



## مقایسه ردیابی حرکات چندگانه با ارائه مدل استنتاجی خودگردان

بهارک امین طینت<sup>۱</sup>

<sup>۱</sup> کارشناسی ارشد مهندسی کامپیوتر - نرم افزار، دانشگاه آزاد اسلامی واحد لاهیجان

bamintinat@yahoo.com

### چکیده :

در این مقاله، به بررسی ردگیری حرکات چندگانه با استفاده از فیلتر کالمن پرداخته شده است، که بر اساس ترکیب مدل پس زمینه هر پیکسل و مجموعه ای از مدل‌های پیش زمینه تک فرضی بر مبنای مدل کلی اندازه شیء، موقعیت، سرعت و توزیع رنگ می باشد. فیلتر کالمن، تخمین موقعیت هدف بر اساس اندازه گیری های قبلی، و تخمین موقعیت بر اساس اندازه گیری های فعلی را برای بدست آوردن تخمین فیلتر شده موقعیت هدف، ترکیب می نماید. این ترکیب بطوری انجام می گردد که تخمینی حاصل شود که مینیمم واریانس را دارا است. عبارت دیگر دارای بهترین دقت می باشد. سیستم، نزدیک به آهنگ فریمهای ویدئویی (ما بین ۲۴ تا ۴۰ فریم در ثانیه) کار میکند. در این جا، دو مدل که شامل ردیابی پرش های انسان و ردیابی موتورسوارها می باشد، بررسی می شود و یک مدل استنتاجی خودگردان پیشنهاد شده است که فاکتورهای مربوط به هر دو مدل را در خود دارد و با توجه به دریافت چند فریم اول از یک تصویر، الگوریتم ردیابی مربوط به تصویر مذکور را با توجه به بلاک دیاگرام های مدلها، انتخاب می نماید. سپس مقدار خطا بین ناحیه اندازه گیری شده و حقیقی در دو مدل محاسبه شده و میزان دقت روش مورد استفاده در مدل‌های مذکور نشان داده می شود.

### کلید واژه :

آشکار سازی حرکت، تحلیل دنباله تصاویر، ردگیری شیء متحرک، فیلتر کالمن، آستانه گیری.

## مقدمه :

سیستم بر روی ARGO که یک وسیله نقلیه خودکار می باشد، نصب گردیده است. این وسیله می تواند جاده را تعقیب نماید و در مسیر صحیح حرکت کند. الگوریتم آشکارسازی آنها بر مبنای چند پایه استوار است [۵]، این که وسائط نقلیه در حالت کلی متقارن هستند و توسط یک مستطیل محیطی با نسبت طول و عرض تا حدی معین، مشخص می گردند.

همچنین اتومبیلها در یک ناحیه مشخص از تصویر واقع می شوند. روش کار به این صورت است که ابتدا با استفاده از موقعیت جاده و محدودیتهای پرسپکتیو، ناحیه مورد علاقه که احتمال دارد اتومبیل در آن جا باشد، معین می شود. به دلیل وجود انعکاسات نوری احتمالی در یکی از طرفین اتومبیل موجود در صحنه، تنها استفاده از سطوح خاکستری جهت یافتن تقارن، نتیجه درستی نمیدهد. زیرا از تقارن در لبه های افقی و عمودی نیز استفاده می شود. پس از آن برای تشخیص مستطیل محیطی جستجو انجام می شود. ردگیری با محاسبه همبستگی بین محتویات مستطیل محیطی در فریم قبلی، با فریم فعلی انجام می شود. از جمله سیستم های پیاده سازی شده می توان به سیستم Koller و همکارانش [۶] اشاره نمود. مزیت استفاده از روش آن که یک ارائه مبتنی بر کانتور از شیء، می باشد، کمتر بودن پیچیدگی محاسباتی نسبت به یک ارائه مبتنی بر ناحیه، است. با این وجود عدم توانایی در تقطیع اشیائی که بطور جزئی رویهم افتادگی پیدا کرده اند وجود خواهد داشت. اگر در هنگام رویهم افتادگی بتوان برای هر شیء، یک کانتور جدید ایجاد نمود، حتی در صورت رویهم افتادگی نیز می توان اشیاء را ردگیری نمود [۷]. با این وجود آغاز یک کانتور جدید، قسمت پیچیده این روش می باشد. Zhong [۸] و دیگر همکارانش با استفاده از محاسبه تفاضل بین دو فریم، جهت ردگیری اشیاء متحرک، کلیشه های تغییر شکل پذیر را برای تطبیق اشیاء ایجاد کردند. در این سیستم ابتدا یک ناحیه بصورت دستی برای ردگیری انتخاب می شد سپس کلیشه به نواحی دارای گرادیان بالا، و مرزهای حرکت جلب شده و با تشابه خود با سطح خاکستری در فریمهای بعدی تغییر شکل میداد. یکی از مزایای آن، وجود اطلاعات سراسری و کلی می باشد که این سیستم را نسبت به ضعفهای ویژگیهای تصویر مقاومتر می نماید. مشکلی که در این سیستم وجود دارد، محدودیت روش

بینایی ماشین یکی از زمینه های پر اهمیت علم کامپیوتر می باشد و سالهاست که محققان به پیشرفت و توسعه آن پرداخته اند. مهمترین هدف بینایی ماشین، شبیه سازی عملکرد سیستم بینایی انسان است و شاید به جرات بتوان گفت که بزرگترین هدف بینایی ماشین ساختن یک ماشین دارای بینایی همه منظوره می باشد. از چنین ماشینی انتظار می رود که بتواند خود را در محیط های دلخواه هدایت کند و اشیاء را در مسیر خود شناسایی نماید و به مقصدی که باید برسد، خود را برساند در واقع امروزه فاصله بینایی ماشین از بینایی انسان بسیار زیاد است و هنوز در این زمینه بسیار جای تحقیق و فعالیت دارد. علاقه به روش های پردازش تصویر از دو محدوده کاربردی اصلی نشأت می گیرد که عبارتند از: بهبود اطلاعات تصویری به منظور تعبیر انسانی، و پردازش داده های صحنه برای ادراک ماشینی مستقل [۱]. در مجموعه الگوریتمهای تصویری ارائه شده طی چند دهه گذشته، هدف عمده، یافتن روش هایی است که قادر به ارتقای اطلاعات تصویری برای تعبیر و تحلیل انسان باشد. از جمله پیشینه ردگیری که توسط محققان صورت گرفته است میتوان به موارد زیر اشاره نمود:

Malik [۲]، سیستمی را جهت صحنه های ترافیک، نظیر تغییر خط حرکت، اتومبیلهای متوقف شده و شمارش کلی اتومبیلهای موجود در صحنه ارائه نمود، که مبتنی بر فیلتر کالمن می باشد و با استفاده از تقطیع حرکت با تفاضل بین زمین و ارائه یک مدل حرکت، برای حرکت اتومبیلها توانست بر مسئله نویز غلبه کند.

Coifman و همکارانش [۳]، آشکارسازی و ردگیری وسائط نقلیه در شرایط نامساعد مانند تغییرات نور، رویهم افتادگی تصاویر اتومبیلها و غیره را بررسی نمودند. در روش آنها از گوشه بعنوان نقاط ویژگی استفاده می شود. بعبارت دیگر ویژگیها نواحی هستند که شدت سطوح خاکستری در آنها بیشتر از یک جهت تغییر کند. آنها از فیلتر کالمن برای ردگیری استفاده کردند. اگر ویژگی یافت شود، فیلتر کالمن به روز در می آید در غیر این صورت این ویژگی، دیگر ردگیری نمیشود. Bertozzi و همکارانش یک سیستم برای آشکارسازی و ردگیری وسائط نقلیه ارائه کردند [۴]. این

تقطیع آنست که بدلیل استفاده از تقطیع مبتنی بر فیلتر سازی زمانی، محیط یا پس زمینه ثابت مورد نیاز می باشد.

### ۱- فیلتر سازی کالمن :

فیلتر سازی کالمن، یک روش تخمین پارامتر بر مبنای مشاهدات نویزی می باشد. در سیستمهای ردگیری متعددی، از این فیلتر جهت تخمین پارامترهای اشیاء متحرک استفاده شده است. پارامترهای تخمین زده شده توسط فیلتر کالمن، معمولاً موقعیت و سرعت اشیاء متحرک می باشد. یک مدل حرکت با سرعت ثابت، ممکن است وجود داشته باشد یا ممکن است هدف در هر لحظه شتاب داشته باشد و یا بچرخد و سرعت یک هدف در همه زمانها ثابت نمی باشد.

کالمن، فیلتر جواب مساله را بصورت روابط (۱-۱) ارائه نمود.

$$X_{n+1,n} = X_{n,n-1} + h_n / T (y_n - X_{n,n-1}) \quad (1-1)$$

$$X_{n+1,n} = X_{n,n-1} + T \dot{X}_{n+1,n} + g(y_n - X_{n,n-1})$$

سرعت شیء در لحظه n

موقعیت واقعی شیء

نشاندهنده موقعیت تخمین زده شده هدف

در لحظه n+1

### ۲- انواع ردیابی های مورد بررسی :

۱- ردیابی پرش های یک انسان که به صورت تصادفی در جهت های مختلف، صورت می گیرد. در این حالت، از روی میانگین شدت سطوح خاکستری و یا توزیع رنگ و همچنین محاسبه مساحت هدف مورد نظر، می توان تخمین وضعیت بعدی هدف نسبت به وضعیت قبلی آن در فریم قبلی را محاسبه نمود.

۲- ردیابی دو وسیله نقلیه (موتورسیکلت)، که توسط یک دوربین ثابت، که درون یک ماشین کار گذاشته شده است، صورت می گیرد. در این ردیابی، از حالت اولیه (در فریم اولیه)، تا چند فریم متوالی، وجود چراغ جلوی موتور سیکلت ها، در توزیع رنگ و شدت سطوح خاکستری وسائط نقلیه، تاثیر گذاشته و همچنین دور نمای هر یک از آنها، در محاسبه صحیح مساحت آنها تاثیر منفی می گذارد. چراکه یکی از آنها، در موقعیت دورتری از دیگری قرار گرفته است و هرچه به فریم بعدی نزدیک تر می شویم موتور سوارها به دوربین نزدیک تر شده و طبعاً، از لحاظ مساحت و اندازه بزرگتر از فریم های قبلی خود تصویر برداری می شوند.



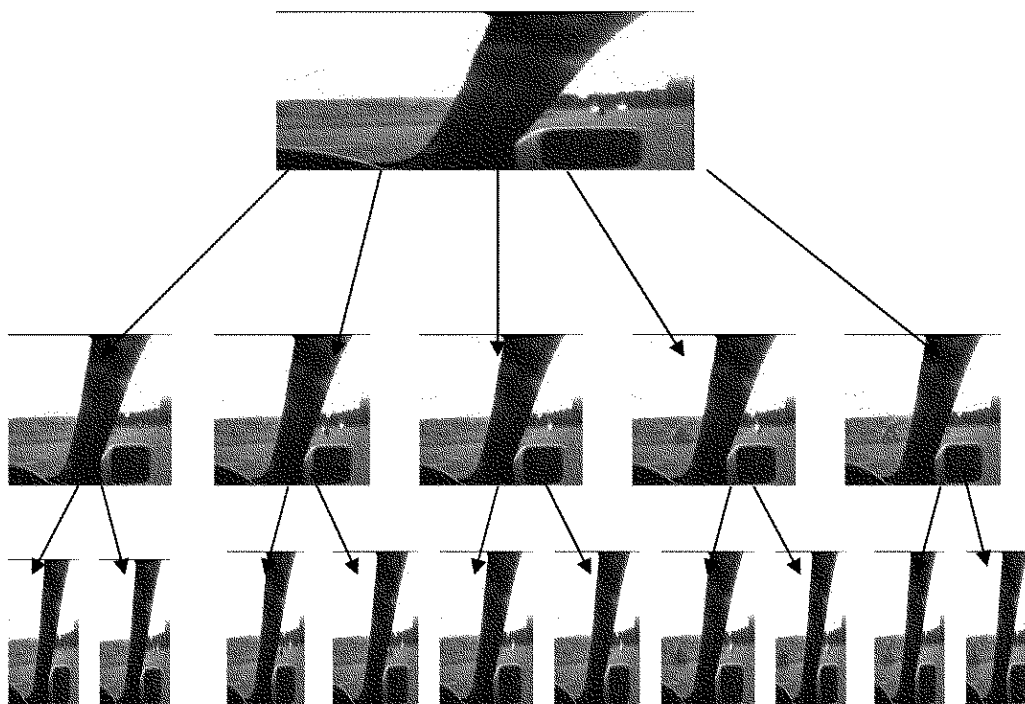
شکل ۱-۲: یک نما از پیاده سازی مطالب بالا می باشد.

راه حل :

لحظه پنجم، برای حل مسئله استفاده می شود. فریمهایی که در آنها هدف ها (موتورسیکلت ها)، از لحاظ مساحت و میانگین شدت رنگ به یک صورت بوده در یک دسته قرار می گیرند و با توجه به دسته بندی آنها و با توجه به محاسبه میانگین رنگ،

فریم بعدی، پیشگویی نمود.

محاسبه نزدیکترین فاصله هدف در فریم قبلی نسبت به فریم بعدی انجام شده و می توان احتمال وجود هدف موردنظر را در



شکل ۲-۲: یک نمونه کلاسه بندی از فریم های ردیابی موتور سیکلت ها را نشان می دهد.

نقلیه، توسط فیلتر کالمن را نشان میدهد. با توجه به میانگین نواحی بدست آمده، مقدار درصد خطای بین این دو ناحیه محاسبه گردیده است.

### ۳- جداول نواحی قابل پیش بینی :

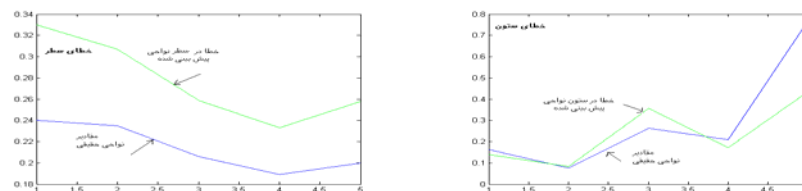
جدول زیر نواحی اندازه گیری شده و نواحی با

مقادیر پیش بینی شده در فریم های مختلف ردیابی وسائط

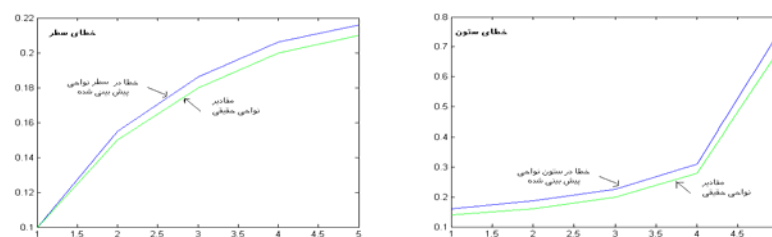
جدول ۳-۱: محاسبه مقدار خطا در ردیابی موتور سوارها نسبت به نواحی اندازه گیری شده

شماره فریم	نواحی اندازه گیری شده (میانگین (ستون، سطر))	نواحی پیش بینی شده (میانگین (ستون، سطر))	مقدار خطا نسبت به نواحی اندازه گیری شده ستون، سطر
فریم ۹	۱-شیء (۸۸و۱۴۵،۵) ۲-شیء (۹۰و۱۳۳)	۱-شیء (۹۷،۵و ۱۶۹،۵) ۲-شیء (۹۷،۵و ۸۱،۵)	۰،۱۶ ۰،۰۸
فریم ۲۲	۱-شیء (۸۱و۱۴۷،۵) ۲-شیء (۸۱،۵و۱۳۰)	۱-شیء (۱۰۶،۵و۱۷۵،۵) ۲-شیء (۱۰۶،۵و ۸۱)	۰،۱۸ ۰،۳۷
فریم ۲۴	۱-شیء (۹۲،۵و ۱۳۲،۵) ۲-شیء (۹۷،۵و۶۵،۵)	۱-شیء (۱۰۹،۵و۱۶۲،۵) ۲-شیء (۱۰۹،۵و ۷۴)	۰،۲۲ ۰،۱۲
فریم ۲۶	۱-شیء (۹۰و۱۲۵)	۱-شیء (۷۲،۵و۸۸،۵)	۰،۱۹
فریم ۳۷	۱-شیء (۹۲،۵و۵۰،۵)	۱-شیء (۷۵و۸۸،۵)	۰،۲۱

#### ۴- نمودار خطای بین نواحی حقیقی و پیش بینی شده :



نمودار ۴-۱: نمودار خطای موجود در ردیابی پرش های انسان نسبت به سطر وستون



نمودار ۴-۲: نمودار خطای موجود در ردیابی حرکت موتورسوارها نسبت به سطر وستون

#### ۶- معیار هماهنگی متغیر:

توزیع های مدل با معیار متغیر برای هر پیکسل به روز می شوند. (تقسیم صحیح معیار جریان تصویر ورودی) پیکسل های ثابت، با معیار پایین تری نسبت به پیکسلهایی که بعنوان پیش زمینه دسته بندی شده اند، به روز می شوند. این معیارها بدین صورت می باشند.

$$R_{input} / N_{bgd}, \quad \text{where} \quad N_{bgd} < N_{max}$$

$$R_{update} = \{$$

$$R_{input} / N_{max}, \quad \text{otherwise} \quad (1-6) \text{ رابطه}$$

$R_{update}$  = Update rate  
 $R_{input}$  = Input Framerate  
 $N_{bgd}$  = No. of consecutive background frames  
 $N_{bgd}$  = Threshold on frame rate division (typically 30 frames / 1 sec)

#### ۵- بررسی مدل در صورت وجود اغتشاش:

یک پیکسل، توسط کمترین فاصله  $D$  با مدل، مرتبط است، اگر این فاصله زیر آستانه معین باشد (معمولا  $2/5$ ). اگر این فاصله بالای آستانه، اما زیر آستانه دوم باشد (معمولا  $4/5$ )، اندازه رنگ پیکسل، با ترکیب گاوسی رنگ در مدل و با گرفتن حداقل فاصله اندازه رنگ از هر میانگین ترکیب گاوسی مقایسه می شود. با توجه به اینکه مدل پیش زمینه ما، شامل نمونه هایی است برای: موقعیت، اندازه، سرعت، توزیع رنگ.



شکل ۵-۱: نمایش تکه شدن پیش زمینه

تغییرات پس زمینه، جهت حرکت شی، را روی  $k$  فریم اولیه ورودی، بررسی نموده و با توجه به شباهت آنها به هر کدام از مدل‌های دیاگرام‌های موجود، مربوط به آنها را انتخاب نموده و تابع مناسب را برای ردیابی حرکات چندگانه مورد بررسی، اجرا می‌نماید. خروجی حاصله، یک روش مناسب جهت ردیابی صحیح با استفاده از فیلتر کالمن و با خطای می نیمم می باشد. گراف ۱-۷ بیانگر مطالب بالا است. بطور مثال این مدل در ۱۰ ثانیه، ۳۰۰ فریم اول را از لحاظ خصوصیات مشخص شده، مورد پردازش قرار داده و ارزیابی می‌نماید. هر یک از مدل‌هایی که این خصوصیات در آنها صدق نماید، تصمیم‌گیری لازم در مورد اجرای الگوریتم مربوطه انجام می‌شود.

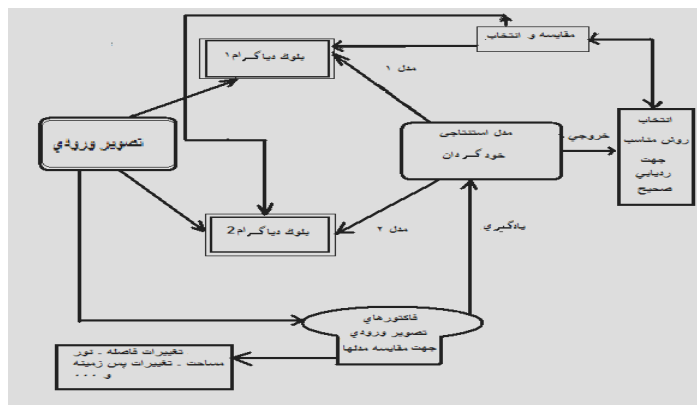
این معیارها، صورتی قابل توجه به کاهش هزینه‌های محاسباتی روش کمک می‌کند. باید توجه داشت که طرح شرح داده شده، تغییرات بین مدل‌های پیش زمینه و پس زمینه را که به صحنه وارد می‌شوند، از هم متمایز نمی‌کند. نتیجه این است که میزان هماهنگی بالایی، برای پیکسل‌های شی پیش زمینه بکار می‌رود. در نتیجه ممکن است که خودروهای با سرعت پایین، جزء مدل پس زمینه به شمار آیند.

## ۷- مدل استنتاجی خودگردان :

در این قسمت مدلی پیشنهاد می‌گردد که می‌تواند تغییرات یکسری فاکتورهای خاص، مانند فاصله هدف از دوربین، شدت روشنایی شی، اندازه شی نسبت به دوربین،

جدول ۱-۷: فاکتورهای مورد استفاده مدل استنتاجی خودگردان

فاکتورها	مدل ۲	مدل ۱
فاصله هدف از دوربین	متغیر	ثابت
شدت روشنایی	متغیر	ثابت
اندازه شی نسبت به دوربین	متغیر	ثابت
تغییرات پس زمینه	متغیر	ثابت
جهت حرکت	معین	نامعین



گراف ۱-۷: نمایش گراف مدل استنتاجی خودگردان

## فهرست منابع

- [1] Gonzalez, R.C., Woods, R.E., "Digital Image Processing", 1987.
- [2] Z. Kim and J. Malik, "High-Quality Vehicle Trajectory Generation from Video Data Based on Vehicle Detection and Description", *Proc. IEEE Intelligent Transportation Systems Conference*, 2003.
- [3] B. Coifman, D. Beymer, P. Mclauchlan, and J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transportation Research* 6C, pp.271-288, 1998.
- [4] A. Broggi, M. Bertozzi, A. fascioli, and G. Conte. Automatic Vehicle Guidance: the experience of the ARGO Autonomous Vehicle world Scientific Publishing Co. Inc., River Edge, 1998.
- [5] M. Bertozzi and A. Broggi, "Gold: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. on Image Processing*, vol. 7, pp. 62-81, 1998.
- [6] D. Koller, K. Danilidis and H.H. Nagel. "Model-based object tracking in monocular image sequences of road traffic". *International Journal of Computer Vision*, 10:257-281, 1993.
- [7] T. Huang, D. Koller, J. Malik, G. Ogasawara, S. Russell and J. Weber. "Automatic symbolic traffic scene analysis using belief networks". In *National Conference on Artificial Intelligence*, AAAI Press, 1994.
- [8] Zhong, Y.[Yu], Jain, A.K.[Anil K.], and Dubuisson-Jolly, M.P.[Marie-Pierre], "Object Tracking Using Deformable Templates", May 2000.

## نتیجه گیری :

در این مقاله، روشهای مختلف ردیابی اشیاء متحرک بررسی گردید، که بر پایه فرض ثابت بودن دوربین می باشد. نتایج کار در بخش پیاده سازی، نشاندهنده ۸۰٪ دقت در ردیابی، و ۲۰٪ خطا بین ناحیه حقیقی و تخمین زده شده با استفاده از فیلتر کالمن می باشد. در نمودار خطای موجود در ردیابی پرش های انسان نسبت به سطر وستون، در هنگام تغییر سریع مختصات (ستون)، ناهمگنی ما بین مقدار پیش بینی شده و حقیقی بوجود می آید، ولی زمانی که این تغییرات سریع مختصاتی قابل رویت نباشد، مقدار خطای مذکور مینیمم می گردد.

برای عمومی سازی کاربرد، دو روش اشاره شده، با استفاده از یک مدل استنتاجی ترکیب و پیشنهاد شده است. با توجه به تحقیقات انجام شده، مبین است که در صورت وجود تغییرات آب و هوایی، این سیستم به خوبی جواب می دهد و در صورت وجود اغتشاش بعلت بکارگیری الگوریتم حذف نویز، که با بررسی وجود آن در همسایگی شیء و با استفاده از یک آستانه معین، نویز حذف میگردد، سیستم مقاوم است و دقت ردیابی مطابق با آنچه گفته شد، می باشد. و در صورت عدم بکارگیری الگوریتم حذف نویز، و وجود اغتشاش در تصویر، تخمین خطا بین ناحیه حقیقی و اندازه گیری شده به ۱۰٪ افزایش می یابد که تفاوت کار این سیستم با سیستم ارائه شده Koller را نشان میدهد که در آن تشخیص و ردگیری بطور یکجا در یک حوزه زمانی و مکان قرار نداشت بطوریکه نویز می توانست باعث تشخیص اشتباه اشیاء گردد. ثابت بودن دوربین در این سیستم، پیچیدگی محاسباتی را کاهش می دهد با توجه به آن، پیچیدگی در سیستم موتور سوارها  $O(\log n)$  در مبنای  $m$  می باشد که  $n$  تعداد فریم های هر دسته و  $m$  تعداد دسته بندی های مورد نظر می باشد در مقابل آن، سیستم ارائه شده Bertozzi است که بعلت نصب دوربین بر روی یک وسیله نقلیه خودکار جهت ردگیری وسائط نقلیه و با توجه به حفظ فاصله قانونی بین اتومبیلها و تنظیم سرعت خود نسبت به سرعت جاده و محیط، محاسبات پیچیده ای دارد.