

آشکارسازی اجسام متحرک و ردیابی در دنباله های تصویر رنگی

نویسنده

مهدی جلالی

دانشکده تحصیلات تکمیلی آزاد تهران جنوب

تلفن : ۰۲۱-۶۶۸۴۰۵۲۲

Jalali.mahdi@gmail.com

چکیده :

در این مقاله الگوریتمی برای آشکارسازی اشیا در دنباله ای از تصاویر رنگی که از یک دوربین متحرک گرفته شده است را ارائه خواهیم داد. اولین مرحله در این الگوریتم تخمین حرکت در صفحه تصویر می باشد. به جای تخمین شار نوری؛ ردیابی نقاط تکی و لبه های تصویر، کلاسترهای حرکت را که توسط گروهی از پیکسل ها در فضای ویژگی موقعیت= رنگ ایجاد شده اند، را تعیین می کنیم. مرحله بعدی قطعه بندی بر پایه حرکت است که کلاسترهای مجاور با ویژگی های مشابهی، جهت ایجاد فرض های هدف با هم ترکیب شده اند. این الگوریتم در صحنه های ترافیکی که شامل اجسام متحرکی مانند خودروها، موتورسیکلت ها و عابران پیاده می باشد به طور موفقیت آمیزی آزمایش شده است.

کلمات کلیدی : کلاسترینگ، رنگ، تخمین حرکت، قطعه بندی حرکت و ردیابی

۱- مقدمه

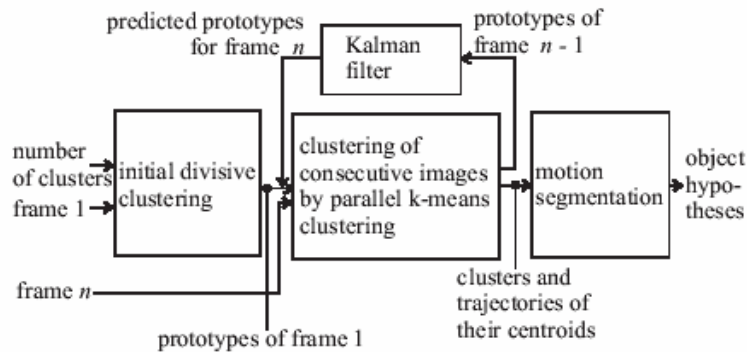
یک روش معمول برای استخراج اهداف متحرک از زمینه، قطعه بندی بر مبنای حرکت تصویر است. برای تخمین حرکت تصویر از تغییرات زمانی-فضایی در تصویر سه بعدی استفاده می شود. در قطعه متحرک بعدی هر تصویر به بخش هایی که متناظر با اجسام با خواص حرکتی مختلفی هستند تقسیم می شود. دو شرط مهم برای این روش وجود دارد: تخمین های حرکت متراکم باید محاسبه شوند و نا پیوستگی های حرکت باید محفوظ باشند.

الگوریتم های زیادی، شار نوری (سرعت میدان در صفحه تصویر) را با محاسبه گرادیان تغییرات وابسته زمانی-فضایی شدت توزیع محلی تخمین می زنند. مشکل اصلی این است که شار نوری در بخش های تصویر با شدت تقریباً یکنواختی نمی تواند بطور قابل اعتمادی تخمین زده شود. برای غلبه بر این مشکل، محدودیت نا هموار بودن را برای کل تصویر که در مراجع معرفی شده اند را در نظر می گیریم، که آن منجر به تخمین چگالی با میدان سرعتی مختلف یکنواختی می شود. با وجود این خطاهای تخمین بزرگی در نا پیوستگی های میدان سرعتی حقیقی اتفاق می افتد. برای حفظ نا پیوستگی های حرکت، محدودیت هایی در ناحیه هایی از تصویر بکار می روند که در تکنیک های قطعه بندی بر پایه چگالی برای تخمین چنین ناحیه هایی استفاده می شود.

۱-۲- بیان الگوریتم

گروه دیگری از تکنیک های تخمین حرکت برپایه، ردیابی ویژگی های تصویر در سراسر یک دنباله تصاویر است. چگالی نقاط ردیابی شده یا لبه ها اغلب برای قطعه بندی بر پایه حرکت کافی نیستند. می دانیم که ناحیه های ردیابی رنگی بوسیله تقسیم بندی رنگ که منجر به نتایج خوبی در آشکار سازی وسایل نقلیه در بزرگراه ها می شوند، تعیین می شوند. علاوه بر این نا پیوستگی های موجود در تصویر متحرک که احتمالاً با خط های لبه ها همزمان می شوند، حفظ می شوند. بالاخره کار قطعه بندی حرکت به گروهی از قطعه های چند صد تایی از رنگ به جای هزاران پیکسل از آنها ساده می شود. هر چند عمل تقسیم بندی رنگ برای هر تصویری جداگانه انجام می شود، گاهی اوقات منجر به تقسیم بندی های نا پایداری در زمان می شود.

دنباله تصویر فرضی با تغییرات فریم به فریم کوچک همراه است. یک الگوریتم متفاوتی برای قطعه بندی رنگ سازگار با زمان مبنی بر کلاستر را پیشنهاد خواهیم کرد. هر تصویر در ابعاد X, Y برای هر پیکسل، به وسیله رنگهای RGB کلاستر بندی می شود. افراز نهایی بوسیله مدل اولیه، که مرکز کلاستر در شکل فضایی مکان- رنگ است توصیف می شود. در فریم های متوالی نمونه های اولیه فریم قبلی در فضای ویژگی بوسیله کلاسترینگ موازی K گونه ای در داده تصویر جدید شیفیت می یابد. بنابراین به یک قطعه بندی دست یافته ایم که در زمان پیوسته است، علاوه بر آن هیچ تطابق آشکاری برای کلاستر های متناظر نیاز نداریم. در اینجا روشی را با اضافه کردن فیلتر کالمن که موقعیت هر نمونه اولیه را بر پایه خط سیر در صفحه تصویر را پیشگویی می کند را توسعه می دهیم. شکل (۱) را ببینید.



شکل (۱): شمای کلی از الگوریتم

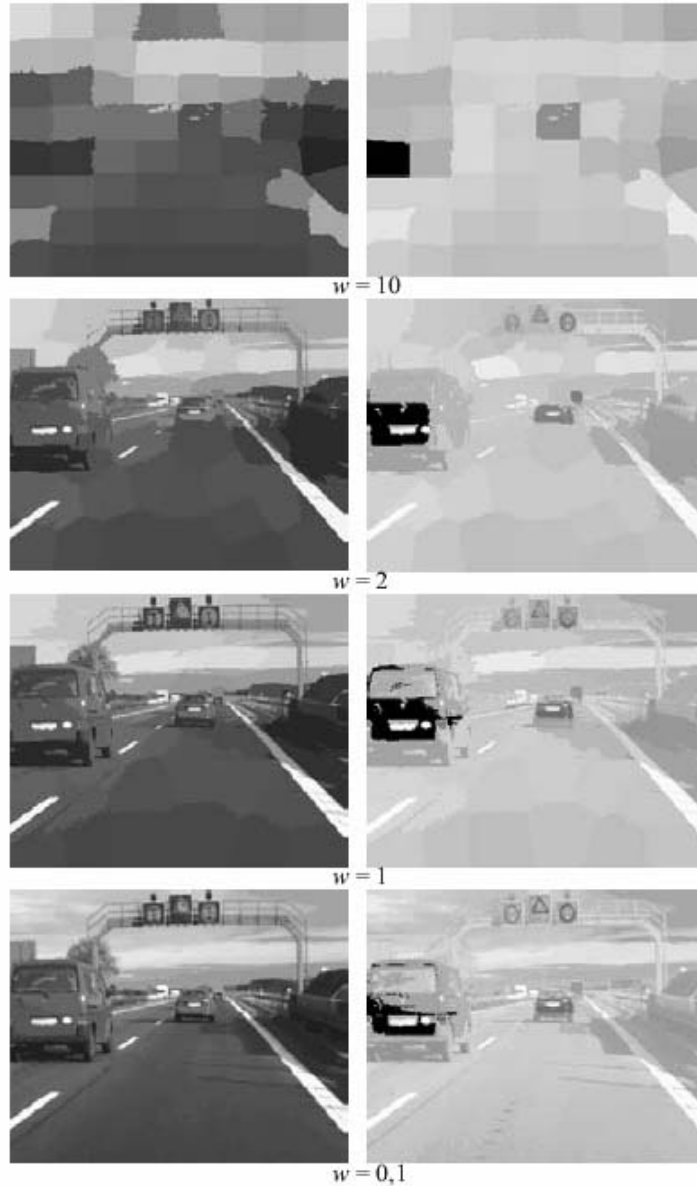
۳-۱- تخمین حرکت با کلاسترینگ:

هر پیکسل n ، در تصویر بوسیله بردار ویژگی f_n توصیف می شود که شامل فضای رنگی RGB در حالت تصویر بصورت $f_n = (R_n, G_n, B_n, \omega.x_n, \omega.y_n)$ که x_n حالت افقی و y_n موقعیت عمودی پیکسل n ام در صفحه تصویر است. برای مقدار اولیه P_r که کمترین مقدار برای خطاهای کوانتیزه شده عبارتست از: $\sum_n \|f_n - P_r(n)\|^2$ که مقدار اولیه $P_r(n)$ به بردار ویژگی f_n در فضای ویژگی موقعیت - رنگ نزدیک است. فاکتور وزن w از نتایج کلاسترینگ شدیداً متأثر است و در شکل (۲) نشان داده شده است. مقادیر بزرگ w منجر به فشردگی می شود. کاهش w کلاستر ها را در صفحه تصویر پخش خواهد کرد. برای هدف ما $w = 1$ به نظر می رسد که مناسب باشد. برای قطعه بندی تصویر اولیه از یک کوانتایزر برداری مقسم که تصویر را به تعدادی از کلاسترهای از قبل انتخاب شده، استفاده می کنیم.

برای هر تصویر بعدی نمونه های اولیه به مقدارهای تصادفی اولیه با کلاستر های موازی درتصویر پیش بینی می شود. این کلاسترینگ مجموعه جدیدی از نمونه های اولیه را برای تصویر بعدی تولید می کند. کلاسترینگ موازی $k - mean$ شامل دو مرحله برای هر تکرار است:

$$C_r(i+1) = \{pix \mid \|f_n - P_r(i)\| \leq \|f_n - P_s(i)\| \forall s\}$$

$$P_r(i+1) = \frac{1}{size[C_r(i+1)]} \sum_{n \in C_r(i+1)} f_n \quad (1)$$



شکل (۳): تاثیر بردار وزن W در ترکیب کلاسترینگ رنگ-حالت، ستون چپ تصاویر کلاستر شده را و ستون راست تاثیر W را با کاهش آن نشان میدهد

بطوریکه c یک مجموعه ای از پیکسل هایی در کلاستر است، i هم شمارنده تکرار است، r شاخص کلاستر و مطابقت دارد با نمونه اولیه و سائز $[C]$ مقدار پیکسل ها در کلاستر C است.

در مرحله افزاز هر پیکسل بوسیله بردار ویژگی f_n مشخص می شود و برای کلاستر $C_r(i+1)$ با نمونه اولیه $P_r(i)$ تعیین می شود. پس از آن نمونه اولیه $P_r(i+1)$ بعنوان متوسط دیتای کلاستر ها دوباره محاسبه می شود. هر یک از دو مرحله متناوبی تا زمانی که تغییرات زیادی در آنها اتفاق نیفتد به مجموعه ای از خطاهای کوانتیزه شده ای تبدیل می شوند. در این مرحله یک مینیمم محلی از مجموع خطاهای کوانتیزه بدست می آید.

فرض کنیم یک حالت متحرک پیوسته ای از کلاستر ها را داشته باشیم، تکنیک های پیش گویی کننده می توانند برای کاهش قابلیت اطمینان بودن تخمین حرکت بکار روند. برای هر نمونه اولیه حالت X, Y در فریم بعدی با فیلتر استاندارد کالمن، ویژگی های رنگ غیر قابل تغییر نگه داشته می شود. از آنجایی که گونه های مختلفی از اجسام در صحنه های توافیکی یک مدل جنبش شناسی فیلتر کالمن پیشرفته تری را انتخاب می کنند [۱]، حرکت در امتداد محور های X, Y با جدا سازی حرکت های X, Y با دو فیلتر جداگانه پیش بینی می شود. فرض می شود که کلاستر متحرک شتاب ثابت باشد. برای محاسبه تغییرات نا چیز در سرعت، شتاب پیوسته زمانی با نویز سفید مدلسازی می شود. معادله حالت گسسته با پیروی نمونه برداری T عبارتست از :

$$s(k+1) = As(k) + w(k) \quad (2)$$

$$A = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \quad (3) \quad \text{با}$$

$$Q = E(w(k)w(k)^T) \quad (4)$$

$$= \begin{pmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ \frac{1}{2}T^2 & T \end{pmatrix} \sigma_w^2$$

معادله مقدار برای حالت یک بعدی :

$$p(k) = Cs(k) + n(k) \quad (5)$$

با

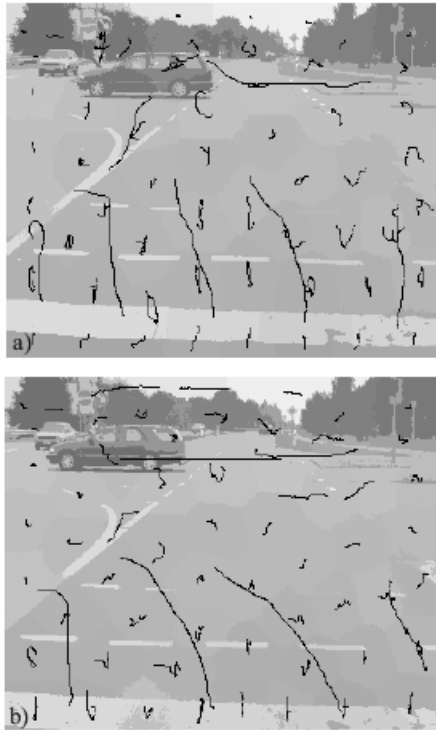
$$C = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad (6)$$

$$E(n^2(k)) = \sigma_n^2 \quad (7)$$

۱-۴- آشکار سازی شیء بوسیله قطعه بندی حرکت

روش پیشنهاد شده برای قطعه بندی، کلاستر های مجاور را با خط سیر های مشابه نسبت به اجسام فرضی ترکیب می کند. آن بر پایه این فرض است که در حین حرکت تصویر، یک کلاستر می تواند بوسیله مرکز ثقل کلاستر و کلاستر های میانی برای جسم مشابهی که خط سیر های مشابهی دارند تقریب زده شود. این فرضیات در بیشتر دنباله های تصویری قابل قبول هستند و در اینجا مورد بررسی قرار می گیرد. ولی این موارد در حضور اجسام غیر سخت ، اجسام چرخشی در مختصات نوری یا اجسام بزرگ ، که در طول مختصات نوری می چرخند معتبر نیستند.

در قطعه بندی حرکت تنها کلاستر هایی را در نظر می گیریم که شرایط زیر را برآورده کنند : طول می نیمم و قابلیت اطمینان خط سیر، شرط اول کلاستر هایی که متناظر با اجسامی در فاصله دوری از دور بین هستند یا نسبت به مشاهده کننده حرکتی ندارند را حذف می کند.



شکل (۳) تقریب عرض خیابان. خط سیرهای کلاسترها با خطوط سیاه رسم شده اند. با پیش بینی مان ماشین متحرک قابل ردیابی نیست. ولی در (b) با بکار بردن فیلتر کالمن این مشکل حل شده.

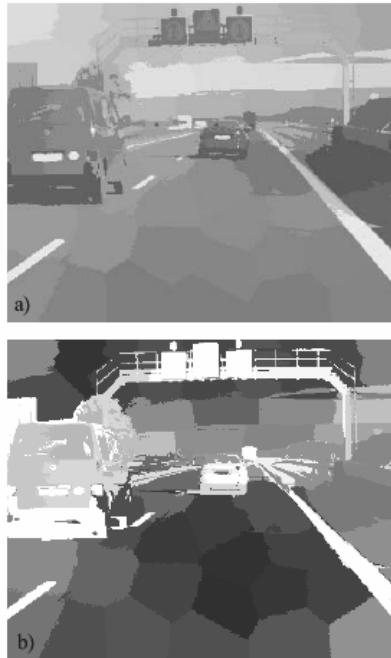
شرط دوم برای مطابقت با مشکل مشهور مان بکار می رود، بطور یکنواخت در بخش های مختلف رنگی (بطور مثال آسمان، جاده و ...) که در مجموعه ای از کلاسترهای مشابه جزء بندی شده است، قابلیت اطمینان کمتری را اختصاص دهیم. برای خط سیر این کلاسترها فاصله میان کلاسترهای موقعیت- رنگ در فضای ویژگی را محاسبه می کنیم. اندازه قابلیت اطمینان خط سیر، بطور خطی با فاصله مرکز ثقل کلاستر در نزدیکترین همسایه افزایش می یابد. یک مثال برای اندازه گیری قابلیت اطمینان در شکل (۴) نشان داده شده است.

شکل (۴- a)، نتایج کلاسترینگ حالت - رنگ فضای ویژگی و شکل (۴- b) اندازه قابلیت اطمینان را برای هر کلاستر را نشان می دهند، مقادیر روشن قابلیت اطمینان بالا را نشان می دهد. اگر در صفحه تصویر دو کلاستر $c1$, $c2$ مجاور هم باشند و شرایط مذکور را برآورده کنند شباهت میان خط سیر هایشان بوسیله اندازه زیر تخمین زده می شود:

$$\rho(c_1, c_2) = \left(1 - \frac{|l_{c1} - l_{c2}|}{l_{c1} + l_{c2}} \right) \cdot \frac{\sum_{k=K}^N (x_{c1}(k) - \bar{x}_{c1})^T (x_{c2}(k) - \bar{x}_{c2})}{\sqrt{\sum_{k=M}^N \| (x_{c2}(k) - \bar{x}_{c2}) \|^2 \sum \| (x_{c1}(k) - \bar{x}_{c1}) \|^2}} \quad (1)$$

$$\bar{x} = \frac{1}{1+N-M} \sum_{k=M}^n x(k) \quad (9)$$

$$l = \sum \|(x(k+1) - x(k))\| \quad (10)$$



شکل (۴) نمایش اندازه قابلیت اطمینان خط سیر ها . نتایج کلاسترینگ در (a) و قابلیت اطمینان در (b) نشان داده شده است. نقاط سفید قابلیت اطمینان بالار دارند.

موقعیت مرکز ثقل کلاستر در فریم k بصورت $x(k) = (x(k), y(k))^T$ ، که بازه $[m, n]$ یک پنجره زمانی برای انطباق خط سیر هاست، می باشد. اولین بخش معادله (۸) اختلاف را در طول محاسبه می کند. بخش دوم همبستگی خطی میان خط سیر هاست. اگر خط سیر ها موازی باشند همبستگی ۱ است، و اگر بر هم عمود باشند صفر است. در حالت ضد چرخشی خط سیر ها (-1) است.

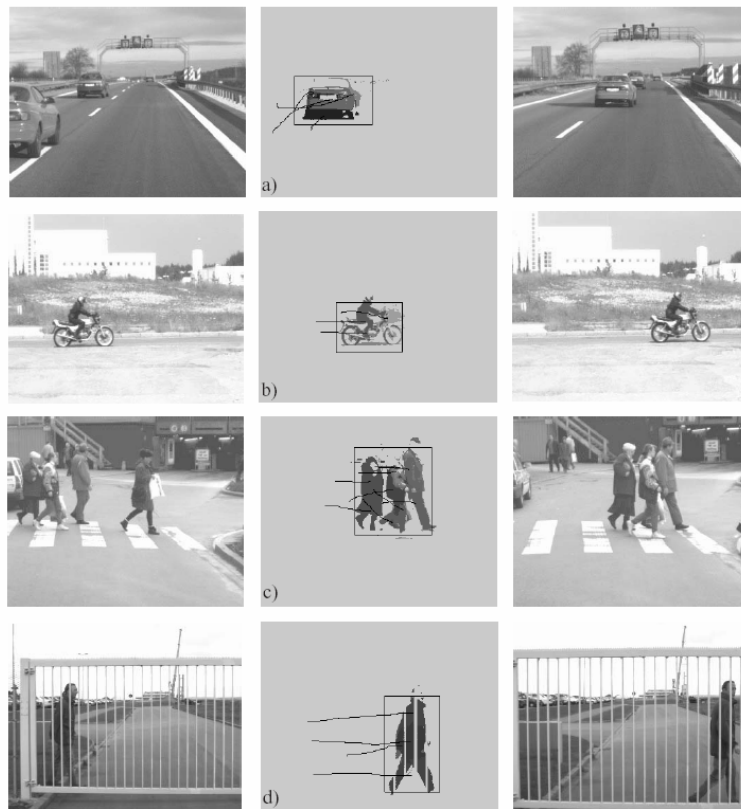
اگر $\rho(c_1, c_2)$ از آستانه ρ_{\min} زیادتر شود، هر دو کلاستر c_1, c_2 در جسم فرضی با هم ترکیب می شوند.

۱-۵- نتایج

نتایج حاصله از صحنه های ترافیکی بدست آمده در بزرگراه ها و شهر ها مورد بررسی قرار می گیرند. یک دور بین CCD که به یک ویدئو رکورد در دیجیتالی وصل شده (Y:U:V, 720×576 pixels, 25 frames/sec) که تصاویر با آن گرفته شده را مورد استفاده قرار می دهیم، دقت تصاویر به نصف مقادیر در سطر و ستون کاهش یافته (pixels ۲۸۸ × ۳۶۰) و از تکنیک های کوانتیزه کننده تقسیم برداری استفاده می کنیم، به منظور رسیدن به انحراف مشابه عدد پیش فرض انتخابی برای کلاستر ها باید به 2^n نزدیک باشد. در آزمایش ما برای پوشش مناسب جزئیات مناسب تصویر ۱۲۸۰ کلاستر مناسب هست. در کلاستر بندی k -means تنها یک تکرار مناسب است. و بطور تقریبی برای یک کامپیوتر پنتیوم II 300MHz، ۱۰۴ ثانیه زمان می برد. آخرین ۵ نقطه خط سیر ها در قطعه بندی مورد استفاده

قرار می گیرد و طول مینیمم خط سیر بالای ۵ فریم برای یک مجموعه ۱۰ پیکسلی است. آستانه ρ_{min} برابر ۰,۹۵ می باشد.

شکل (۵) نتایج الگوریتم بکار رفته در ۴ دنباله را نشان می دهد. در دنباله a ، ماشین مورد آزمایش ما با سرعت تقریبی $80 km/h$ ، حرکت می کند، و از سمت چپ ما عبور می کند و سرعتی برابر $100 km/h$ دارد. ۱۰ فریم از زمانی که ماشین وارد زمینه می شود طول می کشد تا آشکار شود. دلیل دیر آشکار شدن آن اینست که هنگامی که ماشین وارد زمینه می شود تعدادی از کلاسترها از زمینه جدا می شوند. در نتیجه خط سیر ما از حرکت تصویر اصلی متفاوت است. دنباله (b) از یک دوربین ساکن در عرض خیابان گرفته شده است. و عبور یک موتور سیکلت را نشان می دهد. در دنباله (c) ماشین به آهستگی به یک پل عابر نزدیک می شود، پل عابر با اجسام فرضی ترکیب شده، چرا که به آنها نزدیک است و بطور تقریبی سرعت مشابهی دارند. دنباله (d) قدرت کارایی الگوریتم را در حالتی از انسداد جزئی و تغییرات شکل را نشان می دهد.



شکل (۵) تصویرهای چپ و راست در هر ستونی از شروع و پایان دنباله اولیه گرفته شده اند. تصاویر موجود در ستون دوم اجسام آشکار شده هستند. خطوط سیاه موجود در تادرهای کناری اجسام فرضی و خط سیر کلاسترها می باشد.

بطور جالب توجهی، این الگوریتم، عابرهای گذرنده از آنجا را نیز آشکار می کند. تشابه در رنگ باعث می شود هم پاها و هم دست های عابرها با هم در یک کلاستر ترکیب شوند. خط سیر چنین کلاسترهایی منجر به حرکت مرکز ثقل پاها و دست ها که شبیه بخش های دیگر بدن است، می شود. ترکیب دست ها و پاها در کلاستر، فرض قبلی مان را بطور تقریبی نقض می کند. ولی با این الگوریتم، می توانیم عابرها بدون مدل سازی حرکت انسانی قطعه بندی کنیم.

در این روش یک روش تازه برای آشکار سازی اجسام متحرک در دنباله ای از تصاویر رنگی که با یک دور بین متحرک گرفته شده اند، ارائه دادیم. الگوریتم حرکت تصویر را بوسیله ردیابی کلاسترهای تعیین شده در فضای ویژگی حالت - رنگ تخمین می زند. هر تصویر به تعدادی مشخص از کلاسترها بوسیله گروههایی از پیکسل هایی که در رنگ و موقعیت مشابه هستند تقسیم می شوند. برای رسیدن به کلاسترهای مشابه در سراسر زمان، کلاستر بندی هر تصویری بر پایه نتایج کلاستر بندی های تصاویر قبلی قرار می گیرد. در این متن، برای پیشگویی تغییرات دینامیک در موقعیت های کلاستر از فیلترهای کالمن استفاده می شود. بالاخره قطعه بندی حرکت بوسیله ترکیب کلاسترهای مشابه با حرکت یک تصویر مهم و خط سیرهای مشابه مجاور یک جسم مجازی تولید می کند، زمانیکه در صحنه های ترافیکی الگوریتم را تست می کنیم در بخش بندی اجسام متحرکی مانند ماشین ها، موتور سیکلت ها و حتی عابر پیاده ها موفقیت آمیز است.

۱-۶- مراجع

- [۱]- پردازش تصویر رقمی، رافائل سی گونزالس ریچارد ای. وودز، انتشارات دانشگاه فردوسی مشهد
- [2] Brend. Heilise "Motion-Based Object Detection and Tracking in Color Image sequences" IEEE Trans. 2001
- [3] J.H. Duncan and T.C. chon, " On the Deception of the Motion and the Coputation of Optical flow, " IEEE Trans. Patren Analysis mach Intell, Vol. PAMI-14, no. 3 PP-346-352, 1992.