

استفاده از اطلاعات حرکتی و روش تفریق فریم‌های غیر متوالی در مدل‌سازی زمینه

سارا شریف‌زاده، غلامرضا اردشیر پهروستاقی^{*}، حسین میار نعیمی

چکیده

امروزه آشکارسازی و ردیابی اشیا متحرک در بسیاری از سیستمهای مبتنی بر بینایی ماشین برای اهداف کنترلی، نظارتی و امنیتی کاربرد دارد. یکی از روش‌های عمده در این زمینه، تفریق تصویر شامل اشیا متحرک از تصویر زمینه است. بنابراین، مدل‌سازی و تازه کردن زمینه، از بخش‌های اصلی این الگوریتم‌ها محسوب می‌شود. در این مقاله برخلاف اکثر روش‌های بکار رفته در این زمینه، با تکیه بر اطلاعات حرکتی استخراج شده از رشته فریم‌های تصویر، از رویکردی غیر آماری برای بازسازی تصویر زمینه استفاده می‌گردد. در این روش تعداد مناسبی از فریم‌های متوالی براساس سرعت و جهت حرکت اشیا متحرک انتخاب و دسته‌بندی می‌گردد. سپس با ترتیبی مناسب و به صورت غیر متوالی، فریم‌های هر زیر مجموعه از یکدیگر تفریق می‌شوند تا بخش‌های غیر پوشیده و ثابت صحنه آشکار گردد که از کنار هم قرار دادن این بخشها، کل زمینه به جز بخش‌هایی از صحنه که ماهیتی دینامیک دارند همچون امواج دریا قابل بازسازی است. برای مدل کردن این بخشها نیز الگویی مناسب در نظر گرفته شده است که آخرین تغییرات این نواحی را در مدل نهایی درج می‌کند. پس از کامل شدن مدل زمینه، به ازاء هر فریم جدید، اشیا متحرک آشکارسازی می‌شوند و مدل زمینه هم تازه می‌شود. این الگوریتم در صحنه‌های مختلف استاتیک و دینامیک با موفقیت مورد آزمایش قرار گرفته است.

کلمات کلیدی

مدل زمینه، اطلاعات حرکتی، اشیا متحرک، فریم تفریقی، آشکارسازی، ردیابی

Background Modeling Using Movement Information and Nonconsecutive Frames Differencing Method

S.Sharifzadeh, G.Ardeshir, H.Miar Naimi
Faculty of Electrical & Computer Engineering
University of Mazandaran

Abstract

Detection and tracking of moving objects is used in many machine vision systems for monitoring and security purposes. One of the significant methods in this field is differencing. Therefore background modeling and updating is the main part of such algorithms. In our work, we have developed a new method using movement information. First proper number of frames according to velocity and orientation information of moving object are selected and divided in to several groups. Then frames in each subset are differenced in a special nonconsecutive order so that static parts of background are discovered. The whole back ground can be modeled by joining these parts except regions with dynamic nature that are modeled differently considering the most recent changes. The modeled background has been used to detect and track moving objects in different scenes successfully. For every new frame, the model is also updated.

Keywords

Background model, movement information, moving object, differential frame, detection, tracking.

* عضو هیئت علمی دانشگاه مازندران، بابل - خیابان شریعتی - مجتمع آموزش عالی نوشیروانی - دانشکده برق و کامپیوتو دانشگاه مازندران، ۱۱۱-۳۲۳۴۵۰۱.
g.ardeshir@nit.ac.ir.

در یکی از کارهای اخیر، تکنیک On-line Adaboost (نسخه on-line بوستینگ) برای انتخاب ویژگیها و ساختن کلاسیفایرها در یک مرحله بادگیری اولیه به کار رفته‌اند. کلاسیفایرها ضعیف‌تر با هم ترکیب می‌شوند تا کلاسیفایرها قوی‌تر ساخته شوند که در مدل‌سازی زمینه از آنها استفاده می‌شود [۱۳].

در رویکردی که پیش روست، بر خلاف اکثر روش‌های ذکر شده از رویکردی غیر آماری برای مدل‌سازی زمینه استفاده شده است. بطوریکه با تجزیه و تحلیل سرعت و جهت حرکت اشیا متحرک صحنه زیر مجموعه‌هایی از رشته فریم‌های ورودی با طول مناسب انتخاب می‌شوند. زیر مجموعه‌هایی که متحرک‌های هم سودارند، ترکیب شده و زیر مجموعه واحدی را تشکیل می‌دهند. سپس با استفاده از روش تفریق فریم‌های غیر متواലی در هر زیر مجموعه و با ترتیبی مناسب، مدل زمینه به دست می‌آید. همچنین بخش‌هایی که ماهیتی کاملاً دینامیک دارند و به سادگی قابل استخراج از طریق فریم‌های تفریقی نیستند همچون امواج دریا، با روشی خاص بازسازی می‌شوند. بدین ترتیب مدل کامل زمینه با سرعتی مناسب به دست می‌آید. با استفاده از این مدل، اشیا متحرک در فریم‌های بعدی آشکارسازی و ردیابی می‌شوند. همچنین به ازاء هر فریم جدید مدل زمینه تازه می‌شود تا آخرین تغییرات به وجود آمده در زمینه در مدل درج گردد.

در ادامه مقاله، در بخش دوم قالب اصلی و کامل الگوریتم شامل استخراج جهت و سرعت متحرک و به کار بردن این اطلاعات در انتخاب و دسته‌بندی رشته فریم‌ها و فرآیند تفریق فریم‌ها در مدل سازی زمینه تشریح می‌گردد. فرآیند جداسازی و تازه کردن الگوریتم در بخش سوم بررسی خواهد شد. نتایج آزمایش‌ها انجام گرفته و مقایسه با سایر روش‌ها در بخش چهارم نشان داده شده‌اند. بخش پنجم به خلاصه مقاله اختصاص دارد.

۲- تشریح الگوریتم

در رشته تصاویری که از یک صحنه ضبط می‌شود، شی متحرک مقابل زمینه بر اساس سرعت و جهت حرکت خود برای مدت زمان معینی بخشی از صحنه را می‌پوشاند و سپس از آن بخش صحنه خارج می‌گردد. اگر متحرک کاملاً از محل اولیه خود خارج شده باشد و فریم اولیه را از فریم اخیر تفریق نماییم، در تصویر بازیگری که با آستانه گیری براساس آستانه‌ای تعریف شده برای هر فریم تفریقی به دست می‌آید، نقاط سیاه یا همان درایه‌های صفر آرایه تصویر متناظر با بخش‌های زمینه تصویر هستند. لذا این بخش‌های مشترک غیر پوشیده از زمینه را می‌توان از روی نقاط متناظر در دو فریم اولیه شرکت کننده در تفریق استخراج کرد. شکل (۱) چنین فرآیندی را نشان می‌دهد. اگر تعداد کافی از رشته فریم‌هایی در دست باشد که در آنها متحرک از محل اولیه خود در جهتی معین شروع به حرکت می‌کند و از آنجا کاملاً خارج و دور می‌شود، می‌توان با استفاده از رویکرد فوق به مدل‌سازی زمینه پرداخت.

۱- مقدمه

روش تفریق یکی از ساده‌ترین و سریعترین روش‌های جداسازی اشیا متحرک صحنه از پس زمینه است. در این روش معمولاً در مرحله نخست با استفاده از تعداد فریم‌های کافی از صحنه، تصویر زمینه بازسازی می‌گردد تا با استفاده از آن امکان جداسازی اشیا پیش زمینه در مرحله بعد ممکن گردد. اگر چه به نظر می‌رسد که با داشتن تصویر زمینه، استخراج اشیا متحرک به سادگی با آستانه‌گیری فریم تفریقی محقق شود، اما در صورت نداشتن یک طرح مناسب برای تازه کردن مدل زمینه و درج تغییرات صحنه در آن، امکان جداسازی‌های اضافی و حتی حضور بخش‌های ناخواسته ای از زمینه در تصویر حاصل وجود دارد. این مشکل از آنجا ناشی می‌شود که بر خلاف تصور ما، زمینه هرگز کاملاً استاتیک و پایدار نیست. تغییرات زمینه می‌تواند در اثر عوامل مختلفی به وجود بیاید. بعضی صحنه‌ها همچون دریا ماهیتی دائمًا دینامیک دارند. همچنین عواملی همچون باد، رعد و برق، دود و غیره می‌توانند باعث جابجایی، حرکت و یا تغییر اجزا ثابت زمینه گردند. حتی اگر چنین عواملی حضور نداشته باشند، به دلیل تغییرات آرام اما پیوسته روشنایی محیط که گاه بعضی عوامل همچون ابر، مه و غیره آنرا تشخید می‌کنند، بهتر است تا مدل اولیه زمینه بطور پیوسته تازه شود^۱ تا امکان جداسازی‌های اضافی در ضمن استخراج اشیا پیش زمینه حداقل گردد.

برای مدل‌سازی زمینه تاکنون الگوریتمها و روش‌های متنوعی به کار رفته اند که بسیاری از آنها مبتنی بر روش‌های آماری هستند. یکی از نخستین رویکردها در این زمینه، استفاده از توزیع گوسین برای مدل‌سازی زمینه است. در بعضی روشها هر پیکسل به صورت یک توزیع واحد مدل می‌شود(SGM)[۱] و در بعضی دیگر مخلوطی از توزیع‌های گوسین به کار می‌روند(MGM)[۲و۳و۴و۵] که عموماً از الگوریتم(EM) برای بهینه ساختن طرح استفاده می‌کنند. در رویکردی مشابه هر پیکسل به صورت لایه‌هایی از توزیع‌های گوسین نشان داده می‌شود و با استفاده از یادگیری بیزین و تکرار، توزیع احتمال میانگین و کوواریانس هر پیکسل تخمین زده می‌شود[۶].

فیلترهای وفقی همچون کالمون و واینر نیز به منظور پیشگویی شدت روشنایی پیکسلها برای مدل سازی زمینه به کار رفته اند[۷و۸]. متدهای غیر پارامتریک از دیگر روش‌هایی به کار رفته در این زمینه هستند. در این روشها، چندین نمونه از شدت روشنایی هر پیکسل برای تخمین تابع چگالی آن جمع آوری می‌شود. با استفاده از تخمین چگالی کرنل، میزان شباهت نمونه جدید هر پیکسل محاسبه می‌گردد[۹]. در رویکردی مشابه اطلاعات حرکتی با استفاده از تخمین چگالی کرنل وفقی برای مدل‌سازی صحنه‌های دینامیک به کار می‌روند[۱۰].

زنگیره‌های مارکوف (HMM) نیز با توبولوژی ثابت و سیکل تازه کردن پارامترها و یا با توبولوژی دینامیک و متغیر با زمان در مدل‌سازی زمینه به کار رفته‌اند[۱۱و۱۲].



(الف)

(ب)

(ج)

شکل (۲): (الف و ب) نمایش دو فریم از یک صحنه با همپوشانی در محل شخص متحرک. (ج) نمایش فریم تفریقی که در آن نقاط خطأ در محل همپوشانی دیده می‌شود.

دست باشد به طوریکه همه نقاط زمینه از مجموعه آنها قابل استخراج باشد. هر چند در بعضی شرایط امکان آشکار شدن همه نقاط زمینه تنها با این روش ممکن نیست و روشی دیگر برای مدل‌سازی آن نواحی خاص به کار گرفته می‌شود که در بخش‌های بعدی در باره آن توضیح داده خواهد شد.

برای فراهم کردن مجموعه فریم‌های تفریقی با ویژگی ذکر شده لازم است تا زیر مجموعه‌های مناسبی از میان رشته فریم‌های ورودی n انتخاب گردد. به طور مثال اگر زیر مجموعه‌ای از فریم‌ها به طول n به صورت $\{f_1, f_2, \dots, f_n\}$ در دست باشد که می‌دانیم متحرک در آنها در یک جهت معین حرکت کرده و از محل اولیه به اندازه کافی دور شده است. تعداد محدودی از فریم‌ها با این ویژگی در شکل (۳) نشان داده شده‌اند. در نتیجه اگر فریم‌های این مجموعه که به اندازه $\frac{n}{2}$ از هم فاصله دارند دو به دو تفرقی شوند بطوریکه مجموعه‌ای به صورت $\{f_{d1}, f_{d2}, \dots, f_{dn/2}\}$ به دست بیاید، فریم‌های تفریقی حاصله، پس از بسط نواحی پیش زمینه، برای مدل‌سازی زمینه مناسب خواهند بود. نمایش فریم‌های تفریقی به این شیوه و بسط نواحی پیش زمینه آنها برای تصاویر شکل (۳) در شکل (۴) نمایش داده شده‌اند.

برای آنکه آخرین تغییرات ایجاد شده در صحنه زمینه در مدل زمینه درج گردد، بازسازی مدل زمینه با یک ترتیب زمانی از آخرین فریم تفریقی، که از فریم‌های تصویر جدیدتری حاصل شده است، آغاز می‌شود و هر بخشی از زمینه که از طریق اولین فریم تفریقی بازسازی شد، هرگز با فریم‌های بعدی تغییر داده نخواهد شد و لذا فریم‌های تفریقی بعدی به ترتیب از آخر به اول برای بازسازی قسمتهای باقیمانده زمینه مورد استفاده قرار خواهند گرفت.

با توجه به نکات عنوان شده در تهیه فریم‌های تفریقی مناسب، انتخاب زیر مجموعه‌ای موثر از فریم‌ها با ویژگیهای ذکر شده الزامی است.

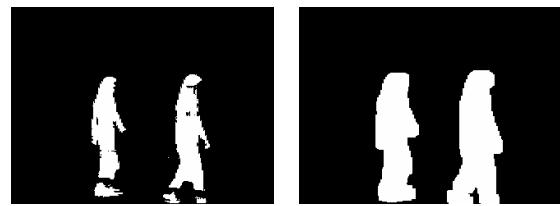
۱-۲- مدل‌سازی زمینه

همانطور که اشاره شد برای داشتن فریم‌های تفریقی با ویژگیهای مطلوب، ابتدا لازم است تا اطلاعات حرکتی اشیا متحرک صحنه از تصاویر ورودی استخراج گردد تا مجموعه مناسبی از آنها برای تهیه



(الف)

(ب)



(ج)

شکل (۱): (الف و ب) نمایش دو فریم از یک صحنه بطوریکه شخصی که راه می‌رود از محل اولیه خود کاملاً خارج شده است. (ج) نمایش تصویر باینری حاصل از تفرقی دو فریم (الف و ب). (د) نمایش بسط نواحی پیش زمینه در فریم تفریقی.

امکان ایجاد خطأ در مدل‌سازی زمینه طی فرآیند فوق در دو حالت وجود دارد. نخست آنکه به دلیل تشابه رنگ زمینه با قسمتی از بافت متحرک، در تصویر باینری حاصله، این بخش همانند سایر نقاط ثابت مربوط به زمینه سیاه دیده می‌شود. در حالیکه این بخش از زمینه در حقیقت در اثر حضور متحرک در یکی از دو فریم اولیه پوشیده شده است. در شکل (۱-ج) این خطأ در قسمتی از پای فردی که از صحنه عبور می‌کند در فریم تفریقی دیده می‌شود. برای غلبه بر این مشکل ابتدا در فریم تفریقی باینری، نقاط منفرد پراکنده حذف می‌شوند. سپس، نواحی آشکار شده پیش زمینه با نرخ خاصی که بطور خودکار و بر اساس ضریبی از مقدار آستانه هر فریم تفریقی تعريف می‌شود، بسط داده می‌شوند تا نواحی خطأ بطور موثر پوشش یابند و به اشتباہ به عنوان بخشی از زمینه وارد مدل نشوند (شکل ۱-۵).

حالت دوم بروز خطأ به محل شی متحرک در دو فریم اولیه‌ای که در تفرقی شرکت می‌کنند، مربوط می‌شود. همانطور که گفته شد، برای استخراج نقاط زمینه از فریم تفریقی لازم است تا متحرک کاملاً از محل اولیه خود خارج شده باشد. هرگاه این مسئله به طور کامل محقق نشود و نوعی همپوشانی در محل قرار گرفتن متحرک در دو فریم اولیه وجود داشته باشد، نقاط قرار گرفته در داخل مرزهای همپوشانی به دلیل تشابه سطوح خاکستری در فریم تفریقی سیاه می‌شوند که به اشتباہ، به عنوان نقاط زمینه غیر پوشیده مشترک در دو فریم اولیه شناسایی می‌شوند. شکل (۲) چگونگی بروز چنین خطایی را برای دو فریم نمونه نشان می‌دهد. برای از بین بردن احتمال بروز چنین خطایی لازم است تا در انتخاب دو فریم اولیه دقت شود و اطلاعاتی از موقعیت متحرک در دو فریم وجود داشته باشد. علاوه بر همه نکات فوق برای مدل کردن کل زمینه لازم است تا فریم‌های تفریقی به اندازه کافی در

طول ($I - N$) به صورت $\{I_{D1}, I_{D2}, \dots, I_{D(N-1)}\}$ ، با استفاده از برچسب‌گذاری عناصر متصل^۱، محل بر جسته‌ترین ناحیه فعال که به متحرک اصلی تصویر تعلق دارد، مشخص می‌شود. دو مجموعه-های *Column* و *Row* شامل مقادیر میانگین سطرها و ستونهای موقعیت متحرک در فریم‌های تفريقي متوالى هستند.

از آنجاييکه گراديان يك روند افزایشي، مثبت است و در مقابل يك مجموعه با روند نزولي، مقادير گراديان منفي دارد، می‌توان از تابع گراديان جهت تشخيص جهت حرکت متحرک استفاده کرد. لذا گراديان گيري نسبت به متغيرهای سطر R و ستون C در مجموعه‌های فوق انجام می‌شود. سپس برای معين شدن راستا و جهت متحرک عناصر مثبت و منفي گراديان هر مجموعه شمرده می‌شوند. روابط به کار رفته به صورت زير هستند:

$$Row = \{R_{M1}, R_{M2}, \dots, R_{M(N-1)}\} \quad (1)$$

$$Column = \{C_{M1}, C_{M2}, \dots, C_{M(N-1)}\} \quad (2)$$

$$\nabla Row = \frac{\partial Row}{\partial R} = \{Gr_{M1}, Gr_{M2}, \dots, Gr_{M(N-1)}\} \quad (3)$$

$$\nabla Column = \frac{\partial Column}{\partial C} = \{Gc_{M1}, Gc_{M2}, \dots, Gc_{M(N-1)}\} \quad (4)$$

$$\nabla Row(p) = \{Gr_{Mi} \mid Gr_{Mi} > 0\} \quad (5)$$

$$\nabla Row(n) = \{Gr_{Mi} \mid Gr_{Mi} < 0\} \quad (6)$$

$$\nabla Column(p) = \{Gc_{Mi} \mid Gc_{Mi} > 0\} \quad (7)$$

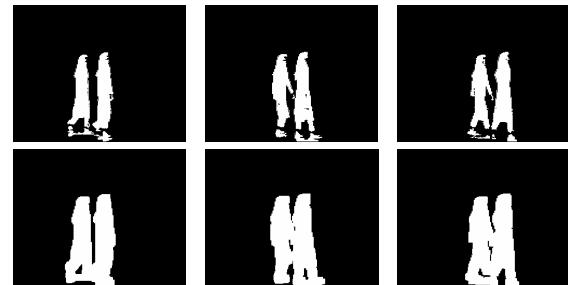
$$\nabla Column(n) = \{Gc_{Mi} \mid Gc_{Mi} < 0\} \quad (8)$$

برچسب p مجموعه گراديان‌های مثبت و برچسب n مجموعه گراديان‌های منفي را مشخص می‌کند. از میان این مجموعه‌ها، مجموعه‌ای که بالاترین تعداد عناصر را دارد، راستا و جهت متناظر با عنصر متحرک را نشان خواهد داد. به عنوان مثال اگر برتری با مجموعه $\nabla Row(n)$ باشد آنگاه مشخص می‌شود که متحرک در فریم‌های مورد بررسی در راستای عمودی و با روال کاهشی یا نزولی سطرها، یعنی از سمت پایین به بالا حرکت می‌کند.

حرکتهاي قطری نیز براساس زاویه مسیر حرکت در زمرة یکی از این مجموعه‌ها دسته‌بندی می‌شوند.



شکل (۳): نمایش فریم‌های متوالى از یک صحنه که شخص متحرک از محل اولیه کاملاً خارج شده است.



شکل (۴): ردیف بالا به ترتیب از چپ به راست، فریم‌های تفريقي (۴)، (۵)، (۶)، (۷)، (۸) از رشته تصاویر شکل (۳). ردیف پایین بسط نواحي پیش زمینه در فریم‌های تفريقي.

فریم‌های تفريقي انتخاب شوند. یکی از ویژگیهای لازم راستا و جهت حرکت اشیا متحرک صحنه است. همچنین برای انتخاب تعداد و طول مناسب برای زیر مجموعه‌ها طوریکه متحرک به اندازه کافی از محل اولیه خارج شود و فریم‌های تفريقي همپوشانی مکانی نداشته باشند و نیز برای آنکه کل زمینه از طریق مجموعه فریم‌های تفريقي قابل استخراج باشد، معیاری از سرعت متحرک لازم است.

۲-۱-۱-۱- یافتن راستا و جهت اشیا متحرک

وقتی متحرکی در راستای معینی حرکت می‌کند شماره سطر و ستون مربوط به موقعیت آن تغییر می‌کند. به عنوان مثال وقتی متحرک در راستای افقی حرکت می‌کند، در رشته فریم‌های متوالى شماره ستونهایی از آرایه تصویر که سطوح خاکستری مربوط به شی در آن قرار گرفته است تغییر می‌کند در حالیکه در حرکت عمودی شماره سطرهای تغییر می‌کند. همچنین جهت حرکت در هر کدام از این راستاهای با توجه به روند افزایيشی یا کاهشی تغیيرات قابل تشخيص است. بطور مثال حرکت از بالا به پایین در راستای عمودی، روند افزایيشی در تغیيرات شماره سطرهای ایجاد می‌کند از این نکته می‌توان در تشخيص راستا و جهت حرکت استفاده نمود.

هر گاه فرض شود که مجموعه رشته فریم‌های اولیه به طول N در دست باشد، برای تعیین محل اشیا متحرک در این فریم‌ها، از تفريق فریم‌های متوالى مجموعه استفاده می‌شود و مجموعه حاصل به

همانطور که در بخش قبل اشاره شد، با فرض آنکه هیچ اطلاعاتی از نوع و بعد متوجه، جهت حرکت یا سرعت آن و حتی فاصله آن نسبت به دوربین وجود ندارد، فریم‌های محدودی در نظر گرفته می‌شوند و در دسته‌های ۲۰ تایی تعیین جهت می‌شوند. سرعت متوسط هر دسته نیز از رابطه (11) محاسبه می‌گردد. با توجه به محدودیت دسته‌های انتخاب شده می‌توان با میانگین‌گیری از متوسط سرعت‌های دسته‌ها، تخمینی از سرعت متوسط متوجه و یا متوجه‌های احتمالی محدود موجود در کل فریم‌های تشکیل دهنده دسته‌ها به دست آورد. این تخمین به انتخاب تعداد مناسبی از فریم‌ها برای انجام تفیق غیر متوازن فریم‌ها که در بخش ۲ تشریح شد، کمک می‌کند. معیار انتخاب تعداد رشته فریم‌های ورودی برای تشکیل دسته‌ها در بخش ۳-۱-۲ شرح داده خواهد شد.

۳-۱-۲- آماده ساختن مجموعه فریم‌های مناسب

با استفاده از سرعت تخمینی به دست آمده برای متوجه یا متوجه‌های احتمالی صحن، می‌توان برآورد مناسبی برای انتخاب مناسب طول رشته فریم‌های لازم جهت تفیق فریم‌های غیر متوازن انجام داد. به فرض فریم‌هایی با ابعاد 240×320 پیکسل در دست باشد. در این صورت اگر متوجهی با سرعت متوسط تخمینی 70 ستون بر فریم در جهت افقی در حرکت باشد، آنگاه حداقل فریم تخمینی لازم برای عبور کامل این متوجه از زاویه دید دوربین، مقدار صحیحی از $320/70 = 4.57$ خواهد بود. برای حرکتهای عمودی این مقدار $240/70 = 3.43$ است. با توجه به تغییرات احتمالی در فاصله جسم تا دوربین و احتمال غیر یکنواخت بودن حرکت شی متوجه و یا حضور بیش از یک متوجه با بعد متمایز، دو برابر مقدار فوق برای زیر مجموعه فریم‌ها انتخاب می‌شود. همچنین برای اطمینان از آشکار شدن کلیه بخش‌های زمینه از فریم‌های تفیقی، سه زیر مجموعه با این تعداد فریم آماده می‌شود تا احتمال پوشیده ماندن بخشی از زمینه در اثر عواملی چون ثابت ماندن موقع متوجه در بخشی از زمینه به حداقل برسد. از آنجا که بالاترین طول رشته فریم‌ها برای کنترین متوجه‌ها به دست می‌آید، با فرض بیشترین طول برابر با L_{\max} حد اکثر طول رشته فریم لازم FL_{mx} برابر است با:

$$FL_{mx} = 3 \times L_{\max} \quad (12)$$

به طور تجربی کنترین متوجه‌ها با سرعت‌هایی در حدود 5 ستون بر فریم در راستای افقی حرکت می‌کنند. می‌توان انتخاب تعداد دسته‌های اولیه 20 تایی فریم‌ها n برای تخمین سرعت و در نتیجه طول مناسب رشته فریم‌های مورد نیاز، براساس نزدیکی آن به مقدار پیش‌بینی شده برای FL_{mx} است. بطوریکه خواهیم داشت:

$$FL_{mx} = n \times 20 \quad (13)$$

از آنجاییکه در یک صحنه ممکن است چندین شی متوجه در جهت‌های متفاوتی در حرکت باشند، برای به حداقل رساندن تعداد متوجه‌ها در روال جهت‌یابی فوق، لازم است تا تعداد محدودی فریم در نظر گرفته شود تا با اطمینان خوبی حداقل تعداد متوجه‌ها در فریم‌ها حضور داشته باشند. با توجه به آنکه فرض می‌شود هیچ اطلاعاتی از نوع و تعداد متوجه‌ها در اختیار نیست، با فرض امکان ضبط 5 فریم در ثانیه، هر بار تعداد 20 فریم از رشته ورودی در الگوی جهت‌یابی قرار می‌گیرد. یعنی هر 4 ثانیه متوجه صحنه جهت‌یابی می‌شود. که به طور تجربی معادل با حرکت کنترین اشیا متوجه از محل خود می‌باشد.

۲-۲- تعیین سرعت حرکت اشیا متوجه

در بخش ۲ توضیح داده شد که برای مدل‌سازی زمینه، تفیق غیر متوازن باید در مجموعه‌ای از فریم‌ها با طول مناسب صورت گیرد. بطوریکه متوجه از محل اولیه خارج و به اندازه کافی دور شود. سرعت اشیا متوجه نقش مهمی در محقق شدن چنین شرایطی ایفا می‌کند. زیرا فریم‌های زیادی مورد نیاز است تا یک متوجه کند از محل اولیه به اندازه کافی دور شود. در حالیکه برای یک متوجه سریع فریم‌های محدودی کفايت می‌کند.

می‌توان سرعت لحظه‌ای متوجه را با استفاده از قدر مطلق گردایان تغییرات سطر یا ستونهای مسیر حرکت آن محاسبه نمود. زیرا متوجه‌های سریعتر تغییرات بزرگتری در شماره سطرهای و ستونهای بر واحد فریم نسبت به متوجه‌های کند دارند. علامت عبارت گردایان نیز همانطور که در بخش قبل برسی شد، با توجه به اینکه گردایان مربوط به تغییرات سطرهای یا ستونهای باشد، جهت حرکت را نشان می‌دهد. با استفاده از روابط (10) و (11) ، مقادیر این گردایان‌ها برای مجموعه‌های 20 تایی از فریم‌ها محاسبه می‌گردد. سپس بر اساس علامت غالب در میان مجموعه‌های (5) تا (8) ، مجموعه متناظر با جهت حرکت اصلی استخراج می‌گردد و میانگین مقادیر این مجموعه، سرعت متوسط متوجه است. به عنوان مثال اگر مجموعه با علامت غالب که جهت را تعیین می‌کند، $\nabla Group$ بنامیم به طوری که داشته باشیم:

$$\nabla Group = \{Gr_1, Gr_2, \dots, Gr_k\} \quad (9)$$

که k تعداد عناصر مجموعه $\nabla Group$ می‌باشد. آنگاه مجموعه سرعت‌های لحظه‌ای و میانگین متوجه در مسیر خود به صورت زیر خواهد بود:

$$|V| = \{|Gr_1|, |Gr_2|, \dots, |Gr_k|\} \quad (10)$$

$$V_m = \frac{1}{k} \sum_{i=1}^k v_i = \frac{1}{k} \sum_{i=1}^k |Gr_i| \quad (11)$$

تفريق شونده فاصله زیادی از هم نداشته باشند و دارای تغییرات کمتری نسبت به هم باشند.
اگر نقاط محدودی پس از پایان این مرحله به جا مانده باشند، توسط نقاط متناظر در آخرین و جدیدترین فریم این زیر مجموعه بازاری می‌شوند.

پس از آماده کردن ۳ زیر مجموعه از فریم‌ها با طول مناسب به دست آمده، در هر کدام از آنها به طور جداگانه، متحرک جهت‌یابی می‌شود و در صورتی که متحرک دو مجموعه متواالی هم جهت باشد، آن مجموعه‌ها با هم ترکیب می‌شوند و تغريف غیر متواالی برای مجموع آنها انجام می‌شود.

۳- آشکار سازی و تازه کردن زمینه

پس از کامل شدن مدل زمینه، برای آشکارسازی و ردیابی اشیا متحرک صحنه در فریم‌های بعدی از آن استفاده می‌شود. در فاز آشکارسازی نیز لازم است تا مدل زمینه به ازاء هر فریم جدید تازه شود تا آخرین تغییرات ایجاد شده در زمینه در مدل درج گردد.

۳-۱- آشکارسازی اشیا متحرک

برای آشکارسازی اشیا متحرک هر فریم جدید از تصویر زمینه تغريف می‌شود و برای هر فریم تغريفی مقدار آستانه معینی تعريف می‌شود. به همین ترتیب حذف نقاط پراکنده فریم تغريفی باینری طی فرآیند خاصی با استفاده از این آستانه انجام می‌گیرد.

۳-۲- تازه کردن زمینه

مدل زمینه به ازاء هر فریم جدید تازه می‌شود تا تغییرات نواحی دینامیک در آن گنجانده شود. این نواحی همان نقاط سیاه رنگ به جا مانده در مدل زمینه پیش از مرحله ترمیم می‌باشند. برای تازه کردن این نقاط ابتدا باید مطمئن شویم که آیا تغییرات جدیدی در این نقاط به وجود آمده است یا خیر. لذا این نقاط دینامیک، پیش از عملیات پس_پردازش بر روی حاصل تغريف باینری که عموماً سبب حذف چنین نقاط پراکنده زائد و آشکارسازی‌های اضافی می‌شود، آزمایش می‌شوند. اگر این نقاط در این تصویر سفید باشند، نشان دهنده وجود تغیير در آن نواحی است. اما برای اطمینان کامل از آنکه تغییرات به وجود آمده از حضور یک شی متحرک در آن نقاط ناشی نشده است، مجدداً نقاط دینامیکی که قبل از پس_پردازش سفید بودند، پس از این عملیات آزمایش می‌شوند. در صورت حذف شدن این نقاط، وجود تغییرات گذرا در آنها مورد تایید است و با مقادیر متناظر شان در فریم جدید جایگزین می‌شوند.

هرگاه تعداد فریم‌های جدید به حد کافی برسد. کل زمینه مجدداً مدل می‌شود.

۴- نتایج آزمایش‌ها

این الگوریتم در شرایط و صحنه‌های مختلف استاتیک و حتی دینامیک همچون شرایط وزش باد و با ساحل دریا با موفقیت مورد آزمایش قرار گرفته است. برای تهیه تصاویر از یک دوربین ساده با

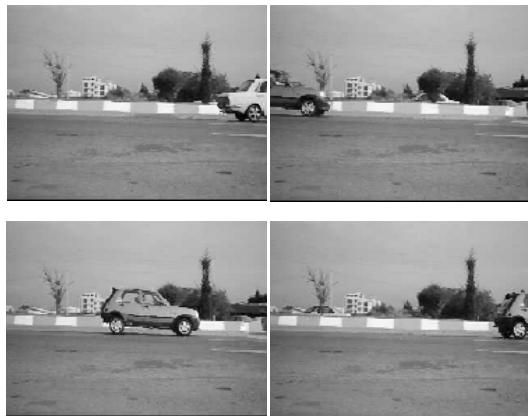
۲-۲- تغريف فریم‌های غیر متواالی

در ابتدا مدل زمینه به صورت آرایه‌ای تمام صفر و سیاه برابر با ابعاد فریم‌های تصویر در نظر گرفته می‌شود. پس از محاسبه نخستین فریم تغريفی و آستانه‌گیری از آن، نواحی پیش‌زمینه برای از بین بردن احتمال خطأ چنانچه در شکل (۱-۵) نشان داده شده است، بسط داده می‌شوند. سپس نواحی سیاه متناظر با زمینه غیر پوشیده مشترک در دو فریم استخراج و در آرایه تمام صفر اولیه قرار می‌گیرند. در صورتی که تمام بخش‌های زمینه باز سازی نشده باشند، فریم تغريفی بعدی محاسبه و مدل سازی بخش‌های به جا مانده زمینه به وسیله آن ادامه می‌باید. تا زمانی که تمام زمینه مدل شود و یا تمام فریم‌های تغريفی ممکن محاسبه شوند، فرآیند فوق ادامه خواهد یافت.

با توجه به استفاده از رشته فریم‌هایی با طول مناسب، در پایان مدل کامل زمینه بازسازی می‌شود. البته در صحنه‌های خاصی با ماهیت دینامیک امکان بجا ماندن نقاط سیاه رنگ وجود دارد. زیرا احتمال دارد که بخش‌های دینامیک زمینه همچون امواج دریا و یا برگهای درختان در هنگام وزش باد، در آستانه‌گیری از تغريف فریم‌ها، همواره به عنوان نقاط متحرک پیش‌زمینه که از جای خود خارج نمی‌شوند، به حساب آمده و سفید باشند. برای بازسازی این بخشها روندی خاص در نظر گرفته شده است.

۲-۲-۱- ترمیم نواحی دینامیک زمینه

نقاط سیاه رنگ باقیمانده در مدل زمینه که به نواحی دینامیک صحنه مربوط می‌شوند، باید به گونه‌ای ترمیم شوند که آخرین تغییرات آن نواحی در مدل بازسازی شود و در فرآیند آشکارسازی و ردیابی اشیا متحرک صحنه، آشکارسازی اضافی در این نقاط حداقل گردد. به همین دلیل در آخرین زیر مجموعه آمده شده با طول مناسب که شامل فریم‌های جدید تری از صحنه است، تغريف فریم‌های غیر متواالی با نرخ انساضی بسیار کمتر از قبل تکرار می‌شود. حتی اگر به دلیل هم سویی متحرک این زیر مجموعه با زیر مجموعه‌های دیگر تلفیق آنها صورت گرفته باشد، در این مرحله تنها از عناصر همین مجموعه به طور مجزا استفاده می‌شود. زیرا مدل زمینه تنها در بخش‌های محدودی کامل نشده است و برخلاف بخش گذشته، به منظور خارج شدن کامل متحرک از محل اولیه، مجموعه‌ای بلند مد نظر نیست. بالعکس برای آنکه نقاط دینامیک هر چه بیشتر به عنوان نواحی زمینه در تصاویر تغريفی ظاهر شده و قابل مدل شدن باشند، بهتر است فریم‌های



شکل (۸): نمایش فریم های ۱۵۶ و ۱۶۲ و ۱۵۹ و ۱۶۵



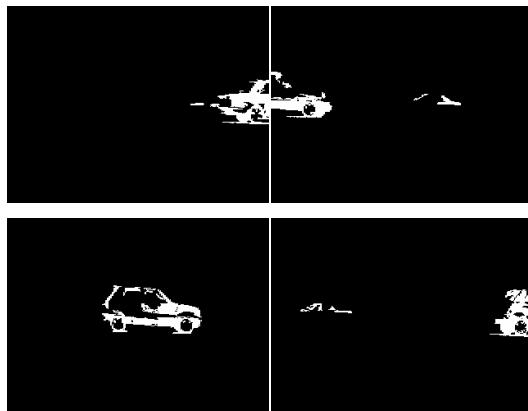
شکل (۵): نمایش فریم های ۸۶۶ و ۸۷۲ و ۸۷۴ و ۸۷۶



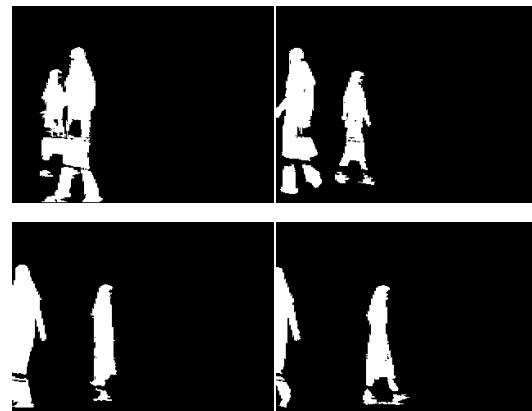
شکل (۹): مدل زمینه بازسازی شده



شکل (۶): مدل زمینه بازسازی شده



شکل (۱۰): نمایش نتایج آشکارسازی برای فریم های ۱۵۶ و ۱۵۹ و ۱۶۲ و ۱۶۵ با استفاده از مدل زمینه



شکل (۷): نمایش نتایج آشکارسازی برای فریم های ۸۶۶ و ۸۷۲ و ۸۷۴ و ۸۷۶ با استفاده از مدل زمینه

شده است. برنامه سرعت متوسط ۲۰ ستون بر فریم را برای متحرک‌ها در رشته فریم‌های ورودی تخمین زده است. مقایسه نتایج جهت‌یابی با جهت حرکت افراد در فریم‌های ورودی نشان می‌دهد که برنامه در تشخیص جهت‌های صحیح موفق عمل کرده است. با توجه به مدل کامل زمینه، این مسئله اثبات می‌شود. نتایج موفق آشکارسازی در شکل (۷) نیز گویای همین موضوع می‌باشد. در آزمایش دوم صحنه

قابلیت ضبط ۵ فریم در ثانیه استفاده شده است. ابعاد تصاویر 320×240 است.

در نخستین آزمایش، از یک صحنه بسته استفاده شده است. که افرادی در مقابل دوربین در حرکت هستند. تعدادی از تصاویر این صحنه در شکل (۵) نشان داده شده‌اند. از آنجایی که صحنه استاتیک بوده است، مدل کامل زمینه در زمان کوتاهی برابر با ۴۱ ثانیه و بدون باقیماندن نقاط سیاه دینامیک به دست آمده است که در شکل (۶) نشان داده

- [7] Koller D, Weber J, Huang T, Malik J, Ogasawara G, Rao B, Russell S. *Towards robust automatic traffic scene analysis in realtime*. In Proceedings of the 33rd IEEE Conference on Decision and Control (Cat. No.94CH3460-3). IEEE, Part vol.4, 1994.
- [8] K.P. Karmann, A. von Brandt, *Moving Object Recognition Using an Adaptive Background Memory*, Time-varying Image Processing and Moving Object Recognition (V. Capel ed.), Elsevier, Amsterdam, pp. 297-307, 1990.
- [9] Elgammal A., Harwood D., Davis L. *non-parametric model for background subtraction*. IEEE ICCV'99 FRAME-RATE WORKSHOP. 1999.
- [10] A. Mittal and N. Paragios, "Motion-Based background subtraction using adaptive kernel density estimation", In Proc. Int'l Conf. on Computer Vision and Pattern Recognition, Washington, DC, volume 2, pp.302-309. 2004.
- [11] J.Rittscher, Jakarta, S.Joga and A.Black. *A Probabilistic background Model for Tracking*, in proc. Europe ,on Conf .on Computer Vision and Pattern Recognition , Fort Collins, CO, volume 2, 1999.
- [12] B. Stenger, V. Ramesh, N. Paragios, F. Coetzee and J.M. Buhmann. *Topology Free Hidden Markov Models: Application to Background Modeling*, in proc. 8th Intl. Conf on Computer Vision, Vancouver, Canada, 2001.
- [13] Helmut Grabner, Peter M. Roth, Michael Grabner, and Horst Bischof. *Autonomous learning of a robust background model for change detection*. In Proc. IEEE Inter-national Workshop on Performance Evaluation of Tracking and Surveillance, pp 39, 46, 2006.

خیابانی که اتومبیلها در دو جهت به سرعت در حرکت هستند، در نظر گرفته شده است.

سرعت تخمینی برای اتومبیلها در حدود ۳۰ سوتون بر فریم محاسبه شده و زمینه در مدت ۴۶ ثانیه پس از ترمیم نقاط دینامیک محدودی همچون بخشایی از درختان دوردست، مدل سازی شده است. نتایج آزمایش در شکلهای (۸ تا ۱۰) نشان داده شده است.

در مقایسه با سایر روشها، این الگوریتم به دور از مدل های آماری پیچیده با سرعتی مناسب به مدل سازی زمینه می پردازد. برخلاف روش های غیر پارامتریک، این روش نیاز به سیکلی برای یادگیری کل صحنه ندارد.

با توجه به تازه شدن الگوریتم به ازاء هر فریم جدید، تغییرات صحنه به خوبی در نظر گرفته می شوند و لذا می توان در صحنه های دینامیک از آن استفاده کرد و آشکار سازی را به صورت موثر تر انجام داد.

۵- نتیجه گیری

در این مقاله الگوریتم جدیدی برای مدل سازی زمینه با استفاده از تجزیه و تحلیل حرکت اشیا و روش تفربیق فریم های غیر متواالی ارائه شده است. بطوریکه با استخراج جهت و سرعت حرکت اشیاء صحنه، زیر مجموعه هایی از فریم ها با طول مناسب انتخاب می شود و بخش های مختلف زمینه از تفربیق غیر متواالی آنها استخراج می گردد. برنامه قابلیت مدل کردن بخش های دینامیک صحنه و تازه کردن آنها را به ازاء هر فریم جدید دارد.

زیرنویس ها

¹ Update

² Connected Component Labeling

مراجع

- [1] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: *Real-time tracking of the human body*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):780–785, July 1997.
- [2] Grimson Wel, Stauffer C, Romano R, Lee L. *Using adaptive tracking to classify and monitor activities in a site*. In Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231). IEEE Comput.Soc. 1998.
- [3] Stauffer C, Grimson W. E. L. *Adaptive background mixture models for real-time tracking*. in Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149). IEEE Comput. Soc. Part Vol. 2, 1999.
- [4] Friedman N., Russell S. *Image Segmentation in Video sequences: A Probabilistic Approach*. in The Thirteenth Conference on Uncertainty in Artificial Intelligence. 1997. Brown University, Providence, Rhode Island, USA: Morgan Kaufmann Publishers, Inc., San Francisco, 1997.
- [5] P. KaewTraKulPong and R. Bowden. *An Improved Adaptive Background Mixture Model for Real time Tracking with Shadow Detection*. In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01. Sept 2001.
- [6] Tuzel, O.; Porikli, F.; Meer, P., *A Bayesian Approach to Background Modeling*, IEEE Workshop on Machine Vision for Intelligent Vehicles (MVIV), Vol. 3, pp. 58, June 2005.