

تشخیص تغییرات صحنه به روش زمینه‌گیری هوشمند

محمود فتحی و حسن شکوری

سیستم و در حالت عکس، باعث افزایش هشدارهای نابجا می‌گردد. یک سیستم تشخیص تغییرات صحنه باید علاوه بر اینکه اشیاء متجاوز به صحنه را مستقل از سرعت حرکت آنها تشخیص می‌دهد باید نسبت به جایجایی، مفقود شدن یا آسیب دیدن اجسام داخل صحنه نیز حساس باشد؛ این در حالی است که حرکت اجسام داخل صحنه (مانند حرکت برگ درختان)، تغییرات شرایط نوری محیط، نویز ناشی از تصویربرداری و مواردی اینچنین که از آنها به عنوان تغییرات ناخواسته یاد می‌شود نباید موجب بروز اشتباه در تصمیم‌گیری شود و سیستم باید در مقابل آنها مقاوم باشد. همچنین به دلیل اهمیت زمان پاسخ در اکثر کاربردهای سیستم‌های نظارتی، باید کلیه ملاحظات مربوط به سیستم‌های بلادرنگ را نیز به مجموعه پیچیدگی‌های این سیستم افزود.

هدف از این مقاله، ارائه روشی است که در برابر اکثر شرایط موجود در محیط‌های واقعی، رفتاری مناسب ارائه کند و سرعت اجرای آن روی کامپیوتر شخصی، برای درصد بالایی از کاربردها مناسب باشد. ادامه مقاله متشکل از این قسمت‌ها است: در بخش دوم به بررسی کارهای انجام شده قبلی پرداخته می‌شود؛ بخش سوم، به تشریح الگوریتم جدید اختصاص دارد و نتایج پیاده‌سازی عملی در بخش چهارم ارائه می‌گردد؛ بخش نتیجه‌گیری نیز پایان‌بخش مقاله خواهد بود.

۲- کارهای انجام شده قبلی

تاکنون محققین زیادی، در رابطه با جداسازی روزمینه از زمینه کار کرده‌اند و الگوریتم‌های متعددی در این رابطه ارائه نموده‌اند که البته با نگاهی کلی‌تر، می‌توان اصول به کار رفته در این الگوریتم‌ها را به چند دسته فراگیر تقسیم کرد.

روش مبتنی بر جریان نوری^۱، از شیوه‌هایی است که در برخی از سیستم‌ها به کار گرفته شده است [۱] و [۲]. مهمترین نقطه ضعف این روش، فرض اولیه آن مبنی بر حرکت دائمی اجسام روزمینه است. این خصوصیت به همراه پیچیدگی محاسباتی بالا [۳] و [۴]، موجبات عدم استقبال محققین از آن را در این کاربرد فراهم آورده است.

روش مبتنی بر ویژگی^۲، عنوان روشی است که به نظر، تنها برای سیستم‌های تک‌منظوره مناسب بوده و استفاده از آن برای تشخیص دایره وسیعی از اجسام روزمینه در محیط‌های تعریف نشده عملی نمی‌باشد؛ نمونه‌ای از این روش در [۵] به کار گرفته شده است.

روش دیگر، روش مبتنی بر تفاضل فریم‌های متوالی است [۶] تا [۱۰]؛ به این ترتیب که در آن از اطلاعات تفاضل دو یا چند فریم متوالی، برای استخراج روزمینه استفاده می‌شود. در این روش نیز همانند روش مبتنی بر جریان نوری، فرض بر حرکت دائمی اجسام متجاوز واقع شده است و در صورتیکه جسم در حین گرفته شدن تصاویر، حرکتی از خود نشان ندهد به عنوان روزمینه شناخته نخواهد شد.

چکیده: جداسازی روزمینه از زمینه، در بسیاری از کاربردهای نظارت تصویری به عنوان اولین و مهمترین قدم شناخته می‌شود و "تفاضل زمینه" روشی است که معمولاً برای این منظور استفاده می‌شود. این روش، هر فریم را با مدلی از صحنه خالی مقایسه کرده و ناحیه‌هایی از آنرا که بطور قابل ملاحظه‌ای متفاوت به نظر می‌رسند به عنوان نواحی روزمینه مشخص می‌کند.

این مقاله روش جدیدی برای تفاضل زمینه ارائه می‌کند که در آن ابتدا هر تصویر به بلوکهای $N \times N$ یکسانی تفکیک شده و سپس ویژگیهای تعیین‌کننده‌ای از بلوک‌ها محاسبه شده و سابقه مقادیر هر یک از این ویژگی‌ها، به صورت ترکیبی از توزیعهای گوسین مدل می‌گردد. با ورود هر فریم جدید، این توزیعها با روش سریعی بهنگام می‌شوند آنگاه توزیعهای گوسین مدل‌های ترکیبی، برای یافتن توزیعهایی که بیان‌کننده زمینه هستند ارزیابی می‌شوند و هر بلوک بر اساس اینکه مقادیر ویژگیهای آن جزو کدامیک از توزیعها باشد در دو کلاس زمینه و روزمینه دسته‌بندی می‌شود.

پیاده‌سازی نرم‌افزاری این روش روی کامپیوتر شخصی، حاکی از عملکرد قابل قبول سیستم در برابر اجسام متجاوز به صحنه (مستقل از سرعت حرکت آنها)، اضافه یا کاسته شدن اجسام داخل صحنه، نویز تصویربرداری و تغییرات ناخواسته صحنه است و سرعت بالا و نیاز به حافظه کم، آنرا برای درصد بالایی از کاربردهای بلادرنگ مناسب می‌سازد.

کلید واژه: بلوک بندی تصاویر، ترکیبی از گوسین‌ها، تفاضل زمینه، جداسازی روز زمینه.

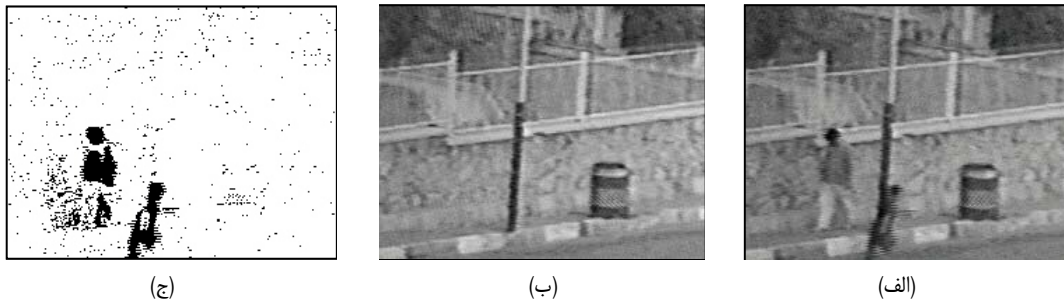
۱- مقدمه

سیستم‌های نظارت تصویری خودکار، معمولاً بسته به نوع کاربرد مراحل مختلفی از تشخیص را شامل می‌شوند؛ که معمولاً در این میان تشخیص تغییرات صحنه، اولین و مهمترین آنها محسوب می‌گردد. هدف از این مرحله، جداسازی نواحی مربوط به روزمینه^۱ و زمینه^۲ در تک‌تک فریم‌های سری تصاویر است. روزمینه، قسمت‌هایی از تصویر است که در حین نظارت، تغییراتی حساسیت‌زا داشته باشند؛ در نقطه مقابل، هر ناحیه‌ای از تصویر که جزو روزمینه نباشد، مربوط به زمینه خواهد بود. این نواحی یا شاهد هیچگونه تغییری نبوده‌اند و یا اینکه با توجه به کاربرد تعریف شده برای سیستم، این تغییرات طبیعی تلقی شده‌اند. اشتباه سیستم در کلاس‌بندی غلط روزمینه به عنوان زمینه، منجر به کاهش دقت کاری

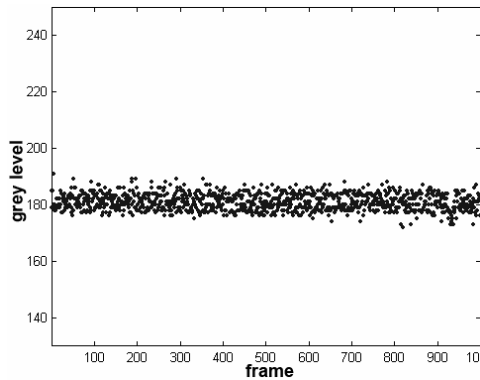
این مقاله در تاریخ ۲۸ آذر ماه ۱۳۸۰ دریافت و در تاریخ ۲۸ تیر ماه ۱۳۸۱ بازنگری شد. این تحقیق به عنوان پایان نامه کارشناسی ارشد مهندسی کامپیوتر در گرایش هوش مصنوعی در دانشگاه علم و صنعت ایران پذیرفته شده است.

محمود فتحی، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، نارمک، کد پستی ۱۶۸۴۴ (email: mahfathy@yahoo.com).

حسن شکوری، گروه پژوهشی اتوماسیون، جهاد دانشگاهی صنعتی شریف، ضلع شمالی دانشگاه صنعتی شریف، تهران کد پستی ۱۴۵۸۸ (email: shakuri@yahoo.com)



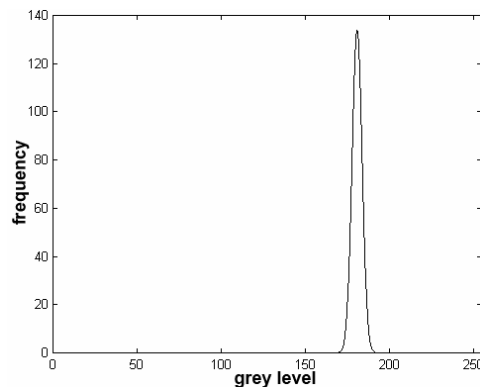
شکل ۱: (الف) تصویر فریم جاری، (ب) تصویر مدل زمینه، (ج) نتیجه حاصل از تفاضل زمینه.



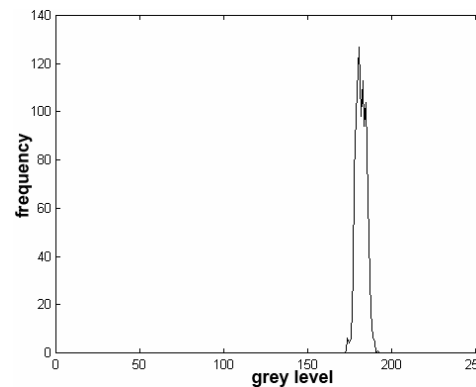
(ب)



(الف)



(د)



(ج)

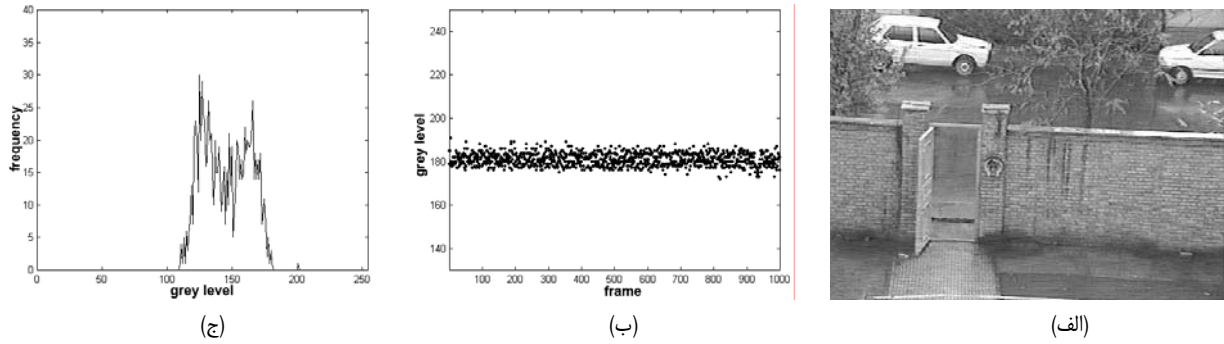
شکل ۲: (الف) یک فریم نمونه از سری تصاویر تهیه شده در شرایط کاملاً ثابت، (ب) مقادیر پیکسل (۱۳۰،۴۰۰) در طول ۱۰۰۰ فریم، (ج) توزیع مقادیر پیکسل مزبور، (د) مدل سازی توزیع مقادیر با یک توزیع نرمال.

برخی روش‌های طرح شده در [۱۲]، [۱۴]، [۱۵]، [۲۳]، [۲۸] و [۳۲] تا [۳۴]، با بکارگیری تصویر به عنوان مدل زمینه، شیوه‌هایی را برای بهنگام‌سازی مقادیر پیکسل‌های آن عنوان نموده‌اند که از آن جمله می‌توان به میانگین‌گیری، میانگین‌گیری و استفاده از فیلترهای پیشگو اشاره نمود. اکثر این روش‌ها، کلاً در شرایطی که سرعت تغییرات زمینه بسیار کندتر از سرعت تغییرات روزمینه باشد و نیز تصویر زمینه در بخش‌های قابل ملاحظه‌ای از زمان قابل رؤیت باشد مفید هستند ولی در شرایطی که اجسام متحرک داخل صحنه، به علت کثرت، مانع از رؤیت زمینه باشند، بخصوص هنگامی که حرکت اجسام نیز کند باشد کارایی این روش‌ها پایین می‌آید [۱۱]. شیوه‌های دیگر، بجای نگهداری مقادیر پیکسل‌ها، آنها را با مقادیر و پارامترهای دیگر مدل سازی می‌کنند.

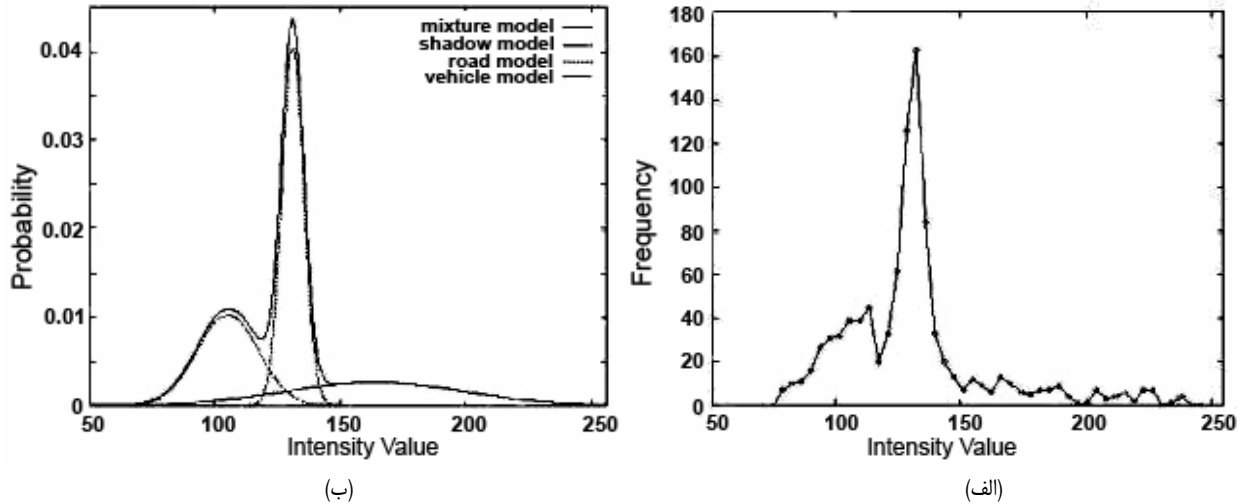
در سیستم W^4 [۲۷]، به ازای هر پیکسل از تصویر، سه مقدار در مدل کردن نقش دارند: حداکثر مقدار سطح خاکستری آن، حداقل مقدار سطح خاکستری آن و حداکثر مقدار تفاضلی که میان دو فریم متوالی

یک روش متداول، روشی موسوم به تفاضل زمینه است که بسیاری از کارهای ارائه شده، بر این پایه طراحی و پیاده‌سازی شده‌اند [۱۱] تا [۳۴]. در این شیوه، اطلاعات صحنه خالی از روزمینه، تحت عنوان مدل زمینه نگهداری می‌شود؛ آنگاه در هر لحظه، تصویر جاری با این مدل مقایسه شده و قسمت‌هایی از تصویر که نسبت به مدل زمینه تفاوت قابل ملاحظه‌ای دارند به عنوان روزمینه مشخص می‌شوند (شکل ۱).

عمده تفاوت روشهای مبتنی بر تفاضل زمینه، به نحوه مدل کردن زمینه و ماهیت آنها مربوط می‌شود. در روش‌های ابتدایی، این مدل متشکل از یک تصویر بود که بحث اصلی در آنها، شیوه انتخاب تصویر و نیز نحوه بهنگام‌سازی پیکسل‌های آن متناسب با تغییرات صحنه بود. در [۲۴]، اولین تصویر بدون هیچ تغییری تا مرحله آخر به عنوان مدل زمینه نگهداری می‌شود. این روش به دلیل ماهیت متغیر اکثر صحنه‌ها، تنها در شرایط کاملاً کنترل شده و برای کاربردهایی با زمان کوتاه مفید می‌باشد.



شکل ۳: (الف) یک فریم نمونه از سری تصاویر در شرایط بارانی، (ب) نمودار تغییرات پیکسل (۲۶۰، ۵۵۰) در طول ۱۰۰۰ فریم متوالی، (ج) توزیع مقادیر پیکسل مزبور، که یک نمودار چند قله‌ای است.



شکل ۴: (الف) توزیع مقادیر یک پیکسل در طول ۱۰۰۰ فریم متوالی تهیه شده از سطح خیابان، (ب) مدل ترکیبی از گوسین‌ها که داده‌ها را مدل می‌کند [۱۳].

مدل‌سازی سابقه مقادیر پیکسل‌ها با ترکیبی از توزیع‌های نرمال بوده است.

گریمن و استافر در آزمایشگاه هوش مصنوعی دانشگاه MIT، با دریافت این ایده توانستند روش Pfunder را برای محیط‌های خارجی توسعه دهند و مشکل توزیع‌های چند قله‌ای پیکسل‌ها را حل کنند. آنها در [۱۱] و [۲۲]، سابقه مقادیر هر پیکسل را با ترکیبی از حداکثر k توزیع گوسین^۲ مدل نموده‌اند ($3 \leq k \leq 5$). هر توزیع گوسین دارای سه پارامتر میانگین، واریانس و وزن می‌باشد که با گرفته شدن هر فریم جدید، این پارامترها بر اساس نسخه سریعی از الگوریتم K-means به‌نگام می‌شوند و در صورت نیاز توزیع جدیدی ایجاد می‌گردد. الگوریتم ارائه شده برای این منظور به گونه‌ای طراحی شده است که به مرور زمان توزیع‌های مربوط به زمینه، صاحب وزن بیشتر و واریانس کمتری نسبت به سایر توزیع‌ها می‌گردند به همین دلیل توزیع‌هایی که نسبت وزن به واریانس در آنها به میزان قابل توجهی بیشتر از سایر توزیع‌ها باشد به عنوان توزیع‌های زمینه انتخاب می‌گردند. در مرحله کلاس‌بندی پیکسل‌ها، هر پیکسل که مقدار آن به یکی از توزیع‌های زمینه تعلق داشته باشد جزو زمینه و در غیر اینصورت در دسته روزمینه کلاس‌بندی می‌شود.

طی گزارشی که از عملکرد روش ترکیبی از گوسین‌ها در [۱۱] ارائه شده است ادعا می‌شود که این روش در برابر شرایط نوری متغیر، سایه‌ها، حرکت کند اشیاء، حرکت اشیاء داخل صحنه، نویز تصویربرداری، و اضافه

مشاهده شده است. سیستم ارائه شده در [۲۵]، هر پیکسل از تصویر زمینه را با واریانس مقادیر پیکسل متناظر آن در طول N فریم قبلی مدل می‌کند.

سیستم Pfunder [۲۰]، مقادیر مربوط به پیکسل‌های یک زمینه کاملاً ثابت را، با یک توزیع نرمال مدل می‌کند (شکل ۲)؛ که این توزیع‌های نرمال، همراه با تغییرات تدریجی نور به‌نگام شده و پارامترهای آنها تنظیم می‌شوند. نتایج حاصل از پیاده‌سازی Pfunder در محیط‌های داخلی، حاکی از موفقیت این سیستم در مقابله با تغییرات تدریجی نور و نویز ناشی از تصویر برداری است اما طبیعی است که این روش در محیط‌های خارجی که اکثراً حاوی نواحی غیرثابت و متحرک هستند چندان مؤثر نباشد چرا که این نوع از حرکت‌ها سبب می‌شوند تا مقادیر پیکسل‌های نواحی مربوطه همراه با زمان تغییر کرده و نمودار توزیع مقادیر آنها چندقله‌ای شود (شکل ۳).

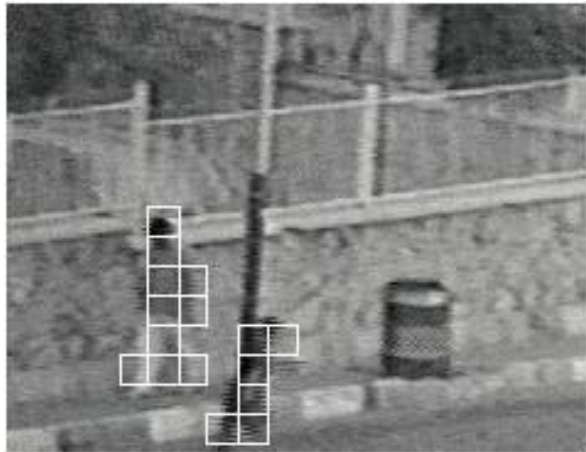
در روش [۱۳]، ترکیبی از سه توزیع نرمال جهت مدل کردن مقادیر پیکسل‌ها، در یک کاربرد نظارت ترافیک استفاده شده است. در این شیوه، سابقه مقادیر هر پیکسل با ترکیب وزن‌دار سه توزیع نرمال جاده، سایه و خودرو مدل می‌شود (شکل ۴) و برای یادگیری و نیز به‌نگام کردن پارامترهای این توزیع‌ها، از یک الگوریتم EM^۱ افزایشی کمک گرفته شده است. این شیوه اگر چه برای کاربردی خاص طراحی شده است و در اکثر شرایط مفید واقع نمی‌شود ولی برای اولین بار، ارائه‌کننده ایده‌ای مبنی بر



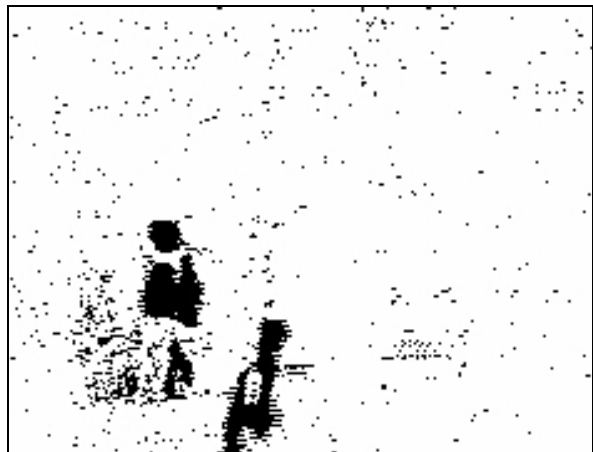
(ب)



(الف)



(د)



(ج)

شکل ۵: (الف) تصویر قبل از حضور افراد در صحنه، (ب) تصویر پس از حضور افراد در صحنه، (ج) تفاضل دو تصویر پیشین، (د) تصویر پس از بلوک‌بندی تصاویر تفاضلی.

۳- روش ترکیبی از گوسین‌های مبتنی بر بلوک

۳-۱ کلیات

در روش جدید، بجای آنکه مقادیر قبلی هر پیکسل با k گوسین مدل شوند، ترتیبی اتخاذ شده است تا تصویر به بلوک‌های $N \times N$ ثابتی تقسیم شده و مقادیر قبلی هر یک از ویژگی‌های یک بلوک، بصورت ترکیبی از گوسین‌ها مدل شوند. با این ترتیب، اگر مقادیر قبلی یکی از ویژگی‌ها را در یک بلوک خاص، به صورت $\{X_1, X_2, \dots, X_t\}$ در نظر بگیریم این مقادیر، با ترکیبی از k توزیع گوسین مدل می‌شوند. هر توزیع گوسین t در لحظه t با دو پارامتر $\mu_{i,t}$ و $\Sigma_{i,t}$ به عنوان میانگین و واریانس توزیع، و نیز یک وزن $\omega_{i,t}$ مشخص می‌شود. وزن هر توزیع، نمایش‌دهنده درصدی از مقادیر قبلی ویژگی است که به این توزیع اختصاص داشته است. با این حساب، احتمال اینکه ویژگی محاسبه شده در لحظه t ، برابر با X_t باشد به صورت زیر محاسبه می‌شود:

$$P(x_t) = \sum_{i=1}^k \omega_{i,t} \times \eta(x_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

که در آن:

$$\eta(x_t, \mu, \Sigma) = \frac{1}{\sqrt{2\pi\Sigma}} \exp\left(-\frac{1}{2}(x_t - \mu)^T \Sigma^{-1} (x_t - \mu)\right) \quad (2)$$

و کم شدن اشیاء صحنه رفتار مناسبی ارائه می‌دهد؛ که البته آزمایشات انجام شده نیز کم و بیش این ادعا را تصدیق می‌کنند. نظیر این روش بعدها در [۱۸]، [۲۶] و [۳۱] نیز به کار گرفته شده است و کار ما نیز در واقع در جهت بهبود این روش ارائه شده است.

بلوک‌بندی تصاویر ایده‌ای است که فتحی و همکاران در [۲۹] مطرح نموده‌اند. در این روش هر تصویر به بلوک‌های $N \times N$ یکسانی تحت عنوان سوپریپیکسل تقسیم می‌شود و بجای پیکسل، این سوپریپیکسل‌ها هستند که مبنای تصمیم‌گیری بوده و در دو کلاس زمینه و روزمینه دسته‌بندی می‌شوند (شکل ۵).

سیستم ارائه شده در [۲۹]، با بکارگیری بلوک‌بندی تصاویر در الگوریتم‌های تشخیص مبتنی بر تفاضل فریم‌ها و تفاضل زمینه، موفقیت‌های چشم‌گیری را نسبت به نمونه‌های مشابه مبتنی بر پیکسل کسب نموده است که از آن جمله می‌توان به افزایش ثبات تصمیم‌گیری، بهبود فوق‌العاده سرعت و کاهش چشمگیر حافظه مصرفی اشاره نمود. این مقاله حاصل بکارگیری ایده تلفیق دو روش بلوک‌بندی تصاویر [۲۹] و روش ترکیبی از گوسین‌ها [۱۱] است که در ادامه، شیوه کار و نتایج آن تشریح می‌شود.

۳-۲ ویژگی‌های انتخاب شده

که در آن، $M_{i,t}$ متغیری است که در صورت تعلق X_t به توزیع i ، یک و در غیر اینصورت صفر است. همچنین α نرخ یادگیری سیستم بوده و مقدار آن در محدوده مابین یک و صفر قرار دارد. مقدار $\sqrt{\alpha}$ به طور تقریبی نشان‌دهنده تعداد فریم‌هایی خواهد بود که در آن مدت، پارامترهای توزیعها تغییر می‌کنند؛ به عنوان مثال اگر یک خودرو وارد صحنه شده و در گوشه‌ای پارک کند، این خودرو به مدت تقریبی $\sqrt{\alpha}$ فریم به عنوان جسم متجاوز شناخته می‌شود ولی در صورت تداوم حضور، حساسیت سیستم از بین رفته و این خودرو تحت عنوان زمینه فراگرفته می‌شود. ρ نیز به صورت زیر محاسبه می‌شود:

$$\rho = \alpha \eta (x_t, \mu_{i,t-1}, \Sigma_{i,t-1}) \quad (6)$$

مطابق با روابط (۵) مشاهده می‌شود که پارامترهای μ و Σ ، برای توزیعهایی که شامل مقدار جدیدی نمی‌شوند، همانگونه دست نخورده باقی می‌مانند؛ ولی در مورد توزیعهایی که شامل مقدار جدید ویژگی هستند، تغییر می‌کنند.

پس از آنکه مقادیر پارامترهای همه توزیعها بهنگام شدند، مقادیر وزنها دوباره به صورت زیر نرمالیزه می‌شوند:

$$\omega_{i,t} = \frac{\omega_{i,t}}{\sum_{i=1}^k \omega_{i,t}} \quad (7)$$

۳-۴ تعیین توزیعهای مربوط به زمینه

براساس روش بهنگام‌سازی ذکر شده، توزیعهایی از مدل ترکیبی که درصد بیشتری از ویژگیهای محاسبه شده به آنها تعلق داشته باشند، به مرور صاحب وزن بیشتر و واریانس کمتری خواهند شد؛ بر همین مبنا برای تشخیص اینکه کدامیک از توزیعها متعلق به زمینه و کدامیک مربوط به روزمینه هستند مطابق زیر عمل می‌شود:

ابتدا همه گوسین‌ها براساس مقدار $\omega_{i,t} / \delta_{i,t}$ مرتب می‌شوند؛ این مقدار هم با افزایش وزن و هم با کاهش واریانس افزایش می‌یابد. با این روش، توزیعهایی که بیشترین مقدار را داشته باشند جزو زمینه و باقی توزیعها به عنوان روزمینه محسوب می‌شوند؛ منتها برای تعیین حد فاصل زمینه و روزمینه، پارامتر T تعریف می‌شود. T نمایانگر حداقل نسبتی از داده‌های یک بلوک است که باید به زمینه تعلق داشته باشد؛ به عبارت دیگر T ، احتمال تعلق هر بلوک از تصویر به زمینه را نشان می‌دهد و مطابق با نظر کاربر تعیین می‌شود. مثلاً در سری تصاویری که از سطح یک خیابان برداشته شده است متناسب با میزان رفت و آمد خودروها، احتمال مشاهده زمینه (کف خیابان) کاهش یا افزایش می‌یابد.

با تعیین مقدار T ، B توزیع بالای لیست به عنوان زمینه شناخته می‌شوند. مقدار B طوری انتخاب می‌شود که مجموع وزن‌های $B-1$ توزیع بالای لیست، کمتر از T ، و مجموع وزن‌های B توزیع بالای لیست، بزرگتر یا مساوی T باشد. به عبارت دیگر:

$$B = \arg \min_b \left(\sum_{i=1}^b \omega_i \geq T \right) \quad (8)$$

۳-۵ کلاس‌بندی بلوک‌ها

با انتخاب توزیعهای مربوط به زمینه، براحتی می‌توان در مورد زمینه یا روزمینه بودن هر بلوک تصمیم گرفت؛ اگر مقدار X_t به هیچیک از توزیعهای زمینه تعلق نداشته باشد، بلوک مربوطه از نظر این ویژگی، جزو

تعداد توزیعهای گوسین (k) و نیز تعداد ویژگیهای هر بلوک، متناسب با قدرت محاسباتی و نیز میزان حافظه، قابل تغییر است ولی ما مقدار k را برابر ۵ و تعداد ویژگیها را برابر با ۳ در نظر گرفته‌ایم. ویژگیهای ما عبارتند از میانه، میانگین و انحراف معیار؛ هر بلوک جداگانه محاسبه می‌شوند. مقادیر این ویژگیها، علی‌رغم مقاومت بالا در برابر نویز و اغتشاشات جزئی، نسبت به تغییرات عمده در بلوک، شدیداً حساس هستند؛ همچنین سازگاری توزیع مقادیر آنها با مدل ترکیبی از گوسین‌ها، به همراه سادگی در محاسبه، از دیگر خصوصیات این ویژگیها هستند که نهایتاً ما را به انتخاب آنها ترغیب نمود.

اگر اندازه بلوک‌ها را $N \times N$ در نظر بگیریم و $I(x, y, t)$ مقدار پیکسل (x, y) بلوک، در لحظه t فرض نماییم، مقدار این ویژگیها، برای لحظه t به صورت زیر محاسبه می‌شوند:

$$\begin{aligned} Mean_t &= \frac{1}{N \times N} \sum_{x=1}^N \sum_{y=1}^N I(x, y, t) \\ Dev_t &= \sqrt{\frac{1}{N-1} \sum_{x=1}^N \sum_{y=1}^N [(I(x, y, t) - Mean_t)^2]} \quad (9) \\ Med_t &= median\{I(x, y, t) | 1 \leq x, y \leq N\} \end{aligned}$$

در محاسبه ویژگیها، تنها بر اطلاعات سطوح خاکستری پیکسل‌ها تکیه کرده‌ایم، در نتیجه مقدار μ و Σ برای توزیعهای گوسین، صرفاً مقادیری عددی خواهند بود.

۳-۳ بهنگام‌سازی پارامترهای توزیعها

بنابراین به ازای هر بلوک از تصویر، $3k$ توزیع گوسین خواهیم داشت که هر یک از این توزیعها، با سه پارامتر μ ، Σ و ω مشخص می‌شوند. طبیعی است که با ورود هر فریم جدید، باید تک‌تک توزیعهای مدل ترکیبی بهنگام شده و پارامترهای هر یک از $3k$ توزیع یاد شده، براساس مقدار محاسبه شده برای ویژگیهای بلوک تنظیم شوند. برای این منظور، از روش برآورد سریعی که در [۱۱] استفاده شده است کمک گرفته‌ایم:

هر مقدار جدید ویژگی X_t ، روی k توزیع گوسین مربوط به خود چک می‌شود تا مشخص شود که این مقدار به کدامیک از توزیعها تعلق دارد؛ اگر مقدار مزبور، در رابطه زیر صدق کند این مقدار به توزیع i تعلق خواهد داشت:

$$|X_t - \mu_{i,t}| < \text{Max}(\sqrt{2/5} \times \delta_{i,t}, \sqrt{2} \times \delta_N) \quad (4)$$

که در آن $\delta_{i,t}^2 = \Sigma_{i,t}$ و δ_N به ترتیب نمایش‌دهنده انحراف معیار نویز گوسی سوار شده بر روی مقادیر ویژگیها، در شرایط کاملاً ثابت می‌باشند. اگر مقدار جدید در هیچیک از k توزیع قرار نگرفت، توزیعی که کمترین احتمال را داراست، با یک توزیع جدید، که مقدار میانگین آن است و دارای واریانس اولیه بالا و وزن اولیه کمی است جایگزین می‌شود. ولی در هر صورت مقادیر پارامترهای هر یک از k توزیع گوسین بصورت زیر بهنگام می‌شوند:

$$\begin{aligned} \omega_{i,t} &= (1-\alpha)\omega_{i,t-1} + \alpha M_{i,t} \\ \mu_{i,t} &= M_{i,t}((1-\rho)\mu_{i,t-1} + \rho X_t) + (1-M_{i,t})\mu_{i,t-1} \\ \Sigma_{i,t} &= M_{i,t}((1-\rho)\Sigma_{i,t-1} + \rho(X_t - \mu_{i,t-1})^2) + (1-M_{i,t})\Sigma_{i,t-1} \end{aligned} \quad (5)$$

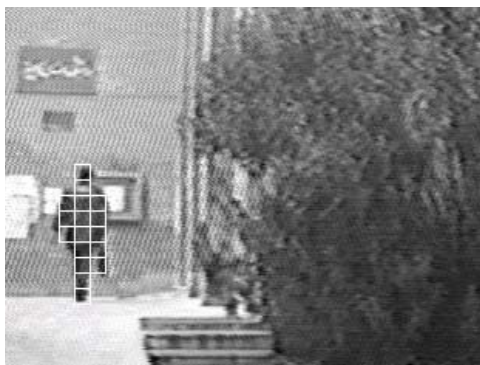


(ب)



(الف)

شکل ۶: (الف) تصویر پیش از مفقود شدن جسم از داخل صحنه، (ب) تصویر پس از تشخیص مفقود شدن جسم از داخل صحنه ($T = 0.95$, $\alpha = 0.1$)



شکل ۸: مقاومت در برابر حرکت برگ درختان ($T = 0.9$, $\alpha = 0.1$)



شکل ۷: صحت عملکرد سیستم در هوای بارانی ($T = 0.9$, $\alpha = 0.1$)

متجاوز محتمل‌الورود به صحنه، لااقل به اندازه دو بلوک از تصویر باشد.

۴- نتایج پیاده‌سازی

به منظور ارزیابی عملی الگوریتم، سیستمی مبتنی بر کامپیوتر شخصی در نظر گرفته شده است که در آن، تصاویر از طریق دوربین CCD و با واسطه برد فریم گرپر وارد کامپیوتر شده، آنگاه الگوریتم تشخیص به صورت نرم‌افزاری روی آنها اعمال می‌گردد و نتایج حاصله نیز روی صفحه مانیتور نمایش داده می‌شود. برای این منظور از کامپیوتری با پردازنده Pentium II (Celeron 350MHz) استفاده شده است.

آزمایش‌های متعددی که روی سیستم مزبور و با سری تصاویر تهیه شده در شرایط مختلف داخلی و خارجی انجام گرفته است حاکی از عملکرد قابل قبول سیستم در برابر اجسام متجاوز به صحنه (بدون توجه به سرعت حرکت آنها)، مفقود شدن و یا آسیب دیدن اجسام داخل صحنه، نویز تصویربرداری، تغییرات تدریجی و ناگهانی نور، اجسام متحرک متعلق به صحنه، زمینه‌های چندحالتی و تغییرات شرایط آب و هوایی است ضمن آنکه در این روش پارامترهای تنظیمی، منحصر به سه مورد N ، α و T شده و نیز سیستم جهت یادگیری زمینه نیازی به تخلیه صحنه از اجسام روزمینه ندارد. نمونه‌هایی از عملکرد سیستم یاد شده در شکل‌های ۶ تا ۱۰ نمایش داده شده است.

۵- نتیجه‌گیری

روش جدید در مقایسه با روش ترکیبی از گوسین‌های مبتنی بر پیکسل، علاوه بر افزایش پایداری سیستم نسبت به اغتشاشات و افزایش مدل‌پذیری داده‌ها، بهبود تشخیص و کاهش هشدارهای اشتباه را به دنبال داشته و افزایش سرعت و کاهش حافظه لازم در آن نمود ویژه‌ای دارد.

روزمینه محسوب می‌شود؛ حال در صورتیکه از میان سه ویژگی مورد بحث حداقل دو ویژگی، بلوکی را جزو روزمینه دسته‌بندی کنند، در کلاس‌بندی نهایی نیز این بلوک، روزمینه خواهد بود؛ ولی اگر حداکثر یک ویژگی، بر روزمینه بودن آن اذعان داشته باشد بلوک مزبور همچنان جزو زمینه محسوب می‌شود.

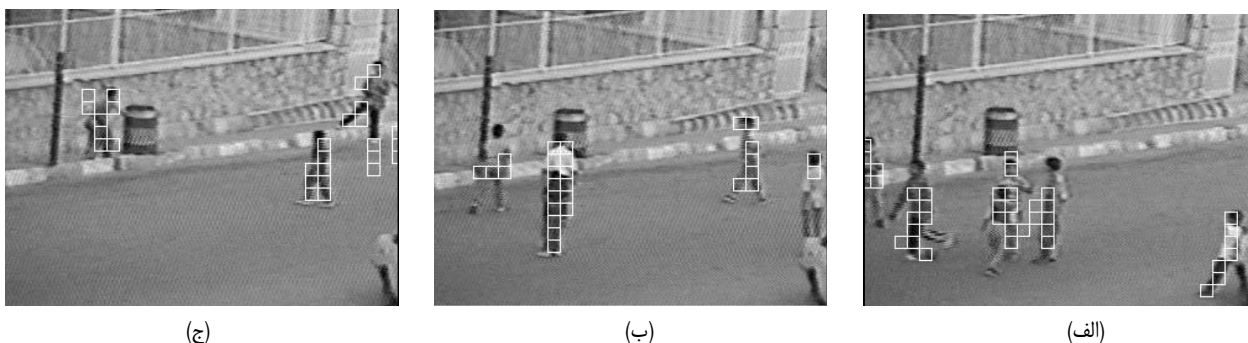
۳-۶ حذف بلوک‌های منفرد

پس از کلاس‌بندی اولیه بلوک‌ها، در مرحله دیگری، بلوک‌های منفردی که به عنوان روزمینه مشخص شده‌اند، از روزمینه حذف می‌شوند و در واقع فیلتر می‌گردند. عمل فیلتر کردن بلوک‌های روزمینه، علاوه بر اینکه اشتباهات ناشی از وجود نویز یا اغتشاشات احتمالی را برطرف می‌نماید، توجه را روی محل تجمع بلوک‌های روزمینه متمرکز می‌کند. روش استفاده شده برای این منظور، به سادگی، همسایگی ۸ تک‌تک بلوک‌های انتخاب شده را بررسی می‌کند و در صورتیکه هیچیک از بلوک‌های همسایگی ۸ آن، جزو روزمینه نباشند، بلوک مزبور نیز جزو زمینه محسوب می‌شود.

۳-۷ انتخاب اندازه بلوک‌ها

از عوامل دخیل در افزایش یا کاهش کارایی این روش، می‌توان به انتخاب اندازه بلوک‌های تصویر اشاره کرد. اندازه بلوک‌ها، باید به نحوی انتخاب شود که اولاً کوچک بودن آنها، منجر به افزایش اشتباهات و کاهش پایداری نشود و ثانیاً بزرگ بودن آنها نیز، افت دقت را به دنبال نداشته باشد.

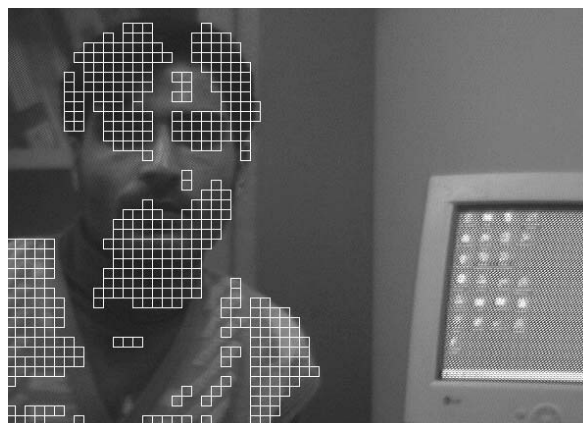
در این روش تنظیم اندازه بلوک‌ها به عهده کاربر گذاشته شده است و توصیه می‌شود که مقدار آن را به نحوی انتخاب کند که کوچکترین جسم



شکل ۹: (الف)، (ب) و (ج) مدل‌سازی صحیح زمینه علیرغم وجود اجسام روزمینه در صحنه ($T = 0.6$ ، $\alpha = 0.5$).

جدول ۱: میزان افزایش سرعت و کاهش حافظه مصرفی الگوریتم جدید نسبت به الگوریتم ترکیبی از گوسین‌های مبتنی بر پیکسل.

ابعاد بلوک	میزان کاهش حافظه مصرفی	
	میزان افزایش سرعت	میزان کاهش حافظه مصرفی
۲×۲	٪ ۱۰	٪ ۲۵
۳×۳	٪ ۱۱۰	٪ ۶۷
۴×۴	٪ ۲۰۰	٪ ۸۱
۵×۵	٪ ۲۷۰	٪ ۸۸
۶×۶	٪ ۳۵۰	٪ ۹۲
۷×۷	٪ ۴۰۰	٪ ۹۴
۸×۸	٪ ۴۴۰	٪ ۹۵
۹×۹	٪ ۴۷۰	٪ ۹۶
۱۰×۱۰	٪ ۵۰۰	٪ ۹۷
۲۰×۲۰	٪ ۵۹۰	٪ ۹۹/۲
۳۰×۳۰	٪ ۶۰۰	٪ ۹۹/۶
۴۰×۴۰	٪ ۶۱۰	٪ ۹۹/۸



شکل ۱۰: مقاومت در برابر جاروب شدن صفحه مانیتور ($T = 0.95$ ، $\alpha = 0.1$).

جدول ۱ نمایش‌دهنده میزان سرعت اجرا و کاهش حافظه لازم الگوریتم جدید نسبت به الگوریتم ترکیبی از گوسین‌های مبتنی بر پیکسل می‌باشد؛ بر اساس این جدول، اگر ابعاد بلوک‌ها 10×10 انتخاب شوند سرعت اجرا ۶ برابر شده و به مقدار ۹۷٪ در حافظه مصرفی صرفه‌جویی می‌شود. نتیجه اینکه این روش می‌تواند بستر مناسبی برای کارهای آتی قرار بگیرد و در حال حاضر، عنوان تکمیلی جدیدی نظیر انتخاب ویژگی‌های بهتر برای مدل‌سازی اطلاعات بلوک‌ها، تعمیم الگوریتم به تصاویر رنگی، وفقی نمودن ضریب محدوده تعلق توزیعها، پارامتر T و نیز ابعاد بلوک‌ها در دستور تحقیقات آتی قرار دارد که به ثمر رساندن هر یک از آنها می‌تواند به بهبود تشخیص کمک کند.

مراجع

- [7] A. T. Ali and E. L. Dagless, "Alternative practical methods for moving object detection," in *Proc. 4th Int. Conf. on Image Processing and Its Applications*, Publication no. 354, Maastricht, 1992.
- [8] A. T. Ali and E. L. Dagless, "Computer vision for security surveillance and movement control," *Electronic Images and Image Processing in Security and Forensic Science*, IEE Colloquium on, pp. 6/1-6/7, 1990.
- [9] P. Voles, M. Teal, and J. Sanderson, "Target identification in complex maritime scene," *IEE Colloquium on Motion Analysis and Tracking*, London, UK, pp. 15/1-15/4, May 10, 1999.
- [10] A. Utsumi and J. Ohya, "Image segmentation for human tracking using sequential-image-based hierarchical adaptation," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 911-916, 1998.
- [11] W. E. L. Grimson and C. Stauffer, "Adaptive background mixture models for real-time tracking," in *CVPR*, 1999.
- [12] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in *Proc. Seventh Int. Conf. on Computer Vision*, pp. 255-261, Sep. 1999.
- [13] N. Friedman and S. Russel, "Image segmentation in video sequences: a probabilistic approach," in *Proc. 13th Conf. on Uncertainty in Artificial Intelligence (UAI)*, pp. 175-181, Aug. 1997.
- [14] S. Arseneau and J. R. Cooperstock, "Real-time image segmentation for action recognition," in *Proc. IEEE PACRIM, Pacific Rim Conf. on Communications, Computers, Visualization and Signal Processing*, pp. 86-89, Victoria, Aug. 1999.
- [15] M. Tsuchikawa, A. Sato, H. Koike, and A. Tomono, "A moving object extraction method robust against illumination level changes for a pedestrian counting system," in *Proc. ISCV 95*, pp. 563-568, 1995.
- [16] G. L. Foresti, "Object recognition and tracking for remote video surveillance," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, no. 7, pp. 10/45-10/62, Oct. 1999.
- [1] E. Durucan, J. Snoeckx, and Y. Weilenmann, "Illumination invariant background extraction," in *Proc. 10th Int. Conference on Image Analysis and Processing (ICIAP '99)*, pp. 1136-1139, 1999.
- [2] A. Iketani, Y. Kuno, N. Shimada, and Y. Shirai, "Real-time surveillance system detecting persons in complex scenes," in *Proc. IAPR Int. Conf. on Image Analysis and Processing*, pp. 1112-1115, 1999.
- [۳] حسن شکوری، برآورد بلادرنگ حرکت، سمینار کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، ۱۳۷۸.
- [4] P. A. Laplant and D. Stoyenko, *Real-time Imaging: Theory, Techniques and Applications*, IEEE PRESS, New York, 1996.
- [5] L. J. Leroux and J. J. D. Van Schalkwyk, "An overview of moving object segmentation in video images," in *Proc. South African Symposium on Communication and Signal Processing*, pp. 53-57, Aug. 30, 1991.
- [6] R. Cucchiara, P. Onfiani, A. Prati, and N. Scarabottolo, "Segmentation of moving objects at frame rate: a dedicated hardware solution," in *Proc. of 7th IEE Conf. on Image Processing and Its Applications (IPA 99)*, Manchester UK, 12-15 Jul. 1999.

- [30] Y. Ivanov, A. Bobick, and J. Liu, "Fast lighting independent background subtraction," in *Proc. IEEE Workshop on Visual Surveillance (VS'98)*, pp. 49-55, Bombay, India, IEEE, New York 1998.
- [31] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," *6th European Conf. on Computer Vision*, Dublin, Ireland, Jun./Jul. 2000.
- [32] Y. Yang and M. D. Levis, "The background primal sketch: an approach for tracking moving objects," *Machine Vision and Application*, vol. 5, no. 1, pp.17-34, 1992.
- [33] D. Hepper and H. Li, "Analysis of uncovered background prediction for image sequence coding," in *Proc. Picture Coding Symposium*, pp. 192-193, 1987.
- [34] K. P. Karmann and A.V. Brandt, "Moving object recognition using an adaptive background memory," *Time-Varying Image Processing and Moving Object Recognition*, V. Cappellini (ed.), pp. 289-296, Elsevier, Amsterdam, The Netherlands, 1990.
- محمود فتحی** در سال ۱۳۶۳ دوره کارشناسی مهندسی برق گرایش الکترونیک را در دانشگاه علم و صنعت ایران به اتمام رساند. در سال ۱۳۶۶ مدرک کارشناسی ارشد خود را در گرایش مهندسی کامپیوتر از دانشگاه برادفورد انگلستان اخذ نمود و دوره دکتری خود را در سال ۱۳۷۰ از دانشگاه یومیسست انگلستان در زمینه مهندسی کامپیوتر و پردازش تصویر به پایان رساند.
- او از سال ۱۳۷۰ در دانشگاه علم و صنعت ایران دانشکده کامپیوتر مشغول به فعالیت می‌باشد و در حال حاضر به عنوان دانشیار در این دانشکده فعالیت دارد. زمینه‌های مورد علاقه نامبرده، پردازش تصویر، معماری کامپیوتر و شبکه‌های کامپیوتری می‌باشد.
- حسن شکوری** در سال ۱۳۷۷ مدرک کارشناسی خود را در رشته مهندسی کامپیوتر (گرایش سخت‌افزار) از دانشگاه صنعتی شریف دریافت نمود و در سال ۱۳۷۹ موفق به اخذ مدرک کارشناسی ارشد مهندسی کامپیوتر در گرایش هوش مصنوعی و رباتیک از دانشگاه علم و صنعت ایران گردید. نامبرده در حال حاضر عضو هیأت علمی جهاد دانشگاهی صنعتی شریف بوده و در زمینه اتوماسیون صنعتی مشغول به کار می‌باشد.
- [17] G. L. Foresti, "A real-time system for video surveillance of unattended outdoor environments," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, no. 6, pp. 697-704, Oct. 1998.
- [18] S. McKenna, S. Jabri, Z. Duric, and H. Wechsler, "Tracking interacting people," *IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp. 348-353, Grenoble, France, Mar. 2000.
- [19] C. Ridder, O. Munakelt, and H. Kirchner, "Adaptive background estimation and foreground detection using kalman-filtering," in *Proc. of Int. Conf. on Recent Advances in Mechatronics, ICRAM'95*, UNESCO Chair in Mechatronics, pp. 193-199, 1995.
- [20] W. Christopher R., A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, Jul. 1997.
- [21] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russel, "Toward robust automatic traffic scene analysis in real-time," in *Proc. Int. Conf. on Pattern Recognition*, pp. 126-131, Jerusalem, Oct. 1994.
- [22] W. E. L. Grimson, C. Stauffer, and R. Romano, "Using adaptive tracking to classify and monitor activities in a site," in *Proc IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 22-29, 1998.
- [23] A. Makarov, "Comparison of background extraction based intrusion detection algorithm," in *Proc. Int. Conf. Image Processing*, pp. 521-524, 1996.
- [24] A. K. Bekhalil, S. S. Ipson, and W. Booth, "Real-time detection and tracking of a moving object using a complex programmable logic device," *IEE colloquium on Target Tracking and Data Fusion*, Austin Court, Birmingham, pp.10/1-10/7, Jun. 1998.
- [25] A. R. Rostampour, O. R. Mitchell, and A. P. Reeves, "The use of temporal variance for moving object extraction," in *Proc. Seventh Annual Phoenix Conf. on Computers and Communications*, Mar. 1988.
- [26] K. W. Lee and J. Kim, "Moving object segmentation based on statistical motion model," *IEE Electronics Letters*, vol. 35, no. 20, pp. 1719-1720, Sep. 1999.
- [27] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Who, when, where, what: a real time system for detecting and tracking people," *Third Int. Conf. on Automatic Face and Gesture*, Nara, Japan, Apr. 1998.
- [28] Y. K. Jung and Y. S. Ho, "Traffic parameter extraction using video-based vehicle tracking," *ITSC'99 Conference Program* (tentative), Oct. 5, 1999.
- [۲۹] م. فتحی، م. عنایتی‌نیا، غ. ر. رمضان، ش. کبیری و ا. ر. خوانساری، "سیستم کنترل و نظارت خودکار تصویری"، آزمایشگاه پردازش سریع دانشکده کامپیوتر دانشگاه علم و صنعت ایران، ۱۳۷۸.