

# ارزیابی روش‌های تفریق زمینه بر پایه الگوریتم سیگمادلتا به منظور تشخیص حرکت

محمدرضا مهویدی<sup>(۱)</sup> - وحید قدس<sup>(۲)</sup>

(۱) کارشناس ارشد - گروه برق و مکترونیک، واحد سمنان، دانشگاه آزاد اسلامی، سمنان، ایران

(۲) استادیار - باشگاه پژوهشگران جوان و نخبگان، واحد سمنان، دانشگاه آزاد اسلامی، سمنان، ایران

تاریخ دریافت: ۱۳۹۶/۹/۲۰ تاریخ پذیرش: ۱۳۹۷/۴/۲۴

**خلاصه:** پردازش توالی تصاویر ویدئویی برای قسمت‌بندی اجسام دارای حرکت (پیشنما) از قسمت‌های ثابت (زمینه)، یک مرحله اساسی در بسیاری از کاربردهای بینایی ماشین به ویژه تشخیص حرکت می‌باشد. یکی از روش‌های مرسوم، بکار بردن رویکرد تفریق زمینه است که اجسام متحرک را از مقایسه هر فریم با فریم زمینه بدست آمده، ایجاد می‌کند. در این مقاله، به بررسی روش‌های تفریق زمینه بازگشتی مبتنی بر الگوریتم سیگمادلتا می‌پردازیم. الگوریتم تفریق زمینه یک تقریب بسیار سریع و ساده از زمینه فراهم می‌آورد و همچنین دارای این مزیت است که به منابع بسیار کمی از حافظه نیاز دارد. به دلیل غیرخطی بودن این الگوریتم، ویژگی جالب آن مقاومت زیاد در مقایسه با میانگین‌های بازگشتی خطی و هزینه محاسباتی بسیار کم است. اما، از طرف دیگر الگوریتم اصلی سیگمادلتا، در صحنه‌های پیچیده و شلوغ با اجسام دارای حرکت آهسته و یا موقتاً متوقف شده، آلوده می‌شود. همچنین، در این الگوریتم اثر روح و اثر روزنه‌ای به وضوح قابل مشاهده است. این مقاله به ارزیابی این الگوریتم و بررسی روش‌های تکمیلی و رویکردهای مختلف ارائه شده برای آن می‌پردازد. در این مقاله، تمام الگوریتم‌ها به صورت گام به گام اجرا و پیاده‌سازی شده است. هدف این مکمل‌ها و رویکردها، رفع و یا کاهش معایب و مشکلات الگوریتم اصلی است. در انتها یک تحلیل کمی بین این رویکردها انجام می‌شود و بهبودهای انجام شده، مزایا و معایب هر الگوریتم مورد ارزیابی قرار می‌گیرد و مقایسه بین الگوریتم اصلی سیگمادلتا و سایر الگوریتم‌های مرتبط ارائه می‌شود.

**کلمات کلیدی:** تشخیص حرکت، تفریق زمینه، الگوریتم سیگمادلتا.

## Background Subtraction Techniques Evaluation based on Sigma=delta Algorithm for Motion Detection

Mohammadreza Mahvidi<sup>(1)</sup> – Vahid Ghods<sup>(2)</sup>

(1) MSc - Department of Electrical Engineering, Semnan branch, Islamic Azad University  
m.r\_mahvidy@yahoo.com

(2) Assistant Professor - Department of Electrical Engineering, Semnan Branch, Islamic Azad University,  
Semnan, Iran.  
v.ghods@semnaniau.ac.ir

Processing a video stream to segment moving objects (foreground) from the static scene (background) is a critical first step in many computer vision applications. One of the common methods is using background subtraction approach, which detects moving objects by comparing each frame with the obtained background frame. In this paper, we examine background subtraction algorithm based on sigma-delta filter. This algorithm provides a simple and very fast approximation of the median and has the advantage of having low memory requirements. The interest of this method lies in the robustness provided by the non-linearity compared to the linear recursive average, and in the very low computational cost. However in the basic sigma-delta algorithm, the background model quickly degrades in complex urban scenes because it is easily “contaminated” by slow-moving or temporarily stopped objects. And in this algorithm ghost effect and aperture effect is clearly visible. This paper is a review to this algorithm and various approaches and improvements proposed for it. In this paper, first basic sigma-delta and then its important approaches is described. The purpose of this approaches and improvements is to eliminate or reduce the defects and disadvantages of the main algorithm. In the end, a quantitative comparison between these algorithms is carried out and improvements and advantage and disadvantages of each algorithm are evaluated.

**Index Terms:** Motion detection, Background subtraction, Sigma-Delta algorithm.

نویسنده مسئول: وحید قدس، استادیار - باشگاه پژوهشگران جوان و نخبگان، واحد سمنان، دانشگاه آزاد اسلامی، سمنان، ایران،  
v.ghods@semnaniau.ac.ir

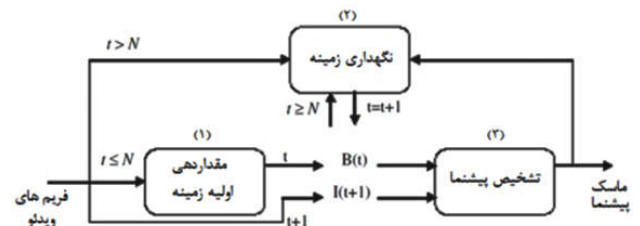
۱- مقدمه

سادگی می‌تواند در حوزه تبدیل کسینوسی گسسته<sup>۶</sup> (DCT) انجام پذیرد. الگوریتم میانگین جاری از چنین توالی ارزیابی خطی‌ای برای تخمین زمینه خود استفاده می‌کند. لذا این روش می‌تواند در حوزه‌ی DCT گسترش یابد که الگوریتم تخمین زمینه RADCT نامیده می‌شود [۲]. از دیگر روش‌های تخمین زمینه می‌توان از روش تفاوت آماری ساده<sup>۷</sup> نام برد. الگوریتم SSD نقاب تشخیص حرکت باینری خود را به وسیله‌ی پیدا کردن مقادیر میانگین و انحراف معیار که با توجه به مقادیر فریم‌های قبلی هر پیکسل محاسبه می‌شوند، تخمین می‌زند. تصویر زمینه از مجموعه‌ی فریم‌های قبلی در بازه زمانی  $[t_0, t_1]$  برای هر پیکسل از طریق محاسبه‌ی میانگین و انحراف معیار تولید می‌شود [۳].

در موضوع دسته الگوریتم‌های تفریق زمینه پارامتری، در ساده‌ترین حالت، فرض شده است که هر پیکسل زمینه می‌تواند با یک تابع چگالی تک کیفیتی مدل شود. این قالب الگوریتم شناخته شده به عنوان میانگین گوسی جاری<sup>۸</sup> (MoG) [۴] و [۵] می‌باشد، که یک الگوریتم بازگشتی است که یک تابع چگالی گوسی برای هر پیکسل گنجانده شده است. فیلتر میانه زمانی<sup>۹</sup> یک استراتژی معمول دیگر است که به منظور عملکرد بهتر نسبت به روش‌های بر اساس میانگین، ارائه شده است. تخمین زمینه برای هر پیکسل به صورت میانه همه مقادیر اخیر آن پیکسل تعریف شده است (در مورد نمونه غیر بازگشتی الگوریتم). فرض بر این است که یک پیکسل زمینه باید بیشتر از ۵۰ درصد دوره در نظر گرفته شده به وضوح قابل دیدن باشد (۶ تا [۹]). مخلوط گوسیها<sup>۱۰</sup> [۱۰] و [۱۱]، یک استراتژی پارامتری دیگر است که به طور گسترده استفاده می‌شود. یک تابع چگالی گوسی تنها برای هر پیکسل، برای رفع مشکل اجسام غیر ثابت زمینه، مانند درختان موج‌دار یا اجسام طبیعی دیگر کافی نیست. ایده MoG چندین مدل‌سازی زمینه برای هر پیکسل می‌باشد. زمینه بدست آمده تلاش می‌کند که شدت‌های روشنایی مختلف را که ممکن است برای هر پیکسل زمینه ظاهر شود، با استفاده از ترکیب n تابع چگالی گوسی مدل کند [۱۲].

دسته‌ی دیگر از الگوریتم‌ها در طبقه الگوریتم‌های غیر پارامتری قرار می‌گیرند. این الگوریتم‌ها زمانی مناسب‌ترند که تابع چگالی پیچیده‌تر فرض شده باشد یا به صورت پارامتری قابل مدل شدن نباشد، در حالیکه یک روش غیر پارامتری قادر به کار کردن با تابع چگالی‌های دلخواه می‌باشد. تخمین چگالی کرنل<sup>۱۱</sup> یک نمونه از روش‌های غیر پارامتری می‌باشد. این روش سعی به رفع مشکل MoG و روش‌های قبلی دیگر دارد. روش‌های قبلی قادرند صحنه‌های با رفتار نرم و تغییرات محدود مانند صحنه‌های با سیر تکاملی تدریجی را به طور مؤثرتری توصیف کنند. اگرچه در حضور صحنه‌های دینامیک با تغییرات سریع و ویژگی‌های غیر ثابت، زمینه نمی‌تواند به طور دقیقی با یک دسته از گوسین‌ها مدل شود. این روش با تخمین احتمالات در هر پیکسل با بسیاری از نمونه‌های اخیر به وسیله‌ی KDE بر این مشکل فائق آمده است [۱۳].

تشخیص اجسام متحرک در یک توالی از تصاویر یک بخش پردازشی مهم و حیاتی در بسیاری از کاربردهای ویدئویی مانند نظارت ویدئویی، شناسایی نظامی و تشخیص اهداف است. بسیاری از این کاربردها احتیاج به یک پایگاه محاسباتی کم مصرف، فشرده، سبک وزن و پرسرعت برای پردازش تصاویر به صورت بیدرنگ دارند. وقتی که دوربین ثابت است، یک دسته از روش‌هایی که معمولاً استفاده می‌شود تفریق زمینه می‌باشد. پایه و اساس این روش‌ها ساختن یک مدل از صحنه ثابت (بدون اجسام متحرک) که زمینه نامیده می‌شود، و سپس مقایسه کردن هر فریم از توالی با زمینه برای تشخیص نواحی دارای حرکت غیر معمول، که پیش‌نما (اشیاء متحرک) نامیده می‌شود، است [۱]. شکل (۱) حالت کلی مراحل تشخیص زمینه را نشان می‌دهد. در شکل (۱) مرحله شماره (۱) مقداردهی اولیه زمینه<sup>۱</sup> با استفاده از N فریم، برای ایجاد اولین تصویر زمینه بدون وجود اجسام متحرک می‌باشد. مرحله شماره (۲) تشخیص اجسام متحرک با تشخیص پیش‌نما<sup>۲</sup> است که به وسیله دسته‌بندی پیکسل‌ها به پیش‌نما یا زمینه از طریق مقایسه تصویر زمینه و فریم جاری انجام می‌شود. مرحله شماره (۳) نگهداری زمینه<sup>۳</sup> است که برای به‌روزرسانی زمینه در طول زمان انجام می‌پذیرد. مراحل ۲ و ۳ در طول زمان به صورت تکرار شونده اجرا می‌شوند [۲].



شکل (۱): فرآیند تشخیص زمینه. N تعداد فریم‌های استفاده شده برای مقداردهی اولیه زمینه می‌باشد. B(t) و I(t) به ترتیب زمینه و فریم جاری در زمان t می‌باشند [۲].

Fig. (1): Background detection process. N is the number of frames used to initialize the background. B(t) and I(t) are the current background and frame at time t, respectively [2].

در ابتدایی‌ترین حالت الگوریتم تفریق زمینه ساده<sup>۴</sup> (SBS) فرض می‌کند که زمینه مرجع موجود باشد. در این صورت هر فریم جاری به منظور استخراج پیش‌نما به سادگی از زمینه مرجع تفریق می‌شود [۳]. معایب روش SBS در الگوریتم میانگین جاری<sup>۵</sup> با استفاده از میانگین همه فریم‌های ویدئو ورودی برای تخمین زمینه، به جای استفاده از یک زمینه مرجع تنها، برطرف شده‌اند. لذا روش میانگین جاری بر معایب استفاده از یک زمینه مرجع فائق آمده است. علاوه بر این اگر یک جسم جدید به زمینه وارد شود، این جسم پس از محاسبات مشخصی در زمینه جدید به‌روزرسانی می‌شود [۳]. در اکثر روش‌های مدل‌سازی زمینه مشاهده شده است که آنها برای ساختن زمینه خود، یک توالی از ارزیابی‌های خطی را بکار می‌برند. این چنین توالی ارزیابی خطی‌ای به

استفاده شده است که برای تشخیص حرکت در تصاویر ویدئویی استفاده می‌شود. یک روش مدل کردن زمینه به نام (-SMSOM BM) [۲۵] در مرجع [۲۶] به ارائه روش یادگیری چندمنظر برای تشخیص پیش‌نما برای کنترل و نظارت ترافیک پرداخته است. نتایج استفاده از مدل کردن روش‌های چند-کیفیتی زمینه در مرجع [۲۷] آورده شده است که بر روی تصاویر HD به صورت بی‌درنگ کار شده است. در مقاله [۲۸] به تشخیص پیش‌نما در کنترل و نظارت ترافیکی با استفاده از (ORPCA) [۲۹] پرداخته شده است. جدا کردن اجزای پیش‌نما و زمینه از مسائل مهم بینایی ماشین است که در مرجع [۲۹] از مدل MAMR [۳۰] برای این مهم در کلیپ‌های ویدئویی استفاده شده است. تشخیص اشیا و دنبال کردن آنها با استفاده از روش (VTSM) [۳۱] در مرجع [۳۰] انجام شده است که یافتن سرعت نیز از اهداف آن است. در این مرجع به دنبال کردن و تعیین سرعت وسیله نقلیه پرداخته شده که در قضاوت درست در تصادفات حائز اهمیت است. در مرجع [۳۱] به تشخیص حرکت اشیا در ویدئوهای با نرخ بیت‌های مختلف با بازده مناسب به منظور نظارت پرداخته شده است. در آن مقاله از شبکه عصبی مصنوعی پخش شمارنده [۳۲] استفاده شده است.

به منظور احترام به حریم شخصی افراد، تشخیص حرکت و دنبال کردن افراد بر روی ویدئوهای رمزنگاری شده انجام می‌شود [۳۲]. در مقاله [۳۲]، ابتدا از ساختار جریان بیت‌ها و طول کلمات رمز، اطلاعات حرکت تخمین زده می‌شود. سپس الگوریتم به روزرسانی ناحیه [۳۳] به منظور کاهش خطای تشخیص حرکت استفاده می‌شود. مقاله [۳۳] به تشخیص شیء متحرک در تغییرات صحنه‌های متفاوت پرداخته است. سه نوع تطبیق با تغییر شرایط صحنه در آن لحاظ شده است که شامل حرکت پس‌زمینه، حرکت پیش‌نما و تغییرات شدت نور است. در این تصاویر ویدئویی، هم زمینه و هم شیء در حرکت هستند در حالی که سطح شدت نور نیز متغیر است. در آن مقاله از به روز رسانی زمینه شرطی صحنه استفاده می‌شود.

در این مقاله، تمرکز بر روی الگوریتم تفریق زمینه بازگشتی غیرخطی سیگمادلتا می‌باشد و اکثر مکمل‌ها و رویکردهای مختلف ارائه شده برای آن بررسی شده است، و در انتها یک مقایسه کمی بین آنها انجام می‌شود. در بخش ۲ به بیان الگوریتم سیگمادلتا در ساده‌ترین حالت آن و مکمل‌های ارائه شده برای آن می‌پردازیم. در بخش ۳ این الگوریتم‌ها به صورت کمی با یکدیگر مقایسه شده و نتایج ارائه می‌شود. در انتها، نتیجه‌گیری کلی بر اساس نتایج بدست آمده در بخش ۴ ارائه می‌شود.

## ۲-۲- الگوریتم تفریق زمینه سیگمادلتا

فیلترهای سیگمادلتا همانطور که از نام آن مشاهده می‌شود از دو مرحله سیگما و دلتا تشکیل شده است. معمولاً مرحله دلتا برای کارایی بهتر اول انجام می‌پذیرد. در مرحله دلتا، عبارت دلتا در ابتدا به صورت اختلاف بین نمونه و مدل زمینه محاسبه می‌شود. در مرحله سیگما، پالسهای دلتای انباشته شده زمینه واقعی را دنبال می‌کنند. وقتی که

یک نمونه تغییر یافته‌ی الگوریتم  $^{12}\text{MoG}$  به وسیله‌ی Porikly و Tuzel برای کار کردن با صحنه‌های دینامیک مطرح شده است [۱۴]. در این الگوریتم، هر پیکسل با یک دسته لایه‌های 3-D از گوسینهای چند متغیری بیان می‌شود. هر لایه سعی به گرفتن ظاهری مختلف از پیکسل مربوط را دارد. روش بیزین مطرح شده نه تنها اجازه تخمین میانه و واریانس برای هر لایه بلکه توزیع احتمال میانه و کواریانس را می‌دهد. یک ویژگی دیگر این الگوریتم این است که تنظیم و انتخاب پارامترها نسبت به MoG اصلی بسیار ساده شده است.

یک روش غیر پارامتری دیگر از الگوریتمی که بر اساس مدلی که کتاب کد [۱۳] نامیده می‌شود، ایجاد شده است [۱۵]. در این مورد، مدل زمینه برای هر پیکسل با تعدادی کلمه کد [۱۴] بیان شده است (بجای پارامترهای معرف تابع‌های احتمال)، که به صورت دینامیکی همراه با روشهای دسته‌بندی تدریجی بکار رفته‌اند. یک موضوع موازی برای درک این روش، مدل کردن رنگ می‌باشد.

هر زمان برای استخراج پیش‌نما از روشهای تفریق زمینه استفاده شود، مساله مهمی که حائز اهمیت است سایه اجسام در فریمها می‌باشد. این سایه معمولاً به عنوان بخشی از پیش‌نما در نظر گرفته می‌شود و باعث به وجود آمدن خطاهای بزرگی در استخراج پیش‌نما می‌شود. به منظور رفع مشکل این تشخیصهای اشتباه (سایه‌هایی که پیش‌نما در نظر گرفته می‌شوند)، از فرضیه لامبرتین که رنگ را به صورت حاصلضرب تابش و بازتاب در نظر می‌گیرد، استفاده شده است. الگوریتم تفریق زمینه بر اساس اعوجاج روشنایی و رنگ (BSBCD) [۱۵] از اعوجاج روشنایی و اعوجاج رنگ برای دسته‌بندی پیکسلها به ترتیب به پیش‌نما، زمینه، سایه و روشن استفاده می‌کند. یک پیکسل به وسیله‌ی مجموعه ۴ تایی  $[E_i, S_i, A_i, B_i]$  مدل می‌شود که  $E_i$  مقدار رنگ مورد انتظار،  $S_i$  انحراف معیار مقادیر رنگ و  $A_i$  و  $B_i$  تغییرات روشنایی و رنگ برای پیکسل  $i$  ام می‌باشند [۳].

یک دسته دیگر از روش‌ها استراتژی پیش‌بینی برای مدل کردن و پیش‌بینی حالت دینامیکی برای هر پیکسل را در نظر می‌گیرند. بعضی از آنها بر اساس فیلتر کالمن [۱۶] بنیان شده‌اند [۱۶] و [۱۷]، که مقادیر شدت و مشتقات مکانی برای تشکیل یک فضای تک حالت، برای دنبال کردن زمینه، با هم ترکیب شده‌اند. متناوباً، آنها می‌توانند بر اساس فیلتر وینر [۱۷] ایجاد شده باشند [۱۸]، یا بر اساس مدل‌های پیچیده‌تر مانند مدل‌های autoregressive [۱۹] و [۲۰] تولید شده باشند. در انتها می‌توان روش‌های بر اساس بیان حالت ویژه [۲۱] که به عنوان زمینه‌های ویژه شناخته شده‌اند، را ذکر کرد. در این روش اجسام جدید به وسیله مقایسه با یک تصویر که به وسیله حالت ویژه ساخته شده است، تشخیص داده می‌شوند.

مرجع [۲۲] به تشخیص و دنبال کردن خودکار هدف با در نظر گرفتن شرایط تغییرات شدت روشنایی، سد (بسته) شدن و حرکت دوربین می‌پردازد. در مرجع [۲۳] به تعقیب خودکار وسایل جراحی داخل بدن انسان که در عروق برای درمان استفاده می‌شود پرداخته شده است. از روش تجزیه پویا [۱۹] (DMD) برای کم کردن زمینه در مرجع [۲۴]

می‌رود. لذا  $V_t$  در این روش دارای بعد انحراف معیار زمانی است که به صورت یک فیلتر سیگمادلتا از اختلاف توالی  $\Delta_t$  محاسبه شده است. این واریانس یک معیار از فعالیت زمانی پیکسل‌ها فراهم می‌کند. از آنجایی که برای تشخیص حرکت پیکسل‌هایی که نرخ تغییرات آنها به مقدار قابل توجهی بیشتر از فعالیت زمانی آنها است مورد توجه هستند، فیلتر سیگمادلتا در توالی  $N$  تایی از اختلافات غیر صفر ( $\Delta_t(x) \neq 0$ ) استفاده می‌شود. در انتها، تشخیص  $D_t$  در سطح پیکسل به‌سادگی با مقایسه  $\Delta_t$  و  $V_t$  فراهم می‌آید [۱].

شکل (۲) نتیجه پیاده‌سازی الگوریتم را برای توالی تصاویر AVSS\_PV\_EVAL از دیتابیس i-LIDS [۳۲] نشان می‌دهد، که چگونه مدل زمینه و واریانس که هر دو بر اساس فیلتر سیگمادلتا ایجاد شده‌اند، کار می‌کنند. مدل زمینه، زمینه را کسر می‌کند، در حالیکه واریانس برای تشخیص پیش‌نما استفاده می‌شود. دیتابیس i-LIDS شامل یک مجموعه از ویدئوهای گرفته شده توسط دوربین‌های مداربسته براساس صحنه‌ها و شرایط لازم در نیازهای دولتی، می‌باشد. این ویدئوها به طور دقیق بیانگر شرایط کاری حقیقی می‌باشند.



شکل (۲): نتیجه اجرای الگوریتم سیگمادلتای پایه برای توالی تصاویر AVSS\_PV\_EVAL از دیتابیس i-LIDS [۳۵] در فریم ۳۴۰۰ (۱) تصویر ورودی (I) (۲) زمینه سیگمادلتا (B) (۳) واریانس (V) (۴) پیش‌نما (D)

Fig. (2): The result of executing the Sigma Delta algorithm for image sequences of AVSS\_PV\_EVAL from i-LIDS database [35] in frame 3400. 1) Input image (I) 2) Sigma Delta background (B) 3) Variance (V) 4) Foreground (D)

تنها پارامتری که در این روش قابل مشاهده است  $N$  می‌باشد، عددی که در محاسبه واریانس  $V_t$  استفاده می‌شود. محدوده مقادیر  $N$  کوچک است (بین ۱ تا ۴)، و معمولاً از یک توان ۲ برای اهداف بهینه‌سازی استفاده می‌شود. به علاوه،  $N$  می‌تواند به صورت خودکار با توجه به تخمین نویز تنظیم شود. این چنین تخمینی می‌تواند با شمارش پیکسل‌های مجزا در نتیجه‌ی آشکارسازی  $D_t$ ، با فرض اینکه این پیکسل‌های مجزا فقط به دلیل نویز هستند، ایجاد شود [۱]. در حقیقت پارامتر دیگری نیز وجود دارد که کمتر قابل مشاهده است: فرکانس به‌روز شدن متغیرهای سیگمادلتا. این فرکانس دارای بعد تعداد سطح‌های روشنایی بر ثانیه است. واضح است که این فرکانس باید بر اساس (۱) دینامیک تصویر (تعداد سطوح روشنایی) و (۲) فرکانس اخذ تصویر (نرخ فریم) انتخاب شود. معمولاً از فرکانس یکسان

زمینه تغییر می‌کند، دلتا تغییر را از طریق علامت خود منعکس می‌کند. اگر مقادیر سطوح خاکستری زمینه بالاتر برود، پالسهای دلتای بیشتری وجود خواهد داشت، و بر عکس.

## ۲-۱- الگوریتم سیگمادلتای پایه

اولین محاسبات تخمین زمینه مشابه مدل ارائه شده توسط McFarlane و Schofield [۳۴] می‌باشد که محاسبات آن در ادامه آمده است.  $I_t$  توالی تصاویر ورودی و  $M_t$  مقادیر تخمین زده شده‌ی زمینه می‌باشند.

مقداردهی اولیه:

برای هر پیکسل  $x$ :

$$M_0(x) = I_0(x)$$

برای هر فریم  $t$  و هر پیکسل  $x$ :

$$M_t(x) = M_{t-1}(x) + \text{sgn}(I_t(x) - M_{t-1}(x)) \quad (1)$$

برای هر فریم  $t$  و هر پیکسل  $x$ :

$$\Delta_t(x) = |M_t(x) - I_t(x)| \quad (2)$$

مقداردهی اولیه:

برای هر پیکسل  $x$ :

$$V_0(x) = \Delta_0(x)$$

برای هر فریم  $t$  و هر پیکسل  $x$  که  $\Delta_t(x) \neq 0$ :

$$V_t(x) = V_{t-1}(x) + \text{sgn}(N \times \Delta_t(x) - V_{t-1}(x)) \quad (3)$$

برای هر فریم  $t$  و هر پیکسل  $x$ :

If  $\Delta_t(x) < V_t(x)$

Then  $D_t(x) = 0$

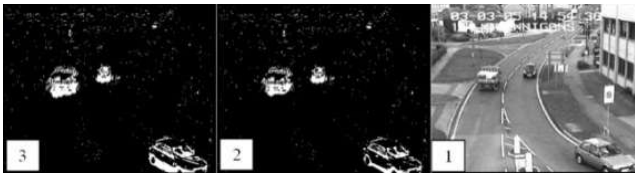
Else  $D_t(x) = 1$

تابع  $\text{sgn}$  (sign) به صورت  $\text{sgn}(a) = -1$  اگر  $a < 0$ ،  $\text{sgn}(a) = 0$  اگر  $a = 0$ ، و  $\text{sgn}(a) = 1$  اگر  $a > 0$ ، بنا براین، برای هر فریم، اگر تخمین از نمونه کمتر باشد به‌طور ساده یک واحد افزایش می‌یابد، یا اگر از نمونه بیشتر باشد یک واحد کاهش می‌یابد. ولی این فیلتر ویژگی‌های جالب دیگری نیز در رابطه با آشکارسازی تغییرات در سیگنال‌های متغیر با زمان دارا می‌باشد. در واقع، این تخمین زمینه به صورت تبدیل یک سیگنال آنالوگ متغیر با زمان به دیجیتال با مدولاسیون سیگمادلتا تفسیر می‌شود (تبدیل آنالوگ به دیجیتال تنها با استفاده از مقایسه و افزایش و کاهش، لذا اسم آن سیگمادلتا فیلتر می‌باشد) [۱].

از آنجایی که دقت مدولاسیون سیگمادلتا محدود به سیگنال‌های با قدرمطلق مشتق زمانی کمتر از ۱ می‌باشد، لذا خطای این مدولاسیون متناسب با نرخ تغییرات سیگنال، که در اینجا مطابق با حدود اندازه‌ی حرکت پیکسل‌ها می‌باشد، است. در نتیجه از قدرمطلق اختلاف بین  $I_t$  و  $M_t$  به عنوان اولین تخمین دیفرانسیلی استفاده می‌شود (اختلاف  $\Delta_t$ ).

همچنین از این فیلتر برای محاسبه واریانس زمانی پیکسل‌ها که بیانگر اندازه فعالیت حرکتی آنها است، استفاده می‌شود. این واریانس برای تصمیم‌گیری در رابطه با متحرک بودن یا ثابت بودن پیکسل به کار

شکل (۳) نتیجه مقایسه الگوریتم سیگمادلتای شرطی را با سیگمادلتای پایه در یک فریم نشان می‌دهد. همانطور که مشاهده می‌شود اثر روح و اثر روزه‌ای به وضوح کاهش یافته است.



شکل (۳): مقایسه سیگمادلتای شرطی و ساده برای توالی تصاویر AVSS\_PV\_EVAL در فریم ۲۰۰. (۱) تصویر ورودی (I) (۲) زمینه سیگمادلتا (B) (۳) زمینه سیگمادلتای شرطی

Fig. (3): Comparison of conditional and simple signaling sequences for AVSS\_PV\_EVAL images in frame 200. 1) Input image (I) 2) Sigma Delta background (B) 3) Background of conditional Sigma Delta

### ۲-۳- سیگمادلتا با پردازشهای مکانی-زمانی

الگوریتم سیگمادلتا مذکور تنها قادر به انجام پردازش زمانی در سطح پیکسل‌ها می‌باشد. پیشرفت‌های اخیر روش‌هایی را برای بهبود این مدل پیشنهاد می‌کند که در آن مقداری پردازش‌های مکانی-زمانی نیز به آن اضافه شده است. هدف از پردازش مکانی-زمانی اضافه شده را می‌توان حذف پیکسل‌های بی‌اهمیت از نقاب آشکارساز و همچنین کاهش اثر «روح» و اثر روزه‌ای دانست. اثر روح در اثر آشکارسازی اشتباه شیء صورت می‌گیرد که ناگهان پس از یک توقف بدون حرکت، شروع به حرکت می‌کند (یک وسیله نقلیه با حرکت کند باعث ایجاد اثری مشابه با اثر دنباله روح مانند می‌شود که در مدل زمینه ظاهر می‌شود). اثر روزه‌ای از اشیای دارای حرکت با نمایش ضعیف، آشکارسازی ضعیفی به عمل می‌آورد (برای مثال حرکت تقریباً عمودی اشیا بر صفحه تصویر).

یک مدل مارکوف برای یکنواخت‌سازی مکانی-زمانی تشخیص سیگمادلتا در سطح پیکسل توسط Manzanera [۱] ارائه شده است. این یک اقتباس از الگوریتم تکرارشونده Caplier و همکاران [۳۸] و Lacassagne و همکاران [۳۹]، است که از تشخیص سطح پیکسل  $D_t$  به عنوان مقدار اولیه، و از اختلاف سیگمادلتا  $\Delta_t$  و واریانس  $V_t$  به عنوان زوج زمینه مشاهدات استفاده شده در طراحی انرژی، استفاده می‌کند.

در اینجا یک استراتژی دیگر برای یکنواخت سازی بیان می‌شود. پردازش مکانی-زمانی که در این بخش مطرح می‌شود هدف سه گانه‌ای دارد:

- حذف پیکسل‌های غیرمهم از تشخیص (نویز، تشخیص اشتباه)، و ارتقای قسمت بندی اشیاء متحرک.

- کاهش اثر روح که باعث ایجاد تشخیص اشتباه در محلی که جسم متحرک بعد از مدت زمان زیادی ثابت مانده‌اند ترک کرده است، می‌شود.

با نرخ فریم برای توالی ۲۵ هرتزی از تصاویر ۸ بیتی استفاده می‌شود، ولی می‌توان از فرکانس کمتر برای تنظیم اندازه و سرعت اشیاء مشاهده شده، در کاربردهای مختلف استفاده کرد. پایداری تخمین سیگمادلتا با به‌روز کردن  $M_t$  فقط وقتی که  $\Delta_t \leq V_t$  می‌تواند با هم بهتر شود [۱].

تخمین زمینه سیگمادلتا یک روش ساده و کارآمد برای تشخیص پیکسل‌های دارای حرکت بارز در صحنه‌های ثابت، به نسبت یک ثابت زمانی وابسته به تعداد سطوح روشنایی و نرخ فریم، فراهم می‌آورد. با این حال این تخمین یک پردازش زمانی خالص است که فقط می‌تواند تشخیص در سطح پیکسل را انجام دهد.

### ۲-۲- سیگمادلتای شرطی

تشخیص سیگمادلتا در صورتی که زمینه برای پیکسل‌های اجسام متحرک تغییر نکند پایداری می‌شود. از این استراتژی برای به‌روز کردن میانگین سیگمادلتا،  $M_t$ ، فقط در پیکسل‌هایی که مقدار پیشنهادی صفر باشد ( $D_t(x)=0$ ) یا به عبارتی فاقد حرکت باشند، استفاده شده است. نسخه شرطی الگوریتم سیگمادلتا از فیدبک مربوط به مکان تخمین زده شده از اجسام متحرک فریم قبل برای به‌روزرسانی میانگین سیگمادلتا استفاده می‌کند [۳۶]. این فیدبک از ورود اجسام متحرک به زمینه جلوگیری می‌کند. الگوریتم سیگمادلتای شرطی در زیر آمده است:

برای هر پیکسل  $x$ :

If  $D_{t-1}(x) = 0$  then

$$M_t(x) = M_{t-1}(x) + \text{sgn}(I_t(x) - M_{t-1}(x))$$

else

$$M_t(x) = M_{t-1}(x)$$

برای هر پیکسل  $x$ :

$$\Delta_t(x) = |M_t(x) - I_t(x)|$$

برای هر پیکسل  $x$  که  $\Delta_t(x) \neq 0$ :

$$V_t(x) = V_{t-1}(x) + \text{sgn}(N \times \Delta_t(x) - V_{t-1}(x))$$

برای هر پیکسل  $x$ :

If  $\Delta_t(x) \leq V_t(x)$

Then  $D_t(x) = 0$

Else  $D_t(x) = 1$

این استراتژی نتایج تشخیص را با به تأخیر انداختن سهم اجسام متحرک در زمینه بهبود می‌بخشد. این روش به طور مؤثری هم اثر روح<sup>۲۶</sup> و هم اثر روزه‌ای<sup>۲۷</sup> ناشی از حرکت در راستای شعاع<sup>۲۸</sup> را کاهش می‌دهد.

همانطور که در [۳۷] گفته شده است در این سطح پایین پردازش، مطمئن تر است که فیدبک ربط را، برای جلوگیری از وارد شدن اجسامی که نادرست تشخیص داده شده‌اند<sup>۲۹</sup> به زمینه، به جای اعمال به هر دوی میانگین ( $M_t$ ) و واریانس ( $V_t$ ) فقط به میانگین اعمال نماییم. اگرچه در صورتی که در تشخیص اجسام متحرک به یک سطح بالا از اطمینان دست یافته باشیم می‌توان فیدبک ربط را به هر دوی  $M_t$  و  $V_t$  اعمال نمود.

اولین ایده استفاده از بازساخت کلاسیک ژئودسیک روی یک سطح می‌باشد  $\text{Rec}^Y(x)$ ، که به وسیله سستی<sup>۳۴</sup> گسترش ژئودسیک عنصر  $B$   $\text{Mfn}(\delta_B^Y(X), Y)$  گسترش شکل‌شناسی،  $X$  ساختمانی ابتدایی معرفی کننده توپولوژی،  $Y$  تصویر مرجع<sup>۳۵</sup>،  $X$  تصویر نشانه<sup>۳۶</sup> تعریف می‌شود. در حقیقت بازساخت ژئودسیک تطبیق‌پذیر نیست، زیرا معیار تک ارتباطی آن به اندازه کافی پایدار نیست و در بیشتر موارد شیء و روح هر دو بازسازی می‌شوند. در این حالت استفاده از عملگر ترکیبی بازساخت براساس عملگر شکل‌شناسی فراموشی معرفی شده توسط Richefeu و Manzanera [۴۳]، ترجیح داده می‌شود. ابتدا گسترش ترکیبی، به عنوان نمونه مکانی گسترش فراموشی، که به وسیله ترتیب سببی زیر محاسبه می‌شود، تعریف می‌شود [۱]:

$$\text{HDil}_\alpha(I)^{(0)}(x, y) = \alpha I(x, y) + (1 - \alpha) \max(I(x, y), \text{HDil}_\alpha(I)^{(0)}(x-1, y)) \quad (5)$$

که با ترتیب غیر سببی زیر پیروی می‌شود:

$$\text{HDil}_\alpha(I)^{(2)}(x, y) = \alpha \text{HDil}_\alpha(I)^{(0)}(x, y) + (1 - \alpha) \max(\text{HDil}_\alpha(I)^{(0)}(x, y), \text{HDil}_\alpha(I)^{(2)}(x+1, y)) \quad (6)$$

و ترتیب سببی عمودی:

$$\text{HDil}_\alpha(I)^{(2)}(x, y) = \alpha \text{HDil}_\alpha(I)^{(2)}(x, y) + (1 - \alpha) \max(\text{HDil}_\alpha(I)^{(2)}(x, y), \text{HDil}_\alpha(I)^{(2)}(x, y-1)) \quad (7)$$

و در انتها

$$\text{HDil}_\alpha(I)(x, y) = \alpha \text{HDil}_\alpha(I)^{(2)}(x, y) + (1 - \alpha) \max(\text{HDil}_\alpha(I)^{(2)}(x, y), \text{HDil}_\alpha(I)(x, y+1)) \quad (8)$$

در اینجا  $1/\alpha$  دارای بعد شعاع مکانی می‌باشد، لذا پارامتر  $\alpha$  جایگزین عنصر ساختمانی می‌شود.

بازساخت ترکیبی براساس طرح مشابهی می‌باشد، که از ترتیب زیر استفاده می‌کند:

$$\text{HRec}_\alpha^I(I)^{(0)}(x, y) = \min(I(x, y), \alpha I(x, y) + (1 - \alpha) \max(I(x, y), \text{HRec}_\alpha^I(I)(x-1, y))) \quad (9)$$

و به همین ترتیب مانند گسترش ترکیبی (روابط ۵، ۶، ۷ و ۸) ادامه می‌یابد.

رفتار بازساخت ترکیبی در شکل (۴) قابل مشاهده است، که اجسام (ماشین‌ها، پیاده‌ها) بعد از ۲۰ فریم مکث حرکت می‌کنند. مزیت بازساخت ترکیبی در این کاربرد فراموش کردن تدریجی (نمایی به صورت دقیقتر) تصویر نشانه می‌باشد که بجای اینکه اکیداً بر اساس ارتباط پیکسلی بنا شده باشد، مانند یک تابع اطمینان عمل می‌کند.

قابل مشاهده است که موفقیت این مرحله‌ی نشانه‌گذاری لبه‌های مشترک به کنتراست خود زمینه وابسته است (شکل (۴) ۷ و ۱۰). ولی در این جا تمرکز بر روی موارد حداکثر بوده است، اجسامی که به‌طور کامل ثابت بوده و در ابتدا توسط زمینه به صورت پوسته در آمده‌اند. به علاوه، همان‌طور که در بخش قبل مشاهده شد، فیدبک ربط بطور مؤثری اثر روح را کاهش می‌دهد.

-کاهش اثر روزه‌ای که باعث ایجاد شدن تشخیص ضعیف در اجسامی که حرکت نمایشی آنها ضعیف است مانند اجسام با حرکت شعاعی، می‌شود [۱].

این پردازش اضافی سعی در بهبود و نظم بخشیدن به آشکارسازی به‌دست آمده از طریق سه مرحله عملیات زیر را دارد. (۱) بازساخت ترکیب لبه-مشترک<sup>۳۷</sup> باز کردن توسط بازساخت<sup>۳۱</sup> و<sup>۳</sup> تائید زمانی<sup>۳۲</sup>. با توجه به تعریف هرکدام از این مراحل، به یادآوری برخی از عملگرهای شکل‌شناسی رایج می‌پردازیم [۴۰] تا [۴۲]:

$\text{Dil}_\lambda(X)$ : گسترش شکل‌شناسی از یک تصویر  $X$  با استفاده از یک توپ شعاع  $\lambda$  به‌عنوان عنصر ساختمانی.

$\text{Ero}_\lambda(X)$ : سایش شکل‌شناسی از یک تصویر  $X$  با استفاده از توپی با شعاع  $\lambda$  به‌عنوان عنصر ساختمانی.

$\tilde{\text{Dil}}_Y^X(X) = \text{Mfn}(\text{Dil}_\lambda(X), Y)$ : گسترش ژئودسیک<sup>۳۳</sup> از یک تصویر نشانه بوسیله توپی با شعاع  $\lambda$  به‌عنوان عنصر ساختمانی و یک تصویر مرجع  $Y$ .

$\tilde{\text{Rec}}^Y(X) = \lim_{k \rightarrow \infty} X(k)$ : بازسازی ژئودسیک از یک تصویر  $X$  (تصویر نشانه) با استفاده از یک تصویر مرجع  $Y$ . در اینجا گسترش ژئودسیک در حالت بازگشتی استفاده شده است، چنان‌که  $X(k) = \tilde{\text{Dil}}_Y^X(X(k-1))$  با مقدار  $X(0) = 0$  می‌باشد. این موضوع بیانگر این است که سری  $X(k)$  به این صورت تعریف شده است که همیشه پس از تکرار یک تعداد متناهی عدد، همگرا می‌شود.

## ۲-۳-۱- بازساخت ترکیبی لبه‌های مشترک

اولین بخش پردازش مکانی-زمانی تشکیل شده از عملیات‌های سطوح روشنایی می‌باشد. ورودی‌ها عبارتند از: (۱) تصویر اصلی  $I_t$  و (۲) اختلاف سیگمادلتا  $\Delta_t$ . هدف این ماژول حذف اثر اجسام روح در  $\Delta_t$  به وسیله تشخیص آنها در  $I_t$  می‌باشد. عملیاتی که انجام داده می‌شود در زیر آمده است:

$$\Delta_t' = \text{HRec}_{\Delta_t}^{\Delta_t}(\text{Mfn}(\|\nabla I_t\|, \|\nabla \Delta_t\|)) \quad (4)$$

به‌طور ساده، این یعنی "بازساخت" در داخل  $\Delta_t$  تصویر مینیمم بین ماژول گرادبان  $\Delta_t$  و ماژول گرادبان  $I_t$ . مفهوم این فرمول این است:  $\Delta_t'$  از اجزای تشکیل دهنده  $\Delta_t$  که همچنین در  $I_t$  هم هستند، ساخته شده است. به جزئیات محاسبات واقعی می‌پردازیم:

گرادبان ماژول  $\Delta_t$  و  $I_t$  با تخمین اجزای مشتق اول به وسیله کانولوشن با نقاب سوبل محاسبه می‌شود، و سپس نرمال اقلیدسی بردار محاسبه می‌شود. سپس تصویر مینیمم  $\text{Min}$  برای هر پیکسل  $x$  به صورت

$$\text{Mfn}(I_t, \Delta_t)(x) = \min(I_t(x), \Delta_t(x))$$

مینیمم  $\text{Min}()$  مانند یک عملگر اشتراک عمل می‌کند با این تفاوت که به‌جای مقادیر باینری، روی سطوح خاکستری عمل می‌کند. اینعملگر مانند یک ترکیب منطقی عمل می‌کند که لبه‌های اشیائی که متعلق به  $\Delta_t$  و هم  $I_t$  می‌باشد، را حفظ می‌کند. به‌منظور بازیافتن کل اشیای موجود در  $\Delta_t$  و  $I_t$  و نه تنها لبه‌ها، یک بازساخت از لبه‌های مشترک

$\text{Mfn}(\|\nabla I_t\|, \|\nabla \Delta_t\|)$  درون  $\Delta_t$  انجام می‌شود [۱].

در قسمت زیر الگوریتم کامل سیگمادلتا با پردازشهای مکانی-زمانی آمده است. ذکر این نکته لازم است که این الگوریتم با فیدبک ربط شرح داده شده در قسمت قبل همراه شده است. این فیدبک برای به روز کردن میانگین سیگمادلتا ( $M_t$ ) فقط در مکانهایی که  $L_t(x)$  برابر صفر می باشد، استفاده شده است (۱) و (۴۴).

مقداردهی اولیه:

$$M_0 = I_0$$

$$V_0 = 0$$

برای هر فریم  $t$ :

$$\Delta_t = |M_t - I_t|$$

به روزسازی واریانس:

$$\text{If } |\Delta_t| \neq 0 : V_t = V_{t-1} + \text{sgn}(N \times |\Delta_t| - V_{t-1})$$

بازساخت هایبریه لبه های مشترک:

$$\Delta'_t = \text{HRec}_{\alpha}^{|\Delta_t|}(\text{Min}(\|\nabla(I_t)\|, \|\nabla(\Delta_t)\|))$$

تشخیص زمانی:

$$\text{If } \Delta'_t \leq V_t \text{ then } D_t = 0 \text{ else } D_t = 1$$

پردازش های مکانی-زمانی باینری:

$$L_t^{(0)} = \text{Rec}^{D_t}(\varepsilon_{\beta}(\Delta_t))$$

$$L_t = \text{Rec}^{L_t^{(0)}}(L_{t-1}^{(0)})$$

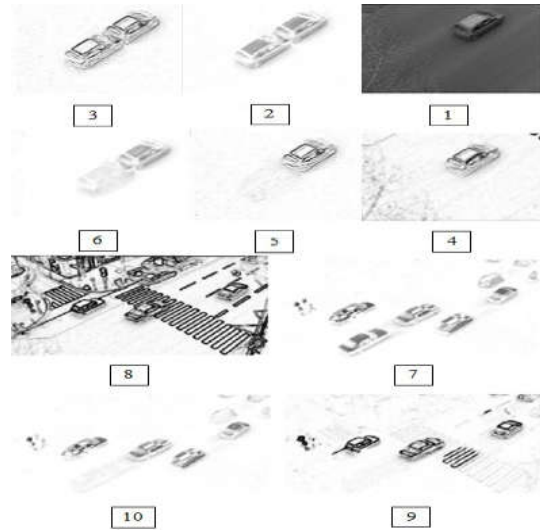
و به روزسازی میانگین با فیدبک ربط:

$$\text{If } L_t \neq 0 : M_t = M_{t-1} + \text{sgn}(S_t - M_{t-1}) \quad (12)$$

همانطور که گفته شد پردازشهای مکانی-زمانی افزوده شده به الگوریتم سیگمادلتا از سه بخش اصلی تشکیل شده است: (۱) بازسازی ترکیبی لبه-مشترک (۲) باز کردن توسط بازساخت و (۳) تأیید زمانی. از آنجایی که محاسبه بازساخت ترکیبی نیاز به مقادیر پیکسلهای همسایه دارد لذا پیاده سازی آن به صورت برداری امکان پذیر نیست. لذا پیاده سازی این عملگر (بازساخت ترکیبی) به صورت پیکسل به پیکسل انجام شده است که به شدت سرعت اجرای این روش را کاهش می دهد و این الگوریتم از حالت بیدرنگ بودن خارج می شود. به همین منظور این روش در دو حالت متفاوت پیاده سازی شده است. حالت اول الگوریتم کامل و در حالت دوم بازساخت ترکیبی از الگوریتم حذف شده است یعنی فقط دو مرحله باز کردن توسط بازساخت و تأیید زمانی انجام شده است. با توجه به حجم زیاد پردازشهای عملگر بازساخت ترکیبی و مقایسه این دو حالت نتیجه گیری می شود که حذف این بخش از الگوریتم از نظر هزینه های محاسباتی مقرون به صرفه می باشد. نتایج اجرای این دو حالت در شکل (۶) آمده است.

#### ۲-۴- تخمین زمینه سیگمادلتا چندگانه

پردازشهای مکانی-زمانی و فیدبک ربط ارائه شده در بخش قبل، باعث یک ارتقاء مشهود در پایداری در موارد آهسته شدن حرکت، ایست و اجسام دارای حرکت شعاعی می شود. با این حال، تخمینگر سیگمادلتا با یک ثابت زمانی توصیف می شود: دوره به روز شدن آن، که دارای بعد تعداد سطوح روشنایی بر ثانیه می باشد. این ثابت زمانی باعث ایجاد



شکل (۴): متلاشی کردن اثر روح به وسیله لبه های مشترک ( $\alpha=1/8$ ). نمونه اول (۱ خودرو): ۱)  $I_t$ ، ۲)  $\Delta_t$ ، ۳)  $\|\nabla(\Delta_t)\|$ ، ۴)  $\|\nabla(I_t)\|$ ، ۵)  $\text{Min}(\|\nabla(I_t)\|, \|\nabla(\Delta_t)\|)$ ، ۶)  $\Delta'_t$ . نمونه دوم (۵ پیاده، ۱ خودرو): ۷)  $I_t$ ، ۸)  $\Delta_t$ ، ۹)  $\|\nabla(\Delta_t)\|$ ، ۱۰)  $\text{Min}(\|\nabla(I_t)\|, \|\nabla(\Delta_t)\|)$ ، ۱۱)  $\Delta'_t$ .  
Fig. (4): Split the ghost effect by the common edges ( $\alpha = 1/8$ ). The first sample (1 car): 1)  $I_t$ , 2)  $\Delta_t$ , 3)  $\|\nabla(\Delta_t)\|$ , 4)  $\|\nabla(I_t)\|$ , 5)  $\text{Min}(\|\nabla(I_t)\|, \|\nabla(\Delta_t)\|)$ , 6)  $\Delta'_t$ . The second sample (5 cars, 1 pedestrian): 7)  $I_t$ , 8)  $\Delta_t$ , 9)  $\|\nabla(\Delta_t)\|$ , 10)  $\text{Min}(\|\nabla(I_t)\|, \|\nabla(\Delta_t)\|)$ , 11)  $\Delta'_t$  [1]

#### ۲-۳-۲- شکل شناسی مکانی-زمانی باینری

بعد از محاسبه بازساخت ترکیبی، یک آستانه گیری قابل تطبیق روی  $\Delta'_t$  اجرا می شود (به جای  $\Delta_t$  در نسخه خالص زمانی):

$$\text{If } \Delta'_t > V_t \text{ then } D_t = 1$$

$$\text{Else } D_t = 0$$

سپس اجزاء متصل کوچک در  $D_t$  با استفاده از بازکردن به وسیله بازساخت حذف می شوند:

$$L_t^{(0)} = \text{Rec}^{D_t}(\varepsilon_{\beta_\lambda}(\Delta_t)) \quad (10)$$

که  $\varepsilon$  سایش شکل شناسی و  $\beta_\lambda$  عنصر ساختمانی یک دایره با شعاع  $\lambda$  می باشد.

در انتها یک تأیید زمانی با محاسبه یک بازساخت دیگر ایجاد می شود:

$$L_t = \text{Rec}^{L_t^{(0)}}(L_{t-1}^{(0)}) \quad (11)$$

$L_t$  نمایشگر آخرین برجسب می باشد، نتیجه تشخیص. مفهوم آن این است: اجسام بزرگتر از  $\lambda$  که در دو فریم متوالی وجود داشته باشند. این مفهوم در شکل (۵) با  $\lambda=1$  که دو آدم در حال راه رفتن را نشان می دهد، به تصویر کشیده شده است [۱].



شکل (۵): شکل شناسی مکانی-زمانی باینری. ۱)  $D_t$ ، ۲)  $L_t^{(0)}$ ، ۳)  $L_t$  [۱]  
Fig. (5): Spatio-temporal binary morphology. 1)  $D_t$ , 2)  $L_t^{(0)}$ , 3)  $L_t$  [1]

If  $t$  is a multiple of  $\alpha_1$  :

$$M_t^i = M_{t-1}^i + \text{sgn}(M_{t-1}^{i-1} - M_{t-1}^i)$$

End if

محاسبه اختلاف ورودی فعلی با  $M^i$ :

$$\Delta_t^i = |I_t - M_t^i|$$

If  $\Delta_t^i = 0$ :

$$V_t^i = V_{t-1}^i + \text{sgn}(N \times \Delta_t^i - V_{t-1}^i)$$

End for

محاسبه میانگین عمومی:

$$M_t = \frac{\sum_{i \in [0, K]} \frac{\alpha_i M_t^i}{V_t^i}}{\sum_{i \in [0, K]} \frac{\alpha_i}{V_t^i}} \quad (13)$$

محاسبه واریانس عمومی:

$$\Delta_t = |I_t - M_t|$$

به روزسازی واریانس عمومی:

If  $|\Delta_t| \neq 0$ :

$$V_t = V_{t-1} + \text{sgn}(N \times |\Delta_t| - V_{t-1})$$

End if

و تشخیص نهایی:

$$\text{If } \Delta_t < V_t \text{ then } D_t = 0 \text{ else } D_t = 0$$

End for

در این روش مدل زمینه به ازای افزایش هزینه‌های محاسباتی نسبت به الگوریتم اصلی سیگمادلتا، بهبود می‌یابد.

## ۲-۵- سیگمادلتا و قانون زیپف

اصالتاً قانون زیپف یک اصل تجربی در حوزه زبان‌شناسی و تئوری اطلاعات می‌باشد [۴۵] که بیان می‌کند، در هر نوع نگارش متن، احتمال رخداد  $n$ مین حرف پرتکرار  $1/n$  احتمال رخداد اولین حرف پرتکرار می‌باشد. بنابراین تابع توزیع زیپف یک تابع نزولی هیپربولیک می‌باشد. اخیراً این تابع در بسیاری از کاربردهای پردازش تصویر به طور ویژه به عنوان یک مدل برای توزیع مکانی محلی استفاده شده است [۴۶]. در اینجا به عنوان یک مدل برای توزیع زمانی (پیکسل به پیکسل) استفاده شده است.

به دلیل دیورژانس مجموع  $1/n$ ، تابع توزیع زیپف شامل یک توان می‌باشد:  $1/n^s$  که  $s > 1$  می‌باشد. شکل کلی توزیع پیوسته متقارن زیپف-مندلبرت را می‌توان به صورت زیر نوشت [۴]:

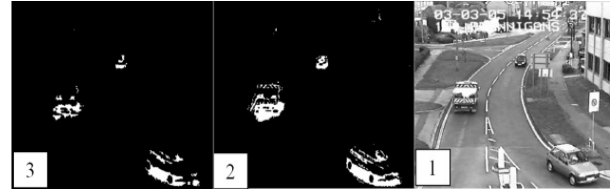
$$Z_{(\mu, k, s)}(x) = \frac{(s-1)k^{s-1}}{2[(k-\mu) + |x|]^s} \quad (14)$$

در این عبارت، پارامتر  $\mu$  بیانگر مد توزیع و  $k$  تعیین‌کننده پراکندگی آن می‌باشد. ویژگی قابل توجه  $Z$ ، که باعث می‌شود به عنوان مدل چگالی  $f_t$  دنباله زمانی  $I_t$  در نظر گرفته شود، شکل تابع افزایشی آن می‌باشد، که نزدیک به تابع پله می‌باشد: اگر  $x < \mu$  باشد، آنگاه  $H_{(\mu, k)}(x) = -K$  و اگر  $x > \mu$  باشد آنگاه  $H_{(\mu, k)}(x) = +K$ ، با  $K = \alpha_{\max} k^s$ . بنابراین قانون زیپف طبیعتاً به مدولاسیون سیگمادلتا که در تبدیل آنالوگ به دیجیتال استفاده می‌شود، مرتبط می‌باشد:

برای هر بازه زمانی  $\Delta_t$ :

$$\text{If } M_{t-\Delta_t} \geq I_t \text{ then } M_t = M_{t-\Delta_t} - \varepsilon$$

یک محدودیت در بنیان ظرفیت تطبیق پذیری با صحنه‌های پیچیده معینی که عموماً اجسام با اندازه‌های مختلف و سرعت‌های مختلف دائماً از آن عبور می‌کنند، می‌شود. در این بخش، یک چارچوب عمومی‌سازی از تخمین سیگمادلتا برای زمینه‌های چندگانه مطرح می‌شود.



شکل (۶): مقایسه دو حالت الگوریتم سیگمادلتا با پردازش‌های مکانی-زمانی برای توالی تصاویر AVSS\_PV\_EVAL در فریم ۱۷۵. (۱) تصویر ورودی (۲) پیشنهاد در حالت اول (۳) پیشنهاد در حالت دوم

Fig. (6): Comparison of two modes of Sigma Delta algorithm with spatial-temporal processing for image sequences of AVSS\_PV\_EVAL images in frame 175. 1) Input image (I) 2) The foreground in the first mode 3) The foreground in the second mode

اساس این چارچوب محاسبه یک مجموعه از  $K$  زمینه  $\{m_t^i\}$   $1 < i < K$  به جای یک زمینه  $M_t$  می‌باشد. هر زمینه  $m_t^i$  با دوره به‌روز شدن آن  $\alpha_i$  و با فازش  $\varphi_i$  مشخص می‌شود. یک مجموعه از  $K$  واریانس  $v_t^i$  هم مانند میانگین سیگمادلتا از اختلافات بین  $I_t$  و  $m_t^i$  محاسبه می‌شود. تصمیم‌گیری زمینه/پیشنما با مقایسه نمونه با هر زمینه که دارای یک مقدار اطمینان که متناسب با  $\alpha_i$  و متناسب با معکوس  $v_t^i$  است، انجام می‌شود [۱].

در زیر یک نمونه از محاسبات زمینه در الگوریتم چندگانه با  $K$  دوره مختلف  $\alpha_1 \ll \dots \ll \alpha_K$  آمده است. در این نمونه فازها در نظر گرفته نشده است. میانگین‌های سیگمادلتا  $m_t^i$  به صورت بازگشتی قابل محاسبه می‌باشند:

$m_t^i = m_{t-1}^i + \text{sgn}(m_{t-1}^{i-1} - m_{t-1}^i)$  با قرارداد  $m_t^0 = I_t$  قاعده و منشاء ضرایب اطمینان متصل به هر زمینه  $m_t^i$ ، وزن بیشتر دادن به زمینه‌های بلند مدت به منظور پشتیبانی کردن از مقداری که بیشتر از همه در دوره‌های طولانی و در یک زمان تکرار می‌شود، می‌باشد. به منظور هماهنگ شدن با یک زمینه متغیر، وزنها بر اساس واریانس کم می‌شوند [۱].

مقداردهی اولیه:

For each  $i \in [1, K]$  :

$$M_0^i = I_0$$

end for

مقداردهی اولیه واریانس عمومی:

$$V_0 = 0$$

for each frame  $t$ :

$$M_t^0 = I_t$$

$$V_t^0 = 0$$

به‌روز کردن بازگشتی مدل‌های زمینه:

For each  $[1, K]$  :



۲-۶- تخمین زمینه سیگمادلتا با سنجش اطمینان

یک مکمل دیگر الگوریتم تفریق زمینه سیگمادلتا توسط Toral و همکارانش در سال ۲۰۱۰ مطرح شده است [۴۴]. هدف از این الگوریتم مناسب ساختن الگوریتم اصلی به طور ویژه برای صحنه‌های ترافیک شهری که شرایط رقابتی مانند جریان فشرده ترافیک، ازدحام ترافیک احتمالی و یا صف‌های وسایل نقلیه در آنها متداول است، با حفظ بهینه بودن محاسباتی الگوریتم اصلی، می‌باشد. در این مضمون، الگوریتم‌های تفریق زمینه باید توانایی کارکردن با اجسام متحرکی که به دلیل توقف موقت در زمینه ادغام شده و سپس دوباره پیش‌نما می‌شوند، را داشته باشند. مکمل‌های زیادی از الگوریتم سیگمادلتا به وسیله یک مرحله پس‌پردازش یا یک مرحله معتبرسازی پیش‌نما بر این مساله فائق آمده‌اند. هدف این الگوریتم کاهش دادن این مرحله پسین به وسیله جلوگیری کردن از ترکیب شدن مدل زمینه با اجسام دارای حرکت آهسته یا متوقف شده برای یک بازه زمانی، می‌باشد. این مکمل الگوریتم سیگمادلتا پایه، معیار اطمینانی را معرفی می‌کند که به هرکدام از پیکسل‌های مدل زمینه گره خورده، میزان اعتماد از مقدار فعلی مورد انتظار برای آن پیکسل را به صورت کمی بیان می‌کند. این کار، روشی را فعال می‌کند که در آن سعی بر این است تا تعادل بهتری بین تطبیق‌پذیری با تغییرات شدت روشنایی با زمینه در صحنه ایجاد شود و همچنین از آلودگی‌های نامطلوب مدل زمینه جلوگیری شود. هدف نهایی، حفظ مدل از خراب‌شدن به وسیله خودروهای دارای حرکت آهسته یا وسایلی است که برای یک بازه زمانی بی‌حرکت هستند. باتوجه به الگوریتم سیگمادلتا پایه، سه تصویر جدید مورد نیاز است: (۱) تصویر شمارنده فریم  $(I_t^{FC})$ ؛ (۲) تصویر شمارنده تشخیص  $(I_t^{DC})$ ؛ و (۳) تصویر اطمینان  $(I_t^{CON})$  [۴۴].

$M_0 = I_0, V_0 = V_{ini}$  // initialize background model, variance

$I_0^{DC} = I_0^{FC} = 0, I_0^{CON} = C_{ini}$  // initialize detection and frame counter and confidence measurement

for each frame t:

$I_t^{FC} = I_t^{FC} + 1$  // increment Frame-counter image

// period evaluation and background updating

if  $I_t^{FC} \geq I_t^{CON}$  // if confidence period expires

if  $I_t^{FC}$  is a multiple of P // if refresh period expire

if  $V_t < V_{th}$  // if not very heavy traffic

if  $(I_t^{DC} / I_t^{FC}) \leq 0.8$  // if not very heavy traffic

$U_t = 1$  // refresh period updating mode

End if

End if

Else // if current confidence period expires

if  $V_t < V_{th}$  // low variance => traffic flow may be evaluated

$I_t^{CON} += \gamma(I_t^{DC} / I_t^{FC})$  // confidence updating from detection ratio

if  $I_t^{CON} = C_{min}$  // if confidence reaches the minimum...

If  $M_{t-\Delta t} < I_t$  then  $M_t = M_{t-\Delta t} + \epsilon$

میانگین افزایش بر واحد زمان در این حالت  $k = \frac{\epsilon}{\Delta t}$  می‌باشد. افزایش ابتدایی  $\epsilon$  از نظر دیجیتال کم ارزش‌ترین بیت<sup>۴۹</sup> این مدل، اگر سیستم عددی استفاده شده اعداد صحیح<sup>۴۰</sup> باشند، می‌باشد. سپس تطبیق‌پذیری با پراکندگی و انحراف مدل با تنظیم دوره به‌روز شدن  $\Delta t$  قابل انجام می‌باشد: هرچه واریانس بزرگتر باشد، مقادیر  $\Delta t$  کوچکتر باید باشد. الگوریتم زیر این رفتار را تولید می‌کند. اساس این الگوریتم چسباندن یک تخمین‌گر انحراف به هر پیکسل، علاوه بر تخمین‌گر حالت  $M_t$  می‌باشد. با فرض اینکه  $V_t \in [0, 2^m - 1]$ ، که بدین معنی است که  $V_t$ ،  $m$  بیتی می‌باشد [۴۷].

For every frame t:

rank = t %  $2^m$ ; pow2 = 1;

do { pow2 = 2 \* pow2; }

while ((rank % pow2 == 0) && (pow2 <  $2^m$ ))

If  $(V_{t-1} > \frac{2^m}{pow2})$

If  $M_{t-1} > I_t$  then  $M_t = M_{t-1} - 1$ ;

If  $M_{t-1} < I_t$  then  $M_t = M_{t-1} + 1$ ;

end if

$\Delta_t = |I_t - M_t|$ ;

If (t % Tv == 0)

If  $V_{t-1} > \max(V_{min}, N \times \Delta_t)$  then

$V_t = V_{t-1} - 1$ ;

If  $V_{t-1} < \max(V_{max}, N \times \Delta_t)$  then

$V_t = V_{t-1} + 1$ ; (15)

end if

end for

در اینجا  $x \% y$  باقیمانده<sup>۴۱</sup>  $x$  بر  $y$  می‌باشد. هدف دو خط اول این الگوریتم (که برای تمام پیکسل‌ها یکبار محاسبه می‌شوند) پیدا کردن بیشترین توان ۲ (Power 2) که بخش‌پذیر بر باقیمانده شاخص زمانی بر  $2^m$  (rank) می‌باشد، است. پس از تعیین این پارامتر، از آن برای تعیین حداقل  $V_{t-1}$  که تخمین زمینه سیگمادلتا به‌روز خواهد شد، استفاده می‌شود. لذا دوره‌ی به‌روز شدن  $M_t$  با انحراف نسبت عکس دارد: اگر  $V_t \geq 2^{m-1}$ ، در هر فریم به‌روز می‌شود، اگر  $2^{m-1} < V_t < 2^{m-1}$  در هر دو فریم به‌روز می‌شود و به همین ترتیب ادامه می‌دهد [۴۷].

در اینجا عامل انحراف  $V_t$  به وسیله تخمین سیگمادلتای قدرمطلق اختلاف  $\Delta_t$  محاسبه شده و با یک پارامتر  $N$  تقویت می‌شود.  $V_t$  می‌تواند مستقیماً به عنوان معیار تشخیص پیش‌نما استفاده شود: نمونه ورودی  $I_t$  پیش‌نما در نظر گرفته می‌شود اگر  $V_{max} \Delta_t \geq V_t$  و  $V_{min} \Delta_t \geq V_t$  برای کنترل سرریز شدن<sup>۴۲</sup> استفاده می‌شوند؛ ۲ و  $2^m - 1$  مقادیر معمول آنها هستند [۴۷].

این نکته قابل توجه است که ثابت‌های زمانی که بیانگر دوره پاسخ الگوریتم تخمین زمینه هستند در این مدل به دینامیک  $V_t$ ، (تعداد بیت‌های دارای ارزش)، و به دوره به‌روز شدن آن  $Tv$  مربوط می‌باشند [۴۷].

تراکم ترافیک در پیکسل‌های مربوطه هستند. در این صورت اگر این چنین نباشد، عمل به‌روزرسانی مجاز است [۴۴].

از سوی دیگر، معمولاً قابلیت تشخیص موثری برای ارزیابی مناسب جریان ترافیک در دست نداریم. در واقع مقادیر بالای واریانس به‌این معنا است که ما این قابلیت را نداریم بنابراین نمی‌توانیم با خیالی آسوده به اطلاعات جمع‌آوری شده مربوط به نرخ تشخیص اعتماد کنیم. در این مورد بهتر است به‌روزرسانی انجام نگیرد.

یک شیوه‌ی موازی برای به‌روزرسانی معیار اطمینان تنظیم شده است. این روش دوم توسط پارامتری به نام «دوره اطمینان» کنترل می‌شود. این دوره یک دوره ثابت زمانی نیست اما برای هر پیکسل به خود اطمینان آن پیکسل بستگی دارد. قانون آن به‌این صورت تعریف می‌شود که هرچه میزان سطح اطمینان بیشتر باشد، به به‌روزرسانی کمتری برای پیکسل مربوطه نیاز است. به طور ویژه طول دوره اطمینان به صورت تعداد فریم‌های آن که برابر با مقدار اطمینان در پیکسل مربوطه است، تعیین می‌شود. هر زمان که دوره اطمینان به انتها برسد، اندازه‌گیری اطمینان با روندی روبه‌رشد و برطبق تابع نمایشی نزولی نرخ تشخیص  $\gamma(d) = \text{round}(\alpha \cdot \exp(-\beta d) - 1)$ ،  $d$ ، به‌روزرسانی می‌شود.

بهره  $\alpha$  به صورت بیشترین افزایش اطمینان تعیین می‌شود (زمانی که نرخ تشخیص به صفر میل می‌کند)، در حالی که  $\beta$ ، بیانگر نرخ تضعیف این افزایش می‌باشد، که باید به نحوی انتخاب گردد که افزایش‌های منفی به نرخ‌های تشخیص بالا محدود شوند. در این‌جا، مقدار  $\alpha$  برابر ۱۱ است، بنابراین حداکثر افزایش اطمینان ۱۰ فریم و  $\beta$  نیز برابر ۴ است، که محدوده‌ی عبور تابع با افزایش ۰/۵- را در محدوده ۰/۸۰-۰/۷۵٪ نرخ تشخیص، تنظیم می‌نماید.

در مواردی که ضریب اطمینان تا یک مقدار حداقل کاهش یابد، الگوریتم زمینه را وادار به به‌روزرسانی می‌کند. این یکی از قوانین مهم اجرایی است چرا که برای مثال در مورد صحنه‌های شلوغ، مدل زمینه ممکن است توسط دوره احیاء به‌روزرسانی نشود. بنابراین در این صورت، این شیوه‌ی به‌روزرسانی سعی دارد از اینکه مدل به طور نامحدود در زمینه‌ای اشتباه یا مهجور قفل شود، جلوگیری کند [۴۴].

به‌عنوان آخرین راه‌حل، حالت دیگری نیز وجود دارد که در آن عمل به‌روزرسانی دستور داده شده است. این همان وقتی است که دوره اطمینان خاتمه می‌یابد اما قابلیت تشخیص، ضعیف تخمین زده می‌شود. در چنین موردی به‌این علت که اطلاعات قابل اعتمادی در دسترس نیست، به‌روزرسانی زمینه ترجیح داده می‌شود. در واقع از راه دیگری قادر به تغییر این موقعیت نیستیم زیرا واریانس به‌روزرسانی نخواهد شد؛ بنابراین این الگوریتم به بن‌بست ختم می‌شود [۴۴].

شکل (۷) نتیجه پیاده‌سازی الگوریتم سیگمادلتا با سنجش اطمینان را نشان می‌دهد. در این شکل تصاویر  $I_{t-1}^{FC}$ ،  $I_t^{FC}$  و  $I_{t-1}^{con}$  به منظور نشان دادن نحوه کارکرد الگوریتم ارائه شده است.

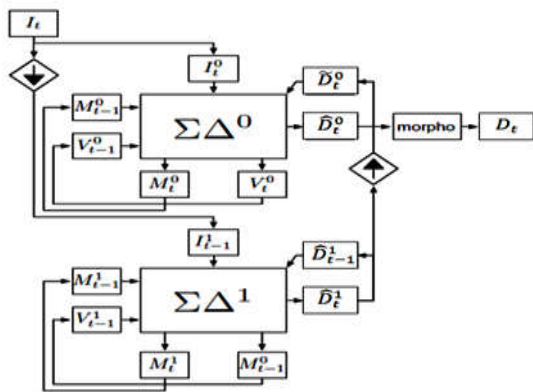
```

 $U_t = 1$  // force updating
End if
Else // we cannot reliably evaluate traffic flow
 $U_t = 1$  // confidence period updating mode (Avoid background deadlock)
End if
 $I_t^{DC} = I_t^{FC} = 0$  // Reset Detection counter and frame counter
End if
// Background updating (if appropriate) and detection
if  $U_t = 1$  // if updating recommended, sigma-delta algorithm
 $M_t = M_{t-1} + \text{sgn}(I_t - M_{t-1})$  // update background model M
 $\Delta_t = |I_t - M_t|$  // compute current Difference
if  $\Delta_t \neq 0$ 
 $V_t = V_{t-1} + \text{sgn}(N \cdot \Delta_t - V_{t-1})$  // update Variance V
End if
 $D_t = (\Delta_t > V_t)$  // compute Detection image D
Else
 $\Delta_t = |I_t - M_t|$ 
 $D_t = (\Delta_t > V_t)$ 
End if
 $I_t^{DC} += (D_t == 1)$  // update the Detection-counter image
End for

```

در الگوریتم پیشنهاد شده، تصویر واریانس تغییرات شدت روشنایی هر پیکسل، زمانی که هیچ جسمی روی آن پیکسل نباشد را نشان می‌دهد. به‌عبارت دیگر، تصویر واریانس به‌تنهایی توسط شدت‌های روشنایی زمینه تعیین می‌شود، به طوری که یک آستانه مناسب باید از آن ایجاد شود. واریانس پایین به معنای دارا بودن یک «مدل زمینه پایدار» است که باید حفظ شود. واریانس بالا نیز به‌این معنا است: «الگوریتم باید به دنبال یک مدل زمینه پایدار باشد». یکی از مشکلات نسخه قبلی الگوریتم سیگمادلتا در محیط‌های ترافیک شهری این است که هنگامیکه واریانس به دلیل عبور وسایل نقلیه افزایش می‌یابد، تشخیص به دلیل اینکه آستانه خیلی بالا می‌رود تنزل می‌یابد. بنابراین ایجاد یک زمینه انتخابی‌تر و به‌روزرسانی واریانس ضروری به‌نظر می‌رسد [۴۴].

شیوه‌ی اصلی به‌روزرسانی زمینه و واریانس به صورت انتخابی با مفهومی با نام «دوره احیاء» مرتبط می‌شود. هر زمان که این دوره خاتمه یابد (برای مثال در هر P فریم)، عمل به‌روزرسانی انجام می‌شود، مشروط به این که شرایط ترافیکی احتمالاً مناسب باشد. نرخ تشخیص را می‌توان به‌عنوان تخمینی از جریان ترافیک دانست. توجه داشته باشید که این تخمین زمانی قابل قبول است که فرض کنیم آستانه واریانس، نوسانات شدت روشنایی زمینه را فیلتر می‌کند. مقادیر بالای ۰/۸۰٪ تشخیص معمولاً مربوط به حضور وسایل نقلیه متوقف‌شده یا



شکل (۸): سیگمادلتای سلسله مراتبی [۳۶]  
Fig. (8): Hierarchical Sigma Delta [36]

رویکردهای مختلفی برای ارزیابی عملکرد الگوریتم‌های تفریق زمینه وجود دارد، از ارزیابی سطح پیکسل گرفته تا در سطح شیء و یا ارزیابی در سطح کاربردی. در مورد ارزیابی مبتنی بر هدف، آشکارسازی زمینه توسط سایر اجزا در سطوح بالاتر برنامه، تحت تأثیر قرار می‌گیرد، برای مثال ماژول حباب استخراج ویژگی، یا ماژول ردیاب که از حیطه کاری این مقاله خارج است. در نتیجه در این بخش، یک ارزیابی پیکسل‌گرا ترجیح داده می‌شود.

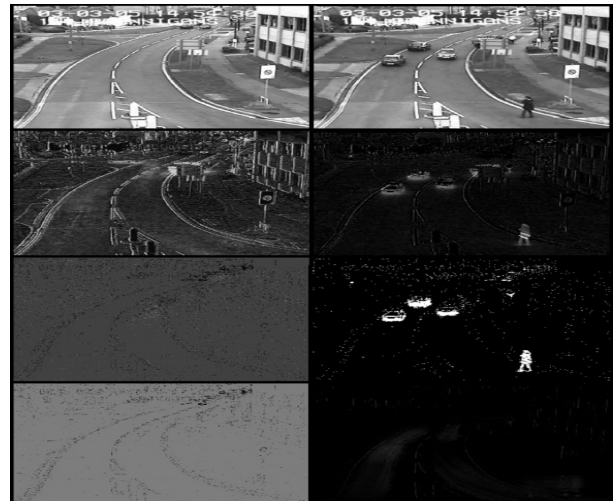
در یک مسئله تصمیم‌گیری باینری، طبقه‌بندی‌کننده نمونه‌ها را برحسب مثبت یا منفی برچسب‌گذاری می‌کند. در این مورد، نمونه‌ها مقادیر پیکسل هستند، «مثبت» به معنای پیکسل جسم پیش‌نما و «منفی» معرفی‌کننده پیکسل جسم زمینه می‌باشد؛ یعنی:

- ۱) مثبت واقعی (TP): پیکسل‌های پیش‌نما به درستی طبقه‌بندی شده؛
  - ۲) منفی واقعی (TN): پیکسل‌های زمینه به درستی طبقه‌بندی شده؛
  - ۳) مثبت کاذب (FP): پیکسل‌های پیش‌نما به اشتباه طبقه‌بندی شده؛
  - ۴) منفی کاذب (FN): پیکسل‌های زمینه به اشتباه طبقه‌بندی شده.
- ماتریس اغتشاش می‌تواند بیانگر یک نقطه تک در فضای مشخصه عملگر گیرنده<sup>۴۵</sup> (ROC) یا فضای فراخوانی دقیق<sup>۴۶</sup> (PR) باشد. منحنی‌های ROC چگونگی تغییر تعداد نمونه‌های مثبت به درستی طبقه‌بندی شده را با تعداد نمونه‌های به اشتباه طبقه‌بندی شده نشان می‌دهند.

منحنی‌های ROC این ریسک را دارند که ارزیابی بهینه‌ای از عملکرد طبقه‌بندی‌کننده را اگر انحراف چشمگیری در توزیع طبقه وجود داشته باشد ارائه کنند. منحنی‌های PR جایگزین خوبی برای این طبقه‌ها هستند. از اینها گذشته، منحنی‌های PR می‌توانند تفاوت بین رفتار الگوریتم‌هایی که در منحنی‌های ROC نمایان نیستند، را آشکار کنند [۴۴]. برای دستیابی به معیارهای سودمندتر، خصوصاً به صورت سنجش‌های اسکالر، می‌توان از مساحت سطح زیر نمودارهای ROC یا PR استفاده کرد. چند مثال در زیر آورده شده است:

میزان F (F-measure):

$$= 2 \left( \frac{PR \cdot RE}{PR + RE} \right) \quad 0 \leq S_F \leq 1S_F$$



شکل (۷): نتایج پیاده‌سازی الگوریتم سیگمادلتا با سنجش اطمینان. از بالا سمت راست به ترتیب، تصویر فریم ورودی، زمینه، تصویر اختلاف تصویر ورودی و زمینه، واریانس، نقاب تشخیصی،  $I^{dc}$ ،  $I^{con}$

Fig. (7): The results of implementing the Sigma Delta algorithm with reliability assurance. From the top right, respectively, the image of the input frame, background, image of the difference in input image and background, variance, diagnostic mask,  $I^{dc}$ ,  $I^{con}$

### ۲-۷- سیگمادلتای سلسله مراتبی

الگوریتم سلسله مراتبی (شکل (۸)) یک نسخه دو مرحله‌ای از فیلتر سیگمادلتا می‌باشد. هر بلوک سیگمادلتا، الگوریتم سیگمادلتای پایه را اجرا می‌کند. هر دو بلوک از به‌روزرسانی شرطی استفاده می‌کنند. در سطح پایین یک به‌روزرسانی شرطی زمانی<sup>۴۳</sup> انجام می‌پذیرد:  $M_t^1$  و  $V_t^1$  بر اساس  $D_{t-1}^1$  به‌روز می‌شوند. در سطح بالا، به‌روزرسانی یک به‌روز-رسانی شرطی مکانی<sup>۴۴</sup> می‌باشد:  $M_t^0$  و  $V_t^0$  بر اساس  $D_t^0$  نقاب باینری نمونه‌گیری شده از  $D_{t-1}^1$ ، به‌روزرسانی می‌شوند. نرخ نمونه‌گیری در بازه [2,10] می‌باشد و بر اساس میزان نویز آشفته‌گی تنظیم می‌شود. در انتها، یک مرحله پس پردازش شکل‌شناسی در دو مرحله به الگوریتم اضافه شده است. مرحله اول پیکسل‌های جدا افتاده و تنها را که نویز فرض می‌شوند را حذف می‌کند، و مرحله دوم یک بستن شکل‌شناسی  $3 \times 3$  می‌باشد [۳۶].

### ۳- تجزیه و تحلیل عملکرد الگوریتم‌ها

در این بخش، مطالعه عملکردی دقیق‌تری از لحاظ فنی با استفاده از برخی ماتریس‌های کیفیت قسمت‌بندی به منظور تعیین کردن صحت آشکارسازی پیش‌نماهایی که بوسیله هر الگوریتم تشخیص داده شده‌اند، انجام شده است. نتایج پیاده‌سازی روش‌های بر پایه تفریق زمینه در انتهای این بخش ذکر خواهد شد.

فریم از توالی تصاویر waving trees از دیتابیس Wallflower ارزیابی کردیم. پارامترهای استفاده شده عبارتند از:  $(N=4)$  برای کلیه الگوریتم‌ها،  $(\lambda=1)$  برای سیگمادلتا با پردازشهای مکانی-زمانی،  $(\alpha=1/8, \lambda=1)$  برای سیگمادلتا با پردازشهای مکانی-زمانی و بازساخت ترکیبی،  $(k=3, a_1=1, a_2=8, a_3=16)$  برای سیگمادلتای چندگانه،  $(T_V=4)$  برای سیگمادلتا و قانون زیف،  $(v_t \in [v_{min}, v_{max}] = [10, 200])$ ،  $[c_{min}, c_{max}] = [10, 125]$  برای سیگما-دلتا با سنجش اطمینان و  $(c_{min} = 8)$  برای سیگمادلتای سلسله مراتبی. نتایج این آزمایش‌ها در شکل‌های (۹) و (۱۱) و نمودارهای نتایج حاصل شده در شکل‌های (۱۰) و (۱۲) آمده است. با بررسی نتیجه بدست آمده از توالی تصاویر waving trees مشهود است: بهترین پاسخ و کمترین خطا در این نمونه، مربوط به سیگمادلتا و قانون زیف می‌باشد و پس از آن سیگمادلتا و سنجش اطمینان در مکان دوم و سیگمادلتای چندگانه و شرطی قرار دارند و سپس سیگمادلتای سلسله‌مراتبی قرار گرفته است. در مورد پایگاه داده SABS بهترین پاسخ را سیگمادلتای چندگانه دارا می‌باشد ولی کمترین خطا در بین الگوریتم‌ها مربوط به سیگمادلتا و قانون زیف می‌باشد.

پس از سیگمادلتای چندگانه، سیگمادلتا و قانون زیف در مکان دوم و سیگمادلتای شرطی و سلسله‌مراتبی در مکانهای بعدی قرار گرفته‌اند. از مشاهدات بالا می‌توان نتیجه گرفت که بهترین پاسخ در حالت کلی مربوط به سیگمادلتای چندگانه بوده و بعد از آن سیگمادلتا و قانون زیف قرار دارد. البته سیگمادلتا و قانون زیف در بین تمام الگوریتم‌ها دارای کمترین خطا نیز می‌باشد. همانطور که مشهود است افزودن فیدبک ربط به الگوریتم سیگمادلتای پایه باعث بهبود عملکرد آن می‌شود. در مورد نتایج الگوریتم سیگمادلتا با سنجش اطمینان باید تحلیل کیفی انجام پذیرد زیرا این الگوریتم دارای ویژگیهای منحصر به فردی است و به طور ویژه برای کار کردن در شرایط ترافیک شهری ایجاد شده است. از جمله ویژگیهای آن می‌توان به مقاومت زیاد زمینه آن در برابر آلوده شدن توسط خودروهای دارای حرکت آهسته یا متوقف شده در زمان کوتاه نام برد که خودروی متوقف شده حدود ۵ دقیقه طول می‌کشد تا شروع به جذب شدن در زمینه نماید، که این ویژگی به طور ویژه در صف‌های ترافیکی و چراغ قرمزها مناسب می‌باشد.

پارامتر دیگری نیز که در مقایسه این الگوریتم‌ها مهم و حیاتی می‌باشد هزینه محاسباتی و حجم پردازش این الگوریتم‌ها می‌باشد. به منظور بررسی این پارامتر زمان لازم برای اجرای هر یک از این الگوریتم‌ها را در شرایط یکسان اندازه‌گیری کرده و با یکدیگر مقایسه کرده‌ایم. قابل توجه است که این زمان بیانگر هزینه محاسباتی این الگوریتم‌ها می‌باشد. نتیجه محاسبه‌ی این زمان برای توالی ۶۰۰ فریم از دیتابیس SABS برای تمام الگوریتم‌ها به در شکل (۱۳) آمده است.

که دقت و فراخوانی را در فرم میانگین هارمونیکشان باهم ترکیب می‌کند و شاخصی را ایجاد می‌کند که بهتر از اندازه‌گیری‌های PR و RE به تنهایی عمل می‌کند. درصد طبقه‌بندی صحیح<sup>۴۷</sup>:

$$S_{CC} = \frac{TP+TN}{TP+TN+FP+FN}$$

از درصد طبقه‌بندی صحیح به‌تنهایی برای ارزیابی عملکرد یک طبقه‌بندی‌کننده استفاده می‌شود. هرچند، این درصد می‌تواند در زمانی که انحراف چشمگیری در توزیع طبقه وجود داشته باشد، تخمینی گمراه‌کننده باشد. به طور ویژه اگر عناصر زمینه تنها در بخش کوچکی از تصویر برای مثال ۵٪، حضور داشته باشند، تفاوت بارزی بین نرخهای بالای به‌دست آمده توسط این ضریب با حالتی که به طور ساده همه چیز را به عنوان زمینه دسته‌بندی کنیم، وجود ندارد. به‌علاوه با استفاده از ضرایب جاکارد<sup>۴۸</sup> و یول<sup>۴۹</sup> [۴۸] در زمانی که حجم زیادی از TN مورد انتظار وجود دارد، می‌توان از این مشکل کاست.

ضریب جاکارد:  $S_J = \frac{TP}{TP+FP+FN}$   $0 \leq S_J \leq 1$   
 ضریب یول:  $S_Y = \frac{TP}{TP+FP} S_J + 1 = PR + PR_N - 1 = \frac{TN}{TN+FN}$   $1 \leq S_Y \leq 2$   
 $PR_N$  را می‌توان به‌عنوان دقت در طبقه‌بندی زمینه (منفی) دانست؛ به‌طریقی مشابه، PR برابر با دقت در طبقه‌بندی پیش‌نما (مثبت) می‌باشد. در فرم اصلی، ضریب یول در بازه‌ی  $[-1, 1]$  تعریف می‌شود. حد پایین این بازه زمانی رخ می‌دهد که هیچ پیکسل منطقی وجود نداشته باشد، درحالی‌که تطبیقی کامل، کران بالای این بازه را می‌سازد. از خطای طبقه‌بندی میانگین<sup>۵۰</sup> برای اندازه‌گیری خطای هر روش بر اساس تعداد زمینه‌های واقعی موجود برای هر توالی داده، استفاده می‌شود.

$$\text{Average Classification Error} = \frac{FP+FN}{N_g}$$

که  $N_g$  تعداد فریم‌هایی می‌باشد که زمینه واقعی موجود است [۳]. در آخر Ilyas و همکارانش یک فاصله اقلیدسی وزن‌دار شده را پیشنهاد کردند [۴۹] که در آن انحرافات FPR و TPR را از مقادیر نسبی ایده‌آل شان ۰ یا ۱ در نظر گرفته و به‌صورت زیر تعریف می‌شود:

$$E_\gamma = \sqrt{\gamma FPR^2 + (1-\gamma)(1-TPR)^2} \quad (0 \leq E_\gamma \leq 1)$$

که  $\gamma$  ( $0 < \gamma < 1$ ) ضریب وزن‌دهی بوده و باید بر اساس مصالحه بین حساسیت و اختصاصی بودن تنظیم گردد. برای مثال زمانی که یک نرخ پایین هشدار تشخیص اشتباه به قیمت دادن حساسیت اولویت داشته باشد، باید مقادیر بالا برای این ضریب انتخاب گردد.

در این بخش، مجموعه شاخص زیر به عنوان کمیتی ارزشمند از عملکرد نسبی هر یک از الگوریتم‌ها در نظر گرفته شده است:

$$S = (S_J, S_Y, 0.5(1 + S_Y)), \quad E = \{(E_{0.25}, E_{0.5}, E_{0.75})\}$$

الگوریتم‌های مطرح شده بر اساس تفریق زمینه را بر روی یک کامپیوتر شخصی (Pentium dual core 3.00 GHz) با سیستم عامل Win7 و Matlab 2012b پیاده‌سازی کردیم. برای تست کردن الگوریتم‌ها دیتابیس‌های SABS<sup>۵۱</sup> و Wallflower<sup>۵۲</sup> برگزیده شدند. این الگوریتم‌ها را برای ۷ فریم برگزیده شده از دیتابیس SABS و یک

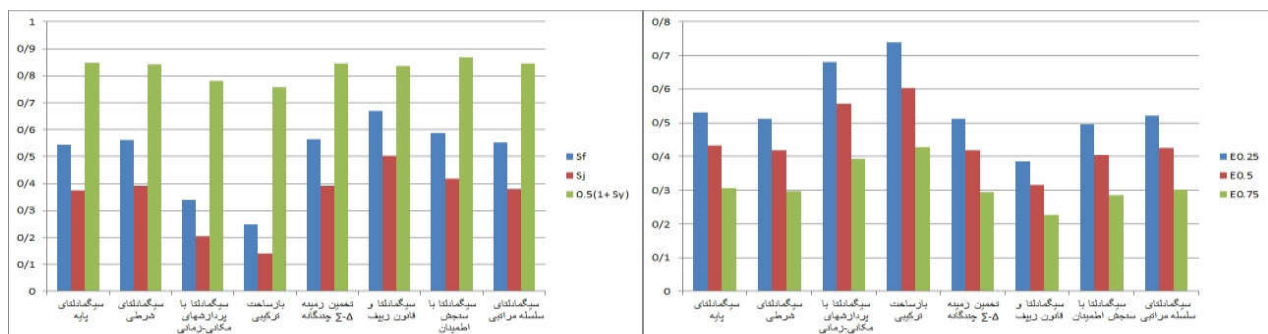


شکل (۹): نتایج مقایسه‌ای نقابهای تشخیص الگوریتم‌های مطرح شده نسبت به زمینه واقعی برای توالی تصاویر Waving trees از دیتابیس Wallflower

Fig. (9): The comparative results of detection masks for the proposed algorithms relative to the actual background for the sequence of Waving trees images from the Wallflower database

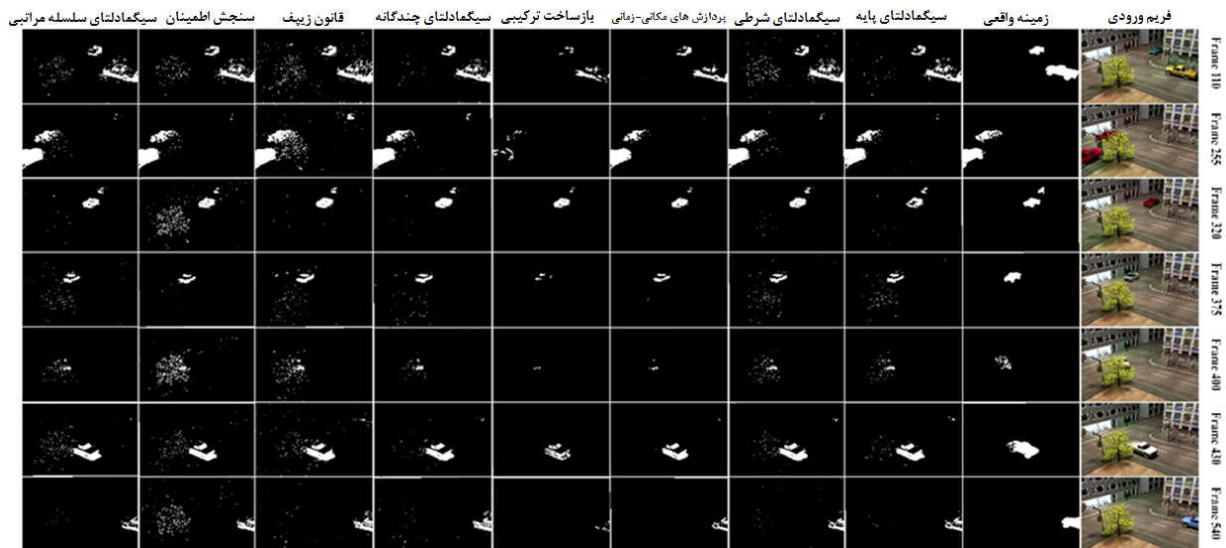
همانطور که در شکل (۱۳) مشاهده می‌شود الگوریتم سیگمادلتهای چندگانه در بین تمام الگوریتم‌ها بیشترین هزینه محاسباتی و پایین‌ترین بازده محاسباتی را دارا است که با توجه به ماهیت الگوریتم نیز کاملاً بدیهی می‌باشد. نکته قابل توجه این است که افزودن قانون زیف به الگوریتم سیگمادلتهای تنها باعث بهبود پاسخ آن می‌شود و خطای آن را کاهش می‌دهد، بلکه به وضوح باعث کاهش هزینه‌های محاسباتی الگوریتم و افزایش بهینگی محاسباتی آن می‌شود. در ادامه مقایسه الگوریتم‌های سیگمادلتهای پایه و الگوریتم سیگمادلتهای چندگانه به عنوان نماینده الگوریتم‌های سیگمادلتهای با الگوریتم‌های تفریق زمینه ساده SBS، میانگین جاری RA، میانگین جاری در حوزه DCTRADCT، تفاوت آماری ساده SSD، تفریق زمینه بر اساس اعوجاج روشنایی و رنگ BSBCD و مخلوط گوسیها MoG در [۲] انجام شده است. نتایج بدست آمده در جدول (۱) برای میزان  $F(S_F)$  و در شکل (۱۴) برای خطای طبقه‌بندی میانگین آورده شده است [۲]. [۳]

همانطور که در شکل (۱۴) و جدول (۱) مشاهده می‌شود الگوریتم سیگمادلتهای چندگانه به طور میانگین نتایج بسیار خوبی در بین تمامی الگوریتم‌های مورد مقایسه ارائه کرده است. همچنین نتایج این الگوریتم در بین توالی تصاویر مختلف دچار تغییرات و افت محسوسی نشده است که نشان می‌دهد این الگوریتم و در کل الگوریتم سیگمادلتهای در شرایط محیطی مختلف به خوبی عمل می‌کند و این یکی دیگر از مزیت‌های این الگوریتم می‌باشد.

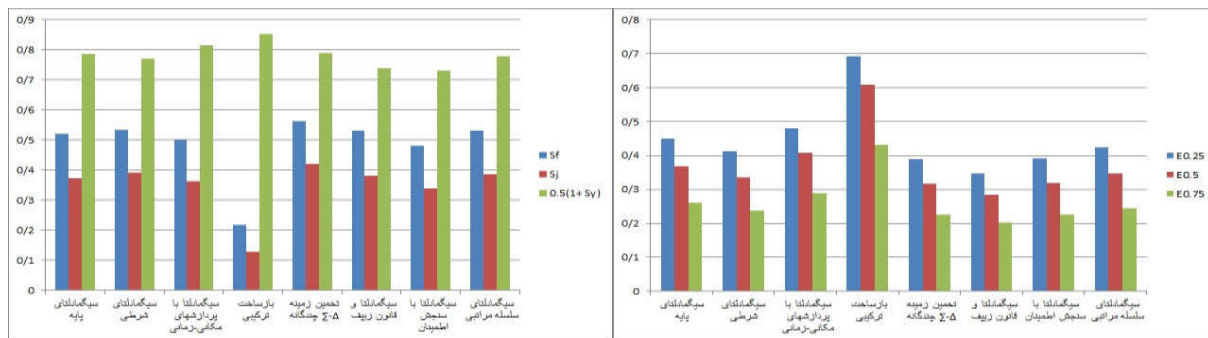


شکل (۱۰): نمودارهای  $S_F, S_J, 0.5(1 + S_F)$  (سمت چپ) و  $E_{0.25}, E_{0.5}, E_{0.75}$  (سمت راست) برای الگوریتم‌های مطرح شده در توالی تصاویر Waving trees از دیتابیس Wallflower

Fig. (10): Graphs of  $S_F, S_J, 0.5(1 + S_F)$  (left side) and  $E_{0.25}, E_{0.5}, E_{0.75}$  (right side) for the proposed algorithms in the Waving trees images from the Wallflower database



شکل (۱۱): نتایج مقایسه‌ای نقاب‌های تشخیص الگوریتم‌های مطرح شده نسبت به زمینه واقعی برای فریم‌های مختلف توالی تصاویر دیتابیس SABS  
 Fig. (11): Comparison of the mask detection of the proposed algorithms in relation to the actual background for the various frameworks of images sequences of the SABS database



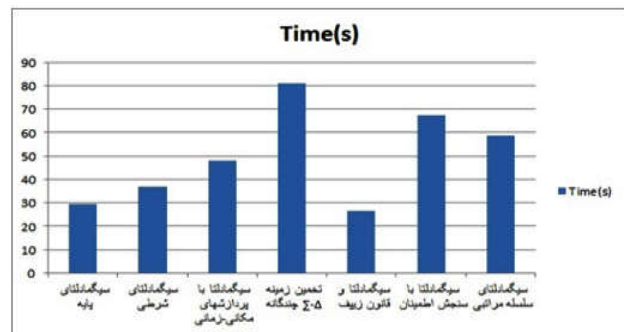
شکل (۱۲): نمودارهای  $S_F, S_J, 0.5(1 + S_J)$  (سمت چپ) و  $E_{0.25}, E_{0.5}, E_{0.75}$  (سمت راست) برای الگوریتم‌های مطرح شده برای میانگین فریم‌های ارزیابی شده در توالی تصاویر دیتابیس SABS  
 ارزیابی شده در توالی تصاویر دیتابیس SABS

Fig. (12): Graphs of  $S_F, S_J, 0.5(1 + S_J)$  (left side) and  $E_{0.25}, E_{0.5}, E_{0.75}$  (right side) for the proposed algorithms for the mean of evaluated frames in the SABS database image sequence

جدول (۱): نتایج محاسبه‌ی میزان  $F(S_F)$  برای الگوریتم‌های به ترتیب از بالا تفریق زمینه ساده، میانگین جاری، میانگین جاری در حوزه‌ی DCT، سیگما-دلتهای پایه، سیگمادلتهای چندگانه، تفاوت آماری ساده، تفریق زمینه بر اساس اعوجاج روشنایی، رنگو مخلوط گوسیها در توالی تصاویر مختلف.

Table (1): The results of calculating the F ( $S_F$ ) for the algorithms were from above, simple background, current average, current average in the DCT domain, basic Sigma Delta, multifrequency Sigma Delta, simple statistical difference, background differentiation based on light distortion, color and mixing Gaussians in a sequence of different images.

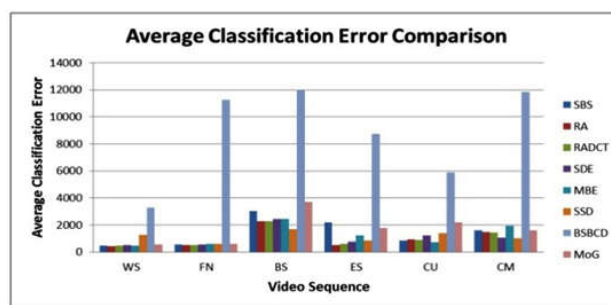
توالی ویدئو						روش
CM	CU	ES	BS	FN	WS	
۰,۲۲۰	۰,۵۶۲	۰,۰۸۶	۰,۲۰۴	۰,۵۴۹	۰,۷۷۸	SBS
۰,۲۲۲	۰,۵۶۶	۰,۳۰۲	۰,۱۸۲	۰,۴۷۲	۰,۷۸۴	RA
۰,۲۲۷	۰,۵۵۸	۰,۲۹۵	۰,۱۸۴	۰,۷۷۶	۰,۴۷۴	RADCT
۰,۱۵۰	۰,۲۰۸	۰,۱۸۴	۰,۰۶۰	۰,۱۵۳	۰,۷۸۸	SDE
۰,۱۷۵	۰,۶۵۴	۰,۱۵۳	۰,۲۳۸	۰,۴۹۸	۰,۷۹۲	MBE
۰,۲۱۵	۰,۳۷۵	۰,۰۱۵	۰,۴۸۵	۰,۲۵۴	۰,۷۲۰	SSD
۰,۰۶۷	۰,۳۳۵	۰,۰۲۲	۰,۰۷۲	۰,۰۴۸	۰,۴۴۸	BSBCD
۰,۳۲۰	۰,۱۶۱	۰,۱۷۸	۰,۳۶۶	۰,۶۹۰	۰,۷۸۸	MOG



شکل (۱۳): زمان لازم برای اجرای ۶۰۰ فریم از دیتابیس SABS برای تمام الگوریتم‌ها

Fig. 13: The required time to execute 600 frames from the SABS database for all algorithms

11. Kernel Density Estimation
12. Heneceforth MoG
13. Codebook model
14. Codewords
15. Background Subtraction based on Brightness and Chromaticity Distortion
16. Kalman filter
17. Wiener filter
18. Eigenspace
19. Dynamic Mode Decomposition
20. Stacked Multilayer Self-Organizing Map Background Model
21. Online Robust Principal Component Analysis
22. Motion-Assisted Matrix Restoration
23. Vehicle Tracking and Speed Measurement system
24. Counter-propagation artificial neural network
25. Region update (RU)
26. Ghost effect
27. Aperture effect
28. Radial movement
29. False detection
30. Common-edge Hybrid Reconstruction
31. Opening by reconstruction
32. Temporal validation
33. Geodesic
34. Relaxation
35. Reference Image
36. Marker image
37. Adaptive thresholding
38. Opening by reconstruction
39. Least significant bit
40. Integer
41. Modulo
42. Overflow
43. Conditional temporal update
44. Conditional temporal update
45. Receiver Operator Characteristic
46. Precision Recall
47. Percentage of correct classification
48. Jaccard coefficient
49. Yule coefficient
50. Average Classification Error
51. <http://www.vis.uni-stuttgart.de/index.php?id=sabs>
52. <http://research.microsoft.com/en-us/um/people/jckrumm/wallflower/testimages.htm>



شکل (۱۴): نتایج محاسبه‌ی خطای طبقه‌بندی میانگین برای الگوریتم‌ها به ترتیب جدول (۱) در توالی تصاویر مختلف.

Fig. (14): The results of the calculation of the average classification error for the algorithms are given in Table 1, in the sequence of different images, respectively.

#### ۴ - نتیجه گیری

در این مقاله، ارزیابی الگوریتم تفریق زمینه سیگمادلتا و مکمل‌های ارائه شده برای آن ارائه شده است. در میان الگوریتم‌های بررسی شده، الگوریتم سیگمادلتا و قانون زیف کمترین خطا و کمترین هزینه محاسباتی را دارا می‌باشد یا به عبارتی بیشترین بازدهی را دارد. در کاربردهایی که هزینه‌های محاسباتی چندان برای ما مهم نیست و بهترین پاسخ‌دهی مهم می‌باشد، الگوریتم سیگمادلتای چندگانه نتایج مطلوبی ارائه می‌دهد. همچنین برای استفاده در محیط‌های ترافیک شهری و نظارت ترافیکی، الگوریتم سیگمادلتا و سنجش اطمینان که به طور ویژه برای کار در این محیط‌ها ارتقا یافته است، به دلیل ویژگی‌های خاص آن بهترین انتخاب می‌باشد. این بررسی به محققان این امکان را می‌دهد که این مکمل‌ها را از لحاظ پارامترهایی مانند پیچیدگی، هزینه‌های محاسباتی، دقت پاسخ و خطای الگوریتم‌ها با یکدیگر و با الگوریتم اصلی سیگمادلتا مقایسه کنند و متناسب با شرایط تحقیق بهترین را انتخاب نمایند.

پی‌نوشت:

1. Background Initialization
2. Foreground Detection
3. Background Maintenance
4. Simple Background Subtraction (SBS)
5. Running Average Method
6. Discrete Cosine Transform (DCT)
7. Simple Statistical Difference
8. Running Gaussian average
9. Temporal Median Filter
10. Mixture of Gaussians (MoG)

#### References

- [1] A. Manzanera, J.C. Richefeu, "A new motion detection algorithm based on  $\Sigma-\Delta$  background estimation", Pattern Recognition Letters, Vol. 28, No. 3, pp. 320-328, 2007.
- [2] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview", Computer Science Review, Vol. 11, pp. 31-66, 2014.
- [3] A. Gandhamal, S. Talbar, "Evaluation of background subtraction algorithms for object extraction", Proceeding of the IEEE International Conference on Pervasive Computing (ICPC), pp. 1-6, 2015.
- [4] C.R. Wren, A. Azarbajehani, T.J. Darrell, A.P. Pentland, "Pfinder: Real-time tracking of the human body", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 780-785, Jul. 1997.

- [5] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, S. Russell, "Towards robust automatic traffic scene analysis in real time", *Proceeding of the ICPR*, pp. 126–131, Nov. 1994.
- [6] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, "Detecting moving objects, ghosts, and shadows in video streams", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 25, No. 10, pp. 1337–1342, Oct. 2003.
- [7] S.C. Cheung, C. Kamath, "Robust techniques for background subtraction in urban traffic video", *Proceeding of the SPIE Electron. Imaging Video Commun. Image Process.*, pp. 1–12, Jan. 2004.
- [8] B.P.L. Lo, S.A. Velastin, "Automatic congestion detection system for underground platforms", *Proceeding of the ISIMP*, pp. 158–161, May 2001.
- [9] Q. Zhou, J. Aggarwal, "Tracking and classifying moving objects from videos", *Proceeding of the IEEE Workshop Perform. Eval. Tracking Surveillance*, pp. 52–59, 2001.
- [10] C. Stauffer, W. Grimson, "Adaptive background mixture models for real-time tracking", *Proceeding of the IEEE Conf. Comput. Vis. Pattern Recog*, Vol. 2, pp. 246–252, 1999.
- [11] M. Harville, "A framework for high-level feedback to adaptive, per-pixel, mixture-of-Gaussian background models", *Proceeding of the Eur. Conf. Comput. Vis.*, Vol. 3, pp. 543–560, May 2002.
- [12] W. Power, J.A. Schoonees, "Understanding background mixture models for foreground segmentation", *Proceeding of the IVCNZ*, pp. 267–271, Nov. 2002.
- [13] A. Elgammal, D. Harwood, L. Davis, "Non-parametric model for background subtraction", *Proceeding of the IEEE ICCV Frame-Rate Workshop*, pp. 1–15, Sep. 1999.
- [14] F. Porikli, O. Tuzel, "Bayesian background modeling for foreground detection", *Proceeding of the ACM Vis. Surveillance Sens. Netw.*, pp. 55–58, 2005.
- [15] K. Kim, T.H. Chalidabhongse, D. Harwood, L. Davis, "Real-time foreground–background segmentation using codebook model", *Real-Time Imaging*, Vol. 11, No. 3, pp. 172–185, Jun. 2005.
- [16] K.P. Karmann, A. Brandt, "Moving object recognition using an adaptive background memory", *Time-Varying Image Processing and Moving Object Recognition*, in V. Cappellini, Ed. Amsterdam, The Netherlands: Elsevier, pp. 289–307, 1990.
- [17] D. Koller, J. Weber, J. Malik, "Robust multiple car tracking with occlusion reasoning", *Proceeding of the ECCV*, pp. 189–196, Sweden, May 1994.
- [18] K. Toyama, J. Krumm, B. Brumitt, B. Meyers, "Wallflower: Principles and practice of background maintenance", *Proceeding of the ICCV*, pp. 255–261, Greece, Sep. 1999.
- [19] A. Monnet, A. Mittal, A. Paragios, V. Ramesh, "Background modeling and subtraction of dynamic scenes", *Proceeding of the ICCV*, pp. 1305–1312, France, Oct. 2003.
- [20] J. Zhong, S. Sclaroff, "Segmenting foreground objects from a dynamic, textured background via a robust Kalman filter", *Proceeding of the ICCV*, pp. 44–50, France, Oct. 2003.
- [21] N.M. Oliver, B. Rosario, A.P. Pentland, "A Bayesian computer vision system for modeling human interactions", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 22, No. 8, pp. 831–843, Aug. 2000.
- [22] H. Bhaskara, K. Dwivedic, D.P. Dograd, M. Al-Muallaa, L. Mihaylova, "Autonomous detection and tracking under illumination changes, occlusions and moving camera", *Signal processing*, Vol. 117, pp. 343–354, 2015.
- [23] D. Volpi, M.H. Sarhan, R. Ghotbi, N. Navab, D. Mateus, S. Demirci International, "Online tracking of interventional devices for endovascular aortic repair", *Journal of Computer Assisted Radiology and Surgery*, Vol. 10, No. 6, pp. 773–781, 2015.
- [24] N.B. Erichson, C. Donovan, "Randomized low-rank Dynamic Mode Decomposition for motion detection", *Computer Vision and Image Understanding*, Vol. 146, pp. 40–50, 2016.
- [25] Z. Zhao, X. Zhang, Y. Fang, "Stacked Multilayer Self-Organizing Map for Background Modeling", *IEEE Transactions on Image Processing*, Vol. 24, No. 9, pp. 2841–2850, 2015.
- [26] K. Wang, Y. Liu, C. Gou, F.Y. Wang, "A multi-view learning approach to foreground detection for traffic surveillance applications", *IEEE Trans. on Vehicular Technology*, Vol. 65, No. 6, pp. 4144–4158, June 2016.
- [27] D.D. Bloisi, A. Pennisia, L. Iocchia, "Parallel multi-modal background modeling", *Pattern Recognition Letters*, Vol. 96, pp. 45–54, 2017.
- [28] G. Han, J. Wang, X. Cai, "Background subtraction based on modified online robust principal component analysis", *International Journal of Machine Learning and Cybernetics*, Vol. 8, No. 6, pp. 1839–1852, 2017.
- [29] X. Ye, J. Yang, X. Sun, K. Li, C. Hou, Y. Wang, "Foreground–background separation from video clips via motion-assisted matrix restoration", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 25, No. 11, pp. 1721–1734, Jan. 2015.



- [30] D. Jeyabharathi, D. Deje, "Vehicle Tracking and Speed Measurement system (VTSM) based on novel feature descriptor: Diagonal Hexadecimal Pattern (DHP)", *Journal of Visual Communication and Image Representation*, Vol. 40, pp. 816–830, 2016.
- [31] B.H. Chen, S.C. Huang, J.Y. Yen, "Counter-propagation artificial neural network-based motion detection algorithm for static-camera surveillance scenarios", *Neurocomputing*, Vol. 273, pp.481-493, 2018.
- [32] J. Guo, P. Zheng, J. Huang, "An efficient motion detection and tracking scheme for encrypted surveillance videos", *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol.13, No.4, pp.1-23, 2017.
- [33] K. Yun, J. Lim, J.Y. Choi, "Scene conditional background update for moving object detection in a moving camera", *Pattern Recognition Letters*, Vol. 88, pp. 57-63, 2017.
- [34] N. McFarlane, C. Schofield, "Segmentation and tracking of pigglets in images", *Mach. Vis. Appl.*, Vol. 8, No. 3, pp. 187–193, May 1995.
- [35] i-LIDS Dataset for AVSS 2007. [Online]. Available: <ftp://motinas.elec.qmul.ac.uk/pub/iLids>
- [36] L. Lacassagne, A. Manzanera, A. Dupret, "Motion detection: Fast and robust algorithms for embedded systems", *Proceeding of the ICIP*, pp. 3265-3268, 2009.
- [37] J. Denoulet, G. Mostafaoui, L. Lacassagne, A. Merigot, "Implementing motion markov detection on general purpose processor and associative mesh", *Proceeding of the CAMP*, 2005.
- [38] A. Caplier, C. Dumontier, F. Luthon, P. Coulon, "Mrf based motion detection algorithm image processing board implementation", *Traitement du signal*, Vol.13, No. 2, pp. 177–190 (in French), 1996.
- [39] L. Lacassagne, M. Milgram, P. Garda, "Motion detection, labeling, data association and tracking in real-time on risc computer", *Proceeding of the IEEE ICIAP*, pp. 520–525, 1999.
- [40] H.J.A.M. Heijmans, "Connected morphological operators for binary images," *Comput. Vis. Image Understand.*, Vol. 73, No. 1, pp. 99–120, Jan. 1999.
- [41] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms", *IEEE Trans. on Image Processing*, Vol. 2, No. 2, pp. 176–201, April 1993.
- [42] P. Salembier, J. Ruiz, "Connected operators based on reconstruction process for size and motion simplification", *Proceeding of the IEEE Int. Conf. Acoust., Speech, Signal Process*, Vol. 4, pp. 3289–3292., 2002.
- [43] A. Manzanera, J.C. Richefeu, "A robust and computationally efficient motion detection algorithm based on sigma-delta background estimation", *Proceeding of the ICVGIP*, pp. 46–51, Dec. 2004.
- [44] M. Vargas, J.M. Milla, S.L. Toral, F. Barrero, "An enhanced background estimation algorithm for vehicle detection in urban traffic scenes", *IEEE Transactions on Vehicular Technology*, Vol. 59, No. 8, pp. 3694-3709, 2010.
- [45] G.K. Zipf, "Human behavior and the principle of least-effort", Addison-Wesley, 1949.
- [46] Y. Caron, P. Makris, N. Vincent, "A method for detecting artificial objects in natural environments", *Proceeding of the Int. Conf. in Pattern Recognition*, pp. 600–603, 2002.
- [47] A. Manzanera, " $\Sigma$ - $\Delta$  Background Subtraction and the Zipf Law", Berlin, Germany: Springer-Verlag, pp. 42–51, 2008.
- [48] P. Sneath, R. Sokal, "Numerical Taxonomy: The Principle and Practice of Numerical Classification", San Francisco, CA: Freeman, 1973.
- [49] A. Ilyas, M. Scuturici, S. Miguet, "Real time foreground-background segmentation using a modified Codebook model", *Proceeding of the 6th IEEE Int. Conf. AVSS*, pp. 454–459, 2009.

