابی انسان از زمینه‌های تحقیق در حوزه بینایی ماشین بوده است. دلیل اهمیت این موضوع، کاربردهای فراوان آن در سامانه‌های امنیتی و کنترل رفت‌وآمد مردم در مکان‌های عمومی، نظیر ادارات، ایستگاه‌های قطار و اتوبوس و... است.

کارهایی که تاکنون در این زمینه صورت گرفته‌اند، به دلیل هم‌پوشانی جزئی یا کامل، تغییرات نوری محیط، وجود سایه و فاصله نسبت به دوربین، معمولاً ضعف‌هایی داشته‌اند. یکی از راه‌حل‌های برطرف کردن ضعف‌های گذشته، استفاده از چند دوربین، به خصوص برای هم‌پوشانی، است؛ در این حالت، نقاط صحنه از چند زاویه دیده می‌شوند.

در این پژوهش، از تصاویر دوربین‌های استریو استفاده شده است. دوربین‌های استریو، که ساده‌ترین شکل این سامانه‌ها هستند، تصویرهایی می‌دهند که شدت روشنایی تصویر با توجه به فاصله شیء تا دوربین متغیر است؛ یعنی اشیای نزدیک‌تر، روشن‌تر هستند. معمولاً از دوربین‌های استریو برای شناسایی و ردیابی انسان به صورت «نقشه کف»، «دید دوربین» و «ترکیب دید دوربین با نقشه کف» استفاده می‌شود.

احتمالاً برای اولین بار، بایمر در سال ۲۰۰۰ از بینایی استریو به صورت از بالا به پایین، یعنی «نقشه کف»، استفاده کرده است [۴]؛ او از دوربین استریو برای شناسایی و شمردن افراد استفاده کرد. بایمر به جای بررسی تصاویر استریوی محل دید دوربین، از تلفیق نقشه محل قرار گرفتن افراد با اورتوگرافیگ محیط (تصویری که به طور عمودی از بالای سر گرفته می‌شود) تصویری به دست آورد که آن را «نقشه کف» نامید. او از مدل مخلوط گوسی و فیلتر کالمن جهت ردیابی افراد استفاده کرد که در این الگوریتم فقط از اطلاعات محل قرار گرفتن افراد در «نقشه کف» برای ردیابی استفاده شد.

موناخ-سالیناز الگوریتم شناسایی خاصی برای شناسایی و ردیابی چند نفر با استفاده از «نقشه کف» ارائه داد [۵]. او علاوه بر نقشه کف و ارتفاع، از «نقشه رنگ» نیز برای شناسایی و ردیابی افراد استفاده کرد؛ همچنین، از فیلترینگ جزئی برای ردیابی استفاده کرد. موناخ-سالیناز و همکارانش بعداً الگوریتم توسعه داده شده‌ای نسبت به کار قبلی‌شان معرفی کردند [۶]. آن‌ها نقشه کف جدیدی، که «نقشه اطمینان» نامیده می‌شد، معرفی کردند که اطلاعات استریوی هر پیکسل را ثبت می‌کرد. نقشه اطمینان برای بهبود شناسایی اشیا (کدام شیء نسبت به دوربین نزدیک است) استفاده می‌شد؛ در این حالت، شناسایی و ردیابی به طور مستقیم روی تصویر صورت می‌گیرد و با استفاده از اطلاعات مربوط به عمق اشیا می‌توان به مسائلی مانند هم‌پوشانی جزئی و ابهام در ردیابی رسیدگی کرد.

ناندا و همکارانش الگوریتمی ارائه دادند که براساس اطلاعات عمق (استریو) بالاتنه و دست افراد را شناسایی می‌کند [۷]. مشکلی الگوریتم آن‌ها استفاده از دوربین استریوی





بود که فقط افراد ثابت را در فاصله کم شناسایی می‌کرد؛ اما ایده آن‌ها هنوز مورد توجه محققان است.

مونا-ز-سالیناز و همکارانش الگوریتم شناسایی انسانی براساس «دید دوربین» و فناوری استریو ارائه دادند [۸]. این الگوریتم برای کاربردهای رباتیک، که دوربین درحال حرکت است، پیشنهاد شد. در این الگوریتم نیازی به مدل پس‌زمینه نبود. آن‌ها با استفاده از روش فیلترینگ جزئی، توأم با مدلی از انسان، که شامل دو بیضی برای سر و نیم‌تنه بالایی بود، افراد را شناسایی می‌کردند. آن‌ها در این روش از فیلترینگ جزئی برای شناسایی افراد به‌صورت مجزا استفاده می‌کردند؛ یعنی افراد را جداگانه پردازش می‌کردند.

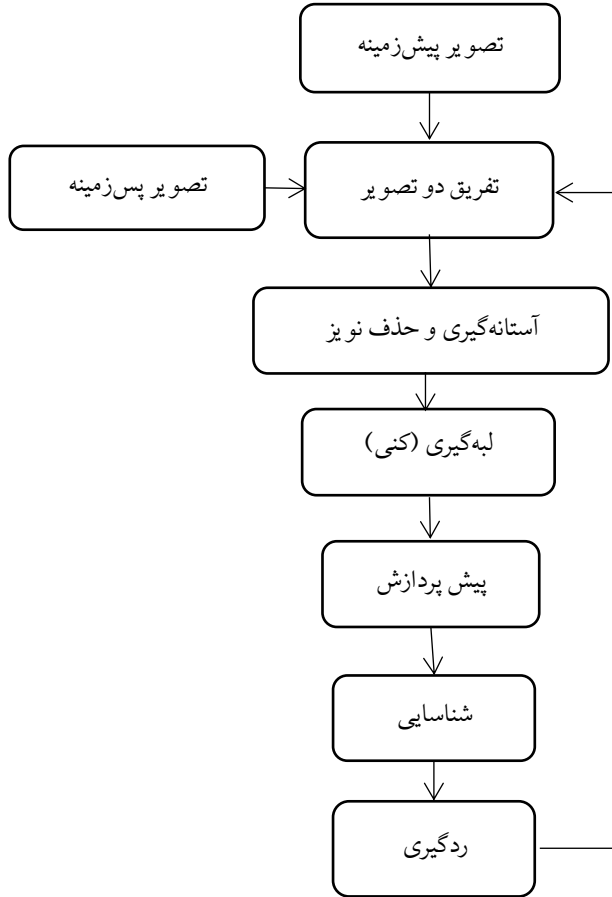
سونگمین جیا و همکارانش ردیابی را با استفاده از فضای دید دوربین انجام داده بودند؛ این الگوریتم، که برای کاربردهای رباتیک طراحی شده بود، مناسب فضای بیرون و داخل ساختمان بود [۲]. آن‌ها به جای استفاده از یکی از فضاهای ردگیری، با استفاده از هر دو فضای ردگیری، یعنی دید دوربین و نقشه کف، توانستند نتایج بهتری در کاربردها به دست آورند. مونا-ز-سالیناز و همکارانش یک الگوریتم استریو، که کاربرد رابط بین ماشین و انسان داشت، ارائه دادند [۹]. آن‌ها همچنین الگوریتم استریویی که بدون مدل پس‌زمینه عمل می‌کرد، ارائه دادند [۸]؛ این الگوریتم برای کاربردهای رباتیک، که دوربین‌ها در آن درحال حرکت هستند، استفاده می‌شد. این مطالعه به جای ردیابی روی نقشه کف، مانند کارهای قبلی‌شان، با استفاده از فیلترینگ جزئی صورت می‌گرفت. یانگ زو در مطالعه خود از ترکیب دید دوربین با نقشه کف به همراه مدل بندی پس‌زمینه فعال استفاده کرد [۳].

۲. روش تحقیق

در این پژوهش، شناسایی و ردیابی انسان با استفاده از روش جدیدی، براساس لبه‌گیری، انجام می‌شود. به دلیل استفاده از تصاویر استریویی، نتایج الگوریتم لبه‌گیری به‌طور کامل واضح و مطلوب هستند که بر مشکلاتی نظیر انسداد غلبه می‌کنند؛ همچنین اطلاعات عمقی افراد، که در کاربردهای مختلف استفاده می‌شود، قابل دستیابی هستند. سامانه‌ی طراحی شده می‌تواند برای شناسایی و ردیابی افراد در محیط‌های نسبتاً شلوغ استفاده شود.

تصاویر استفاده‌شده در این تحقیق از وبسایت [۱] استخراج شده‌اند و شبیه‌سازی‌ها با matlab 2018 صورت گرفته‌اند.

بخش بندی کلی سامانه شناسایی و ردگیری به صورت زیر است که هر بخش به ترتیب بررسی خواهد شد.



شکل ۱: نمودار سامانه اصلی

۳. تفریق پیش زمینه از پس زمینه

مانند بیشتر روش های شناسایی انسان، ابتدا تصویر پیش زمینه را از تصویر پس زمینه تفریق می کنیم؛ اشیای پیش زمینه (افراد) در تصویر حاصل، به دلیل وجود نویز و تغییرات نوری، به طور کامل از پس زمینه جدا نیست؛ یعنی پیکسل های پس زمینه، صفر نیستند. با استفاده از حد آستانه (در اینجا ۱۱ در نظر گرفته شده است) می توان پیکسل هایی که سطح روشنایی کمتر از حد آستانه دارند، صفر و بقیه را خودشان در نظر گرفت؛ نتیجه حاصل از این عملیات ها در شکل ۲ قابل مشاهده هستند. حد آستانه گیری لازم است؛ زیرا بیشتر پیکسل ها پس از تفریق به طور کامل صفر نمی شوند.





(ب)



(الف)



(ج)



(د)

شکل ۲: الف. تصویر پس‌زمینه؛ ب. تصویر پیش‌زمینه؛ ج. تصویر تفریق پیش‌زمینه از پس‌زمینه؛ د. تصویر حاصل از آستانه‌گیری

۴. حذف نویز و پیش پردازش

تصویری که در مرحله گذشته به دست آمد، برای مراحل بعدی قابل استفاده نیست و باید عملیات‌هایی جهت حذف نویز و حفره‌های احتمالی به‌وجودآمده، انجام شوند؛ این عملیات‌ها با استفاده از عمل باز و بسته کردن [۱۳] انجام می‌شوند. هرکدام از اعمال باز و بسته کردن، ترکیبی از دو عمل گسترش و سایش هستند.

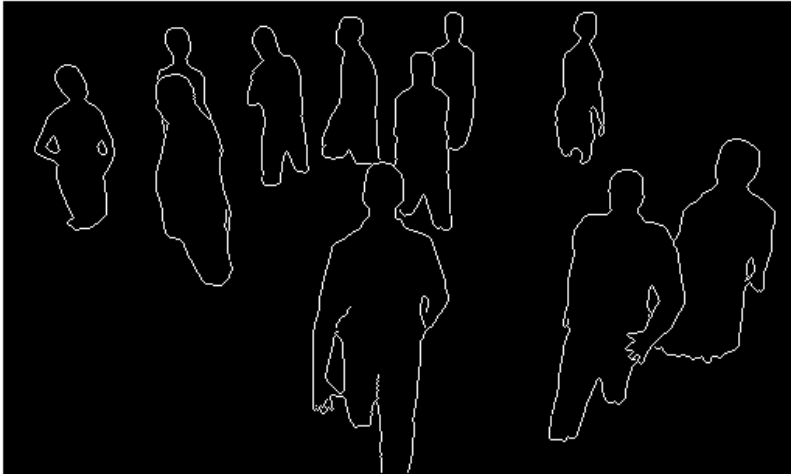
۵. لبه‌گیری

برای تشخیص لبه از آشکارساز لبه [۱۳] استفاده شده است. داده‌هایی که در آشکارساز لبه‌کنی استفاده شده‌اند، به‌صورت زیر هستند:

- انحراف معیار استفاده‌شده در فیلتر گوسی: $\sigma = 0.3$ ؛
- حد آستانه بالایی = 0.1 ؛
- حد آستانه پایینی = 0.04 .



با توجه به نتیجه (شکل ۳)، نواحی (افراد) به‌خوبی شناسایی و از یکدیگر مجزا هستند؛ بنابراین، شناسایی نواحی به‌خوبی صورت می‌گیرد و به مرحله بعد (شناسایی) کمک می‌کند تا نتیجه بهتری حاصل شود.



شکل ۳: آشکارسازی لبه‌ها توسط لبه‌گیرکنی

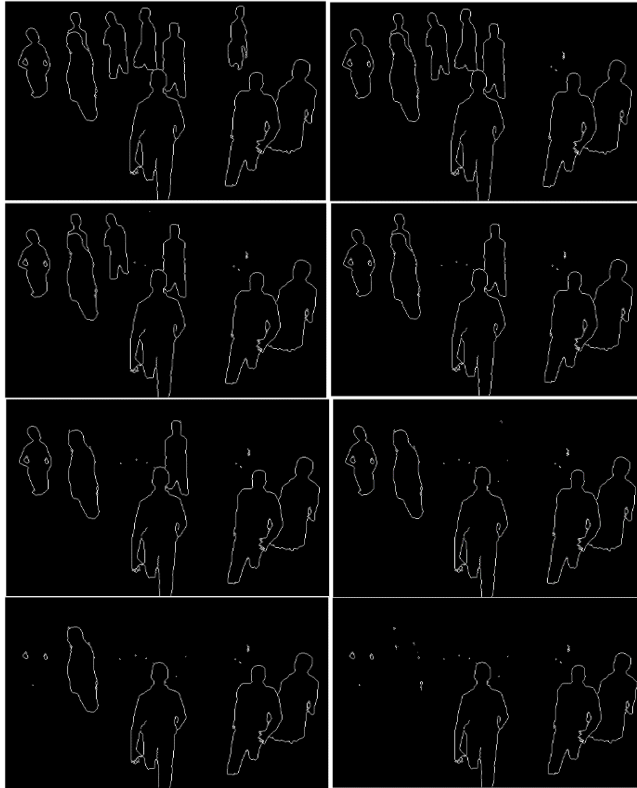
تصویر حاصل را با عملیات باز و بسته کردن، ترمیم می‌کنیم؛ یعنی حفره‌های مرزی را از بین می‌بریم. اگر در مرزها گسستگی باشد، در مرحله شناسایی - که مرزها دنبال می‌شوند - نتیجه مطلوب حاصل نمی‌شود؛ یعنی باعث عدم شناسایی یا حذف نواحی، که مرز ناحیه مربوطه نیست، می‌شود. شناسایی اشتباه یا عدم شناسایی، روند کلی الگوریتم را دچار مشکل می‌کند. پس از هر مرحله شناسایی، این مرحله تکرار می‌شود تا در مرزها گسستگی به وجود نیاید. نتیجه این عملیات در شکل ۴ دیده می‌شود. تصویر حاصل برای مرحله بعد (شناسایی) آماده است.

۶. شناسایی

الگوریتم شناسایی به این طریق عمل می‌کند:

- از بالا و سمت چپ، شروع به اسکن می‌کند؛
- اولین پیکسل غیر صفر را پیدا می‌کند؛
- در صورت اتصال چند پیکسل غیر صفر به یکدیگر وارد مرحله بعد می‌شود؛
- از سمت چپ و سمت راست، مرزها را دنبال می‌کند؛
- در صورت اتصال ناحیه و غیر صفر بودن مرکز ناحیه در تصویر اصلی، فرد، شناسایی و از تصویر حذف می‌شود و در غیر این صورت فقط ناحیه حذف می‌شود؛
- این روند تا پایان تصویر ادامه می‌یابد.

مزیت این الگوریتم، حذف نواحی مربوط به فرد مورد نظر و حذف نکردن نواحی مشترک است؛ همچنین نقاط و خطوطی که ممکن است حین اجرای الگوریتم باقی مانده باشند، بدون اینکه فرد جدیدی به شمار آید، از تصویر حذف می‌شوند (شکل ۴). مختصات و شدت روشنایی مرکز ثقل و مختصات بالاترین نقطه هر ناحیه (فرد) در بردار X ذخیره می‌شود.

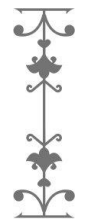


شکل ۴: شناسایی

۷. ردیابی

ردیابی افراد رابطه مستقیمی با مرحله شناسایی دارد. اگر شناسایی دقیقی صورت گیرد، ردیابی با مشکل کمتری مواجه می‌شود. الگوریتم ردیابی استفاده شده براساس توزیع احتمال گوسی عمل می‌کند که روند کار آن به شرح زیر است:

- ابتدا بردار X ، که در مرحله شناسایی فرم اول یا فرم قبلی به دست می‌آید، براساس شدت روشنایی مراکز ثقل افراد مرتب می‌شود؛
- شناسایی در فرم جدید صورت می‌گیرد و مختصات افراد در بردار A ذخیره می‌شود؛
- با استفاده از توزیع احتمال گوسی $N(x, 0, \sigma^2)$ ، احتمال قرار گرفتن هر فرد در مکان x را با توجه به بردار X (محل افراد در مرحله قبل) محاسبه و در بردار D که سطرهای

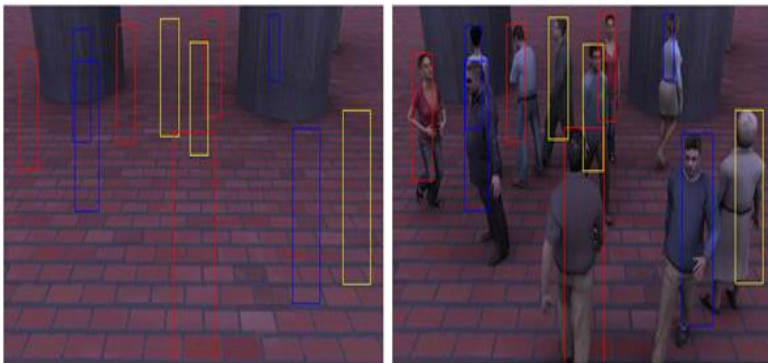


آن برابر با تعداد افراد شناسایی شده در مرحله قبل و ستون‌های آن برابر تعداد افراد شناسایی شده در فرم جدید است- قرار می‌دهد؛

- پس از تکمیل بردار D ، برای هر ستون این بردار (به ترتیب) بررسی می‌شود که مقدار کدام سطر بزرگتر است. در صورتی که شدت روشنایی این نقطه با شدت روشنایی مرکز ثقل قبلی اش متناسب باشد (اختلاف زیادی نداشته باشد)، انتخاب و احتمال‌های متناظر این نقطه در سطرهای بعدی صفر می‌شود؛
- مختصات متناظر A در X جایگزین می‌شود؛
- این روند برای تمام ستون‌های D انجام می‌شود؛
- اگر ستونی وجود داشت که برای آن سطری (مکان جدید) پیدا نشده بود، محل این فرد همان مکان قبلی اش در نظر گرفته می‌شود و اگر این عمل چندبار تکرار شد، این فرد از فهرست حذف می‌شود.
- در صورتی که سطری وجود داشت که ستونی از آن غیر صفر بود، این فرد، جدید است و هنوز شناسایی نشده که به فهرست افراد شناسایی شده افزوده می‌شود و مختصات آن در X اضافه می‌شود.

- این روند برای هر فرم جدید تکرار می‌شود.
- داده‌های زیر در این مرحله استفاده شده‌اند:

- انحراف معیار مورد استفاده در تابع گوسی: $\sigma = 6$.
- حد اکثر تغییر شدت روشنایی قابل قبول: ۵.



(الف)

(ب)

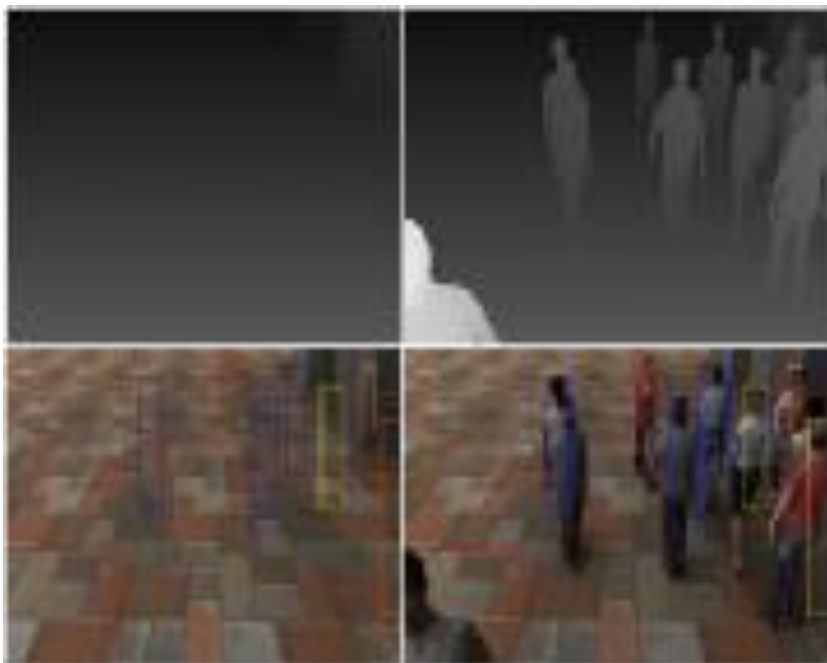
شکل ۵: الف. ردگیری فرم اول؛ ب. ردگیری فرم دوم

۸. ویژگی‌های الگوریتم

- ✓ افراد می‌توانند از هر طرف قاب وارد صحنه شوند؛
- ✓ این الگوریتم، مرزهایی را که فقط متعلق به فرد مورد ردیابی است، حذف می‌کند؛ یعنی نواحی بسته را خراب نمی‌کند؛
- ✓ تعداد افراد مورد ردیابی می‌توانند نسبتاً زیاد باشند (نسبت به الگوریتم‌های دیگر)؛
- ✓ هر زمان که فرد جدیدی وارد صحنه شده باشد، شناسایی می‌شود؛
- ✓ ردیابی به ترتیب انجام می‌شود؛ یعنی اول افرادی ردیابی می‌شوند که نسبت به دوربین نزدیک‌تر هستند.
- ✓ فاصله نسبی افراد نسبت به یکدیگر و دوربین مشخص است.
- ✓ مستطیل‌ها متناسب با فاصله افراد نسبت به دوربین هستند.

۹. نتیجه‌گیری

در این مقاله، روش جدیدی مبتنی بر لبه‌گیری برای شناسایی و ردگیری انسان ارائه شده است. این روش، نتایج خوبی در شناسایی در محیط‌های می‌دهد و مکان‌هایی که افراد هم‌پوشانی دارند، می‌دهد. الگوریتم‌های لبه‌گیری روی تصاویر استریو، به دلیل شدت روشنایی متفاوت هر ناحیه، نتیجه مطلوبی می‌دهند.



شکل ۶: ردگیری (یک نمونه دیگر)



- کارهایی می‌توانند در آینده صورت گیرند که می‌توان به موارد زیر اشاره نمود:
- استفاده از دوربین‌های با حوزه دید بیشتر (با پیشرفت تکنولوژی این کار محقق می‌شود).
 - استفاده از چند دوربین استریو که هم‌زمان پردازش می‌شوند و می‌توانند حوزه وسیعی را تحت پوشش قرار دهند.
 - ما شناسایی و ردیابی انسان را بررسی کردیم؛ این کار می‌تواند برای اشیاء، نظیر ماشین، حیوان یا محل و جابجایی اشیاء در کارخانه‌ها صورت گیرد.
- متأسفانه نتوانستیم تصاویر استریویی مناسبی، که به‌صورت متوالی و فریم‌به‌فریم باشند، دسترسی داشته باشیم؛ بنابراین سعی کردیم این نقیصه را با استفاده از تصویر پس‌زمینه جبران کنیم. برای این کار ابتدا روی تصویر اصلی، الگوریتم شناسایی را انجام دادیم؛ سپس با توجه به روند حرکتی افراد، مختصات آن‌ها را در فرم قبلی پیدا کردیم. با توجه به مختصات افراد، شناسایی و سپس ردگیری الگوریتم در فرم اصلی انجام شده است که همان‌طور که دیده می‌شود، شناسایی و ردگیری به‌خوبی انجام شده است.
- در پایان لازم است از وبسایت [۱] بابت تصاویر این مطالعه تشکر نمایم.



- [1] http://www.cdvp.dcu.ie/datasets/pedestrian_detection.
- [2] Songmin Jia, Liang Zhao, Xiuzhi Li, Wei Cui. Autonomous Robot Human Detecting and Tracking Based on Stereo Vision. Proceedings of the 2011 IEEE International Conference on Mechatronics and Automation August 7 - 10, Beijing, China.
- [3] Yong Zhao . Detection, Tracking, and Identification of People using Stereo Vision . Ph.D., Brown University, May 2011.
- [4] M. Harville. Stereo person tracking with adaptive plan-view templates of height and occupancy statistics. IVC, 22(2):127–142, February 2004.
- [5] Rafael Munoz-Salinas. A bayesian plan-view map based approach for multiple-person detection and tracking. Pattern Recogn., 41(12):3665–3676, 2008.
- [6] Rafael Munoz-Salinas, R. Medina-Carnicer, F. J. Madrid-Cuevas, and A. Carmona-Poyato. People detection and tracking with multiple stereo cameras using particle filters. J. Vis. Comun. Image Represent., 20(5):339–350, 2009.
- [7] Harsh Nanda and Kikuo Fujimura. Visual tracking using depth data. In CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 3, page 37, Washington, DC, USA, 2004. IEEE Computer Society.
- [8] Rafael Munoz-Salinas, Miguel Garc'ia-Silvente, and Rafael Medina Carnicer. Adaptive multimodal stereo people tracking without background modelling. J. Vis. Comun. Image Represent., 19(2):75–91, 2008.
- [9] Rafael Munoz-Salinas, Eugenio Aguirre, and Miguel Garc'ia-Silvente. People detection and tracking using stereo vision and color. Image Vision Comput., 25(6):995–1007, 2007.
- [10] Rafael Munoz-Salinas, Eugenio Aguirre, Miguel Garc'ia-Silvente, and Gonzalez Antonio. People Detection and Tracking Through Stereo Vision for Human-Robot Interaction. Springer Berlin / Heidelberg, 2005.
- [11] Hamid Laga, Sofiane Yous , Kunihiro Chihara . People Detection and Tracking with World- Z Map from a Single Stereo Camera .
- [12] Rafael c.gonzalez, Richard e.woods, steven l.eddins. Digital image processing using matlab. Second edition.
- [13] Rafael c.gonzalez, Richard e.woods, steven l.eddins. Digital image processing using matlab. Second edition.

