



فصل اول

مقدمه

یکی از مباحث مهم در بینایی ماشین، تجزیه و تحلیل دنباله تصویر ویدیویی به منظور درک صحنه می باشد. برای درک یک صحنه لازم است، تمامی اشیاء و رویدادها شناسایی و به روش مناسب بازنمایی شوند. با این وجود، ساختن یک سیستم عمومی برای درک تمامی اشیاء و رویدادهای هر نوع صحنه ویدیویی، کاری بسیار طاقت فرسا است. چون پردازش بر روی هر شیء و رویدادی نیازمند الگوریتم خاص خود می باشد، بدون داشتن اطلاعات کافی در مورد اشیاء موجود در یک صحنه ویدیویی و رویدادهای آن، امکان درک آن صحنه میسر نیست. بنابراین اغلب سیستم های موفق در زمینه بینایی ماشین، علاوه بر ویژگی های سطح پایین تصویر، بر روی دانش موجود در یک حوزه خاص تمرکز دارند.

از جمله حوزه هایی که تاکنون مورد تجزیه و تحلیل قرار گرفته اند، می توان به سیستم های نظارت و مراقبت، کنترل ترافیک، کنترل کیفی، سنجش از راه دور، اتوماسیون صنعتی، تحلیل حرکات ورزشی و بسیاری از حوزه های دیگر اشاره نمود. همان گونه که اشاره شد در تمامی حوزه های مورد بحث، دانش حوزه در درک صحنه از اهمیت خاصی برخوردار است. در این پایان نامه بر روی تجزیه و تحلیل صحنه های ویدیویی مسابقه فوتبال که در بین مسابقات ورزشی طرفداران زیادی را به خود جذب کرده، تمرکز شده است.

در مسابقه فوتبال، تمامی اشیاء مهم از جمله بازیکنان، داور و توپ بر روی زمین چمن قرار گرفته اند و شکل تقریبی آنها در تصویر از قبل مشخص است. حرکات این اشیاء محدود بوده و تا حد زیادی قابل پیش بینی است. به عنوان مثال، بازیکنان به صورت یک حباب متحرک بر روی سطح 2 بعدی زمین چمن و با سرعتی محدود حرکت می کنند. با داشتن مجموعه ای از این گونه اطلاعات، پردازش تصاویر ویدیویی ساده تر و با دقت بیشتری امکان پذیر خواهد بود.

مدت زمان یک مسابقه فوتبال حداقل 90 دقیقه می باشد. در این فاصله زمانی، تنها بازه های زمانی کوتاهی از کل مسابقه از اهمیت خاصی برخوردار می باشند و بیشتر زمان مسابقه بیان گر استراتژی کلی تیم ها بوده که نسبت به صحنه هایی چون صحنه های گل از حساسیت کمتری برخوردار است. از طرف دیگر، در مسابقه فوتبال بیش از هر چیز دیگر، نتیجه مسابقه اهمیت دارد؛ به عبارت دیگر صحنه های گل نقش سرنوشت ساز تیم ها را تعیین می کند. به همین علت معمولاً ارزیابی و تجزیه و تحلیل تیم ها بر روی صحنه های گل صورت می گیرد. از این رو، در این پایان نامه به جای تجزیه و تحلیل کل مسابقه فوتبال که هزینه محاسباتی آن بالا است، تنها تجزیه و تحلیل صحنه های گل مورد مطالعه قرار گرفته است.

به منظور تجزیه و تحلیل تصاویر ویدیویی صحنه های گل در مسابقات فوتبال، لازم است صحنه های گل به طور خودکار از تصاویر ویدیویی استخراج و سپس اشیاء و رفتار موجود در صحنه شناسایی گردند. از جمله اشیاء مهم درون زمین فوتبال، می توان به بازیکنان، داور، توپ و نشانه های خطی اشاره نمود. مسیر حرکت بازیکنان در تحلیل استراتژی تیم نقش مهمی را ایفا

می‌کند. همچنین به منظور مشخص نمودن نقاط ضعف و قوت تیم در یک صحنه گل، لازم است مسیر حرکت هر یک از بازیکنان به دقت مورد بررسی قرار گیرد. بدین منظور لازم است بازیکنان در تصویر شناسایی و موقعیت آن‌ها نسبت به نشانه‌های خطی که بیان‌گر مدل میدان هستند، به دست آید. البته به منظور تحلیل کامل صحنه‌های گل و تشخیص به ثمر رسیدن گل، علاوه بر تعیین مسیر بازیکنان، لازم است مسیر حرکت توپ نیز شناسایی گردد. همچنین برای تشخیص موارد داوری مانند خطاهای کارت‌دار نیز، شناسایی و تشخیص رفتار داور امری بدیهی است.

تماشاچی حاضر در ورزشگاه قادر است علاوه بر جزئیات صحنه گل (مانند بازیکنان و توپ و همچنین حرکت آن‌ها)، نمای وسیع‌تری از آن صحنه (حرکت اشیاء نسبت به یکدیگر و مدل میدان) را مشاهده کند، در صورتی که یک بیننده تلویزیونی در هر لحظه، فقط بر روی قسمت کوچکی از میدان بازی دید دارد. در این پایان‌نامه برای نمایش بهتر دنباله تصویر، همانند آنچه که یک تماشاچی حاضر در ورزشگاه می‌بیند، دنباله تصویر در کنار هم قرار داده شده و نمای دوری از بازیکنان و حرکت آن‌ها در صحنه مورد نظر نمایش داده می‌شود.

در این فصل ابتدا ضرورت انجام این تحقیق در بخش 1-1 مورد بحث قرار می‌گیرد. سپس به تعریف کامل و دقیق مساله مورد مطالعه در بخش 2-1 پرداخته می‌شود. مشکلات و موانع موجود در این مساله در بخش 3-1 بیان می‌گردد. در بخش 4-1 اشاره‌ای بر سیستم‌های مشابه موجود می‌شود. بخش 5-1 طرح کلی پایان‌نامه را شامل می‌شود.

1-1- ضرورت انجام پروژه

انجمن محققان ورزشی با توجه به گسترش پردازش تصاویر ویدیویی در شاخه‌های مختلف ورزشی، به تازگی تجزیه و تحلیل خودکار تصاویر ویدیویی مسابقات فوتبال را مورد توجه قرار داده است. این در حالی است که افزایش قابل ملاحظه حجم اطلاعات ویدیویی مسابقات فوتبال که روزمره در گوشه و کنار جهان برگزار می‌گردد، اهمیت موضوع را بیش از پیش آشکار می‌نماید. همچنین نیاز به دقت و سرعت بیشتر جهت تجزیه و تحلیل مسابقه فوتبال نیز وابسته به خودکار شدن آن است. در اکثر بازیهای ورزشی مانند مسابقه فوتبال، قسمتهای حساس و جذاب بازی را تنها زمان‌های کوتاهی از کل مسابقه که معرف خلاصه‌ای از بازی می‌باشد، تشکیل می‌دهد. این لحظات حساس در مسابقه فوتبال بیشتر در صحنه‌های درون محوطه جریمه که موقعیت گل را ایجاد می‌کنند، دیده می‌شود. علاوه بر این، تمرکز اصلی تماشاچیان و مربیان برای تجزیه و تحلیل بازی معطوف به صحنه‌های گل می‌باشد. با توجه به مطالب یاد شده، لزوم خودکار شدن بازیابی و تجزیه و تحلیل صحنه‌های گل به خوبی احساس می‌شود.

1-2- تعریف مساله

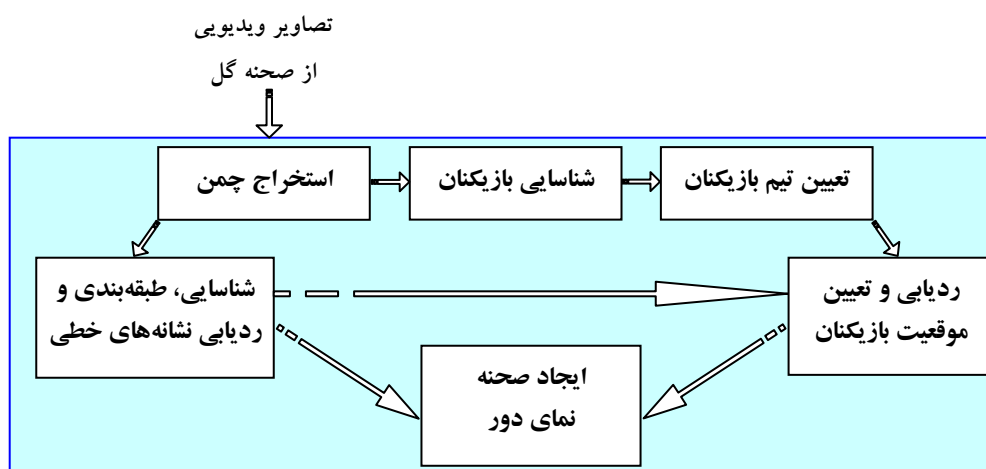
به منظور بازیابی و تجزیه و تحلیل صحنه‌های گل 3 موضوع اصلی زیر مورد بحث قرار می‌گیرند:

1- تشخیص صحنه‌های گل

2- تعیین مسیر پیموده شده توسط بازیکنان

3- نمایش نمای دور¹ صحنه گل به منظور درک بهتر صحنه

در این پایان‌نامه با فرض آنکه صحنه‌های گل از قبل استخراج شده‌اند، بر روی 2 موضوع اخیر تمرکز می‌شود. باید توجه داشت که در بسیاری از موارد، سیستم‌های تجاری موفق بر مبنای بینایی ماشین نیست؛ به عنوان مثال سیستم TRAKUS [1] برای ردیابی بازیکنان، گیرنده‌های مایکروویو² را مورد استفاده قرار می‌دهد، به این صورت که سیستم به تجزیه و تحلیل سیگنال منتشر شده از روی فرستنده‌های خاص که به هر کدام از بازیکنان متصل است، می‌پردازد. با این وجود، چون در بسیاری از ورزش‌ها امکان حمل چنین وسایلی برای بازیکنان میسر نیست، لذا روش فوق همواره قابل اجرا نمی‌باشد. واضح است مسیر بازیکنان در یافتن استراتژی تیم‌ها و تجزیه و تحلیل آن نقش مهمی را ایفا می‌کند. به منظور تعیین مسیر هر بازیکن، لازم است ابتدا بازیکن شناسایی و ردیابی و سپس مسیر حرکت آن در مدل میدان بازی تعیین گردد. چون انجام این مهم به صورت دستی مشکل و همراه با خطا است، در این پایان‌نامه روش‌های خودکار بدین منظور مورد مطالعه قرار گرفته است. علاوه بر این، صحنه نمای دور نمایش وسیع‌تری از میدان بازی را در اختیار بیننده قرار می‌دهد و به وی این امکان را می‌دهد تا در یک لحظه تصویر بزرگتری از موقعیت و حرکت بازیکنان را مشاهده نماید. در ساختن صحنه نمای دور، روش‌های کنار هم قرار دادن ایستا³ و پویا⁴ معرفی شده‌اند که هر کدام از معایب و محاسنی برخوردارند و در این پایان‌نامه هر 2 روش مورد بررسی قرار گرفته است. شکل (1-1) مراحل کلی در نظر گرفته شده در این پایان‌نامه را برای تجزیه و تحلیل خودکار تصاویر ویدیویی مسابقه فوتبال نشان می‌دهد.



شکل 1-1- مراحل در نظر گرفته شده در این پایان‌نامه برای تجزیه و تحلیل خودکار تصاویر ویدیویی مسابقه فوتبال

در این پایان‌نامه مفروضاتی در نظر گرفته شده است که عبارتند از:

- داده‌های مساله، تصاویر ویدیویی واقعی از صحنه‌های گل در مسابقات فوتبال است.
- مسابقات فوتبال از قوانین فیفا تبعیت می‌کنند؛ به عنوان مثال میدان بازی کاملاً پوشیده از چمن سبز رنگ است. همچنین اندازه میدان بازی بر طبق استاندارد فیفا می‌باشد.

¹ Panoramic View

² Microwave

³ Static Image Mosaicking

⁴ Dynamic Image Mosaicking

- رنگ لباس‌های بازیکنان کاملاً متمایز است.
- صحنه‌هایی مورد تجزیه و تحلیل قرار خواهند گرفت که از کنار زمین تصویربرداری شده باشند و نه از پشت دروازه.
- ردیابی داور مد نظر نیست.
- هم‌پوشانی حداکثر 3 بازیکن در نظر گرفته می‌شود.
- سیستم پیاده‌سازی شده به صورت غیر بلادرنگ عمل می‌کند.
- تجزیه و تحلیل صحنه‌های نمای دور مد نظر است، به عبارت دیگر صحنه‌هایی که شامل نشانه‌های خطی در دنباله تصویر باشد.

1-3- مشکلات و موانع

برای رسیدن به اهداف یاد شده، به خصوص در صحنه‌های گل، مشکلات و موانع متعددی وجود دارد که برخی از آن‌ها عبارتند از:

- پویایی (تغییرات پس‌زمینه و اشیاء)
 - وجود سایه حاصل از بازیکنان و محیط
 - دقت پایین بازیکنان (کوچک بودن آن‌ها) در صحنه‌های نمای دور
 - ازدحام (تعداد زیاد بازیکنان حاضر در صحنه)
 - فعل و انفعالات دوربین (بزرگ‌نمایی¹ و چرخش افقی و عمودی) به طور متناوب و سریع
 - صلب نبودن اشیاء موجود در صحنه و تغییر شکل آن‌ها به طور مداوم
 - وجود اعوجاجات لنز دوربین به خصوص در موارد بزرگ‌نمایی
 - تغییر سرعت بازیکنان و همچنین تغییر مسیر و جهت آنان به صورت غیر قابل پیش‌بینی
- همچنین بر خلاف تشخیص صحنه‌های گل که به دقت بالایی نیازمند نیستند، عملکرد صحیح ردیابی و تعیین مسیر بازیکنان و همچنین نمایش نمای دور به شدت متکی به دقت تصاویر گرفته شده است.

1-4- سیستم‌های مرتبط

گانگ² به همراه همکارانش در سال 1995 [2] سیستمی را برای طبقه‌بندی صحنه‌های مختلف مسابقات فوتبال ارائه داد. آن‌ها دنباله‌ای از فریم‌های مسابقه فوتبال را که نشان‌دهنده موقعیت خاصی از میدان بازی می‌باشد، طبقه‌بندی کردند؛ مانند موقعیت کرنر در جناح چپ و کناره پایین میدان، محوطه پناستی در جناح راست، موقعیت بازی در میانه میدان و غیره. آن‌ها بر اساس 4 مولفه اصلی که عبارتند از منطقه بازی، توپ، بازیکنان و جهت بردارهای حرکتی به طبقه‌بندی صحنه‌ها پرداخته و از این طریق موقعیت‌های مشابه را در یک گروه قرار دادند.

¹ Zoom

² Gong

یو^۱ و همکارانش در سال ۱۹۹۵ [۳] سیستمی خودکار را برای شناسایی و استخراج صحنه‌های جالب توجه فوتبال ارایه نمودند. این سیستم به تجزیه و تحلیل محتوای درون تصویر می‌پردازد و سپس صحنه‌های جالب توجه، مانند صحنه‌های گل را با ایجاد نمای دور از روی صحنه‌های منتخب و حرکت بازیکنان و ردیابی توپ در این نما تکرار می‌کند. تجزیه و تحلیل این سیستم شامل شناسایی ویژگی‌های پایدار بازی، ردیابی توپ، اصلاح حرکت دوربین و ایجاد نمای دور می‌باشد.

در سال ۱۹۹۵ اینتیل^۲ و بابیک^۳ [۴] سیستمی را برای ردیابی بازیکنان فوتبال آمریکایی ارایه دادند. این سیستم از تجزیه و تحلیل دنیای بسته^۴ برای به کار بردن دانش حوزه در ردیابی سطح پایین استفاده می‌کند. دنیای بسته به صورت یک ناحیه مکانی و یک فاصله زمانی تعریف می‌شود که محتوای آن کاملاً مشخص و قابل پیش‌بینی است. بعضی از روش‌های به کار رفته همچون تجزیه و تحلیل دنیای بسته و ردیابی در حوزه خاص، می‌تواند در بازی فوتبال نیز به کار گرفته شود. از جمله معایب سیستم یاد شده در نظر گرفتن بازیکنان هم‌پوشان شده به عنوان یک بازیکن است.

سو^۵ و همکارانش در سال ۱۹۹۷ [۵]، ردیابی مبتنی بر رنگ را برای شناسایی توپ و بازیکنان فوتبال مورد استفاده قرار دادند. آن‌ها برای به دست آوردن موقعیت بازیکن در مدل میدان در صورتی که دایره مرکزی مشاهده شود، نقاط ویژگی دایره مرکزی و در غیر این صورت، روش کنار هم قرار دادن دنباله تصویر را به کار گرفتند. در روشی که در ادامه کار آن‌ها در سال ۱۹۹۸ ارایه شد [۶]، موقعیت ۳ بعدی توپ را در دنباله تصویر به دست آوردند. این روش با به کارگیری رابطه‌های هندسی مثلثاتی نقاط شروع و انتهای توپ بر روی زمین، ارتفاع توپ را نسبت به یک بازیکن مرجع به دست می‌آورد. در مواردی نیز که نقطه انتهایی توپ بر روی زمین به دلیل برخورد با بازیکن یا هم‌پوشانی مشخص نمی‌باشد، قانون فیزیکی سهمی وار بودن حرکت توپ برای به دست آوردن ارتفاع توپ به کار گرفته شده است.

واندنبروک^۶ و همکارانش در سال ۱۹۹۷ [۷] روشی را برای ردیابی بازیکنان فوتبال بر اساس مدل پیرامون فعال و بر روی تصاویر گرفته شده از یک دوربین ثابت ارایه دادند. این مدل، پیکسل‌های طبقه‌بندی شده لباس بازیکنان را برای محاسبه انرژی مورد استفاده قرار می‌دهد. طبقه‌بندی پیکسل‌های لباس بازیکنان به ۲ طبقه برای هر تیم و به صورت یادگیری با نظارت^۷ بر روی ویژگی‌های مولفه‌های رنگی در چندین سیستم رنگی انجام شد. آن‌ها همچنین در ادامه کار خود در سال ۲۰۰۰ [۸] سیستمی را برای طبقه‌بندی با نظارت پیکسل‌های افراد داخل زمین به ۵ طبقه (یک طبقه برای داور، ۲ طبقه برای هر یک از تیم‌ها و ۲ طبقه برای ۲ دروازه‌بان) و با دقت بهتر معرفی نمودند که برای رسیدن به هدف خود، ویژگی بافت مولفه‌های رنگی پیکسل‌های همسایه در چندین سیستم رنگی را مورد استفاده قرار دادند.

میورا^۸ و همکارانش در سال‌های ۱۹۹۹ و ۲۰۰۰ [۹، ۱۰] سیستم‌هایی را برای ردیابی بازیکنان و توپ در تصاویر ویدیویی مسابقه فوتبال معرفی کردند. آن‌ها برای ردیابی بازیکنان، رنگ شورت و پیراهن آن‌ها را به کار بردند. سیستم‌های یاد شده، قادرند هم‌پوشانی حداکثر ۳ بازیکن را جواب‌گو باشد. آن‌ها همچنین برای ردیابی توپ، پارامترهای حرکتی را مورد استفاده قرار دادند. در سیستمی که توسط یامادا^۹ و همکاران وی در سال ۲۰۰۲ [۱۱] در ادامه این کار ارایه شد، علاوه بر

^۱ Yow

^۲ Intille

^۳ Bobick

^۴ Closed-World

^۵ Seo

^۶ Vandenbroucke

^۷ Supervised Learning

^۸ Miura

^۹ Yamada

ردیابی بازیکنان، پارامترهای دوربین و موقعیت 3 بعدی توپ با استفاده از برازش یک مدل فیزیکی 3 بعدی به مسیر توپ، تخمین زده شد.

در سال 2000 لفره¹ و همکارانش [12] روشی مبتنی بر مدل‌های پیرامون فعال² و انرژی بالونی را برای ردیابی بازیکنان بازیکنان ارایه نمودند. مزیت این روش، عدم نیاز به پیش پردازش و پیش‌بینی موقعیت بازیکنان می‌باشد. اما این روش در مواردی که بازیکنان به یکدیگر نزدیک می‌شوند، به درستی عمل نمی‌کند و بازیکنان نزدیک به هم را، حتی هنگامی که هم‌پوشانی ندارند و کاملاً از یکدیگر جدا هستند، به عنوان یک بازیکن تلقی می‌کند.

در سال 2001 پرس³ و همکارانش بازی اسکواش 2 نفره و هندبال را مورد مطالعه قرار دادند. آن‌ها برای ردیابی بازیکنان بازیکنان اسکواش، روش حباب اختلاف حرکت⁴ را مورد استفاده قرار دادند [13]. در بازی هندبال نیز با استفاده از 2 دوربین که در حفره‌های سقف سالن هندبال قرار می‌گرفت، به ردیابی بازیکنان پرداختند [14, 15]. آن‌ها برای انجام این کار روش‌های حباب اختلاف حرکت، ردیابی مبتنی بر رنگ و ردیابی کلیشه را به کار گرفتند. آن‌ها همچنین در همین سال دقت سیستم ارایه شده خود را مورد تجزیه و تحلیل قرار دادند [16].

برانکا⁵ در سال 2001 به همراه همکارانش [17] سیستمی را برای تخمین فاصله توپ از صفحه دروازه معرفی کرد. سیستم آن‌ها مبتنی بر بینایی 2 چشمی⁶ است و برای به دست آوردن فاصله توپ از دروازه، اختلاف منظر⁷ بین 2 دیدگاه نسبت نسبت به صفحه مسطح فیزیکی صفحه گل را مورد استفاده قرار می‌دهد. در این سیستم مفروضاتی مانند مشخص بودن موقعیت 3 بعدی نشانه‌های خطی محدوده دروازه و موقعیت توپ در تصویر، در نظر گرفته شده است.

سیستمی که توسط زو⁸ و همکاران وی در سال 2001 [18] ارایه شد به طبقه‌بندی صحنه‌های بازی به 2 کلاس در حال اجرا و متوقف می‌پردازد. آن‌ها بر اساس درصد پیکسل‌های سبز رنگ فضای تصویر، فریم‌ها را به 3 دسته نمای دور⁹، نمای میانه¹⁰، و نمای نزدیک¹¹ طبقه‌بندی و سپس طبق قواعد فیلم‌برداری، وضعیت بازی را (در حال اجرا یا متوقف) تعیین کردند. در سال 2001 کیم¹² و هونگ¹³ [19] روشی مبتنی بر درجه‌بندی خودکار دوربین و ردیابی نشانه‌های خطی را برای ایجاد نمای دور مورد استفاده قرار دادند. در این روش پارامترهای دوربین و ماتریس تبدیل فریم‌های دنباله تصویر به اولین فریم، با روش درجه‌بندی خودکار به دست می‌آید. برای اصلاح درجه‌بندی خودکار نیز نشانه‌های خطی موجود در زمین فوتبال به کار گرفته شده است. همچنین در این روش ردیابی نشانه‌های خطی در فریم بعدی با استفاده از پارامترهای به دست آمده دوربین در فریم فعلی انجام می‌شود.

یون¹⁴ و همکارانش در سال 2002 [20] برای ایجاد نمای دور، سیستمی را برای کنار هم قرار دادن دنباله تصویر مسابقه فوتبال معرفی کردند که در صورت فقدان نشانه‌های خطی در دنباله تصویر، تابلوهای تبلیغاتی را مورد استفاده قرار می‌دهد. علاوه بر این، در این سیستم بازیکنان از طریق قوانینی چون مساحت، کوچکترین مستطیل محیطی، فشردگی، نسبت طول

¹ Lefevre

² Active Contour

³ Perš

⁴ Motion Difference Blob

⁵ Branca

⁶ Stereo Vision

⁷ Parallax

⁸ Xu

⁹ Global Shot

¹⁰ Zoom-In

¹¹ Close-Up

¹² Kim

¹³ Hong

¹⁴ Yoon

عمودی به افقی و ویژگی‌های رنگی، استخراج و با استفاده از هیستوگرام مشترک، تیم هر بازیکن مشخص می‌شود. همچنین سیستم یاد شده به ردیابی بازیکنان با به کارگیری تطبیق کلیشه همبستگی می‌پردازد.

در سال 2002 اوتسومی¹ و همکارانش [21] روشی مبتنی بر خصوصیات کمیت رنگی و لبه محلی را برای شناسایی بازیکنان معرفی کردند. همچنین روشی برای ردیابی بازیکنان در طی 4 مرحله زیر ارائه دادند:

(الف) شناسایی هم‌پوشانی بین 2 بازیکن

(ب) ردیابی بازیکنان هم‌پوشان نشده

(ج) شناسایی هم‌پوشانی بین بازیکن و غیر بازیکن

(د) ردیابی بازیکنان هم‌پوشان شده

در سال 2002 دی ارازیو² و همکاران وی [22] با اصلاح تبدیل دایره جهتی هاف³، روشی را برای شناسایی توپ ارائه نمودند. آن‌ها برای آزمایش روش ارائه شده 2 شرایط نوری طبیعی و مصنوعی را به کار گرفتند. متأسفانه روش آن‌ها در شرایط هم‌پوشانی از دقت کافی برخوردار نیست.

اکین⁴ در سال 2002 به همراه تکالپ⁵ [23] پس از مرزبندی صحنه‌ها، با توجه به پیکسل‌های سبز رنگ در هر فریم و اندازه و تعداد اشیاء شناخته شده در یک صحنه، به طبقه‌بندی صحنه‌های مختلف مسابقه فوتبال به 3 گروه صحنه‌های نمای دور، صحنه‌های میانه داخل زمین، و صحنه‌های دیگر مانند صحنه‌های خارج از زمین یا نمای نزدیک پرداختند. در فعالیت دیگری که در ادامه کار آن‌ها و دیگر همکارانش در سال 2003 [24] ارائه شد، سیستمی برای تجزیه و تحلیل و استخراج صحنه‌های جالب توجه مسابقه فوتبال بر اساس ویژگی‌های مبتنی بر شیء و خصوصیات سینمایی ارائه گردید. در این سیستم ابتدا قسمت‌های حرکت آهسته و صحنه‌های گل در بازی بر اساس خصوصیات سینمایی استخراج و سپس قسمت‌های حرکت آهسته با توجه به ویژگی‌های مبتنی بر شیء تجزیه و تحلیل می‌گردد.

1-5- طرح کلی پایان‌نامه

در این پایان‌نامه، سیستمی برای تجزیه و تحلیل خودکار تصاویر ویدیویی صحنه‌های گل در مسابقات فوتبال معرفی شده است. این سیستم با شناسایی و تعیین تیم بازیکنان، آن‌ها را ردیابی و موقعیت هر یک را نسبت به مدل میدان به دست می‌آورد. ساختن صحنه نمای دور از روی دنباله تصویر ویدیویی مربوط به یک صحنه و نمایش حرکت بازیکنان در این نما از دیگر قابلیت‌های این سیستم به شمار می‌رود.

در فصل دوم روش‌های موجود برای انجام هر یک از مراحل موجود در شکل (1-1) به تفصیل تشریح شده و محدودیت‌ها و معایب هر یک ذکر می‌گردد. در فصل سوم روش پیشنهادی برای شناسایی بازیکنان که شامل مراحل استخراج زمین چمن، شناسایی و تعیین تیم بازیکنان می‌باشد، آورده شده است. در فصل چهارم مشکلات موجود در ردیابی بازیکنان در تصاویر ویدیویی صحنه‌های گل مورد بحث قرار گرفته است. سپس چند روش کلی برای ردیابی اشیاء متحرک به همراه مزایا

¹ Utsumi

² D'Orazio

³ Directional Circle Hough Transform

⁴ Ekin

⁵ Tekulp

و معایب هر یک از آنها آمده است. در پایان این فصل، روش پیشنهادی برای ردیابی بازیکنان، یعنی ردیابی بازیکنان مبتنی بر مدل میدان که دانش موجود بر میدان بازی را مورد استفاده قرار می‌دهد، تشریح شده است.

فصل پنجم ساخت صحنه نمای دور از صحنه گل را مورد بحث قرار می‌دهد. در این فصل ابتدا شناسایی و ردیابی نشانه‌های خطی با ترکیب الگوریتم “تبدیل هاف” و “فیلتر کالمن توسعه یافته” توضیح داده شده و سپس کنار هم قرار دادن دنباله تصویر به صورت ایستا و پویا برای نمایش نمای دور آمده است. فصل ششم نیز نتایج گردآوری شده از آزمایش‌های انجام شده بر روی تصاویر واقعی را نشان می‌دهد. در پایان، فصل هفتم به جمع‌بندی مطالب ارایه شده و ارایه پیشنهادات جهت ادامه کار می‌پردازد.

فصل دوم

پیشینه مطالعاتی

در این فصل به بررسی روش‌های موجود در هر یک از مراحل زیر پرداخته می‌شود:

- استخراج چمن میدان بازی
- شناسایی بازیکنان
- تعیین تیم بازیکنان
- ردیابی بازیکنان
- شناسایی و طبقه‌بندی نشانه‌های خطی
- ایجاد نمای دور

2-1- استخراج چمن میدان بازی

استخراج چمن از مهمترین مراحل برای آنالیز و درک صحنه می‌باشد، زیرا تمامی اشیاء مهم در صحنه مانند بازیکنان، داور، توپ، دروازه، نشانه‌های خطی و بردارهای حرکت، همگی بر روی میدان بازی واقع هستند. همچنین روش‌هایی که خصوصیات سطح بالا مانند خصوصیات سینمایی را برای آنالیز صحنه (مثل مرزبندی بین صحنه‌ها) مورد استفاده قرار می‌دهند [18, 23, 24]، متکی بر مقدار چمن موجود در صحنه هستند.

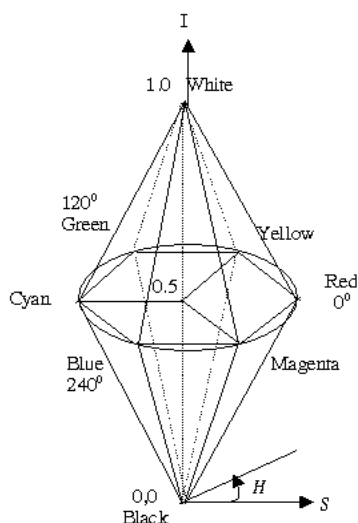
زو و همکارانش [18] به منظور در نظر داشتن اختلاف احتمالی در میزان روشنایی و شرایط چمن در قسمت‌های مختلف میدان بازی، رنگ چمن هر مسابقه را از روی مجموعه‌ای تصادفی از فریم‌های فیلم ویدیویی (50 فریم از یک نماآهنگ 5 دقیقه‌ای یعنی 600 فریم) تعیین کردند. آن‌ها نقطه اوج هیستوگرام فام¹ مجموع فریم‌های انتخاب شده را به عنوان مقدار فام چمن در نظر گرفتند. این آزمایش 8 بار بر روی بخش‌های متفاوت تکرار و میانگین و واریانس مقدار فام آن‌ها محاسبه شد. طبق نتایج به دست آمده، واریانس مقدار فام چمن در طول زمان از واریانس آن در بخش ابتدایی مسابقه کمتر بود. بنابراین با این فرض که تغییرات مقدار فام چمن به خاطر میزان روشنایی و شرایط زمین بازی نسبت به زمان تدریجی است، به صورت دوره‌ای با مدت زمان طولانی، مقدار فام چمن را به هنگام می‌نمودند.

¹ Hue

در بسیاری از سیستم‌های موجود، تنها مقدار فام در سیستم رنگی HSI^1 برای تعیین چمن میدان بازی مورد استفاده قرار گرفته است. زو و همکارانش [18] ادعا کرده‌اند مقدار فام چمن در شرایط متفاوت در فاصله 54 الی 90 درجه تغییر می‌کند. اکین و همکارانش [23] این بازه را در فاصله 65 الی 85 درجه در نظر گرفته‌اند. اوتسومی و همکارانش [21] بازه 60 الی 150 درجه را برای مقدار فام چمن انتخاب کرده‌اند. به نظر می‌رسد در هر یک از کارهای انجام شده، برای تصاویر ویدیویی مختلف مقدار فام چمن در بازه‌های مختلفی در نظر گرفته شده است. در هر صورت، نتایج بیان گر کم بودن نرخ تغییرات مقدار فام چمن به دلیل شرایط زمانی و مکانی مختلف است.

اکین و همکارانش [24] برای استخراج چمن زمین بازی به جای استفاده از مقدار فام به تنهایی، مقدار فام، درجه اشباع² و شدت روشنایی³ را همراه با هم مورد استفاده قرار دادند؛ به این صورت که در ابتدا داده‌های آماری رنگ غالب میدان بازی در فضای HSI یاد داده شده و سپس با گذشت زمان به صورت خود کار تطبیق داده می‌شود.

فضای رنگی HSI با 3 ویژگی مقدار فام که بیان گر نوع رنگ است، درجه اشباع که به میزان رنگ اشاره می‌کند و شدت روشنایی که میزان تاریکی و روشنی رنگ را نشان می‌دهد، مشخص می‌شود. فضای HSI را همان گونه که در شکل (1-2) مشاهده می‌شود، می‌توان به صورت مختصات استوانه‌ای مدل نمود. در این مدل مقدار فام به صورت زاویه که از 0° الی 360° تغییر می‌کند، نشان داده شده است. درجه اشباع نیز با شعاع مشخص می‌شود که از صفر الی یک متغیر است. همچنین شدت روشنایی در بازه صفر (برابر با رنگ سیاه) و یک (برابر با رنگ سفید) قرار می‌گیرد.



شکل 1-2- مدل استوانه‌ای فضای رنگی HSI

در سیستم اکین، رنگ غالب میدان بازی با محاسبه مقدار میانگین هر مولفه رنگی (درجه اشباع، مقدار فام و شدت روشنایی) به دست می‌آید. این مقادیر در نزدیکی نقطه اوج هیستوگرام متناظر می‌باشد؛ بنابراین ابتدا اندیس نقطه اوج یعنی i_{peak} برای هر هیستوگرام تعیین می‌گردد (برای به دست آوردن نتایج بهتر پیشنهاد شده است از چندین فریم برای این منظور

¹ Hue-Saturation-Intensity

² Saturation

³ Intensity

استفاده گردد) و سپس بازه $[i_{\min} \ i_{\max}]$ در حول نقطه اوج هر هیستوگرام به گونه‌ای تعیین می‌شود که شرایط زیر را برقرار کند:

$$H[i_{\min}] \geq K \times H[i_{peak}] \quad (1-2)$$

$$H[i_{\min} - 1] < K \times H[i_{peak}] \quad (2-2)$$

$$H[i_{\max}] \geq K \times H[i_{peak}] \quad (3-2)$$

$$H[i_{\max} + 1] < K \times H[i_{peak}] \quad (4-2)$$

$$i_{\min} \leq i_{peak} \quad (5-2)$$

$$i_{\max} \geq i_{peak} \quad (6-2)$$

که $i_{\min}$ و $i_{\max}$ اندیس کمینه و بیشینه بازه برای هیستوگرام رنگی H است. شرایط فوق اندیس‌های $i_{\min}$ و $i_{\max}$ را به گونه‌ای تعیین می‌کند که تعداد پیکسل‌ها در این اندیس‌ها حداقل به نسبت K از مقدار پیکسل موجود در نقطه اوج هیستوگرام باشد. در سیستمی که توسط اکین معرفی شده، این نسبت برابر 20٪ در نظر گرفته شده است. رنگ میانگین در این بازه برای هر مولفه رنگی طبق رابطه زیر محاسبه می‌گردد:

$$\text{Color mean} = \frac{\sum_{i=i_{\min}}^{i_{\max}} H[i] \times i}{\sum_{i=i_{\min}}^{i_{\max}} H[i]} \times Q_{size} \quad (7-2)$$

که Q_{size} اندازه کوانتیزاسیون برای هر مولفه رنگی می‌باشد (فرض شده است مقادیر فام، درجه اشباع و شدت روشنایی مقیاس‌های متفاوتی داشته باشند).

پیکسل‌های چمن در هر فریم با پیدا کردن فاصله استوانه‌ای¹ هر پیکسل تا رنگ میانگین به دست آمده، بدین صورت محاسبه می‌گردد:

$$d_{\text{cylindrical}} = \sqrt{(d_{\text{intensity}}(j))^2 + (d_{\text{chroma}}(j))^2} \quad (8-2)$$

که در آن فاصله شدت روشنایی و فاصله رنگی به ترتیب به صورت زیر تعریف می‌گردند:

$$d_{\text{intensity}}(j) = |I_j - \bar{I}| \quad (9-2)$$

$$d_{\text{chroma}}(j) = \sqrt{(S_j)^2 + (\bar{S})^2 - 2 \cdot S_j \cdot \bar{S} \cdot \cos(\theta(j))} \quad (10-2)$$

که زاویه $\theta(j)$ برابر است با:

$$\theta(j) = \begin{cases} \Delta(j) & \text{if } \Delta(j) \leq 180 \\ 360^\circ - \Delta(j) & \text{otherwise} \end{cases} \quad (11-2)$$

$$\Delta(j) = |\overline{Hue} - Hue_j| \quad (12-2)$$

در رابطه‌های فوق، Hue_j ، S_j و I_j به ترتیب مقدار فام، درجه اشباع و شدت روشنایی برای پیکسل j ام و $\overline{Hue}$ ، $\bar{S}$ و $\bar{I}$ مقدار رنگ غالب آن‌ها را که از رابطه (7-2) به دست می‌آید، نشان می‌دهد. لازم به ذکر است برای پیکسل‌های غیر

¹ Cylindrical Metric

رنگی¹ تنها از فاصله شدت روشنایی یعنی رابطه (2-9) استفاده شده است و در غیر این صورت رابطه‌های (2-9) و (2-10) همراه با هم به کار گرفته شده‌اند. در نهایت، پیکسل‌هایی که شرط زیر را دارا باشند به عنوان چمن میدان بازی انتخاب شدند:

$$d_{\text{cylindrical}} < T_{\text{color}} \quad (13-2)$$

که T_{color} مقدار آستانه‌ای از پیش تعیین شده است و امکان تنظیم بهتر آن با گذر زمان وجود دارد. روش‌های دیگری نیز معرفی شده‌اند که از فضای RGB² برای یافتن پیکسل‌های چمن استفاده کرده‌اند [5, 20]. یون و همکارانش [20] با فرض اینکه حداقل 65٪ پیکسل‌های فریم مورد نظر را چمن سبز رنگ پوشانده است، قواعد زیر را در فضای RGB برای استخراج میدان بازی به کار بردند:

$$O(x, y) = \begin{cases} 0, & \text{if } \begin{cases} I_G(x, y) > I_R(x, y), & (I) \\ I_G(x, y) > I_B(x, y), & (II) \\ |I_R(x, y) - R_{\text{peak}}| < R_t, & (III) \\ |I_G(x, y) - G_{\text{peak}}| < G_t, & (IV) \\ |I_B(x, y) - B_{\text{peak}}| < B_t, & (V) \\ GL(x, y) < GL_t & (VI) \end{cases} \\ 1, & \text{otherwise} \end{cases} \quad (14-2)$$

که $O(x, y)$ تصویر خروجی باینری شده، I_R ، I_G و I_B به ترتیب مقادیر رنگ قرمز، سبز و آبی پیکسل (x, y) ، $GL(x, y)$ مقدار سطح خاکستری آن پیکسل و R_{peak} ، G_{peak} و B_{peak} مقادیر نقطه اوج هیستوگرام رنگی را نشان می‌دهند. همچنین R_t ، G_t و B_t به ترتیب مقادیر آستانه برای رنگ‌های قرمز، سبز و آبی و GL_t نیز به مقدار آستانه برای سطح خاکستری، $GL(x, y)$ اشاره می‌کند. در رابطه (2-14) شرایط (I) و (II) وجود رنگ سبز را تصدیق می‌کند. علاوه بر این، شرایط (III) الی (V) وجود یک رنگ غالب را بررسی می‌کند و شرط (VI) نیز تمایز بین نشانه‌های خطی و میدان بازی را آشکار می‌نماید، چرا که خطوط سفید سطح خاکستری بالاتری نسبت به زمین چمن دارند.

یون و همکارانش برای مقادیر آستانه مولفه‌های رنگی، مقادیر ثابتی را اختیار کردند ($R_t = B_t = 10$ ، $G_t = 15$) و سپس برای در نظر گرفتن شرایط نوری مختلف، مقادیر آستانه را به صورت تطبیقی و با توجه به انحراف معیار هیستوگرام رنگی از مقادیر R_{peak} ، G_{peak} و B_{peak} تنظیم نمودند؛ به این ترتیب که با پیشی گرفتن مقادیر انحراف معیار از مقادیر از پیش تعیین شده، مقادیر آستانه را افزایش می‌دادند. همچنین برای مقدار آستانه GL_t مقدار ثابت 150 را اختصاص دادند.

2-2- شناسایی بازیکنان

بازیکنان در تصاویر ویدیویی و به خصوص در تصاویر بزرگ‌نمایی شده به صورت حباب مشاهده می‌شوند. معمولاً بازیکنان بر روی زمین چمن میدان بازی قرار دارند و وجه تمایزشان با زمین چمن، رنگ لباس‌های آن‌ها می‌باشد؛ به همین علت برای شناسایی بازیکنان در فریم اول، اطلاعات رنگ لباس بازیکنان به کار می‌رود و سپس در فریم‌های متوالی ردیابی می‌گردند.

¹ Achromatic

² Red-Green-Blue

یون و همکارانش [20] با فرض آنکه نشانه‌های خطی موجود بر روی زمین بازی، سطح خاکستری بالاتری نسبت به لباس بازیکنان دارد، با اعمال مجموعه قوانین زیر بر روی تصویر حاصل از جداسازی چمن، به شناسایی بازیکنان پرداختند:

• **قانون 1:** مساحت هر بازیکن (تعداد پیکسل‌هایی که اشغال می‌کند) باید حداقل برابر 50 و حداکثر برابر 1500 باشد.

• **قانون 2:** کوچکترین مستطیل محیطی¹ هر بازیکن باید در محدوده 130 الی 5000 واقع شود.

• **قانون 3:** نسبت فشردگی² هر بازیکن که به صورت زیر تعریف می‌شود، باید حداقل برابر 0/3 باشد.

$$\text{فشردگی} = \frac{(\text{محیط})^2}{4 * \pi * (\text{مساحت})}$$

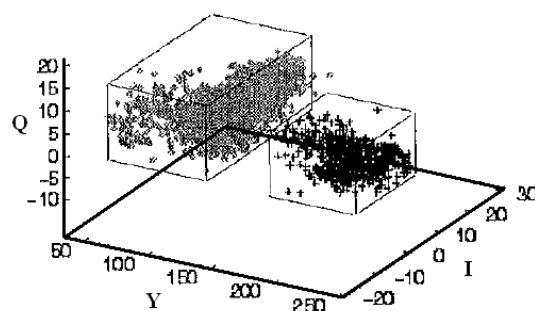
• **قانون 4:** مقدار مولفه رنگی سبز هر پیکسل، باید کمتر از مقدار مولفه‌های رنگی قرمز و آبی برای آن پیکسل باشد.

• **قانون 5:** سطح خاکستری هر پیکسل باید کمتر از 150 باشد.

• **قانون 6:** نسبت طول عمودی به افقی باید حداکثر برابر 3 باشد.

مشکل این روش در مواجهه شدن با تصاویری است که بازیکنان لباس‌هایی با سطح خاکستری بالا (سفید رنگ) پوشیده‌اند. علاوه بر این، در صورتی که بازیکنان لباس‌های سبز رنگ پوشیده باشند، قوانین یاد شده باعث حذف بازیکن و در نتیجه الگوریتم یاد شده با شکست مواجه می‌گردد.

میورا و همکارانش [9-11] برای شناسایی بازیکنان، اطلاعات رنگ و روابط موقعیتی بین شورت و پیراهن بازیکنان را مورد استفاده قرار دادند؛ بدین صورت که در ابتدا رنگ لباس 2 تیم از نمونه‌های تصویری، استخراج و سپس در صفحه YIQ تصویر می‌شود. نمونه‌ای از توزیع رنگ لباس بازیکنان در یک فیلم ویدیویی، توسط میورا و همکارانش در شکل (2-2) نمایش داده شده است. سرانجام، به وسیله مستطیل متساوی‌السطوحی که توزیع هر رنگ لباس را می‌پوشاند، یک مدل رنگی از هر لباس به دست آمده و از این طریق نواحی منتخب لباس بازیکنان استخراج می‌گردند.



شکل 2-2- نمونه‌ای از توزیع رنگ لباس بازیکنان توسط میورا و همکارانش [9]

برای حذف نواحی منتخب که به اشتباه به دست آمده‌اند، روابط مکانی بین آن‌ها مورد استفاده قرار گرفته است. در این سیستم با فرض آنکه بازیکن در حالت ایستاده باشد، پیراهن وی در بالای شورتش واقع است و لذا سیستم با این فرض به دنبال شورت و پیراهنی می‌گردد که پیراهن بالای شورت بوده و فاصله بین مراکز ثقل این 2 از یک حد آستانه که بر حسب موقعیت

¹ Minimum Bounded Rectangle

² Compactness

عمودی آن در تصویر تغییر می کند، کمتر باشد. کاربرد آستانه برای در نظر گرفتن تناظر بین اندازه بازیکن و موقعیت وی در تصویر (در محور عمودی) می باشد؛ زیرا بازیکنی که به دورین نزدیکتر باشد، فضای بیشتری از تصویر و همین بازیکن در صورتی که از دورین فاصله بگیرد، قسمت کوچتری از تصویر را اشغال می کند و به همین نسبت فاصله مرکز ثقل شورت و لباس بازیکن تغییر خواهد کرد.

روش میورا در صحنه های گل کارایی چندان مناسبی ندارد، زیرا همواره و به خصوص در صحنه های گل، فرض واقع شدن لباس بازیکن در بالای شورتش صحیح نیست. علاوه بر این، روش فوق در موارد هم پوشانی به علت پنهان شدن قسمتی از لباس یا شورت بازیکن دچار مشکل خواهد شد.

اوتسومی و همکارانش [21] برای شناسایی بازیکنان، ویژگی های کمیت رنگی¹ و لبه محلی² را مورد استفاده قرار دادند. آن ها با فرض آنکه رنگ اشیاء داخل زمین در مقایسه با رنگ چمن، متفاوت و فضای کوچکی از تصویر را اشغال می کند، ویژگی کمیت رنگی را در فضای هیستوگرام مورد استفاده قرار دادند. آن ها همچنین برای در نظر گرفتن ویژگی لبه محلی اشیاء داخل زمین، 2 ویژگی یاد شده را با توابع فازی ترکیب کردند.

تابع ارزیابی اوتسومی برای هر پیکسل $p(i, j)$ به صورت زیر تعریف می گردد:

$$O(i, j) = \max(R(i, j), \min(E_R(i, j), E_G(i, j), E_B(i, j))) \quad (15-2)$$

که R ویژگی کمیت رنگی و E_R ، E_G و E_B به ترتیب ویژگی لبه محلی برای مولفه های رنگی R ، G و B را نشان می دهد. مقادیر R ، E و O در بازه $[0, 1]$ قرار می گیرند و احتمال متعلق بودن پیکسل به بازیکن داخل زمین را نشان می دهند. عمل مینیمم گیری در رابطه فوق بین E_R ، E_G و E_B با فرض آن است که ویژگی لبه در تمامی مولفه های رنگی R ، G و B زیاد بوده، ولی این امکان وجود دارد که به علت وجود نویز، مقدار به دست آمده از ویژگی لبه محلی در برخی از این مولفه های رنگی واقعی نباشد. به منظور در نظر گرفتن هر 2 ویژگی کمیت رنگی و لبه محلی، ماکزیمم R و $\min(E_R, E_G, E_B)$ مورد استفاده قرار می گیرد.

ارزیابی کمیت رنگی R برای هر پیکسل $p(i, j)$ به صورت زیر تعریف می شود:

$$R(i, j) = \begin{cases} 0 & H(c(i, j)) > 2\bar{H} \\ 1 - \frac{H(c(i, j))}{2\bar{H}} & otherwise \end{cases} \quad (16-2)$$

که $c(i, j)$ رنگ پیکسل $p(i, j)$ ، H هیستوگرام رنگی کل تصویر و $\bar{H}$ میانگین توزیع هیستوگرام H را نشان می دهد. ارزیابی ویژگی لبه محلی E برای هر پیکسل $p(i, j)$ به صورت زیر تعریف می گردد:

$$E_K(i, j) = \frac{1}{8} \sum_{n=1}^8 e_K^n(i, j) \quad (17-2)$$

که در آن $K = \{R, G, B\}$ و e_K^n بدین صورت تعریف می گردد:

$$d_K^n(i, j) = |c_K^n(i, j) - c_K(i, j)| \quad where \quad n = 1, 2, \dots, 8 \quad (18-2)$$

$$e_K^n(i, j) = \begin{cases} 1 & d_K^n(i, j) > 2\bar{d}_K \\ \frac{d_K^n}{2\bar{d}_K} & otherwise \end{cases} \quad (19-2)$$

¹ Color Rarity

² Local Edge

که c_K^n و $\overline{d_K}$ به ترتیب نشان‌دهنده مقدار رنگ RGB در پیکسل $p(i, j)$ مقدار رنگ RGB در پیکسل همسایه m از هر یک از 8 همسایه و میانگین کلی d_K است.

پس از به دست آوردن تابع ارزیابی O برای هر پیکسل، فیلتر گاوسی-لاپلاسی¹ برای کاهش نویز به کار گرفته شده است. برای به دست آوردن نقاط منتخب اشیاء داخل زمین نیز آستانه‌سازی مورد استفاده قرار گرفته است. بالاخره، شرط زیر برای به دست آوردن نواحی بازیکنان از بین نواحی منتخب اعمال گردیده است:

$$15 \leq h \leq 50, 6 \leq w \leq 45, w \leq h \quad (20-2)$$

که h و w به ترتیب ارتفاع و عرض ناحیه اشیاء را مشخص می‌کند.

برای حذف نواحی بازیکنان نیز که به اشتباه به دست می‌آید، میانگین مساحت بازیکنان در نیمه بالایی و پایینی تصویر محاسبه و در صورتی که مساحت بازیکن از $1/5$ برابر میانگین مساحت نیمه مربوطه بیشتر باشد، حذف می‌گردد. ایده این معیار از یک طرف، در نظر گرفتن مساحت بازیکنان تقریباً به صورت هم اندازه و از طرف دیگر، در نظر گرفتن مساحت بزرگتر در حالت نزدیک بودن به دوربین (نیمه پایینی تصویر) می‌باشد.

سو و همکارانش [5] پس از استخراج زمین چمن، ماسک بازیکنان را از رابطه زیر به دست آوردند:

$$P(x, y) = \begin{cases} 1: & \text{if } (x, y) \in \text{field and } \text{Grass}(x, y) = 0 \\ 0: & \text{otherwise} \end{cases} \quad (21-2)$$

که field به محوطه زمین چمن استخراج شده اشاره می‌کند و $\text{Grass}(x, y)$ تابعی است که مشخص کننده متعلق بودن پیکسل (x, y) به چمن می‌باشد. سرانجام، با اعمال الگوریتم "برچسب گذاری اجزای متصل"² بر روی ماسک بازیکنان، نواحی بازیکنان را مشخص کردند.

2-3- تعیین تیم بازیکنان

از آنجایی که رنگ لباس بازیکنان هر تیم به گونه‌ای انتخاب می‌شود که به راحتی قابل تمایز باشد، ویژگی رنگ و بافت لباس بازیکنان در تعیین تیم نقش مهمی را ایفا می‌کند. واندنبروک و همکارانش [7, 8] با یادگیری با نظارت بافت پیکسل‌های افراد داخل زمین، به طبقه‌بندی آن‌ها پرداختند. سو و همکارانش [5] با استفاده از توزیع عمودی³ مولفه‌های رنگی لباس بازیکنان به جداسازی تیم‌ها پرداختند. میورا و همکارانش [9-11] همان‌گونه که در فصل قبل اشاره شد، از نمونه‌های استخراج شده و روابط موقعیتی بین شورت و پیراهن بازیکن برای تعیین تیم استفاده کردند. همچنین روش‌هایی مبتنی بر توزیع هیستوگرام مولفه‌های رنگی لباس بازیکنان معرفی گردیده است [2, 20, 25]. در ادامه به بررسی بیشتر روش‌های ذکر شده پرداخته می‌شود.

واندنبروک و همکارانش [8] برای طبقه‌بندی پیکسل‌های افراد داخل زمین به 5 طبقه داور، هر یک از بازیکنان 2 تیم و هر یک از 2 دروازه‌بان، ابتدا برای هر طبقه حالت‌های مختلف (دویدن، راه رفتن، پاس دادن، دریبل و غیره) و موقعیت‌های مختلف (رو به دوربین، پشت به دوربین و غیره) را استخراج کردند. سپس برای هر طبقه، ویژگی‌های بافت رنگی افراد داخل

¹ Laplacian of Gaussian

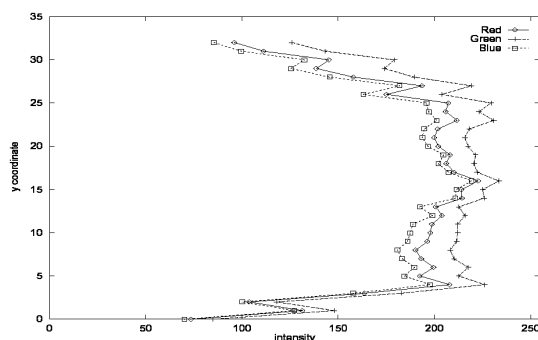
² Connected Component Labeling Algorithm

³ Vertical Distribution

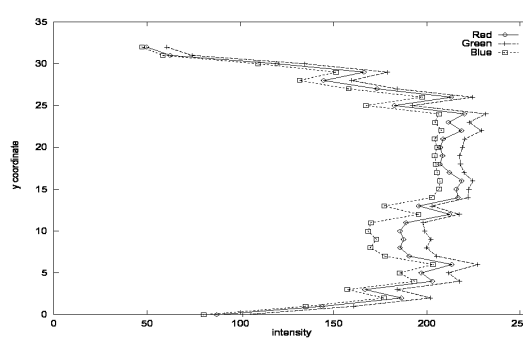
زمین از جمله میانگین، میانه، مد، واریانس و میزان کجی (درجه نامتقارن بودن مقادیر پیکسل‌ها حول مقدار مرکزی) را برای هر مولفه رنگی و در چندین سیستم رنگی محاسبه کردند.

واندنبروک برای طبقه‌بندی یک شیء مجهول، فاصله اقلیدسی ویژگی‌های بافت به دست آمده میان شیء مجهول و هر یک از طبقه‌ها را مورد استفاده قرار دادند. آن‌ها همچنین برای کاهش درجه ویژگی‌ها و از این رو کاهش محاسبات، به جای انتخاب تمام ویژگی‌ها، تنها آن‌هایی را انتخاب کردند که فاصله درون کلاسی^۱ آن کم، و فاصله بین کلاسی^۲ آن زیاد باشد. مشکل این روش در انتخاب دستی افراد داخل زمین در حالت‌ها و موقعیت‌های مختلف و بار محاسباتی بالای آن می‌باشد.

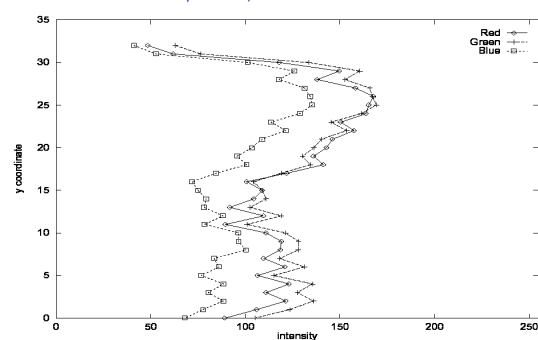
سو و همکارانش [5] توزیع عمودی مولفه‌های رنگی لباس بازیکنان را برای تعیین تیم مورد استفاده قرار دادند؛ بدین صورت که برای هر سطر در الگوی بازیکن، مقادیر میانگین مولفه‌های رنگی قرمز، سبز و آبی را محاسبه کردند. آن‌ها قبل از هر چیز، برای به دست آوردن توزیع‌های هم اندازه، طول توزیع را به صورت نرمال در آوردند. شکل (2-3) نمونه‌ای از توزیع عمودی نرمال شده برای لباس بازیکنان را نشان می‌دهد. همان‌طور که دیده می‌شود بازیکنان تیم اول و همچنین بازیکنان تیم دوم توزیع عمودی مشابهی دارند.



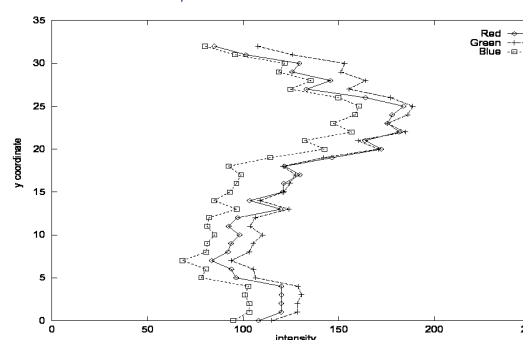
(ب) بازیکن دوم از تیم اول



(الف) بازیکن اول از تیم اول



(د) بازیکن دوم از تیم دوم



(ج) بازیکن اول از تیم دوم

شکل 2-3- نمونه‌ای از توزیع عمودی نرمال شده برای لباس بازیکنان [5]: محور x، شدت رنگ و محور y، ارتفاع الگوی بازیکن را نشان می‌دهد. بازیکنان تیم اول و همچنین بازیکنان تیم دوم توزیع عمودی مشابهی دارند.

¹ Intra Class

² Inter Class

آن‌ها برای مقایسه توزیع عمودی بازیکن مجهول با مدل تیم‌ها، به جای اعمال کانولوشن¹ در تمام قسمت‌های توزیع عمودی، تنها در محدوده کوچکی از توزیع عمودی عمل کانولوشن را انجام دادند تا اثر تغییرات کوچک الگو نسبت به ارتفاع را کاهش دهند. با این وجود، در صورتی که الگوی بازیکن به خاطر حالت‌های مختلف (دویدن، راه رفتن و غیره) و موقعیت‌های مختلف نسبت به دورین دچار تغییر نسبتاً زیادی شود، الگوریتم با شکست مواجه خواهد شد.

روش‌هایی نیز مبتنی بر توزیع هیستوگرام رنگی معرفی شد که در ابتدا توسط گانگ و همکارانش [3] مورد استفاده قرار گرفت. گانگ با محاسبه فاصله نقطه اوج هیستوگرام مولفه‌های رنگی مدل تیم‌ها با نقطه اوج هیستوگرام متناظر بازیکنان، اقدام به شناسایی تیم بازیکنان نمود. وی برای تعیین هیستوگرام مولفه‌های رنگی مدل تیم‌ها، فریم‌هایی را به کار برد که بازیکنان در آن بخش وسیعی از تصویر را اشغال کرده باشند. علاوه بر این، روش‌های دیگری مبتنی بر توزیع هیستوگرام مشترک² و توزیع هیستوگرام باز تصویر شده³، برای یافتن تیم بازیکنان معرفی شده است [25, 20] که در ادامه به توضیح بیشتر این روش‌ها پرداخته می‌شود.

با فرض آنکه توزیع هیستوگرام مولفه‌های رنگی برای مدل تیم‌های A و B به ترتیب $H_{ModelA}(R, G, B)$ و $H_{ModelB}(R, G, B)$ و توزیع هیستوگرام مولفه‌های رنگی بازیکن مجهول $H_p(R, G, B)$ باشد، توزیع هیستوگرام مشترک، H_I ، و باز تصویر شده، H_{BP} ، برای تیم‌های A و B به ترتیب به صورت زیر تعریف می‌شود [26]:

$$H_{I_A}(R, G, B) = \min(H_{ModelA}(R, G, B), H_p(R, G, B)) \quad (22-2)$$

$$H_{I_B}(R, G, B) = \min(H_{ModelB}(R, G, B), H_p(R, G, B))$$

$$H_{BP_A}(R, G, B) = \min\left(\frac{H_{ModelA}(R, G, B)}{H_p(R, G, B)}, 1\right) \quad (23-2)$$

$$H_{BP_B}(R, G, B) = \min\left(\frac{H_{ModelB}(R, G, B)}{H_p(R, G, B)}, 1\right)$$

بنابراین تیم متناظر هر بازیکن با مقایسه میزان شباهت توزیع هیستوگرام‌های فوق با تیم‌های A و B به صورت زیر مشخص می‌گردد:

$$Similarity_{I_A} = \sum H_{I_A}(R, G, B) \quad (24-2)$$

$$Similarity_{I_B} = \sum H_{I_B}(R, G, B)$$

$$Similarity_{BP_A} = \sum H_{BP_A}(R, G, B) \quad (25-2)$$

$$Similarity_{BP_B} = \sum H_{BP_B}(R, G, B)$$

که $Similarity_{BP_X}$ و $Similarity_{I_X}$ به ترتیب میزان شباهت بازیکن مجهول به تیم $X \in \{A, B\}$ در هیستوگرام مشترک و هیستوگرام باز تصویر شده را نشان می‌دهد.

¹ Convolution

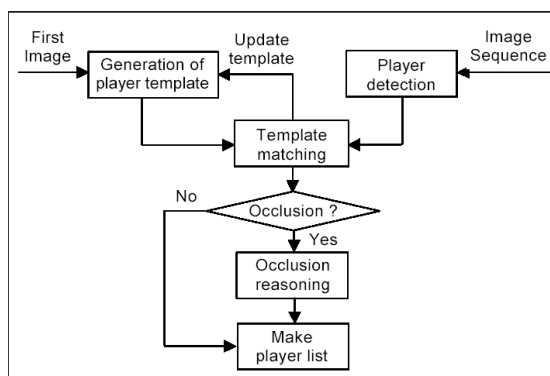
² Histogram Intersection

³ Histogram Back-Projection

2-4- ردیابی بازیکنان

اغلب سیستم‌ها برای ردیابی بازیکنان مبتنی بر کلیشه‌های همبستگی¹ هستند [20,5,3,2]. به عنوان مثال، در سیستم یون [20] کلیشه بازیکنان از روی بازیکنان شناخته شده در هر فریم به هنگام می‌گردد. موقعیت بازیکنان در فریم بعدی نیز بر اساس نقاط نزدیکترین همسایگی و اطلاعات تیم تخمین زده می‌شود.

در کار انجام شده توسط یون و همکارانش برای رفع مشکل هم‌پوشانی² تنها هم‌پوشانی بین 2 بازیکن از تیم‌های مختلف در نظر گرفته شده است؛ بدین صورت که با هم‌پوشانی 2 بازیکن، کلیشه‌ای جدید شامل هر 2 بازیکن ایجاد و تا جدا شدن آن‌ها از یکدیگر ردیابی این کلیشه ادامه می‌یابد. با جدا شدن بازیکنان از یکدیگر نیز 2 کلیشه مجزا برای هر یک ساخته می‌شود. شکل (2-4) رویه ردیابی بازیکنان توسط یون و همکارانش را نشان می‌دهد.



شکل 2-4- رویه ردیابی بازیکنان توسط یون و همکارانش [20]

روش به کار رفته توسط سو و همکارانش [5] برای ردیابی بازیکنان، استفاده از تطبیق کلیشه همبستگی و فیلتر کالمن است. در این روش، کلیشه هر بازیکن با استخراج نواحی متصل به یکدیگر (پیکسل‌های بازیکنان) به دست می‌آید. در ابتدا بازیکنان جدیدی که با بازیکنان واقع در لیست ردیابی هم‌پوشانی نداشته باشند، یافت شده و به لیست ردیابی اضافه می‌شوند. موقعیت بازیکنان در هر فریم نیز با فیلتر کالمن پیش‌بینی می‌شود و سپس یافتن بازیکن در موقعیت پیش‌بینی شده با استفاده از تطبیق کلیشه همبستگی صورت می‌گیرد. سرانجام، با مشخص شدن ناحیه جدید بازیکن، کلیشه همبستگی به هنگام می‌گردد. سو همچنین تنها هم‌پوشانی بین تیم‌های مختلف را مورد بررسی قرار داد. وی برای رفع مشکل هم‌پوشانی هیستوگرام باز تصویر شده [26] را انتخاب نمود. چون در صحنه‌های گل، مسیر حرکت بازیکنان ناشناخته و غیر قابل پیش‌بینی است، به نظر می‌رسد به کارگیری فیلتر کالمن در چنین مواردی با مشکل مواجه خواهد شد.

اوتسومی و همکارانش [21] برای ردیابی بازیکنان انجام مراحل زیر را به ترتیب پیشنهاد دادند:

• شناسایی هم‌پوشانی بین 2 بازیکن

بدین منظور فریم‌های همسایه مورد استفاده قرار گرفته است؛ بدین صورت که اگر مرکز ثقل بازیکنان (روش شناسایی بازیکنان توسط اوتسومی در بخش قبلی آورده شده است) در فریم بعدی به یک ناحیه تعلق داشته باشد، هم‌پوشانی در نظر گرفته می‌شود.

¹ Correlation Templates

² Occlusion Reasoning

• ردیابی بازیکنان غیر هم پوشان

برای هر یک از بازیکنان غیر هم پوشان، 2 فریم همسایه مقایسه شده و بزرگترین ناحیه روی هم افتاده در 2 فریم به آن بازیکن اختصاص داده می شود.

• شناسایی هم پوشانی بین بازیکن و غیر بازیکن

باقیمانده بازیکنان از 2 مرحله قبلی به عنوان بازیکنان هم پوشان با اشیاء غیر بازیکن در نظر گرفته می شود.

• ردیابی بازیکنان هم پوشان

برای ردیابی بازیکنان هم پوشان که از مراحل (الف) و (ب) به دست آمده اند، روش تطبیق کلیشه مبتنی بر رنگ به کار گرفته شده است.

واندنبروک و همکارانش [7] روش پیرامون فعال را برای ردیابی بازیکنان انتخاب کردند. آن ها برای محاسبه انرژی داخلی پیرامون فعال، خصوصیات کشیدگی و خمیدگی منحنی را مورد استفاده قرار دادند؛ بدین منظور انرژی داخلی E_{int} به صورت زیر محاسبه می گردد:

$$E_{int} = \alpha \cdot \int_0^1 \|v_i'\|^2 ds + \beta \cdot \int_0^1 \|v_i''\|^2 ds \quad (26-2)$$

که در آن v_i' و v_i'' به ترتیب مشتقات مرتبه اول و دوم منحنی بسته پیرامون v_i و α و β به ترتیب ضرایب وزنی کشیدگی و خمیدگی را نشان می دهد. برای محاسبه انرژی خارجی نیز مقدار گرادیان تصویر اصلاح شده مورد استفاده قرار گرفته است. در این سیستم برای هر تیم، یک تصویر از روی پیکسل های طبقه بندی شده بازیکنان (طبقه بندی پیکسل های بازیکنان توسط واندنبروک در بخش (2-3) آورده شده است) ایجاد و سپس برای محاسبه گرادیان ناحیه هر بازیکن، تصویر مربوط به تیم آن بازیکن به کار گرفته می شود. از این رو، برای تیم های مختلف، پیرامون های متفاوتی ایجاد شده و در نتیجه مشکل هم پوشانی بین 2 بازیکن از تیم های مختلف در این سیستم رفع می گردد.

لفوره و همکارانش [12] نیز مدل پیرامون فعال را برای ردیابی بازیکنان به کار بردند. انرژی داخلی تعریف شده توسط آن ها به این صورت می باشد:

$$E_{int} = \alpha \cdot E_{con} + \beta \cdot E_{bal} + \gamma \cdot E_{cur} \quad (27-2)$$

که E_{con} ، E_{bal} و E_{cur} به ترتیب پیوستگی، انرژی بالونی و انرژی خمیدگی را نشان می دهند. انرژی پیوستگی بر روی انحنای پیرامون تاثیر دارد. این انرژی نقاط پیرامون را در فاصله ای یکسان قرار می دهد، به طوری که پیرامون به سمت یک دایره میل کند. انرژی $E_{con}(n_i)$ برای نقطه n_i در همسایگی یک نقطه پیرامون مانند v_i به صورت قدر مطلق تفاضل بین 2 فاصله زیر تعریف شده است:

• فاصله بین 2 نقطه متوالی کانطور (که در هر تکرار محاسبه می شود)

• فاصله نقطه کانطور قبلی v_{i-1} تا نقطه n_i

انرژی بالونی $E_{bal}(n_i)$ در سیستم لفوره به صورت حاصل ضرب بردار نرمال پیرامون در نقطه v_i با بردار $\overrightarrow{v_i n_i}$ تعریف شده است. این انرژی باعث کوچک شدن پیرامون فعال می گردد. انرژی خمیدگی نیز به صورت مشتق مرتبه دوم در نظر گرفته شده است تا انسجام¹ شکل حفظ گردد.

¹ Coherent

انرژی خارجی در نظر گرفته شده توسط لافوره بر اساس گرادیان در تمامی مولفه‌های رنگی و شدت مولفه سبز رنگ به صورت زیر می‌باشد:

$$E_{ext} = \lambda \cdot E_{gra} + \mu \cdot E_{gre} \quad (28-2)$$

که E_{gre} و E_{gra} به ترتیب گرادیان‌های مولفه‌های رنگی و شدت مولفه سبز رنگ را نشان می‌دهد. انرژی E_{gra} از ترکیب گرادیان‌ها در 3 مولفه رنگی RGB محاسبه می‌شود و باعث می‌شود تا پیرامون فعال در پیرامون واقعی بازیکنان متوقف شود. برای محاسبه گرادیان، اپراتور لبه‌یاب سوبل همراه با آستانه‌سازی به کار رفته است. انرژی E_{gre} از روی اختلاف شدت روشنایی مولفه سبز رنگ نقطه n_i با یک مقدار آستانه که نشان‌دهنده رنگ چمن است، به دست می‌آید. این انرژی برای میل کردن پیرامون فعال به سمت پیکسل‌های سبز رنگ (چمن) مورد استفاده قرار گرفته است تا به این ترتیب پیرامون فعال، پیرامون واقعی بازیکن را کاملاً در بر گیرد.

اکین و تکالپ [23] ردیابی نقاط ویژگی و استفاده از مستطیل محیطی را برای ردیابی بازیکنان به کار بردند. در این روش، موقعیت بازیکنان در فریم اول به صورت دستی و از طریق مستطیل محیطی تعیین می‌شود. حرکت نقاط ویژگی از فریم قبلی به فریم فعلی برای تخمین گوشه‌های بالایی سمت چپ و پایینی سمت راست مستطیل محیطی به کار گرفته شده است. سپس در ناحیه‌ای بزرگتر از مستطیل محیطی تخمینی، نقاط ویژگی با روش طبقه‌بندی رنگ به 2 کلاس بازیکن و غیر بازیکن به دست می‌آید؛ بدین صورت که پیکسل‌هایی که رنگ مشابه با میانگین رنگ بازیکن ردیابی شونده دارند، به عنوان پیکسل‌های بازیکن شناخته شده و آنگاه بر اساس الگوریتم "برجسب گذاری اجزای متصل" به صورت نواحی متمایز مجتمع می‌شوند. سپس مستطیل محیطی بزرگترین ناحیه به دست آمده به عنوان محدوده فعلی بازیکن انتخاب می‌گردد. در موارد هم‌پوشانی نیز الگوریتمی مشابه با فیلتر کالمن مورد استفاده قرار گرفته است.

2-5- شناسایی و طبقه‌بندی نشانه‌های خطی

گانگ و همکارانش [2] در اولین گام جهت شناسایی و طبقه‌بندی نشانه‌های خطی، مراحل زیر را به ترتیب پیشنهاد دادند:

الف) شناسایی لبه‌های بارز تصویر

گانگ فیلتر لبه‌یاب گاوسی-لاپلاسن را برای شناسایی لبه‌های تصویر مورد استفاده قرار داد. در این فیلتر ترکیبی، انحراف معیار موجود در فیلتر گاوسی، حساسیت به لبه را مشخص می‌کند؛ بدین صورت که با انتخاب مقادیر بزرگ برای آن، تنها لبه‌های عریض و با انتخاب مقادیر کوچک، هم لبه‌های عریض و هم لبه‌های نازک انتخاب می‌شوند. سپس با توجه به آنکه نشانه‌های خطی سطح خاکستری بالایی دارند، لبه‌هایی با سطح خاکستری پایین حذف می‌گردد. در پایان برای داشتن لبه‌هایی با عرض یک پیکسل، اپراتور نازک‌سازی لبه به کار می‌رود.

ب) اصلاح لبه‌های شناسایی شده

گانگ 2 فیلتر لبه‌یاب گاوسی-لاپلاسن با انحراف معیارهای کوچک و بزرگ را برای شناسایی لبه‌ها مورد استفاده قرار داد. سپس برای اصلاح لبه‌ها، نشانه‌های خطی به دست آمده با مقادیر بزرگ انحراف معیار را با ارجاع به نشانه‌های خطی به دست آمده با مقادیر کوچک انحراف معیار (که می‌تواند همراه با نویز و لبه‌های نازک باشد) و با به کارگیری روش دنبال کردن نشانه‌های خطی به یکدیگر وصل نمود.

ج) شناسایی الگوی نشانه‌های خطی

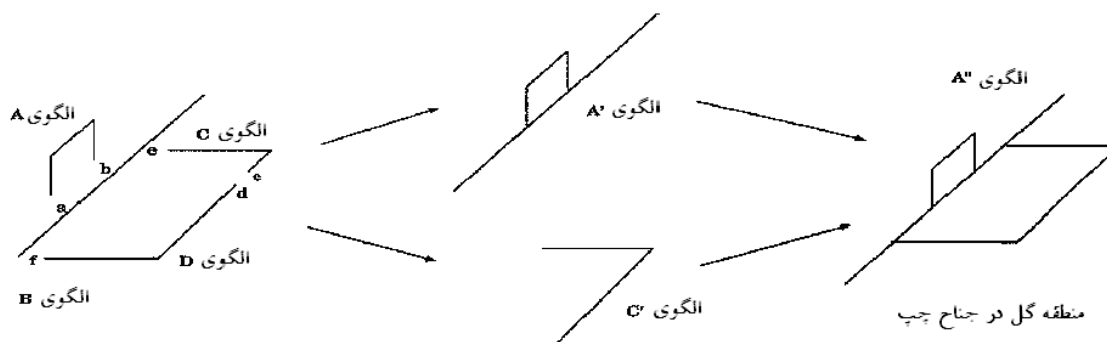
پس از فرآیند اصلاح لبه، تمامی الگوهای خطی موجود در زمین فوتبال به چند جزء از 4 جزء اصلی که در شکل (2-5) نشان داده شده است، تقسیم می‌شوند. به همین جهت، گانگ¹ را برای تشخیص اجزای اصلی و الگوهای خطی مشتق شده از آن‌ها به کار برد. پس از تعیین اجزای اصلی، الحاق اجزای الگو طبق 2 اصل زیر انجام شده تا الگوهای نهایی به دست آیند:

- نقطه پایان جزء اصلی یک الگو در همسایگی نقطه پایان جزء اصلی الگوی دیگر باشد.
- دو جزء اصلی الگو به یکدیگر وصل و یا در موقعیت‌های نزدیک به یکدیگر واقع شده باشند.



شکل 2-5-4 جزء اصلی نشانه‌های خطی زمین فوتبال (خط صاف، جعبه مستطیلی شکل، گوشه و بیضی)

شکل (2-6) نمونه‌ای از فرآیند الحاق اجزای الگو برای شناسایی الگوی نشانه‌های خطی توسط گانگ و همکارانش را نشان می‌دهد. در این فرآیند، چون مستطیل A با خط مستقیم B در نقطه a و b متصل است، با یکدیگر الحاق می‌شوند و الگوی A' را می‌سازند. از طرف دیگر نقطه پایانی الگوی C (در نقطه c) در همسایگی نقطه پایان الگوی D (در نقطه d) واقع شده است، بنابراین با الحاق آن‌ها الگوی C' تشکیل می‌شود. در پایان، چون این 2 الگو در نقاط e و f به یکدیگر متصل هستند، الگوهای A' و C' با یکدیگر الحاق و الگوی A'' را تولید می‌کند. الگوی A'' با مدل سیستم تطبیق و در نتیجه به عنوان منطقه گل در جناح چپ شناخته می‌شود.



شکل 2-6- نمونه‌ای از فرآیند الحاق اجزای الگو برای شناسایی الگوی نشانه‌های خطی توسط

گانگ و همکارانش [2]

¹ Signature

یون و همکارانش [20] برای تشخیص نشانه‌های خطی روی زمین، الگوریتم هاف و سطح خاکستری بالای خطوط را به کار بردند. آن‌ها برای طبقه‌بندی خطوط تنها در حالتی که یک خط افقی و یک خط عمودی در تصویر شناسایی شود، آن‌ها را به ترتیب، خط کناری¹ و خط مرکزی² در نظر گرفته و در صورتی که تعداد خطوط شناسایی شده در تصویر زیاد باشد، اطلاعات را به صورت دستی تنظیم کردند.

2-6- ایجاد نمای دور

یو و همکارانش [3] کنار هم قرار دادن ایستا را برای ایجاد نمای دور به کار بردند. در این روش رنگ هر پیکسل در نمای دور از روی اعمال فیلتر میانه بر روی رنگ پیکسل‌های هم موقعیت در دنباله تصویر مورد نظر به دست می‌آید. نمای دور به دست آمده، تصویری از پس‌زمینه فریم‌های کنار هم قرار داده شده را نشان می‌دهد. برای نمایش حرکت اشیاء بر روی نمای دور پس‌زمینه نیز هر فریم بر روی فریم قبلی قرار داده شده است. مشخص کردن پیکسل‌های هم موقعیت در بین 2 فریم از طریق پارامترهای حرکتی دوربین انجام گردیده است؛ به این صورت که نگاشت زیر برای محاسبه پارامترهای بزرگ‌نمایی و چرخش دوربین در نظر گرفته شده است:

$$I_{k+1}(X, Y) = I_k(A(X, Y)) \quad (29-2)$$

که در آن

$$A(X, Y) = (a_1 X + a_2, a_1 Y + a_3) \quad (30-2)$$

که $I_k(X, Y)$ شدت روشنایی در پیکسل (X, Y) و ضرایب a_1 ، a_2 و a_3 به ترتیب ضریب بزرگ‌نمایی، ضریب چرخش افقی و ضریب چرخش عمودی را نشان می‌دهد.

برای محاسبه پارامترهای یاد شده از یک مدل روشنایی مبتنی بر جمله سری تیلور مرتبه دوم استفاده گردیده است. در این رابطه تنها قسمت‌هایی مد نظر قرار می‌گیرد که به اندازه کافی بافت و اختلاف روشنایی داشته باشند؛ زیرا قسمت‌های مختلف چمن میدان بازی که اغلب فضای تصویر را شامل می‌شود، نسبت به یکدیگر اختلاف چندانی ندارند.

یون و همکارانش [20] تبدیل آفین را برای کنار هم قرار دادن دنباله تصویر به کار بردند، بدین صورت که با مشخص کردن 3 نقطه متناظر در تصویر و مدل میدان، پارامترهای تبدیل آفین را به دست آوردند. نقاط در نظر گرفته شده، تقاطع نشانه‌های خطی یا نقاط معلوم در دایره مرکزی می‌باشند. در تبدیل آفین نگاشت زیر تعریف می‌شود:

$$x = a_{11}u + a_{21}v + a_{31} \quad (31-2)$$

$$y = a_{12}u + a_{22}v + a_{32} \quad (32-2)$$

که با داشتن 3 نقطه متناظر بین تصویر و مدل که بر روی یک خط راست قرار ندارند، رابطه خطی زیر به دست می‌آید:

$$\begin{bmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = \begin{bmatrix} u_0 & v_0 & 1 \\ u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix} \quad (33-2)$$

¹ Side Line

² Center Line

در این رابطه‌ها x و y مختصات نقاط مقصد (مدل) و u و v مختصات نقاط مبدا (تصویر) را نشان می‌دهد. ضرایب اندیس‌دار a نیز پارامترهای آفین می‌باشند. بنابراین پارامترهای آفین با حل رابطه زیر محاسبه می‌شوند:

$$\begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix} = \begin{bmatrix} u_0 & v_0 & 1 \\ u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} \quad (34-2)$$

پس از تعیین پارامترهای آفین برای هر پیکسل در نمای دور، متناظر آن در تصویر با استفاده از رابطه زیر به دست آمده و به این ترتیب رنگ پیکسل‌های نمای دور با رنگ پیکسل متناظر در تصویر مشخص می‌شود:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix}^{-1} \quad (35-2)$$

یون در صورتی که نشانه‌های خطی به اندازه کافی در تصویر وجود نداشته باشد، به منظور یافتن 3 نقطه مشخص در تصویر برای تبدیل آفین، تابلوهای تبلیغاتی را برای ساخت نمای دور مورد استفاده قرار داد. در این روش تابلوهای تبلیغاتی از تصویر استخراج و تابع خطای زیر محاسبه می‌گردد:

$$E^y(t) = \sum_{y=1}^n ((I(x', y', t+1) - I(x, y, t))^2 \quad (36-2)$$

که در آن تابع $E^y(t)$ مجموع اختلاف شدت روشنایی پیکسل‌های تابلو تبلیغاتی بین فریم‌های t و $t+1$ را نشان می‌دهد و I میزان شدت روشنایی و n پهنای تابلو تبلیغاتی می‌باشد. بنابراین تناظری که تابع $E^y(t)$ را به حداقل رساند، برای کنار هم قرار دادن 2 فریم مورد استفاده قرار می‌گیرد؛ زیرا پیکسل‌های متناظر در فریم‌های t و $t+1$ باید شدت روشنایی مشابهی داشته باشند؛ به عبارت دیگر

$$I(x', y', t+1) \approx I(x, y, t) \quad (37-2)$$

سو و همکارانش [5] برای ترازبندی 2 تصویر از روش ثبت سلسله مراتبی¹ استفاده نمودند. آن‌ها به جای در نظر گرفتن خود تصویر برای انجام تطابق بین 2 فریم، تصویر لبه را بدین منظور به کار بردند؛ بدین صورت که در ابتدا تصویر لبه با اپراتور سوبل به دست آمده و آنگاه تصویر لبه به گونه‌ای نرمالیزه می‌شود که 85٪ لبه‌ها را در بر گرفته باشد. همچنین نواحی بازیکنان از تصویر لبه حذف می‌گردد. سپس تابع زیر برای میزان تطابق بین 2 تصویر لبه I_p^1 و I_p^2 مورد استفاده قرار می‌گیرد:

$$E(\bar{u}) = \sum_{\bar{x}} (I_p^1(\bar{x}, t) - I_p^2(\bar{x} - \bar{u}(\bar{x}), t-1))^2 \quad (38-2)$$

که $\bar{x} = (x, y)$ موقعیت تصویر، t فریم فعلی و $\bar{u}(\bar{x}) = (u(x, y), v(x, y))$ سرعت تصویر را نشان می‌دهد. این تابع نسبت به 8 پارامتر تبدیل پرسپکتیو² بهینه می‌شود:

$$u(x, y) = \frac{p_1 x + p_2 y + p_5}{p_7 x + p_8 y + 1} \quad (39-2)$$

$$v(x, y) = \frac{p_3 x + p_4 y + p_6}{p_7 x + p_8 y + 1} \quad (40-2)$$

¹ Hierarchical Direct Registration

² Perspective Transformation

با به دست آمدن پارامترهای تبدیل پرسپکتیو بین تصاویر و مدل میدان، تمامی فریم‌ها به مدل نگاشت شده و نمای دور حاصل می‌گردد.

کیم و هونگ [19, 27, 28] برای کنار هم قرار دادن دنباله تصویر، روش درجه‌بندی خودکار دوربین و ردیابی نشانه‌های خطی را به کار بردند. آن‌ها ابتدا با استفاده از تکنیک درجه‌بندی خودکار، مدل دوربین را تخمین زده و سپس بر طبق پارامترهای تخمین زده شده دوربین، به ردیابی نشانه‌های خطی پرداختند. در ادامه به توضیحات بیشتر این روش پرداخته می‌شود.

کیم، ماتریس تصویر¹ زیر را که بر اساس ماتریس چرخش و ماتریس دوربین به دست می‌آید، برای تعیین پارامترهای دوربین در فریم k ام در نظر گرفت:

$$P_k = K_k [R_k | 0] \quad (41-2)$$

که R_k چرخش k امین فریم نسبت به فریم صفر (فریم مرجع) و K_k ماتریس دوربین فریم k ام را نشان می‌دهد. ماتریس دوربین K_k به صورت زیر تعریف می‌شود:

$$K_k = \text{diag}(f_k, f_k, 1) \quad (42-2)$$

که f_k فاصله کانونی دوربین فریم k ام می‌باشد.

در مدل فوق ماتریس تصویر 2 بعدی زیر برای تبدیل هر نقطه مانند u_0 (یا هر خط مانند $\vec{l}_0$) در فریم مرجع به نقطه متناظرش، u_k (یا خط متناظرش $\vec{l}_k$)، در فریم k ام به کار گرفته می‌شود $(u_k = H_k u_0, \vec{l}_k = H_k \vec{l}_0)$:

$$H_k = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = K_k R_k K_0^{-1} = \begin{bmatrix} \cos \beta_k & \sin \alpha_k \sin \beta_k & -f_0 \cos \alpha_k \sin \beta_k \\ 0 & \cos \alpha_k & f_0 \sin \alpha_k \\ \frac{\sin \beta_k}{f_k} & -\frac{\sin \alpha_k \cos \beta_k}{f_k} & \frac{f_0 \cos \alpha_k \cos \beta_k}{f_k} \end{bmatrix} \quad (43-2)$$

که α_k و β_k به ترتیب زوایای چرخشی حول محور x و y در مختصات دوربین فریم مرجع را نشان می‌دهد و در نتیجه:

$$\alpha_k = \frac{h_{23}h_{33}}{|h_{23}h_{33}|} \tan^{-1} \sqrt{\frac{h_{23}h_{33}}{|h_{23}h_{33}|}} \quad (44-2)$$

$$\beta_k = \frac{h_{13}h_{33}}{|h_{13}h_{33}|} \tan^{-1} \sqrt{\frac{h_{13}h_{33}}{|h_{13}h_{33}|}} \quad (45-2)$$

فاصله‌های کانونی f_0 و f_k نیز به صورت زیر قابل محاسبه می‌باشند:

$$f_0 = \begin{cases} \sqrt{|h_{13}h_{33}/(h_{11}h_{31} + h_{12}h_{32})|} & \text{if } |\alpha_k| \leq |\beta_k| \\ \sqrt{|h_{23}h_{33}/(h_{21}h_{31} + h_{22}h_{32})|} & \text{Otherwise} \end{cases} \quad (46-2)$$

$$f_k = \frac{1}{2} \left(\sqrt{\frac{f_0^2(h_{11}^2 + h_{12}^2) + h_{13}^2}{f_0^2(h_{31}^2 + h_{32}^2) + h_{33}^2}} + \sqrt{\frac{f_0^2(h_{21}^2 + h_{22}^2) + h_{23}^2}{f_0^2(h_{31}^2 + h_{32}^2) + h_{33}^2}} \right) \quad (47-2)$$

¹ Projective Matrix

کیم برای اصلاح پارامترهای دوربین، تطبیق نشانه‌های خطی شناخته شده در فریم مرجع و فریم k ام را مورد استفاده قرار داد. خطای تطابق خطوط $\vec{l}_0 = \{\vec{l}_0^1, \dots, \vec{l}_0^M\}$ و $\vec{l}_k = \{\vec{l}_k^1, \dots, \vec{l}_k^M\}$ بدین صورت به دست می‌آید:

$$E = \frac{1}{M} \sum_{i=1}^M \left(\frac{d(\vec{l}_k^i, H_k^{-T} \vec{l}_0^i) + d(\vec{l}_0^i, H_k^T \vec{l}_k^i)}{2} \right)^2 \quad (48-2)$$

که $d(\vec{l}_a, \vec{l}_b)$ فاصله بین خطوط $\vec{l}_a$ و $\vec{l}_b$ را نشان می‌دهد. در صورتی که 2 نقطه پایانی خط $\vec{l}_a$ ، e_a^1 و e_a^2 بوده و مختصات همگن $\vec{l}_b$ به صورت $(l_{b,1}, l_{b,2}, l_{b,3})^T$ باشد، فاصله بین 2 خط بدین صورت محاسبه می‌گردد:

$$d(\vec{l}_a, \vec{l}_b) = \sqrt{\frac{|\vec{l}_b^T e_a^1|^2 + |\vec{l}_b^T e_a^2|^2}{(l_{b,1}^2 + l_{b,2}^2)}} \quad (49-2)$$

کیم برای کمینه کردن خطای E ، روش لیونبرگ-مارکورت¹ را به کار گرفت. ردیابی نشانه‌های خطی نیز با استفاده از پارامترهای به دست آمده دوربین و بهینه سازی غیر خطی لیونبرگ-مارکورت انجام گرفت که توضیحات کامل آن در [27] آورده شده است. به این ترتیب در این روش، نشانه‌های خطی برای تخمین پارامترهای دوربین و از طرف دیگر پارامترهای دوربین برای ردیابی نشانه‌های خطی در فریم‌های بعدی مورد استفاده قرار گرفته است. بنابراین با به دست آمدن پارامترهای دوربین و ماتریس تبدیل از نقاط مرجع به نقاط فریم k ام و برعکس، عمل کنار هم قرار دادن دنباله تصویر صورت می‌گیرد. این روش همان‌طور که ملاحظه می‌گردد، بار محاسباتی بالایی دارد؛ زیرا برای هر فریم باید روش حداقل کردن خطا که نیازمند محاسبات زیادی است، به کار گرفته شود.

7-2- جمع‌بندی

در این فصل، روش‌های به کار رفته در مراحل استخراج چمن میدان بازی، شناسایی بازیکنان، تعیین تیم بازیکنان، ردیابی بازیکنان، شناسایی و طبقه‌بندی نشانه‌های خطی و ایجاد نمای دورمورد مطالعه قرار داده شد. به منظور استخراج چمن، سیستم‌های رنگی RGB و HSI مورد استفاده قرار گرفته است. متأسفانه در تمامی روش‌های موجود، استخراج مطلوب ناحیه چمن با تعیین دستی تعدادی پارامتر امکان‌پذیر است. در بخش دوم سیستم‌های موجود برای شناسایی بازیکنان مورد مطالعه قرار گرفت. این سیستم‌ها اطلاعات رنگی، روابط مکانی شورت و پیراهن، ویژگی‌های محلی و اعمال محدودیت‌ها را برای رسیدن به هدف به کار بردند. بخش سوم به روش‌های ارایه شده در تعیین تیم بازیکنان که بیشتر بر اساس مدل هیستوگرام رنگی بازیکنان می‌باشد، می‌پردازد. در بخش چهارم روش‌هایی که به منظور ردیابی بازیکنان به کار رفته است، معرفی گردید. در روش‌های ردیابی، معمولاً اطلاعات حرکتی بازیکنان نسبت به تصویر برای تخمین موقعیت فعلی بازیکنان مورد استفاده قرار می‌گیرد. بخش پنجم نیز به بررسی روش‌های موجود برای شناسایی و طبقه‌بندی نشانه‌های خطی پرداخته است. در پایان، روش‌هایی که برای کنار هم قرار دادن دنباله تصویر صحنه‌های ویدیویی مسابقه فوتبال به منظور ایجاد نمای دور ارایه شده است، معرفی گردید.

¹ Levenberg-Marquardt

فصل سوم

شناسایی و تعیین تیم بازیکنان

به منظور شناسایی بازیکنان بهتر است در ابتدا چمن میدان بازی استخراج شده و آنگاه از روی پیکسل‌های غیر چمن، بازیکنان شناسایی گردند. در این فصل ابتدا روش ارایه شده در استخراج چمن میدان بازی تشریح می‌گردد. سپس روش پیشنهادی برای شناسایی بازیکنان از روی پیکسل‌های غیر چمن معرفی می‌گردد. همچنین، روش پیشنهادی برای تعیین تیم بازیکنان، مبتنی بر ماشین بردار پشتیبان یا SVM در بخش 3 ارایه می‌گردد. در پایان، مطالب در بخش 4 جمع‌بندی شده است.

3-1- استخراج چمن میدان بازی

نتایج تحقیقات اکین و همکارانش [24] نشان می‌دهد که رنگ چمن در ورزشگاه‌های مختلف با یکدیگر متفاوت است و رنگ چمن یک ورزشگاه نیز می‌تواند تحت تاثیر شرایط نوری و آب و هوایی متفاوت قرار گیرد. علاوه بر این، در تمامی قسمت‌های میدان بازی رنگ چمن یکسان نمی‌باشد و به همین دلیل، همان‌طور که در فصل قبل اشاره شد، معمولاً مقدار خاصی برای تعیین رنگ چمن مورد استفاده قرار نمی‌گیرد. شکل (3-1) هیستوگرام‌های رنگی به دست آمده از 45 دقیقه فیلم ویدیویی، در انواع حالت‌های یاد شده را که توسط اکین و همکارانش گردآوری شده است، نشان می‌دهد.

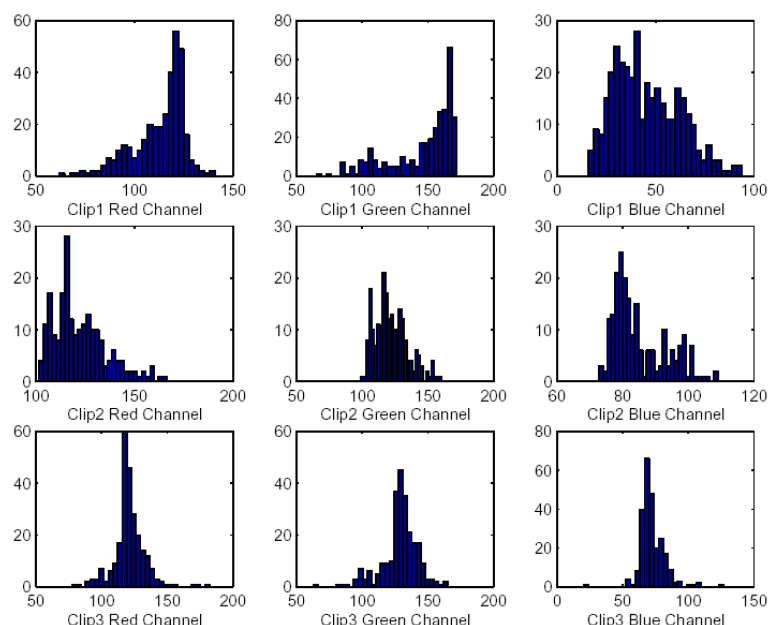
معمولاً صحنه‌های نمای دور و نمای میانه، بخش اعظم تصاویر ویدیویی مسابقه فوتبال را تشکیل می‌دهند. بنابراین رنگ چمن در یک فیلم ویدیویی مسابقه فوتبال بر رنگ‌های دیگر غالب است. به همین جهت در اغلب روش‌ها [3-6] فقط فرضیه وجود یک رنگ غالب منحصر به فرد که تعیین کننده چمن میدان بازی است، مورد استفاده قرار می‌گیرد.

روش به کار گرفته شده در این پایان‌نامه برای استخراج چمن، استفاده از فضای RGB و اصلاح روش یون و همکارانش [20] است. بدین منظور مقادیر آستانه R_t ، G_t و B_t در فرمول (2-14) به ترتیب برابر با انحراف معیار هیستوگرام رنگی از مقادیر R_{peak} ، G_{peak} و B_{peak} انتخاب شده است. همچنین مقدار آستانه GL_t بر حسب نقطه اوج هیستوگرام سطح خاکستری (GL_{peak}) و مقدار انحراف معیار هیستوگرام سطح خاکستری از این نقطه به دست می‌آید:

$$GL_t = GL_{peak} + \alpha \cdot std(GL(x, y)) \quad (3-1)$$

که در آن ضریب α فاصله نقطه GL_t تا نقطه اوج GL_{peak} را بر حسب انحراف معیار هیستوگرام سطح خاکستری از نقطه اوج نشان می‌دهد. در سیستم ارایه شده مقدار α برابر 0/75 در نظر گرفته شده است. با این وجود، برای هر فیلم ویدیویی مسابقه فوتبال مقدار بهینه α را می‌توان به صورت سعی و خطا به دست آورد؛ ولی مهم آن است که α در مقایسه با GL_t

راحت تر به دست می آید، چرا که وابستگی کمتری به تصاویر ویدیویی دارد. واضح است برای دست یافتن به دقت بیشتر در تعیین پارامترهای فوق می توان به جای یک فریم، چندین فریم را به کار برد.



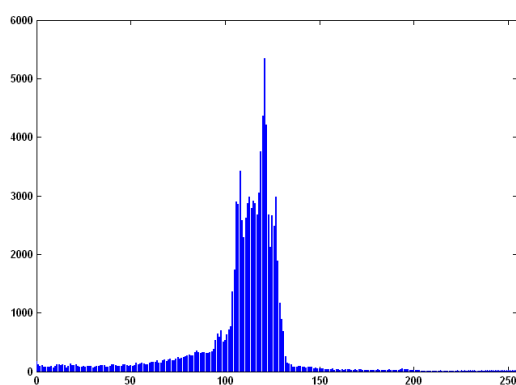
شکل 3-1- هیستوگرام رنگی چمن گردآوری شده از 45 دقیقه بازی [9]: Clip1 و Clip2 در یک ورزشگاه و به ترتیب با شرایط نوری طبیعی و مصنوعی و Clip3 از ورزشگاه دیگر با شرایط نوری طبیعی گرفته شده است.

شکل (3-2) هیستوگرام های مولفه های رنگی و سطح خاکستری یک تصویر واقعی را نشان می دهد. مقادیر نقاط اوج در هیستوگرام مولفه های رنگی قرمز، سبز و آبی به ترتیب برابر با 97، 122 و 27 و مقادیر آستانه برای آن ها به ترتیب برابر با 23، 28 و 21 می باشد. به عبارت دیگر، بازه های قابل قبول در استخراج چمن زمین برای مولفه رنگی قرمز برابر [74 120]، برای مولفه رنگی سبز برابر [94 150] و برای مولفه رنگی آبی برابر [6 48] است. همچنین مقدار آستانه سطح خاکستری 119 شده است.

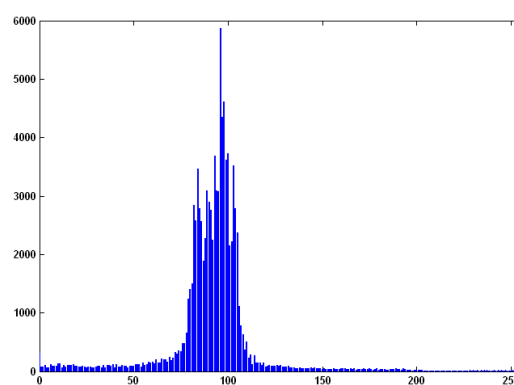
همان طور که در شکل (3-2) مشاهده می شود، همواره توزیع هیستوگرام های یاد شده نسبت به نقطه اوج هیستوگرام، متقارن نیست و در نتیجه انتخاب نقطه اوج به عنوان نقطه مرکز بازه نمی تواند معیار دقیقی برای دستیابی به بازه قابل قبول برای زمین چمن و همچنین تعیین نقطه آستانه برای نشانه های خطی باشد. به عنوان مثال، بازه قابل قبول به دست آمده برای مولفه رنگی قرمز زمین چمن [74 120] است که انتظار می رود این بازه به سمت چپ متمایل گردد تا پیکسل هایی را در بر گیرد که دفعات تکرار آن ها در تصویر با این بازه رنگ آبی زیاد باشد. همچنین نقطه آستانه به دست آمده برای نشانه های خطی بر اساس این روش 119 است که با مشاهده هیستوگرام سطح خاکستری، ملاحظه می گردد این مقدار کمی بیش از مقدار واقعی برای جداسازی نشانه های خطی از زمین چمن است. به همین دلیل نقطه مرکز بازه، به جای نقطه اوج هیستوگرام، نقطه مرکز ثقل آن که البته در نزدیکی نقطه اوج هیستوگرام قرار می گیرد، انتخاب می گردد.



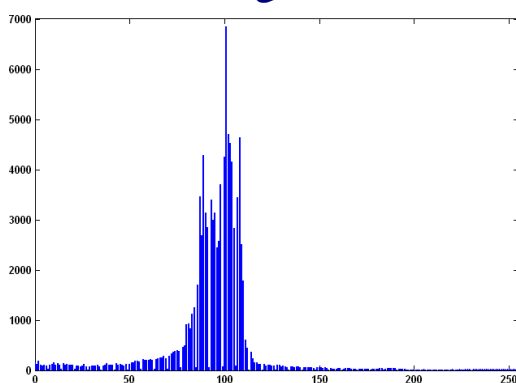
(الف)



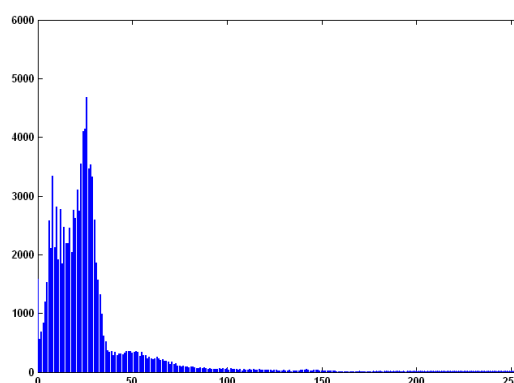
(ب)



(ج)



(د)



(و)

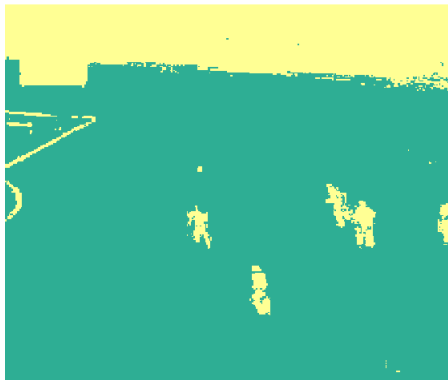
شکل 3-2- هیستوگرام‌های مولفه‌های رنگی و سطح خاکستری یک تصویر واقعی: (الف) تصویر واقعی (ب) هیستوگرام مولفه رنگ قرمز (در روش یون: $R_t = 23$ و $R_{peak} = 97$ - در روش ارایه شده: $R_t = 22$ و $R_{peak} = 94$ (ج) هیستوگرام مولفه رنگ سبز (در روش یون: $G_t = 28$ و $G_{peak} = 122$ - در روش ارایه شده: $G_t = 27$ و $G_{peak} = 117$ (د) هیستوگرام مولفه رنگ آبی (در روش یون: $B_t = 21$ و $B_{peak} = 27$ - در روش ارایه شده: $B_t = 22$ و $B_{peak} = 21$ (و) هیستوگرام سطح خاکستری (در روش یون: $GL_t = 119$ و $GL_{peak} = 102$ - در روش ارایه شده: $GL_t = 115$ و $GL_{peak} = 102$)

از آنجایی که برخی از مقادیر هیستوگرام پیکسل‌های اندکی دارند و این مقادیر نمی‌توانند نماینده چمن سبز رنگ باشند، از آن‌ها صرف نظر می‌گردد. با توجه به مطالب یاد شده، نقطه مرکز بازه برای هر هیستوگرام بدین صورت اصلاح می‌گردد:

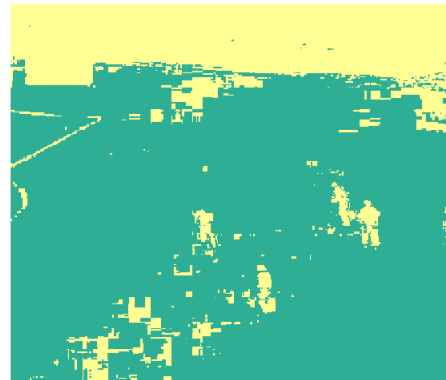
$$i'_{peak} = \frac{\sum_{H(i) \geq \beta \cdot i_{peak}} i \cdot H(i)}{\sum_{H(i) \geq \beta \cdot i_{peak}} H(i)} \quad (2-3)$$

که در آن i'_{peak} مقدار اصلاح شده نقطه اوج، i_{peak} نقطه اوج هیستوگرام، i مقدار (اندیس) هیستوگرام و $H(i)$ تعداد پیکسل‌ها در اندیس i است. همچنین شرط مجموع در رابطه فوق، اندیس‌هایی را بر می‌گزیند که تعداد پیکسل‌ها در آن حداقل به نسبت β از تعداد پیکسل‌های نقطه اوج هیستوگرام باشد؛ به عبارت دیگر β پیکسل‌هایی را مشخص می‌کند که می‌توانند نماینده چمن سبز رنگ باشند. در آزمایش‌های انجام شده این نسبت برابر 10٪ انتخاب گردید.

با اعمال نقطه اوج اصلاح شده به جای نقطه اوج در قوانین استخراج چمن زمین، پارامترها بدین صورت حاصل شد: مقادیر نقاط اوج اصلاح شده حاصل از هیستوگرام مولفه‌های رنگی قرمز، سبز و آبی به ترتیب برابر با 94، 117 و 21 و مقادیر آستانه برای آن‌ها به ترتیب برابر با 22، 27 و 22 به دست آمد و مقدار آستانه برای سطح خاکستری نیز برابر با 115 تعیین گردید. شکل (3-3) خروجی حاصل از استخراج چمن میدان بازی را در 2 حالت به کارگیری نقطه اوج هیستوگرام و به کارگیری نقطه اوج اصلاح شده برای شکل (3-2-الف) نشان می‌دهد. در این تصاویر پیکسل‌های تیره نشان‌دهنده چمن استخراج شده می‌باشد. همان گونه که مشاهده می‌شود، در شکل (3-3-ب) نویز به مراتب کمتر از شکل (3-3-الف) می‌باشد.



(ب)



(الف)

شکل 3-3- چمن میدان بازی استخراج شده از شکل (3-2-الف): (الف) از روی نقطه اوج (ب)

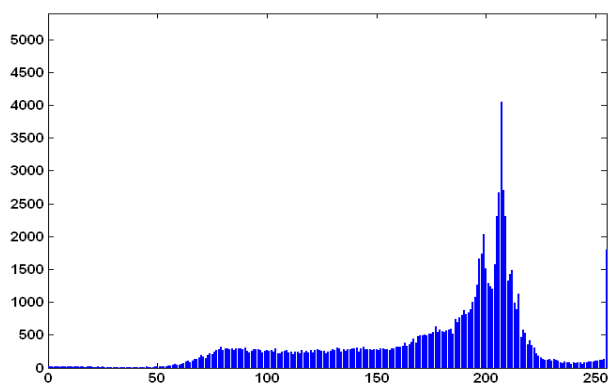
از روی نقطه اوج اصلاح شده - چمن استخراج شده با پیکسل‌های تیره نشان داده شده است. در

شکل (ب) نویز به مراتب کمتر از شکل (الف) می‌باشد.

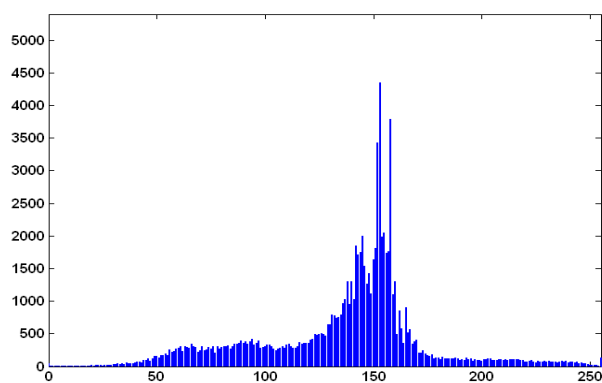
برای واضح تر شدن مطلب، هیستوگرام‌های مولفه‌های رنگی و سطح خاکستری یک تصویر واقعی دیگر در شکل (3-4) نشان داده شده است. شکل (3-5) نیز نتایج حاصل از استخراج چمن میدان بازی را در 2 حالت به کارگیری نقطه اوج هیستوگرام و به کارگیری نقطه اوج اصلاح شده برای شکل (3-4-الف) نشان می‌دهد. همان گونه که در این شکل مشاهده می‌شود، نشانه‌های خطی موجود در شکل (3-5-ب) دقت بیشتری نسبت به شکل (3-5-الف) دارند.



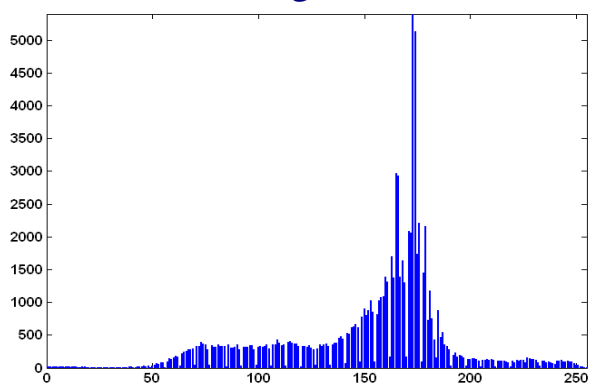
(الف)



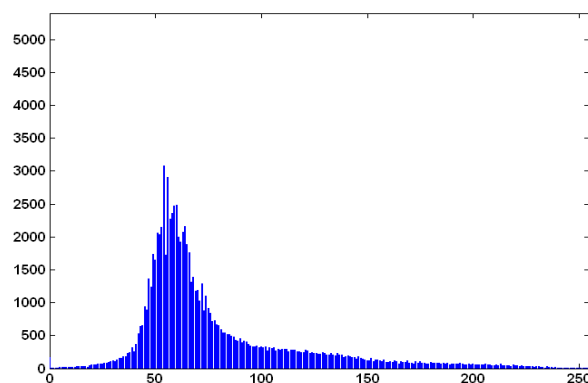
(ج)



(ب)



(و)



(د)

شکل 3-4- هیستوگرام‌های مولفه‌های رنگی و سطح خاکستری یک تصویر واقعی: الف) تصویر واقعی

ب) هیستوگرام مولفه رنگ قرمز (در روش یون: $R_t = 42$ و $R_{peak} = 154$ - در روش ارایه شده:

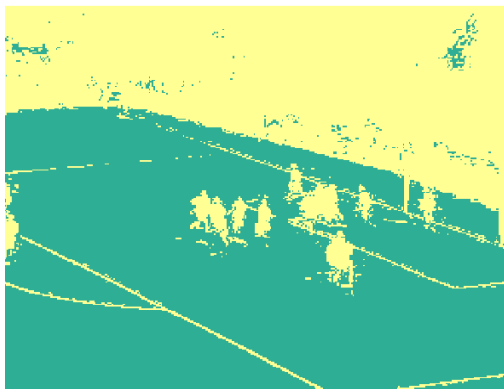
ج) هیستوگرام مولفه رنگ سبز (در روش یون: $G_t = 56$ و $G_{peak} = 208$ و $R_{peak} = 149$ و $R_t = 40$)

- در روش ارایه شده: $G_t = 51$ و $G_{peak} = 200$) د) هیستوگرام مولفه رنگ آبی (در روش یون:

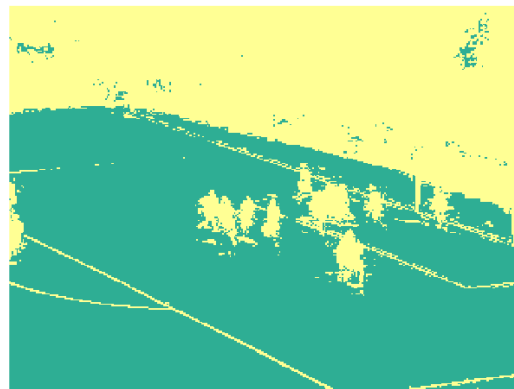
$B_t = 45$ و $B_{peak} = 55$ - در روش ارایه شده: $B_t = 41$ و $B_{peak} = 65$) ه) هیستوگرام سطح

خاکستری (در روش یون: $GL_t = 207$ و $GL_{peak} = 174$ - در روش ارایه شده: $GL_t = 199$ و

$GL_{peak} = 174$)



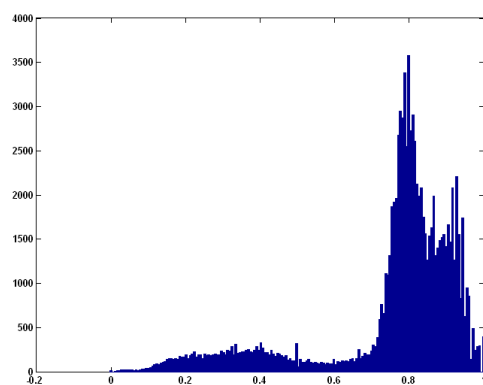
(ب)



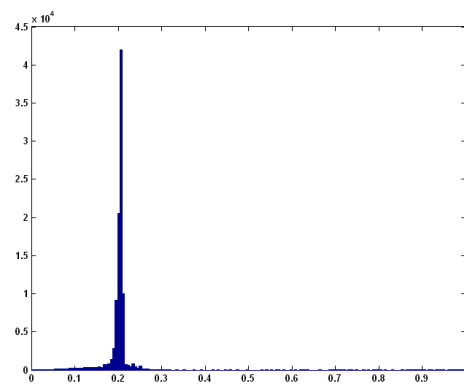
(الف)

شکل 3-5- چمن میدان بازی استخراج شده از شکل (3-4-الف): الف) از روی نقطه اوج ب) از روی نقطه اوج اصلاح شده - نشانه‌های خطی در شکل (ب) با دقت بیشتری در مقایسه با شکل (الف) به دست آمده‌اند.

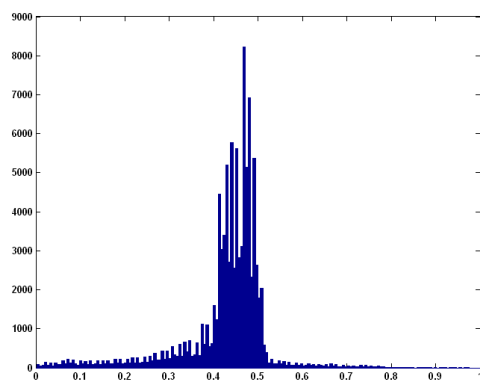
در شکل (3-6) هیستوگرام‌های مولفه‌های فام، درجه اشباع و شدت روشنایی شکل (3-2-الف) به نمایش درآمده است. علاوه بر این، چمن میدان بازی استخراج شده از شکل (3-2-الف) با اختصاص مقدار فام چمن در بازه 54 الی 90 درجه [18]، 65 الی 85 درجه [23]، و 60 الی 150 درجه [21] و نیز با به کارگیری روش اکین در [24] با 2 مقدار اختصاص داده شده $T_{color} = 0.18$ و $T_{color} = 0.2$ در شکل (3-7) نشان داده شده است. همان‌طور که در شکل (3-7-الف) تا (3-7-ج) دیده می‌شود، در نظر گرفتن مقدار فام چمن در بازه خاص نمی‌تواند در استخراج چمن میدان بازی برای تمامی تصاویر ویدیویی مفید واقع شود. همچنین با مقایسه شکل (3-7-د) و (3-7-و) با شکل (3-3-ب) مشاهده می‌شود که روش ارایه شده با دقت بیشتری به جداسازی چمن از دیگر اشیاء داخل زمین می‌پردازد. در شکل (3-7-د) 2 ناحیه علامت‌گذاری شده، به صورت اشتباه به ناحیه غیر چمن تخصیص داده شده‌اند که این ناحیه‌ها ممکن است در مرحله شناسایی بازیکنان به عنوان بازیکن شناخته شوند. همچنین همان‌طور که در شکل (3-7-و) نشان داده شده است، نشانه‌های خطی به درستی تشخیص داده نشده‌اند. شکل (3-8) چمن میدان بازی استخراج شده از 3 تصویر ویدیویی متفاوت را با به کارگیری روش معرفی شده، نشان می‌دهد.



(ب)



(الف)



(ج)

شکل 3-6- هیستوگرام‌های فضای HSI برای شکل (3-2-الف): (الف) هیستوگرام مولفه فام ب)

هیستوگرام مولفه درجه اشباع ج) هیستوگرام مولفه شدت روشنایی



(ب)



(الف)



(ج)

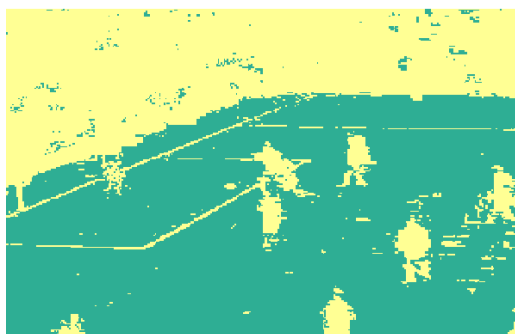


(و)



(د)

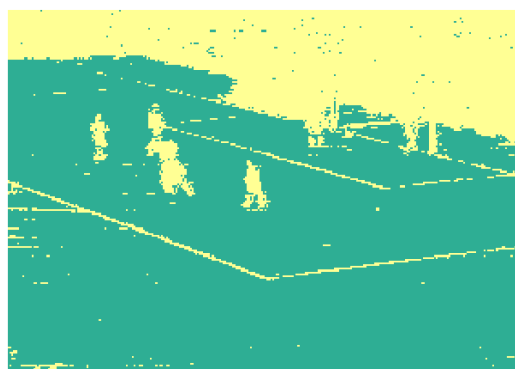
شکل 3-7- چمن میدان بازی استخراج شده از شکل (3-2-الف): الف) با اختصاص مقدار فام چمن در بازه 54 الی 90 درجه ب) با اختصاص مقدار فام چمن در بازه 65 الی 85 درجه ج) با اختصاص مقدار فام چمن در بازه 60 الی 150 درجه د) با به کارگیری روش اکین با اختصاص $T_{color} = 0.18$ و) با به کارگیری روش اکین با اختصاص $T_{color} = 0.2$ - با مقایسه نتایج به دست آمده از شکل های فوق با شکل (3-3-ب)، مشاهده می شود که روش ارایه شده با دقت بیشتری به جداسازی چمن از دیگر اشیاء داخل زمین می پردازد.



(الف-2)



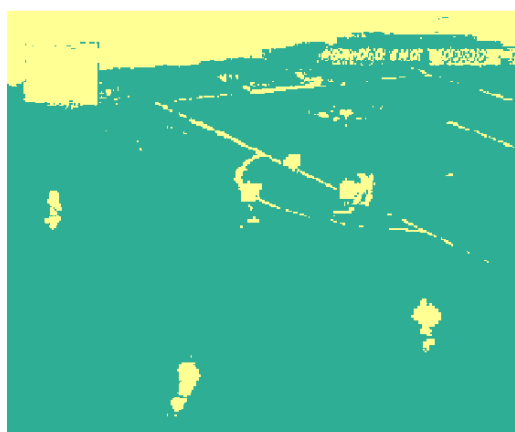
(الف-1)



(ب-2)



(ب-1)



(ج-2)



(ج-1)

شکل 3-8- چمن میدان بازی استخراج شده از 3 تصویر ویدیویی متفاوت با به کارگیری روش معرفی شده

3-2- شناسایی بازیکنان

پس از استخراج چمن میدان بازی، پیکسل‌های غیر چمن به دست آمده (در شکل‌های بخش 3-1 با پیکسل‌های روشن نمایش یافته است، متعلق به بازیکنان، نشانه‌های خطی، تابلوهای تبلیغاتی و نویز می‌باشد. در این فصل روشی جدید مبتنی بر فیلتر میانه و عملگر مورفولوژی باز کردن معرفی می‌گردد که با دقت بالا به جداسازی بازیکنان از دیگر اشیاء داخل زمین می‌پردازد.

در این پایان‌نامه، ابتدا فیلتر میانه بر روی تصویر حاصل از مرحله قبل اعمال شده و پس از آن الگوریتم "برچسب‌گذاری اجزای متصل" برای متمایز کردن بازیکنان به کار گرفته شده است. فیلتر میانه باعث حذف نشانه‌های خطی نازک و نویز می‌گردد. همچنین فیلتر میانه موجب یکپارچه شدن اشیایی می‌شود که اجزای آن با فاصله‌های بسیار کم از یکدیگر مجزا هستند.

آزمایش‌ها نشان می‌دهد در اکثر مواردی که بازیکنان در مجاور یکدیگر قرار گرفته‌اند از طریق باریکه‌های نازک با یکدیگر اتصال دارند. بنابراین در صورت حذف این باریکه‌های نازک، شناسایی بازیکنان با دقت بیشتری صورت خواهد گرفت. بدین منظور عملگر مورفولوژی باز کردن به کار گرفته شده است، زیرا این عملگر قادر است باریکه‌های نازک را قطع و برآمدگی‌های نازک را حذف نماید.

عملگر مورفولوژی باز کردن¹ تصویر A به وسیله عنصر سازنده S که با $A \circ S$ نشان داده می‌شود، به صورت زیر تعریف می‌گردد [29]:

$$A \circ S = (A \ominus S) \oplus S \quad (3-3)$$

که A و S مجموعه‌ای از مختصات پیکسل‌ها با مقادیر 1 در فضای 2 بعدی می‌باشند (مقادیر 1 که در اینجا پیکسل‌های روشن می‌باشند، نشان‌دهنده شکل موجود در تصویر است). در این مدل، بر طبق قرارداد، مبدا مختصات با مقدار (0,0) بر روی پیکسل گوشه بالای سمت چپ تصویر قرار دارد و هر عنصر در این مجموعه به صورت (سطر، ستون) نشان داده می‌شود. به عبارت دیگر، هر تصویر مانند X به صورت زیر ارایه می‌گردد:

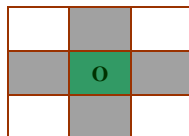
$$X = \{ \underline{x} \mid \underline{x} \in E^2 \} \quad (4-3)$$

در رابطه (3-3) نماد $\ominus$ نشان‌دهنده عملگر سایش² و نماد $\oplus$ نشان‌دهنده عملگر گسترش³ می‌باشد که بدین صورت تعریف می‌شوند:

$$A \ominus S = \{ \underline{z} \in E^2 \mid (\underline{z} + \underline{s}) \in A \quad \forall \underline{s} \in S \} \quad (5-3)$$

$$A \oplus S = \{ \underline{z} \in E^2 \mid \underline{z} = \underline{a} + \underline{s} \quad \forall (\underline{a} \in A \ \& \ \underline{s} \in S) \} \quad (6-3)$$

در سیستم معرفی شده عنصر سازنده S به صورت شکل (3-9) در نظر گرفته شده است.



شکل 3-9- عنصر سازنده S در فضای E^2 : نقاط تیره پیکسل‌های فعال و نقاط روشن پیکسل‌های غیر فعال را نشان می‌دهد. همچنین پیکسل مرکزی O مبدا مختصات عنصر سازنده می‌باشد.

برای شناسایی بازیکنان، یک بار فیلتر میانه به تنهایی و بار دیگر به همراه عملگر مورفولوژیکی باز کردن بر روی شکل (2-3 الف) اعمال شده که نتایج آن در شکل (3-10) نشان داده شده است. همان‌طور که دیده می‌شود در صورت به کار

¹ Opening Morphology

² Erosion

³ Dilation

بردن عملگر مورفولوژیکی باز کردن به همراه فیلتر میانه، بازیکنان نزدیک به هم (در ناحیه علامت گذاری شده با مستطیل) به صورت مجزا شناسایی می شوند. علاوه بر این، در شکل (3-11) نتایج شناسایی بازیکنان برای شکل (3-4-الف) با 2 حالت یاد شده نشان داده شده است. نتایج به دست آمده در این شکل نیز دلیل دیگری بر کارا بودن روش اعمال عملگر مورفولوژی باز کردن به همراه فیلتر میانه است.

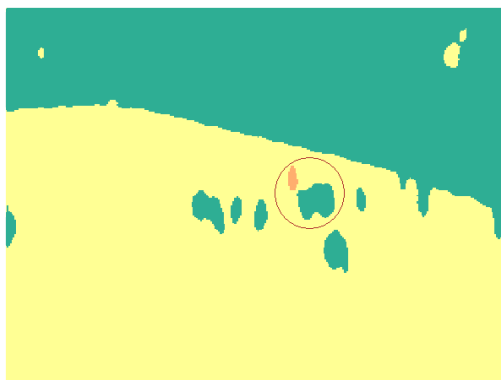


(ب)



(الف)

شکل 3-10- نتایج شناسایی بازیکنان بر روی شکل (3-2-الف): الف) با روش اعمال فیلتر میانه (ب) با روش اعمال فیلتر میانه به همراه عملگر مورفولوژی باز کردن - برخلاف شکل (الف) که یک بازیکن به صورت هم پوشان شناسایی شده، در شکل (ب) آن بازیکن مجزا شناسایی شده است.



(ب)



(الف)

شکل 3-11 نتایج شناسایی بازیکنان بر روی شکل (3-4-الف): الف) با روش اعمال فیلتر میانه (ب) با روش اعمال فیلتر میانه به همراه عملگر مورفولوژی باز کردن - برخلاف شکل (الف) که یک بازیکن به صورت هم پوشان شناسایی شده، در شکل (ب) آن بازیکن مجزا شناسایی شده است.

همان‌طور که در شکل (3-10) و (3-11) دیده می‌شود، در صورت هم‌پوشانی و حتی مجاورت چند بازیکن، تمامی آن‌ها به صورت یک شیء واحد شناسایی می‌گردند. همچنین در حالتی که قسمتی از بدن بازیکن در جلوی تابلوهای تبلیغاتی ورزشگاه قرار گیرد، آن بازیکن به عنوان قسمتی از تابلو تبلیغاتی شناخته می‌شود.

واضح است که نتایج می‌تواند همراه با نویز و خطا باشد. از این رو، اصلاحات زیر برای کاهش خطا اعمال شده‌اند:

• **اصلاح 1:** حذف اشیایی که مساحت آن کمتر از 50 است. این قانون باعث حذف اشیاء کوچک که به احتمال بسیار زیاد نویز هستند، می‌گردد.

• **اصلاح 2:** حذف اشیایی که نسبت محور اصلی به محور فرعی آن بیشتر از 6 است. این مورد می‌تواند بیشتر در صورت باقی ماندن نشانه‌های خطی (کلفت) پس از اعمال فیلتر میانه پیش آید.

• **اصلاح 3:** حذف اشیایی که اکثر پیکسل‌های آن، هم‌رنگ با رنگ چمن میدان بازی است. این قانون باعث حذف قسمت‌هایی از زمین چمن می‌گردد که به خاطر اختلاف اندک با دیگر قسمت‌های زمین، به عنوان بازیکن انتخاب شده است.

• **اصلاح 4:** حذف اشیایی که مولفه‌های رنگی اغلب فضای آن، کمی سیاه‌تر از مولفه‌های رنگی زمین چمن میدان بازی است. این حالت می‌تواند برای سایه بازیکنان رخ دهد.

• **اصلاح 5:** نواحی متعلق به تابلوهای تبلیغاتی حذف شوند.

• **اصلاح 6:** نواحی متعلق به آرم‌های تلویزیونی حذف شوند.

همان‌طور که دیده می‌شود، در اصلاحات فوق محدودیتی برای حداکثر مساحت بازیکن در نظر گرفته نشده است، زیرا این حالت می‌تواند باعث حذف بازیکنانی گردد که با یکدیگر هم‌پوشانی داشته و مساحت زیادی را به خود اختصاص داده‌اند. برای اعمال اصلاح 3، نسبت پیکسل‌هایی که شرایط نشان داده شده در (3-9) را داشته باشند نسبت به تعداد کل پیکسل‌های بازیکن محاسبه و در صورتی که این نسبت از یک حد آستانه که در این پایان‌نامه مقدار 0/8 در نظر گرفته شده است، بیشتر باشد، به عنوان ناحیه هم‌رنگ با چمن تشخیص داده می‌شود. لازم به ذکر است مقدار اختصاص داده شده در این پایان‌نامه برای th_{expand} برابر با 1/3 می‌باشد. همچنین برای در نظر گرفتن اصلاح 4، در صورتی که شرایط در نظر گرفته شده برای اصلاح 3 با جایگزینی R_{peak} ، G_{peak} و B_{peak} به ترتیب با $R_t - R_{peak}$ ، $G_t - G_{peak}$ و $B_t - B_{peak}$ مهیا باشد، ناحیه مربوطه به عنوان سایه مد نظر قرار می‌گیرد.

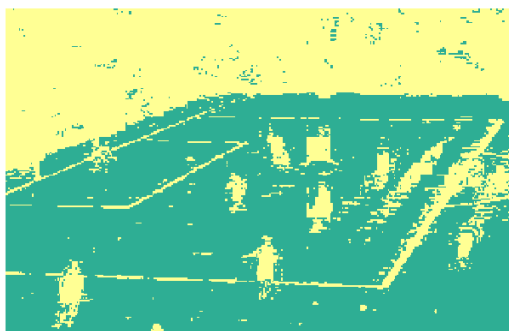
$$\begin{cases} I_G(x, y) > I_R(x, y), & (I) \\ I_G(x, y) > I_B(x, y), & (II) \\ |I_R(x, y) - R_{peak}| < R_t \times th_{expand}, & (III) \\ |I_G(x, y) - G_{peak}| < G_t \times th_{expand}, & (IV) \\ |I_B(x, y) - B_{peak}| < B_t \times th_{expand}, & (V) \\ GL(x, y) < GL_t \times th_{expand} & (VI) \end{cases} \quad (7-3)$$

برای حذف تابلوهای تبلیغاتی که به عنوان بازیکن شناخته شده‌اند، اندازه و موقعیت آن در تصویر مورد استفاده قرار گرفته است. تابلوهای تبلیغاتی در گوشه‌های تصویر ظاهر می‌گردند و در صحنه‌های گل بسته به موقعیت بازی در سمت چپ یا راست تصویر واقع می‌شوند (شناسایی جناح بازی بر اساس زاویه نشانه‌های خطی انجام‌پذیر است). همچنین معمولاً تابلوهای تبلیغاتی فضای وسیعی از تصویر را به خود اختصاص می‌دهند. از این رو، برای شناسایی تابلوهای تبلیغاتی در صورتی که بازی

در سمت چپ در حال گردش باشد، گوشه‌های بالایی و پایینی سمت چپ تصویر و در غیر این صورت گوشه‌های بالایی و پایینی سمت راست تصویر بررسی شده و در صورتی که نواحی منتخب بازیکن در این مناطق وجود داشته باشد، به طوری که مساحت زیادی را در بر گیرد، به عنوان تابلوهای تبلیغاتی در نظر گرفته می‌شود.

برای تشخیص آرم‌های تلویزیونی نیز از ثابت بودن موقعیت آن در تمامی فریم‌ها استفاده می‌شود. بدین منظور واریانس مقادیر رنگی پیکسل‌ها در بین چندین فریم با فواصل زمانی طولانی محاسبه و سپس نتایج، آستانه‌سازی می‌شوند به طوری که پیکسل‌های با واریانس پایین، مقدار یک و پیکسل‌های با واریانس بالا، مقدار صفر داشته باشند. واضح است که نواحی به دست آمده با مقدار یک، نشان‌دهنده تغییرات اندک در آن نقطه می‌باشد؛ به عبارت دیگر در تصویر همواره ثابت بوده‌اند. سرانجام، با به کارگیری فیلتر میانه، نویزهای موجود حذف و در نهایت نواحی باقیمانده به عنوان آرم‌های تلویزیونی شناخته می‌شوند.

در شکل (3-12) نتایج حاصل از شناسایی بازیکنان در 2 حالت اعمال اصلاحات فوق و بدون لحاظ کردن آن بر روی یک تصویر نمونه که ناحیه چمن آن به درستی استخراج نشده، آورده شده است. برای مشخص نمودن نوع اصلاح به کار رفته در حذف ناحیه، شماره اصلاح آن مورد استفاده قرار گرفته است. همان‌طور که مشاهده می‌شود، اعمال اصلاحات، کاهش خطا را به میزان قابل توجه به همراه داشته است. شکل (3-13)، نمونه‌ای دیگر از نتیجه اصلاحات معرفی شده بر روی یک تصویر واقعی را نشان می‌دهد. علاوه بر این، در شکل (3-14) نتایج حاصل از شناسایی بازیکنان برای تصاویر شکل (3-8)، با به کارگیری روش معرفی شده، نشان داده شده است.



(ب)



(الف)



(د)



(ج)

شکل 3-12- نتایج شناسایی بازیکنان با توجه به اصلاحات کاهش خطا: (الف) تصویر اصلی ب) چمن استخراج شده ج) بازیکنان شناخته شده بدون اعمال اصلاحات کاهش خطا د) بازیکنان شناخته شده با اعمال اصلاحات کاهش خطا - شماره‌های نشان داده شده در شکل (ج) شماره اصلاح به کار رفته برای حذف آن ناحیه را نشان می‌دهد. در شکل (د) خطا به میزان قابل توجهی در مقایسه با شکل (ج) کاهش یافته است.



(الف)



(ج)



(ب)

شکل 3-13- نتایج شناسایی بازیکنان با توجه به قوانین کاهش خطا: الف) تصویر اصلی
 ب) بازیکنان شناخته شده بدون اعمال قوانین کاهش خطا ج) بازیکنان شناخته شده با
 اعمال قوانین کاهش خطا - شماره‌های نشان داده شده در شکل (ب) شماره اصلاح به کار
 رفته برای حذف آن ناحیه را نشان می‌دهد. در شکل (ج) خطا به میزان قابل توجهی در
 مقایسه با شکل (ب) کاهش یافته است.



(الف)



(ب)



(ج)

شکل 3-14- نتایج حاصل از شناسایی بازیکنان برای تصاویر شکل (3-8) با به کارگیری روش معرفی شده

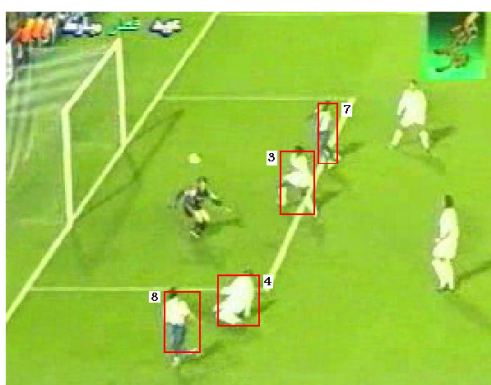
3-3- تعیین تیم بازیکنان

شکل (3-15) 2 تصویر واقعی با بازیکنان شماره گذاری شده و شکل (3-16) هیستوگرام‌های مولفه قرمز برای بازیکنان مشخص شده در این شکل را نشان می‌دهد. شایان ذکر است که هیستوگرام‌ها به صورت زیر از قبل بر حسب فراوانی نرمال شده‌اند:

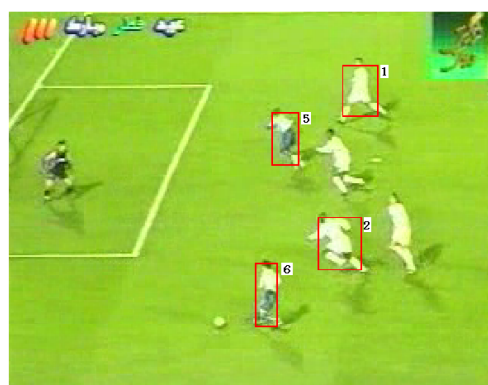
$$H(i) = H(i) \times \frac{C_{Normal}}{\sum_j H(j)} \quad \text{for } i = 1:256 \quad (8-3)$$

که H هیستوگرام، i اندیس هیستوگرام و C_{Normal} ضریب نرمال می‌باشد و در این پایان‌نامه برای آن مقدار 1000 اختیار شده است؛ به عبارت دیگر هیستوگرام به گونه‌ای نرمال شده است که مجموع مقادیر هیستوگرام برابر 1000 باشد.

در شکل (16-3) هیستوگرام‌های شکل‌های (16-3-الف) تا (16-3-الف-4) متعلق به یک تیم و هیستوگرام‌های شکل‌های (16-3-ب-1) تا (16-3-ب-4) متعلق به تیم دیگر است. همان‌طور که دیده می‌شود، شکل هیستوگرام‌های بازیکنان متعلق به یک تیم، با یکدیگر مشابه و با تیم دیگر متفاوت می‌باشد. بنابراین هیستوگرام بازیکنان می‌توان به عنوان معیاری مناسب برای تعیین تیم بازیکنان به کار گرفته شود.



(ب)



(الف)

شکل 3-15-2 تصویر واقعی با بازیکنان شماره گذاری شده

همان‌گونه که در فصل دوم اشاره شد، از جمله روش‌هایی که برای تعیین تیم بازیکنان به کار گرفته شده است، روش‌های معرفی شده توسط سویین و بالارد [26] یعنی "مقایسه توزیع هیستوگرام مشترک" و "مقایسه توزیع هیستوگرام باز تصویر شده" می‌باشد. با دقت بر روی هیستوگرام‌های نشان داده شده در شکل (16-3) می‌توان انتظار داشت که این روش‌ها می‌توانند با خطا مواجه شوند، زیرا با وجود آنکه شکل کلی هیستوگرام‌ها برای بازیکنان هر تیم مشابه است، ولی بر روی یکدیگر تطابق ندارند (مانند شکل (16-3-ب-3) و (16-3-ب-4)). از این رو، در این پایان‌نامه روش ماشین بردار پشتیبان¹ مبتنی بر ویژگی هیستوگرام مولفه‌های رنگی بازیکنان، مورد استفاده قرار گرفته است.

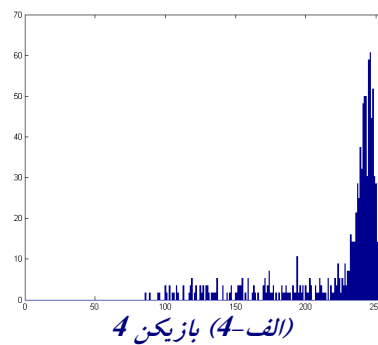
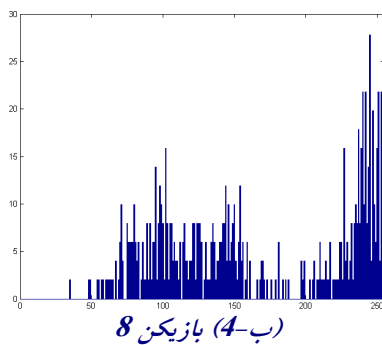
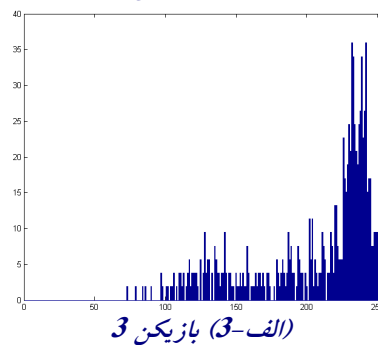
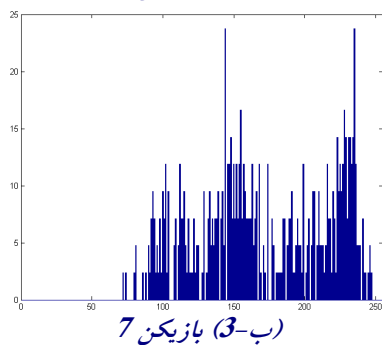
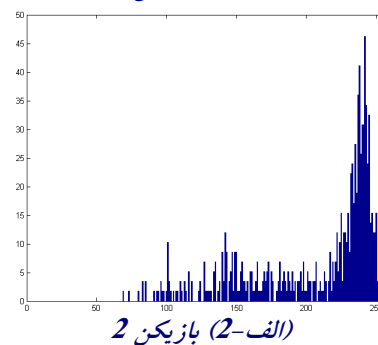
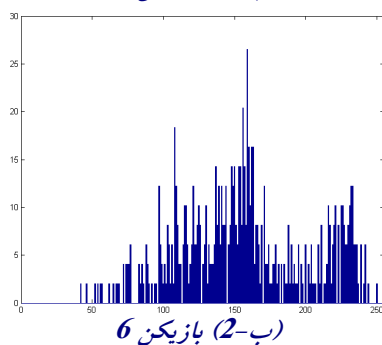
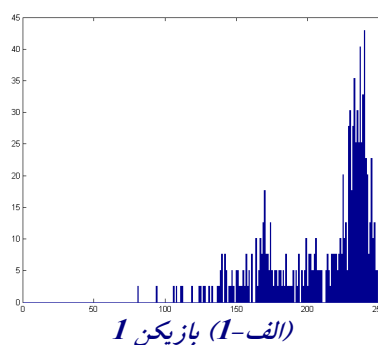
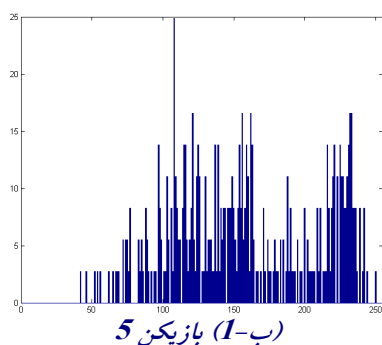
ماشین بردار پشتیبان یا SVM در انواع دسته‌بندی‌ها همچون تشخیص ارقام دستنویس، دسته‌بندی انواع صداها [30-32] و مانند آن مورد استفاده قرار گرفته است که در مقایسه با روش‌های دیگر از کارایی قابل ملاحظه‌ای برخوردار است. نتایج حاصل از به کارگیری این روش در کاربردهای متفاوت حاکی از کارایی بالای این روش است. ماشین بردار پشتیبان دارای قابلیت‌های سودمندی است که آن را نسبت به دیگر روش‌های موجود برتر ساخته است؛ به عنوان مثال ماشین بردار پشتیبان در دسته‌بندی نمونه‌ها از قابلیت تعمیم² برخوردار است و مشکل بهینه‌های محلی³ در مرحله آموزش را ندارد؛ به عبارت دیگر دسته‌بندی کننده را به صورت بهینه تعیین می‌نماید. همچنین توابع تمایز غیر خطی را به سادگی و با محاسبات کم، با استفاده از

¹ Support Vector Machine

² Generalization

³ Local Optima

هسته‌های^۱ غیرخطی و مفهوم حاصل ضرب داخلی^۲ در فضای هیلبرت^۳، تشکیل می‌دهد [30]. در بخش‌های بعدی انواع SVM یعنی SVM خطی جدایی‌پذیر، SVM خطی جدایی‌ناپذیر، و SVM غیرخطی مورد بررسی قرار می‌گیرد.



شکل 3-16- هیستوگرام‌های مولفه قرمز برای بازیکنان مشخص شده در شکل 3-15: (ب-1) تا (ب-4) متعلق به یک تیم و هیستوگرام‌های شکل‌های (الف-1) تا (الف-4) متعلق به یک تیم دیگر است. شکل هیستوگرام‌های بازیکنان متعلق به یک تیم با یکدیگر مشابه و با تیم دیگر متفاوت می‌باشد.

¹ Kernel

² Dot Product

³ Hilbert Space

3-3-1- ماشین بردار پشتیبان خطی جدایی پذیر

فرض می شود مجموعه نمونه های آموزشی $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ در دست است که $\bar{x}_i \in R^d$ و هر کدام از نمونه ها با یکی از برچسب های $y_1, y_2, \dots, y_m$ نسبت داده شده اند که $y_i \in \{-1, 1\}$. این برچسب ها نشان دهنده تعلق نمونه $\bar{x}_i$ به یکی از 2 کلاس می باشد. با این تعریف، ابر صفحه $\bar{w} \cdot \bar{x} + b$ داده ها را از یکدیگر متمایز می کند اگر و تنها اگر [34, 33]:

$$\begin{cases} \bar{w} \cdot \bar{x}_i + b > 0 & \text{if } y_i = 1 \\ \bar{w} \cdot \bar{x}_i + b < 0 & \text{if } y_i = -1 \end{cases} \quad (9-3)$$

حال می توان با مقیاس کردن بردار $\bar{w}$ و مقدار b به رابطه زیر رسید:

$$\begin{cases} \bar{w} \cdot \bar{x}_i + b \geq 1 & \text{if } y_i = 1 \\ \bar{w} \cdot \bar{x}_i + b \leq -1 & \text{if } y_i = -1 \end{cases} \quad (10-3)$$

و یا

$$y_i(\bar{w} \cdot \bar{x}_i + b) \geq 1 \quad (11-3)$$

برای پیدا کردن ابر صفحه متمایز کننده بهینه، لازم است صفحه ای که فاصله بین ابر صفحه و نزدیک ترین نمونه ها به این ابر صفحه را بیشینه می کند، پیدا نمود. فاصله نزدیک ترین نمونه برابر است با:

$$d(\bar{w}, b) = \min_{\{\bar{x}_i | y_i = 1\}} \frac{\bar{w} \cdot \bar{x}_i + b}{|\bar{w}|} - \max_{\{\bar{x}_i | y_i = -1\}} \frac{\bar{w} \cdot \bar{x}_i + b}{|\bar{w}|} \quad (12-3)$$

با توجه به رابطه (11-3) دیده می شود که مقدار بیشینه و کمینه مناسب به ترتیب برابر 1 و -1 می باشد. بدین ترتیب رابطه (12-3) به صورت زیر در می آید:

$$d(\bar{w}, b) = \frac{1}{|\bar{w}|} - \frac{-1}{|\bar{w}|} = \frac{2}{|\bar{w}|} \quad (13-3)$$

بنابراین مساله به کمینه کردن $\frac{|\bar{w}|^2}{2}$ با قید رابطه (11-3) تبدیل می شود. با تشکیل رابطه لاگرانژ و حل دوگان مساله¹، رابطه زیر که یک مساله بهینه سازی مقید² است، حاصل می شود:

$$\text{minimize } \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \bar{x}_i \bar{x}_j \quad (14-3)$$

با فرض آنکه:

$$\begin{cases} \alpha_i \geq 0 \\ \sum_i \alpha_i y_i = 0 \end{cases} \quad (15-3)$$

که α_i ضرایب لاگرانژ³ می باشند. برای هر نمونه آموزشی، یک ضریب لاگرانژ وجود دارد. نمونه های آموزشی با ضرایب لاگرانژ مخالف صفر، بردار پشتیبان⁴ نامیده می شوند. شایان ذکر است، نمونه هایی با ضرایب لاگرانژ صفر می توانند بدون آن که موقعیت ابر صفحه نهایی را تحت تاثیر قرار دهند از مجموعه آموزشی حذف شوند.

¹ Dual Problem

² Constrained Optimization

³ Lagrange Multipliers

⁴ Support Vectors

3-3-2- ماشین بردار پشتیبان خطی جدایی ناپذیر

در عمل تقریباً تمامی مسایل جدایی پذیر نیستند و نمی توان همواره روش قبلی را برای حل مسایل مورد استفاده قرار داد. به همین جهت رابطه (3-10) به صورتی اصلاح می شود که قیدها ضعیف تر شده و به عبارت دیگر یک ریسک¹ برای دسته بندی غلط لحاظ شود:

$$\begin{cases} \bar{w} \cdot \bar{x}_i + b \geq 1 - \xi_i & \text{if } y_i = 1 \\ \bar{w} \cdot \bar{x}_i + b \leq \xi_i - 1 & \text{if } y_i = -1 \\ \xi_i \geq 0 & \forall i \end{cases} \quad (16-3)$$

در صورتی که $\bar{x}_i$ به اشتباه دسته بندی شود باید رابطه $\xi_i > 1$ برقرار باشد؛ لذا تعداد خطاها همواره کمتر از $\sum_i \xi_i$ می باشد. بنابراین می توان با اضافه کردن یک مقدار ریسک به نمونه های آموزشی به تابع هدف زیر رسید:

$$\text{minimize } \frac{|\bar{w}|^2}{2} + C(\sum_i \xi_i) \quad (17-3)$$

که C ضریبی است که برای دقت برازش داده های آموزشی به کار می رود. حال می توان دوگان لاگرانژ را با رابطه زیر بیان کرد:

$$\text{minimize } \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \bar{x}_i \bar{x}_j \quad (18-3)$$

با فرض آنکه:

$$\begin{cases} 0 \leq \alpha_i \leq C \\ \sum_i \alpha_i y_i = 0 \end{cases} \quad (19-3)$$

3-3-3- ماشین بردار پشتیبان غیر خطی

ماشین های بردار پشتیبان ذکر شده در قسمت های قبلی برای دسته بندی الگوهای یک مساله 2 کلاسه، مرزهای جدا کننده خطی و یک ابر صفحه را مورد استفاده قرار می دهند. با این وجود، می توان مسایلی که به صورت خطی جدایی پذیر نیستند را با نگاشت نقاط نمونه ها به فضایی با بعد بیشتر با استفاده از یک نگاشت $\phi: R^d \rightarrow H$ یعنی یک نگاشت انتخاب شود به گونه ای که بعد H از d بیشتر باشد. سپس ابر صفحه ای متمایز کننده در این فضا یافت می شود که معادل سطح غیر خطی متمایز کننده در فضای R^d است.

با یافتن ابر صفحه متمایز کننده، داده های آموزشی در رابطه (3-18) که به صورت حاصل ضرب داخلی ظاهر شده اند به شکل $\phi(\bar{x}_i) \cdot \phi(\bar{x}_j)$ تبدیل می شوند. با این وجود، با تعریف یک تابع هسته به صورت $k(\bar{x}_i, \bar{x}_j) = \phi(\bar{x}_i) \cdot \phi(\bar{x}_j)$ و استفاده از آن به جای $\bar{x}_i \cdot \bar{x}_j$ در مساله بهینه سازی مورد نظر، دیگر نیازی به دانستن ϕ نمی باشد. نمونه هایی از توابع هسته که به دفعات مورد استفاده قرار می گیرند، عبارتند از:

$$k(\bar{x}_i, \bar{x}_j) = (\bar{x}_i \cdot \bar{x}_j + 1)^p \quad (20-3)$$

¹ Penalty

$$k(\bar{x}_i, \bar{x}_j) = \exp(-g \|\bar{x}_i - \bar{x}_j\|^2) \quad (21-3)$$

$$k(\bar{x}_i, \bar{x}_j) = \tanh(\bar{x}_i \cdot \bar{x}_j + \theta) \quad (22-3)$$

که به ترتیب هسته چند جمله‌ای^۱، هسته گاوسی^۲، و هسته تانژانت هایپربولیک^۳ نامیده می‌شوند. مساله دسته‌بندی با رابطه دوگان لاگرانژ در حالت جدایی‌ناپذیر و غیر خطی به صورت زیر خواهد بود:

$$\text{minimize } \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(\bar{x}_i, \bar{x}_j) \quad (23-3)$$

با فرض آنکه:

$$\begin{cases} 0 \leq \alpha_i \leq C \\ \sum_i \alpha_i y_i = 0 \end{cases} \quad (24-3)$$

3-4- ماشین بردار پشتیبان در تعیین تیم بازیکنان

همان‌گونه که بیان شد، ویژگی به کار رفته برای تعیین تیم بازیکنان در این پایان‌نامه، ویژگی هیستوگرام مولفه‌های رنگی می‌باشد. به عبارت دیگر هر نمونه آموزشی و آزمایشی در ماشین بردار پشتیبان شامل 256 عنصر برای هر مولفه رنگی یعنی 768 عنصر می‌باشد. واضح است با زیاد شدن تعداد نمونه‌های آموزشی، دقت سیستم بالاتر می‌رود. همچنین تعداد کلاس‌های در نظر گرفته شده 4 طبقه به ازای تیم اول، تیم دوم، دروازه‌بان و داور می‌باشد. با وجود آنکه در ابتدا، ماشین بردار پشتیبان برای مسایل 2 کلاسه طراحی شده است، ولی با تعریف مجدد توابع هدف و هزینه می‌توان در حالت چند کلاسه نیز عمل دسته‌بندی را به خوبی انجام داد [34]. نتایج به دست آمده با آزمایش‌های مختلف به صورت خطی و غیر خطی و با ضرایب ریسک متفاوت در فصل 6 آورده شده است. این نتایج بیان‌گر قابلیت بالای SVM در دسته‌بندی کلاس بازیکنان می‌باشد.

3-4- جمع‌بندی

در این فصل، مراحل لازم برای استخراج بازیکنان و تعیین تیم آن‌ها با به کارگیری روش‌های پیشنهادی مورد بحث قرار گرفت. ابتدا زمین چمن با فرض غالب بودن آن در فضای تصویر و با به کارگیری هیستوگرام مولفه‌های رنگی در فضای RGB شناسایی گردید. در روش ارایه شده بر خلاف روش یون، پارامترهای شناسایی چمن به طور بهینه‌تر و در عین حال با وابستگی کمتر به تصاویر ویدیویی و شرایط مختلف زمین چمن به دست آمدند. سپس روشی نوین مبتنی بر فیلتر میانه و عملگر مورفولوژی باز کردن برای شناسایی بازیکنان معرفی شد. علاوه بر این، با اعمال اصلاحات به منظور کاهش خطا در مرحله شناسایی بازیکنان، نتایج بهبود داده شد. در پایان، برای تعیین تیم بازیکنان، روش نوین ماشین بردار پشتیبان که از جمله روش‌های مناسب در دسته‌بندی الگو به شمار می‌رود، ارایه شد.

¹ Polynomial Kernel

² Guassian Kernel

³ Tangent Hyperbolic Kernel

فصل چهارم

ردیابی بازیکنان

هم زمان با شکل گیری سیستم های مبتنی بر بینایی ماشین، ردیابی اشیاء متحرک در دنباله تصویر ویدیویی، جزء زمینه های مهم و کاربردی در مبحث بینایی ماشین شناخته شد. موارد کاربردی ردیابی اشیاء متحرک بسیار متنوع و از آن جمله می توان به کاربرد آن در سیستم های خودکار هدایت کننده اتومبیل [35]، سیستم های خودکار کنترل کننده ترافیک [36]، ردیابی هدف [37]، ردیابی اعضای بدن [38] و تحلیل حرکات ورزشی اشاره نمود.

به منظور ردیابی موفق بازیکنان در دنباله تصویر ویدیویی مسابقات فوتبال، پیشنهاد می گردد اطلاعات پیش زمینه درباره حوزه مورد بحث جهت اعمال محدودیت ها و انتخاب ویژگی های مناسب به کار گرفته شود؛ زیرا ردیابی اشیاء کوچک و غیر صلب در محیطی پیچیده¹ و پویا² تنها متکی بر ردیابی سطح پایین شدت روشنایی می تواند با خطا همراه باشد. با توجه به اینکه اشیاء و رویدادهای مسابقه فوتبال محدود به یک محیط بسته با ابعاد مشخص است، می توان عمل ردیابی سطح پایین را با در نظر گرفتن خصوصیات مدل میدان بازی مانند موقعیت های مکانی نشانه های خطی نسبت به مدل میدان، محدودتر و در نتیجه با دقت بیشتری انجام داد.

در مطالعات پیشین تنها اطلاعات رنگی لباس بازیکنان برای ردیابی بازیکنان مورد استفاده قرار گرفته است [2,3,7,12,21] و در برخی موارد اطلاعات حرکتی بازیکنان با فرض ثابت بودن سرعت حرکت بازیکنان در یک بازه زمانی کوتاه، جهت مشخص نمودن ناحیه جستجوی بازیکن در فریم های متوالی به کار گرفته می شود [5,20,23,39]. با توجه به اینکه بازیکنان در صحنه های گل به طور ناگهانی سرعت و جهت حرکت خود را تغییر می دهند، به کارگیری روش های قبلی در چنین شرایطی با شکست مواجه خواهد شد.

در این پایان نامه روشی جدید و کارآمد متکی بر دانش حاکم بر خصوصیات مکانی نشانه های خطی موجود در مدل میدان بازی، برای ردیابی بازیکنان معرفی گردیده است. در این روش موقعیت متناظر بازیکنان در فریم قبلی نسبت به مختصات زمین فوتبال مبنای عملیات ردیابی بازیکن در فریم فعلی قرار می گیرد.

از آنجا که معمولاً فیلم برداران در صحنه های گل برای ردیابی توپ مجبور به حرکت سریع دوربین می باشند، موقعیت بازیکنان در تصویر در بین 2 فریم متوالی می تواند با تغییرات زیادی مواجه شود. با این وجود، تغییر موقعیت بازیکنان در مدل میدان در بین 2 فریم متوالی چندان مشهود نخواهد بود، زیرا معمولاً نرخ نمونه برداری برای تصاویر ویدیویی بین 20 الی 30

¹ Complex

² Dynamic

فریم در ثانیه است و در نتیجه موقعیت بازیکن در $1/30$ الی $1/20$ ثانیه نمی تواند تغییر چندانی داشته باشد. نتایج به دست آمده در ردیابی بازیکنان برای دنباله تصویر ویدیویی که با حرکت سریع دوربین همراه است، دلیل دیگری بر کارا بودن روش ارایه شده می باشد. در این فصل ابتدا مشکلات ردیابی در تصاویر ویدیویی مسابقات فوتبال بیان می گردد. سپس روش های عمده برای ردیابی اشیاء متحرک تشریح و مزایا و معایب به کارگیری آن ها در تصاویر ویدیویی صحنه های گل تشریح می شود. در پایان به جزییات روش پیشنهادی برای ردیابی بازیکنان یعنی ردیابی مبتنی بر مدل میدان پرداخته می شود.

4-1- مشکلات ردیابی

تصاویر ویدیویی مسابقه فوتبال در هر صحنه از یک دوربین ثابت و بر روی حرکت بازیکنان، داور، توپ و دیگر اشیاء موجود در ورزشگاه می باشد. این تصاویر ویدیویی معمولاً با دقت¹ پایین و همراه با فعل و انفعالات دوربین گرفته شده اند. اشیاء موجود داخل زمین در تصاویر ویدیویی گرفته شده کوچک و غیر صلب بوده و تعداد زیاد اشیاء و حرکت سریع و غیر قابل پیش بینی² آن ها محیطی پویا و پیچیده را ایجاد می کند. علاوه بر این، بازیکنان همواره با یکدیگر تصادم و هم پوشانی دارند که ردیابی بازیکنان را دشوارتر می سازد. در ادامه به بررسی بیشتر مشکلات یاد شده، پرداخته می شود.

4-1-1- پویا و پیچیده بودن محیط

چون معمولاً در صحنه های گل تجمع بازیکنان زیاد است، تصاویر ویدیویی چنین صحنه هایی با اشیاء متعددی همراه است. علاوه بر این، وجود اشیاء دیگر مانند توپ، نشانه های خطی، دروازه، تابلوهای تبلیغاتی و تماشاچیان، محیط را پیچیده تر می کند. بازیکنان که اشیاء اصلی این محیط می باشند، همواره با تغییر شکل و جابجایی مکانی مواجه هستند. از طرف دیگر، فعل و انفعالات دوربین باعث تغییرات دایمی صحنه قابل رویت می شود. این مسایل باعث پویا و پیچیده شدن دنباله تصویر ویدیویی می شود و در نتیجه ردیابی اشیاء در چنین صحنه هایی به سادگی امکان پذیر نخواهد بود.

4-1-2- فعل و انفعالات دوربین

دوربین می تواند به صورت ثابت یا همراه با فعل و انفعال باشد. فعل و انفعالات دوربین به 2 دسته تقسیم می شوند: الف) بدون تغییر موقعیت دوربین و ب) همراه با تغییر موقعیت دوربین. در نوع اول، 3 حالت چرخش افقی³، چرخش عمودی⁴ و تغییر طول کانونی دوربین یا بزرگ نمایی⁵ امکان پذیر است. نوع دوم نیز شامل انتقال های افقی⁶، عمودی⁷ و افقی-جانبی⁸ می باشد.

¹ Resolution

² Erratic

³ Panning

⁴ Tilting

⁵ Zooming

⁶ Horizontal Translation

⁷ Vertical Translation

⁸ Horizontal Lateral Movement

تصاویر ویدیویی مورد استفاده در این پایان‌نامه از کنار میدان بازی گرفته شده است و همراه با فعل و انفعالات دوربین به صورت چرخش افقی، چرخش عمودی و بزرگ‌نمایی می‌باشد. از آنجایی که صحنه‌های گل معمولاً با حرکات سریع توپ و بازیکنان به طرف دروازه و تغییر موقعیت بازی به صورت ناگهانی همراه است، کارگردان تلویزیونی در برخی مواقع، مجبور به چرخش سریع دوربین به همراه بزرگ‌نمایی آنی می‌شود که این امر ردیابی بازیکنان را با مشکل مواجه می‌سازد. شکل (4-1) چندین فریم از یک صحنه گل را نمایش می‌دهد. همان‌طور که در شکل دیده می‌شود در این صحنه تغییرات موقعیت نمایش و حرکات سریع دوربین در بین فریم‌ها کاملاً آشکار است.

به منظور ردیابی دقیق بازیکنان فوتبال لازم است حرکت دوربین شناخته شود. رویه ردیابی یا باید حرکت دوربین را در هر فریم در نظر بگیرد و یا اینکه حرکت دوربین، قبل از شروع به ردیابی حذف گردد؛ زیرا تخمین موقعیت بازیکنان تنها در صورتی کارآمد است که حرکت دوربین شناخته شود. همان‌طور که اشاره شد در این پایان‌نامه ویژگی مدل میدان یا به عبارت دیگر موقعیت متناظر نشانه‌های خطی در تصویر و مدل میدان برای در نظر گرفتن حرکت دوربین مورد استفاده قرار گرفته است.

4-1-3- حرکت سریع و غیر قابل پیش‌بینی بازیکنان

در صحنه‌های گل، بازیکنان با سرعت زیاد در حال دویدن هستند و در برخی لحظات به طور غیر قابل پیش‌بینی مسیر خود را عوض می‌کنند. در اغلب الگوریتم‌های ردیابی فرض بر آن است که حرکت اشیاء یکنواخت و در یک مسیر است. علاوه بر این، حرکت بازیکنان با حرکت دوربین ترکیب شده و در نتیجه انتخاب یک نقطه مرجع از بازیکن برای محاسبه سرعت امکان‌پذیر نمی‌باشد. شکل (4-2) دنباله‌ای از تصاویر ویدیویی از دویدن یک بازیکن را با نرخ نمونه‌برداری 30 فریم در ثانیه نشان می‌دهد. همان‌گونه که در این شکل دیده می‌شود بازیکن در هر فریم با تغییر شکل و جابجایی نسبت به فریم قبلی همراه است که در برخی از فریم‌ها این تغییرات نمایان‌تر است.

4-1-4- دقت پایین تصاویر ویدیویی

متأسفانه تصاویر ویدیویی به کار گرفته شده در این پایان‌نامه کیفیت مناسبی نداشته‌اند. همچنین صحنه‌های در نظر گرفته شده، صحنه‌های نمای دور می‌باشند؛ به عبارت دیگر تصاویر ویدیویی به صورت بزرگ‌نمایی بوده به طوری که فضای وسیعی از میدان بازی را می‌پوشانند. از این رو، معمولاً اندازه بازیکنان در دنباله تصویر ویدیویی در نظر گرفته شده بسیار کوچک است. علاوه بر این، در مواقع حرکت سریع دوربین، ممکن است بازیکنان به صورت مات در تصویر ظاهر شوند. نمونه‌هایی از تصاویر چندین بازیکن در حالت‌ها و موقعیت‌های مختلف در شکل (4-3) نشان داده شده است. همان‌طور که مشاهده می‌شود، معمولاً تصویر بازیکنان دقت پایین و شکل پیچیده‌ای دارد؛ به طوری که مدل کردن آن‌ها امکان‌پذیر نمی‌باشد. همچنین در برخی از این تصاویر به علت مات شدگی، بازیکنان به درستی قابل تشخیص نمی‌باشند. واضح است روش‌های پردازش تصویر که نیازمند لبه‌هایی با تباين¹ بالا یا دقت مناسب می‌باشند، در این گونه تصاویر کاربرد ندارند.

¹ Contrast



فریم 45



فریم 25



فریم 50



فریم 30



فریم 55



فریم 35

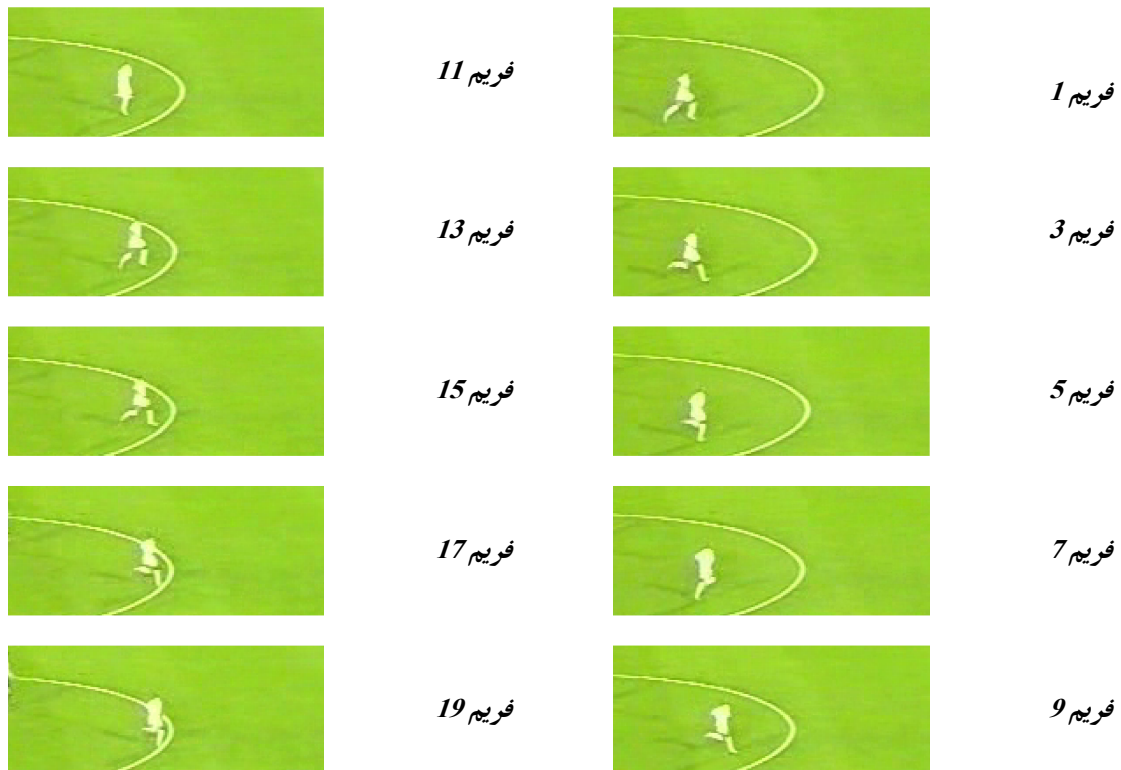


فریم 60



فریم 40

شکل 4-1- چندین فریم از یک صحنه گل همراه با چرخش افقی، چرخش عمودی و بزرگ‌نمایی دوربین: تغییرات موقعیت نمایش و حرکات سریع دوربین در بین فریم‌ها کاملاً آشکار است.



شکل 4-2- دنباله‌ای از تصاویر ویدیویی از دویدن یک بازیکن با نرخ نمونه‌برداری 30 فریم در ثانیه: موقعیت و شکل بازیکن در بین فریم‌ها به سرعت در حال تغییر است.



شکل 4-3- نمونه‌هایی از تصاویر چندین بازیکن در حالت‌ها و موقعیت‌های مختلف: معمولاً تصویر بازیکنان دقت پایین و شکل پیچیده‌ای دارد و در برخی از تصاویر به علت مات شدگی، بازیکنان به درستی قابل تشخیص نمی‌باشند.

4-1-5- غیر صلب بودن بازیکنان

بازیکنان فوتبال به صورت غیر صلب می‌باشند، به خصوص در هنگام دویدن که پاها و دستهای بازیکن به طور مرتب در حال حرکت است. غیر صلب بودن بازیکنان به همراه حرکت سریع آنها، حرکت دوربین، تصاویر بزرگ‌نمایی شده و دقت پایین تصاویر ویدیویی، بازیکنان را به صورت یک حباب در حال حرکت نشان می‌دهد. همان‌طور که در شکل (4-3) دیده می‌شود، شکل بازیکنان در حالت‌های مختلف با تغییرات محسوس همراه است، به طوری که ویژگی‌های فردی بازیکنان مانند شماره آنها قابل تشخیص نمی‌باشد.

در شکل (4-2) نیز تغییرات شکل یک بازیکن دوندۀ در فریم‌های متوالی نشان داده شده است. واضح است در نظر گرفتن مدل‌های ساده برای شکل بازیکن (به عنوان مثال استفاده از کلیشه همبستگی ثابت یا مدل‌های هندسی) عملی نیست. علاوه بر این، مدل‌های حرکتی ساده نمی‌توانند حرکات پیچیده دست و پا را مدل کنند و این در حالی است که استفاده از مدل‌های حرکتی پیچیده به دلیل کافی نبودن پیکسل‌های یک بازیکن مقدور نیست.

4-1-6- هم‌پوشانی

هم‌پوشانی بازیکنان فوتبال در مسابقات فوتبال و به خصوص در صحنه‌های گل به دفعات روی می‌دهد. هم‌پوشانی می‌تواند بین 2 یا چند بازیکن و همچنین بین اعضای یک تیم یا اعضای تیم‌های متقابل رخ دهد. از بین تمامی مشکلات یاد شده برای ردیابی بازیکنان، مساله هم‌پوشانی جدی‌تر و در اغلب موارد مساله‌سازتر می‌باشد. بنابراین در این پایان‌نامه با در نظر داشتن هم‌پوشانی حداکثر بین 3 بازیکن، برای حالت‌های مختلف هم‌پوشانی روش‌های متفاوتی معرفی شده است که در ادامه به جزئیات آنها پرداخته می‌شود.

4-2- انواع روش‌های ردیابی

وب¹ و اگرآوال² [40] روش‌های ردیابی را به 2 دسته مبتنی بر مدل‌های آیکونی³ یا کلیشه‌های همبستگی و مبتنی بر مدل‌های ساختاری یا مدل‌های ویژگی طبقه‌بندی کرده‌اند. کلیشه همبستگی قسمتی از فریم را با قسمتی از فریم دیگر تطابق می‌دهد. مدل ساختاری نیز ویژگی‌های خاص (مانند خط) را بین 2 فریم با استفاده از مشخصه‌های آن (مانند طول و جهت) تطابق می‌دهد. 2 روش عمده دیگر عبارتند از روش‌های مبتنی بر کلیشه‌های فرم‌پذیر⁴ که بر اساس مدل‌های انرژی نسبت به زمان تغییر می‌یابند و ردیاب‌های حرکتی که شار نوری ناحیه‌ای از تصویر را به منظور پیش‌بینی موقعیت جدیدش تخمین می‌زنند.

¹ Webb

² Aggrawal

³ Iconic Model

⁴ Deformable Templates

4-2-1- ردیابی کلیشه همبستگی

در کلیشه همبستگی استاندارد، تکه‌ای از تصویر با الگوی کوچکی مقایسه می‌شود. معمولاً برای راحتی کار، این ناحیه را به صورت مستطیلی شکل در نظر می‌گیرند و الگوی اولیه را نیز به طور مستقیم از اشیاء شناخته شده در اولین تصویر استخراج می‌کنند. همبستگی استاندارد به اندازه الگو، شدت روشنایی و مقدار تباین تصویر حساس است. برای رفع این مشکل، همبستگی نرمالیزه که از تقسیم مقدار همبستگی بر انحراف معیار آن ناحیه به دست می‌آید، به کار گرفته می‌شود. این همبستگی نیز همچنان به نسبت سیگنال به نویز¹ حساس خواهد بود، ولی اندازه پنجره و مقدار میانگین شدت روشنایی و تباین تصویر نرمالیزه خواهد شد.

کلیشه‌های همبستگی به هم‌پوشانی و تغییرات کوچک در تکه‌ای از تصویر که باید در فریم‌های بعدی ردیابی شود، بسیار حساس است. این حساسیت در صحنه‌هایی که تغییر شکل اشیاء به صورت پویا می‌باشد و ممکن است با یکدیگر برخورد یا هم‌پوشانی داشته باشند، مساله برانگیز خواهد بود. ردیابی کلیشه در صورتی که حرکت اشیاء یا حرکت دوربین باعث تغییرات در شکل شیء ردیابی شونده گردد، با شکست مواجه خواهد شد. روشی که در این حالت مورد استفاده قرار می‌گیرد، استفاده از کلیشه همبستگی تطبیقی خواهد بود. در این روش هر بار که کلیشه با تصویر بعدی مطابقت داده شد، کلیشه جدیدی که در مرکز موقعیت تطبیقی فعلی واقع می‌گردد، استخراج می‌شود. بنابراین با استفاده از این روش، کلیشه می‌تواند تدریجاً به گونه‌ای تنظیم شود که نشان‌دهنده نواحی تصویر متغیر با زمان باشد.

روش کلیشه همبستگی تطبیقی، در صورتی که تغییر شکل اشیاء بسیار سریع باشد، دچار مشکل می‌شود. علاوه بر این، به دلیل عدم وجود مدل² در روش کلیشه تطبیقی، احتمال انحراف تدریجی³ وجود دارد. چون در هر مرحله کلیشه تطبیق داده می‌شود، خطاهای کوچک تطبیق در کلیشه انباشته شده و در نهایت اگر محدودیت‌هایی برای این روش در نظر گرفته نشود، ممکن است کلیشه از ردیابی هدف خود منحرف شده و به ردیابی ناحیه‌ای کاملاً متفاوت بپردازد.

در حوزه فوتبال، همان‌طور که در بخش قبلی نشان داده شد، تغییر شکل بازیکنان از فریمی به فریم دیگر بسیار زیاد است. همچنین تنوع موقعیت بدنی بازیکنان در مسابقه فوتبال و به خصوص در صحنه‌های گل به دفعات دیده می‌شود. از طرف دیگر بزرگ‌نمایی دوربین و احتمال هم‌پوشانی بازیکنان، استفاده از روش کلیشه همبستگی را بیش از پیش دشوار می‌کند.

در صورت استفاده از کلیشه همبستگی تطبیقی در حوزه فوتبال، در هنگام عبور بازیکن از روی نشانه‌های خطی، احتمال انحراف تدریجی کلیشه وجود دارد؛ چرا که در هر فریم کلیشه همبستگی از روی نواحی استخراج شده برای بازیکن ساخته می‌شود و همان‌طور که در فصل قبلی توضیح داده شد، بازیکن از روی نواحی غیر چمن استخراج می‌گردد. چون نشانه‌های خطی در کنار بازیکن نیز به عنوان نواحی غیر چمن شناسایی می‌گردند، قسمتی از نشانه‌های خطی به عنوان لباس بازیکن و در نتیجه به عنوان قسمتی از کلیشه شناخته می‌شود. شکل (4-4) یک بازیکن (شکل (4-4-الف)) و کلیشه تخصیص داده شده (شکل (4-4-ب)) را نشان می‌دهد. در شکل (4-4-ب) پیکسل‌های رنگی، کلیشه بازیکن را نشان می‌دهد. همان‌طور که دیده می‌شود، کلیشه با ناحیه نسبتاً بزرگی از پیکسل‌های اشتباه پر شده است و این موضوع می‌تواند ردیابی بازیکن را به تدریج از هدف منحرف و به ردیابی نشانه‌های خطی بپردازد.

¹ Signal-to-Noise Ratio² Model-Free Method³ Drift



(ب) کلیشه بازیکن



(الف) بازیکن در تصویر اصلی

شکل 4-4- کلیشه یک بازیکن واقع بر روی نشانه‌های خطی: پیکسل‌های رنگی، کلیشه بازیکن را نشان می‌دهند. قسمتی از نشانه‌های خطی به اشتباه به عنوان بازیکن شناخته شده است.

4-2- ردیابی حرکت¹

اغلب الگوریتم‌های ردیابی، یک مدل حرکتی را مکمل ردیابی داده‌های خام تصویری قرار می‌دهند. بیشتر مدل‌های حرکتی معمول، فرض می‌کنند شیء با سرعتی ثابت و تغییراتی هموار حرکت می‌کند و به ندرت شتاب مدل می‌شود، زیرا داده‌های نویزی تصویر و نرخ پایین نمونه برداری مکانی و زمانی، تخمین شتاب را دشوار می‌سازد [41].

سیستم‌های دیگری نیز وجود دارند که الگوریتم اختلاف حرکت شار نوری² را برای پیش‌بینی موقعیت شیء متحرک مورد استفاده قرار می‌دهند، بدون آنکه به طور مستقیم از تطبیق ویژگی‌های تصویر استفاده نمایند. الگوریتم‌های حرکت شار نوری در هنگام هم‌پوشانی با شکست مواجه می‌شوند، زیرا این الگوریتم‌ها فرض می‌کنند تغییرات پیکسل‌ها بین 2 فریم، هموار و کوچک است. اشیاء هم‌پوشان شده و برخورد داده شده با دقت پایین و حباب مانند مدل‌های ضعیفی هستند، زیرا 2 قید اصلی شار نوری یعنی حرکت هموار³ و مسطح⁴ را نقض می‌کنند.

فرضیات حرکات هموار و کوچک و صلبی بودن شیء در حوزه فوتبال رعایت نمی‌شود؛ لذا ردیابی حرکتی در این حوزه به درستی قابل اجرا نمی‌باشد. علاوه بر این، دقت پایین و حرکات‌های داخلی بازیکنان مانند حرکت دست و پای بازیکن در حین حرکت باعث مشکل‌تر شدن این مساله می‌شود.

برای مشکل صلبی نبودن اشیاء شاید بتوان با میانگین‌گیری تمامی بردارهای حرکتی حاصل از اندام بازیکن، بردار حرکتی دقیق شیء را محاسبه کرد. هر چند دست و پای بازیکن در هنگام دویدن مخالف بدن وی حرکت می‌کند، ولی میانگین سرعت تقریباً دقیق است. با این وجود، این روش با مشکلی همانند مساله انحراف تدریجی در روش همبستگی تطبیقی مواجه می‌شود. الگوریتم برای انتخاب صحیح ناحیه‌ای که بردارهای حرکتی باید میانگین‌گیری شوند، بدون داشتن ایده‌ای مناسب از موقعیت بازیکن، راهی نخواهد داشت. چون به طور معمول بازیابی حرکت در مواقع هم‌پوشانی غیر دقیق است، میانگین‌گیری بر روی این ناحیه اشتباه، باعث نادرستی محاسبه میانگین بردار حرکت می‌شود.

4-3- ردیابی با استفاده از کلیشه فرم‌پذیر

برای ایجاد کلیشه ردیابی با تحمل‌پذیری بالاتر در برابر اعوجاجات هندسی بین فریم‌ها و تغییر شکل شیء نسبت به کلیشه‌های همبستگی، کلیشه‌های فرم‌پذیر که می‌توانند تغییرات کوچک و محدود را با گذشت زمان تحمل کنند، معرفی

¹ Motion Tracking

² Differential Optical Flow Motion

³ Smoothness

⁴ Planar

شده‌اند. یک مدل پارامتری اضافی نیز که معمولاً آفین^۱ است، می‌تواند برای پیش‌بینی چگونگی تغییر شکل کلیشه شیء متحرک در نظر گرفته شود [41].

کلیشه‌های فرم‌پذیر بر مبنای مدل‌های پیرامون فعال یا انرژی اسنیک^۲ بنا شده‌اند و به گونه‌ای طراحی می‌شوند که در صورت اعوجاجات قابل پیش‌بینی تکه‌ای از شیء ردیابی شونده، تحمل پذیر باشند. کلید موفقیت سیستم، انتخاب یک مدل انرژی مناسب است.

مدل پیرامون فعال در ابتدا توسط کاس^۳ و همکارانش [42] معرفی شد. یک اسنیک به صورت یک منحنی پارامتری تعریف می‌گردد:

$$v(t) = [x(t), y(t)] \quad \text{for } t \in [0 \ 1] \quad (1-4)$$

این منحنی پارامتری به صورت تکراری تغییر شکل می‌یابد تا اینکه تابع انرژی را به حداقل برساند. برای هر نقطه v_i از اسنیک، تابع انرژی $E_{snake}(n_i)$ برای تمامی نقاط n_i در همسایگی v_i محاسبه و در صورتی که رابطه $E_{snake}(n_i) < E_{snake}(v_i)$ صحیح باشد، n_i با v_i جایگزین می‌شود. این چرخه تا همگرایی کامل (یکی بودن اسنیک فعلی با اسنیک قبلی) ادامه می‌یابد. تابع انرژی به طور کلاسیک از 2 قسمت انرژی داخلی و انرژی خارجی همانند رابطه زیر تشکیل می‌شود:

$$E_{snake} = \alpha \cdot E_{int} + \beta \cdot E_{ext} \quad (2-4)$$

که E_{int} و E_{ext} به ترتیب انرژی داخلی و انرژی خارجی می‌باشند. انرژی داخلی، خصوصیات فیزیکی پیرامون و انرژی خارجی، خصوصیات تصویر مانند گرادیان یا شدت روشنایی را شامل می‌شود.

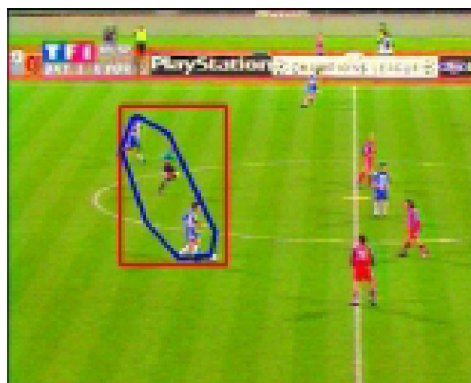
در حوزه فوتبال، وجود سایه بازیکنان که باعث ناپیوستگی و برجستگی شکل می‌شود، همراه با حرکات سریع و پیچیده دست و پای بازیکن باعث دشوارتر شدن محاسبه انرژی مناسب می‌شود. به طور یقین، مدل‌های ساده‌ای مانند آفین در این خصوص دچار شکست می‌شوند. علاوه بر این، حرکت دوربین و حرکت سریع و ناشناخته بازیکنان، استفاده از این مدل را مشکل‌تر می‌سازد. از طرف دیگر، مشکل اصلی اسنیک‌ها وابستگی زیاد آن‌ها به مقداردهی اولیه می‌باشد [21].

تاکنون روشی مبتنی بر کلیشه‌های فرم‌پذیر که بتواند مشکل هم‌پوشانی را حل کند، معرفی نشده است. در سیستمی که توسط لفوره و همکارانش [14] با استفاده از مدل پیرامون فعال معرفی شد، نزدیک بودن بازیکنان حتی بدون هم‌پوشانی و وجود نشانه‌های خطی روی زمین باعث اختلال در ردیابی بازیکنان می‌شود. شکل (4-5) نمونه‌ای از خروجی سیستم لفوره را در هنگام نزدیک بودن بازیکنان نشان می‌دهد. همان‌طور که دیده می‌شود بازیکنان نزدیک به هم، حتی بدون هم‌پوشانی به عنوان یک بازیکن شناخته شده‌اند.

¹ Affine

² Snake

³ Kass



شکل 4-5- نتایج ردیابی سیستم لافوره مبتنی بر مدل پیرامون فعال [12]: در این سیستم بازیکنان نزدیک به هم، حتی بدون هم پوشانی به عنوان یک بازیکن شناخته شده اند.

4-2-4- ردیابی ساختاری

همبستگی روش ساده ای برای مشخص کردن الگوها از طریق شدت روشنایی تصویر می باشد. بر خلاف روش همبستگی که به طور مستقیم مقادیر پیکسل ها را به کار می برد، در روش ردیابی ساختاری ابتدا بر روی تصویر پیش پردازش اعمال شده و سپس ویژگی های تحمل پذیر¹ و تغییرناپذیر² متناظر بین فریم ها به عنوان الگوی ردیابی مورد استفاده قرار می گیرد [41]. یکی از پرکاربردترین ویژگی های به کار رفته، لبه اشیاء می باشد. بدین منظور تصویر را ابتدا با استفاده از یک لبه یاب مانند لبه یاب عبور از صفر³، لبه یاب کنی⁴ و غیره پردازش و آنگاه لبه های استخراج شده بین 2 فریم را برای تطبیق مورد استفاده قرار می دهند. ویژگی هایی که معمولاً برای تطبیق در نظر گرفته می شوند عبارتند از طول خط، جهت، تباین و روابط هندسی آن ها با لبه های دیگر.

دریخه⁵ و فاگراس⁶ [43]، یک سیستم ردیاب لبه معرفی کردند که از فیلتر کالمن برای پیش بینی موقعیت جدید لبه و از فاصله ماهالانوبیس⁷ برای محاسبه امتیاز تطبیق استفاده می کند. ردیابی لبه معمولاً نسبت به کلیشه همبستگی دارای تحمل پذیری بیشتری در برابر تغییرات روشنایی می باشد و اغلب در ردیابی اشیاء صلبی با تباین زیاد به خوبی عمل می کند؛ ولی بر روی تصاویر کم دقت و اشیاء غیر صلبی که به سرعت تغییر شکل می دهند و لبه هایشان در تصویر کاملاً متمایز نیست (مانند تصاویر ویدیویی مسابقه فوتبال)، با خطا مواجه می شوند.

نوع دیگری از ویژگی های پر کاربرد، حباب اختلاف حرکت است [13-15]. در این روش حباب های مختلفی در نتیجه حذف فریم در هر مرحله از فریم پس زمینه (که از پیش تعیین شده است) و آستانه سازی به دست می آید. فریم پس زمینه، معمولاً با استفاده از یک فیلتر میانه زمانی⁸ از تصاویر گرفته شده با یک دوربین ثابت به دست می آید. سپس حباب ها بر اساس اندازه و مشخصه های شکل و در مواردی بر اساس اطلاعات مدل تطبیق داده می شوند. شرط لازم برای عملکرد صحیح الگوریتم، ثابت بودن دوربین است و در صورت چرخش افقی و عمودی و بزرگ نمایی، الگوریتم به درستی عمل نمی کند.

¹ Robustness

² Invariant

³ Zero Crossing

⁴ Canny

⁵ Deriche

⁶ Faugeras

⁷ Mahalanobis Distance

⁸ Temporal Median Filter

ویژگی‌های ساختاری دیگری نیز مورد استفاده قرار گرفته‌اند که می‌توان از آن جمله مدل هندسی، بافت، شدت روشنایی محلی و هیستوگرام رنگی را نام برد. به هر حال، هدف ردیابی ساختاری، یافتن ویژگی‌هایی است که نسبت به تغییرات مورد انتظار بین فریم‌ها تحمل‌پذیر باشند.

3-4- روش پیشنهادی برای ردیابی بازیکنان مبتنی بر مدل میدان

ایده به کار گرفته شده برای ردیابی بازیکنان بر اساس تخمین موقعیت جدید بازیکن از روی مختصات موقعیت قبلی بازیکن نسبت به مدل میدان بازی می‌باشد. همان‌طور که پیش از این اشاره شد، در مطالعات قبلی موقعیت بازیکن در هر فریم از روی موقعیت قبلی بازیکن در فریم‌های قبلی نسبت به تصویر تخمین زده می‌شد، در صورتی که در روش حاضر مدل میدان به عنوان الگویی جهت تخمین موقعیت بازیکن مورد استفاده قرار گرفته است؛ به این صورت که موقعیت بازیکن در فریم قبلی در مدل میدان به عنوان منطقه جستجوی اصلاح شده برای فریم فعلی در نظر گرفته می‌شود.

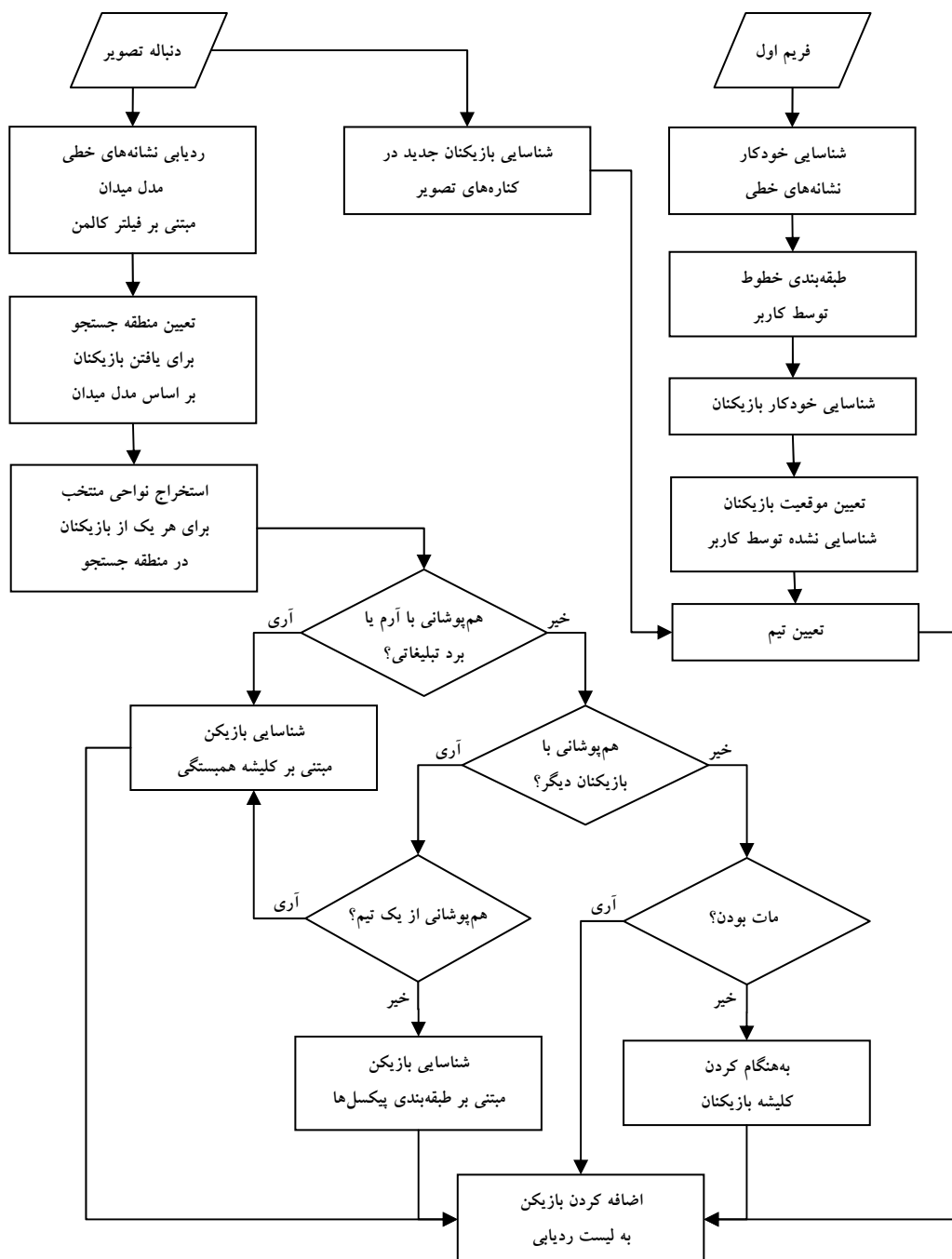
در ابتدا نشانه‌های خطی زمین برای فریم اول با الگوریتم هاف مشخص می‌گردد (جزئیات بیشتر در فصل پنجم تشریح شده است). با در نظر گرفتن این که اطلاعات لازم برای انتساب "خطوط به دست آمده در تصویر" با "نشانه‌های خطی مدل میدان" همواره کافی نمی‌باشد (مانند بروز مشکل در شناسایی برخی خطوط به دلیل هم‌پوشانی با بازیکن یا وضوح پایین خطوط)، در فریم اول، نشانه‌های خطی به صورت دستی توسط کاربر به خطوط متناظر در مدل میدان انتساب داده شده و سپس در فریم‌های متوالی ردیابی می‌گردند (جزئیات روش ردیابی نشانه‌های خطی در فصل پنجم آمده است). همچنین در فریم اول، بازیکنان با استفاده از روش ارایه شده در فصل قبلی شناسایی و تیم آن‌ها مشخص می‌شود. لازم به ذکر است، چون سیستم بدون داشتن اطلاعات قبلی، قادر به شناسایی بازیکنان هم‌پوشان با تابلوهای تبلیغاتی یا آرم‌های تلویزیونی نمی‌باشد، تعیین بازیکنان شناسایی نشده در فریم اول توسط کاربر انجام می‌پذیرد.

در فریم‌های بعدی، موقعیت بازیکنان در فریم قبلی نسبت به مدل میدان محاسبه شده و از روی موقعیت به دست آمده در مدل میدان و همچنین خطوط متناظر بین فریم فعلی و مدل میدان، با استفاده از تبدیل 4 نقطه‌ای پرسپکتیو، موقعیت هر یک از بازیکنان در فریم فعلی مشخص می‌گردد. موقعیت به دست آمده، البته با محدوده‌ای بزرگتر، به منظور احتساب منطقه جستجو برای یافتن بازیکن در فریم فعلی مورد استفاده قرار می‌گیرد. پس از آن، نواحی منتخب برای بازیکنان در منطقه جستجو استخراج می‌شوند. در پایان، با توجه به حالت‌های مختلف هم‌پوشانی برای هر یک از نواحی منتخب برای بازیکنان، یکی از روش‌های لیست شده در جدول (4-1) برای شناسایی بازیکن به کار گرفته می‌شود.

جدول 4-1- روش اعمال شده برای شناسایی بازیکنان در حالت‌های مختلف هم‌پوشانی

حالت هم‌پوشانی	روش اعمال شده برای شناسایی بازیکن
هم‌پوشانی با تابلوهای تبلیغاتی یا آرم‌های تلویزیونی	کلشه همبستگی
هم‌پوشانی با بازیکن از یک تیم	کلشه همبستگی
هم‌پوشانی با بازیکنان غیر هم تیم	طبقه‌بندی پیکسل‌ها
بدون هم‌پوشانی	مبتنی بر ناحیه

لازم به ذکر است چون در حالت هم پوشانی بازیکن، ناحیه مربوط به بازیکن به طور کامل و دقیق قابل تشخیص نمی باشد، از به هنگام کردن کلیشه بازیکن در چنین حالتی صرف نظر می گردد. علاوه بر این، همان گونه که اشاره شد، در برخی شرایط همچون حرکت سریع دوربین، ممکن است بازیکن در تصویر به صورت مات در تعداد معدودی از فریم ها ظاهر شود که پس از آن بازیکن تقریباً به همان شکل اولیه در تصویر ظاهر می گردد. از این رو، تنها در هنگامی کلیشه بازیکن به هنگام می گردد که بازیکن در تصویر به طور کامل و دقیق شناسایی شده باشد. شکل (4-6) نمودار گردش کار روش ارایه شده برای ردیابی بازیکنان در یک دنباله تصویر را نشان می دهد. در بخش های بعدی به جزئیات بیشتر مراحل این نمودار پرداخته می شود.



شکل 4-6- نمودار گردش کار روش ارایه شده برای ردیابی بازیکنان در دنباله تصویر

4-3-1- تعیین منطقه جستجو برای یافتن بازیکنان بر اساس مدل میدان

همان‌طور که پیش از این اشاره شد، عمل ردیابی بازیکن در هر فریم، در محدوده موقعیت واقعی بازیکن (موقعیت بازیکن نسبت به زمین فوتبال) در فریم قبلی انجام می‌شود. واضح است بنا به دلایلی چون تغییر موقعیت بازیکنان در هر فریم نسبت به مختصات مدل میدان به دلیل حرکت مرتب بازیکنان و خطای حاصل از تعیین تناظر نقاط تصویر و مدل میدان، لازم است منطقه جستجوی بازیکن بزرگتر از ناحیه بازیکن در نظر گرفته شود.

برای واضح شدن مطلب، در شکل (4-7) موقعیت جدید بازیکن در حالت‌های مختلف نشان داده شده است. شکل (4-7الف) فریم قبلی را که در آن مستطیل محیطی یک بازیکن علامت‌گذاری شده است، نشان می‌دهد. در صورتی که مستطیل محیطی در موقعیتی یکسان نسبت به تصویر بر روی فریم فعلی قرار داده شود، شکل (4-7ب) حاصل می‌شود. همان‌گونه که دیده می‌شود، مستطیل محیطی به دست آمده در فریم فعلی تنها قسمت کوچکی از بازیکن را شامل می‌شود و این امکان وجود دارد (به خصوص در حالتی که 2 بازیکن هم تیم باشند) که بازیکنان کنار آن به اشتباه به عنوان بازیکن مورد نظر شناخته شوند. شکل (4-7ج) مستطیل محیطی تصویر شده بر روی مدل میدان را از روی مستطیل محیطی بازیکن در فریم قبلی نشان می‌دهد. شکل (4-7د) نیز مستطیل محیطی به دست آمده پس از تصویر کردن ناحیه به دست آمده از مدل میدان بر روی فریم فعلی را نشان می‌دهد.



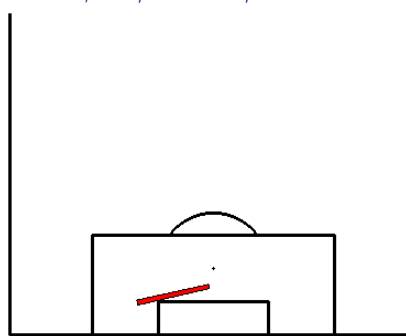
(ب) فریم فعلی (فریم 40م)



(الف) فریم قبلی (فریم 39م)



(د) فریم فعلی (فریم 40م)



(ج) مدل میدان

شکل 4-7- تعیین موقعیت جدید بازیکن در فریم فعلی: الف) موقعیت یک بازیکن در فریم قبلی با مستطیل محیطی علامت‌گذاری شده ب) فریم فعلی همراه با مستطیل محیطی هم موقعیت با فریم قبلی نسبت به تصویر ج) مستطیل محیطی تصویر شده بر روی مدل میدان د) فریم فعلی همراه با مستطیل محیطی هم موقعیت با فریم قبلی نسبت به مدل میدان - با توجه به شکل‌های (ب) و (د)، مستطیل محیطی به دست آمده نسبت به مدل میدان دقت بالاتری نسبت به مستطیل محیطی به دست آمده نسبت به تصویر دارد.

همان‌طور که در شکل (4-7-د) دیده می‌شود، مستطیل محیطی بازیکن که از روی مدل میدان به دست آمده است، با دقت بالا بر روی بازیکن واقع شده است. تنها اختلاف پیکسل‌های در بر گیرنده مستطیل محیطی در بین 2 فریم، به دلیل حرکت بازیکن است که با وجود آنکه بازیکن با سرعت بالا به سمت دروازه حرکت می‌کند، ولی تغییر جابجایی آن نسبت به مدل میدان همان‌طور که در شکل مشاهده می‌شود بسیار اندک است. بنابراین با به کار بردن این روش، محدوده واقعی بازیکن به دست می‌آید.

به منظور محاسبه موقعیت یک بازیکن در تصویر نسبت به مدل میدان و همچنین در حالت برعکس، محاسبه موقعیت متناظر ناحیه‌ای از مدل میدان در تصویر، تبدیل پرسپکتیو مورد استفاده قرار گرفته است. در این روش با داشتن 4 نقطه غیر واقع بر روی یک خط راست، متناظر یک نقطه در تصویر مانند (x, y) ، در مدل میدان، (x', y') ، و یا بر عکس متناظر یک نقطه در مدل میدان در تصویر به دست می‌آید [44]:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \begin{pmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} \quad (3-4)$$

که پارامترهای تبدیل از رابطه زیر تعیین می‌شوند:

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & x_1 x'_1 & y_1 y'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & x_1 y'_1 & y_1 y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & x_2 x'_2 & y_2 y'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & x_2 y'_2 & y_2 y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & x_3 x'_3 & y_3 y'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & x_3 y'_3 & y_3 y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & x_4 x'_4 & y_4 y'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & x_4 y'_4 & y_4 y'_4 \end{pmatrix} \begin{pmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \end{pmatrix} = \begin{pmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \end{pmatrix} \quad (4-4)$$

$$h_{22} = 1 \quad (5-4)$$

از مطالب یاد شده بر می‌آید که برای یافتن موقعیت واقعی بازیکن در فریم قبلی، تبدیل پرسپکتیو بر روی 4 نقطه متناظر در فریم قبلی و مدل میدان محاسبه می‌گردد. پس از مشخص شدن ناحیه بازیکن در مدل میدان، تبدیل پرسپکتیو بر روی 4 نقطه متناظر در فریم فعلی و مدل میدان برای یافتن موقعیت تخمینی بازیکن در فریم فعلی به کار گرفته می‌شود. نقاط متناظر برای انجام تبدیل پرسپکتیو، نقاط تقاطع نشانه‌های خطی انتخاب شده است؛ زیرا مختصات این نقاط در مدل میدان با ثبات و مشخص می‌باشد.

4-3-2- استخراج نواحی منتخب برای بازیکنان در منطقه جستجو

چون اشیاء موجود در میدان بازی می‌تواند چمن، بازیکن، نشانه‌های خطی و توپ باشد، می‌توان به طور کلی پیکسل‌های میدان بازی را به دو کلاس چمن و غیر چمن طبقه‌بندی کرد. بنابراین در حالتی که فقط بازیکن در بخشی از میدان بازی وجود دارد، می‌توان با استفاده از روش شناسایی بازیکن که در فصل سوم توضیح داده شده است، پیکسل‌های غیر چمن یا به عبارت دیگر پیکسل‌های بازیکن را به دست آورد. در حالتی نیز که بازیکن بر روی نشانه‌های خطی واقع شده باشد و یا در کنار توپ قرار گرفته باشد، ممکن است قسمتی از پیکسل‌های آن‌ها به اشتباه به عنوان پیکسل‌های بازیکن در نظر گرفته شود (همانند

شکل 4-4). با این وجود، چون این حالت به ندرت اتفاق می افتد و قسمت عمده ناحیه به دست آمده را پیکسل های بازیکن شامل می شود، می توان از خطای به دست آمده صرف نظر نمود.

با توجه به آنکه ممکن است شناسایی بازیکن در منطقه مورد نظر به دلایلی از قبیل مات شدگی بازیکن در تصویر به علت حرکت سریع دوربین (به شکل 4-8) مراجعه شود) یا کم بودن تعداد پیکسل های بازیکن در تصویر، با شکست مواجه شود، لازم است عملیات شناسایی بازیکن به گونه ای تغییر یابد که امکان شناسایی بازیکن مات شده مهیا گردد. برای این منظور، پارامترهای تعیین کننده زمین چمن یعنی R_t ، G_t ، B_t ، $Gray_t$ به صورت تکراری کوچک تر می شود تا شانس متعلق شدن پیکسل ها به غیر چمن بیشتر شود. علاوه بر این، مساحت مورد نیاز برای بازیکن (همان طور که در فصل سوم آمده است، این مقدار برابر 50 پیکسل می باشد)، به طور تکراری کمتر در نظر گرفته می شود تا امکان شناسایی بازیکنان با تعداد پیکسل های کم فراهم شود. برای روشن شدن موضوع، شکل 4-8) یک فریم از دنباله تصویر و انتخاب پیکسل های یک بازیکن مات شده را نشان می دهد.



(الف)

(ب)

(ج)

شکل 4-8- تشخیص و شناسایی بازیکن در حالت مات شدگی: (الف) محدوده جستجوی یک

بازیکن که به صورت مات شناسایی شده است (ب) نمای بزرگ از محدوده بازیکن مات شده (ج)

پیکسل های شناخته شده برای بازیکن پس از محدود کردن پارامترهای چمن و مساحت مورد نیاز

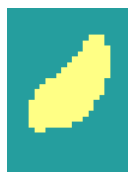
برای بازیکن

شایان ذکر است، در مرحله شناسایی بازیکن در منطقه جستجوی مورد نظر، امکان استخراج چندین بازیکن وجود دارد که این مورد می تواند به دلیل هم پوشانی بازیکنان در فریم های قبلی و جدا شدن آن ها از یکدیگر در فریم فعلی روی داده باشد.

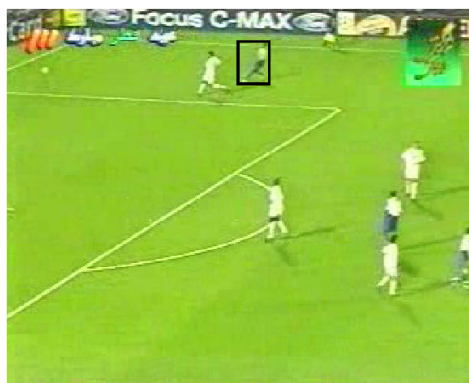
4-3-3- تشخیص هم پوشانی

به منظور تشخیص هم پوشانی بین 2 بازیکن، موقعیت نواحی متعلق به 2 بازیکن مورد بررسی قرار می گیرد. همچنین تشخیص هم پوشانی بازیکن با آرم تلویزیونی، با توجه به ثابت بودن موقعیت آرم در تصویر، با بررسی قرار گرفتن ناحیه بازیکن در ناحیه ای خاص از تصویر امکان پذیر می گردد. برای پی بردن به وجود هم پوشانی یک بازیکن با تابلو تبلیغاتی نیز موقعیت بازیکن در مدل میدان مورد استفاده قرار گرفته است. در صورتی که موقعیت یک بازیکن نسبت به سیستم مختصات مدل

میدان در نزدیکی کناره‌های مدل میدان قرار گیرد (به عبارت دیگر نزدیک به تابلوهای تبلیغاتی باشد) و پیکسل‌های شناسایی شده بازیکن محدود به منطقه جستجو نباشد، آن بازیکن به عنوان بازیکن هم‌پوشان با تابلو تبلیغاتی در نظر گرفته می‌شود. شکل (4-9) یک بازیکن در 2 فریم که در یکی به صورت غیر هم‌پوشان (شکل (4-9-الف)) و در دیگری هم‌پوشان با تابلو تبلیغاتی (شکل (4-9-الف)) شناخته شده است را نشان می‌دهد. همان گونه که دیده می‌شود، در شکل (4-9-الف) با وجود آنکه بازیکن در کناره‌های زمین قرار دارد غیر هم‌پوشان شناسایی شده است؛ زیرا پیکسل‌های بازیکن در مرکز منطقه جستجو قرار گرفته‌اند، در صورتی که در شکل (4-9-ب) پیکسل‌های بازیکن تا کناره‌های بالایی منطقه جستجو ادامه یافته است و در نتیجه این بازیکن به صورت هم‌پوشان با تابلو تبلیغاتی در نظر گرفته می‌شود.



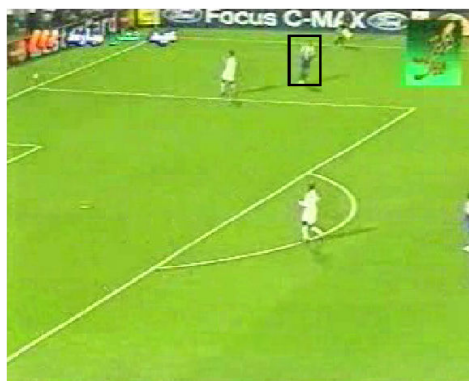
(ب)



(الف)



(د)



(ج)

شکل 4-9- تشخیص هم‌پوشانی بازیکن با تابلوهای تبلیغاتی: الف و ج) محدوده جستجوی یک بازیکن با مستطیل سیاه علامت‌گذاری شده است ب و د) نمای بزرگ از پیکسل‌های شناخته شده بازیکن در محدوده مشخص شده - در شکل (ب) بر خلاف شکل (د)، با وجود اینکه بازیکن در کنار تابلو تبلیغاتی قرار گرفته است، ولی با تابلو تبلیغاتی هم‌پوشانی ندارد.

نکته‌ای که در مورد کلیشه همبستگی باید یاد آور شد آن است که در کلیشه بازیکنان که به صورت مستطیل شکل در نظر گرفته می‌شود، لازم نیست تمامی پیکسل‌ها مد نظر قرار بگیرند، بلکه تنها پیکسل‌هایی که قسمتی از بازیکن می‌باشند، کلیشه بازیکن را تشکیل می‌دهند. از این رو، علاوه بر تصویر مستطیل شکل از ناحیه بازیکن، یک ماتریس که مشخص کننده پیکسل‌های فعال و غیر فعال در کلیشه می‌باشد نیز ذخیره می‌گردد.

4-3-4- شناسایی بازیکنان مبتنی بر طبقه‌بندی پیکسل‌ها

در صورتی که یک بازیکن با بازیکنان غیر هم تیم خود هم‌پوشانی داشته باشد، مساله به جدا کردن ناحیه به دست آمده (ناحیه شناسایی شده به عنوان بازیکن) به 2 کلاس مجزا (یک کلاس برای بازیکن مورد نظر و کلاس دیگر برای بازیکنان دیگر) ساده می‌شود. واضح است در این حالت مناسب‌ترین ویژگی متمایز کننده بین 2 کلاس ویژگی رنگ می‌باشد. به همین علت در این پایان‌نامه برای جداسازی پیکسل‌ها، طبقه‌بندی پیکسل‌ها مورد استفاده قرار گرفته است.

به منظور جداسازی پیکسل‌ها به 2 کلاس مختلف با استفاده از روش طبقه‌بندی پیکسل‌ها، برای هر پیکسل فاصله اقلیدسی رنگ آن با رنگ پیکسل‌های هر یک از 2 کلاس محاسبه و در صورتی که فاصله اقلیدسی برای کلاس بازیکن مورد نظر کمتر باشد، به آن بازیکن تعلق می‌گیرد. سرانجام، به منظور یافتن ناحیه بازیکن مورد نظر، بر روی پیکسل‌های به دست آمده فیلتر میانه اعمال می‌گردد و بزرگترین ناحیه متصل بدین منظور اختصاص داده می‌شود. همان‌گونه که اشاره شد، چون شناسایی بازیکنان می‌تواند به صورت ناقص و همراه با خطا باشد، کلیشه بازیکن تا جدا شدن بازیکن از دیگر بازیکنان به هنگام نمی‌گردد.

شکل (4-10) دو تصویر از یک دنباله تصویر را نشان می‌دهد که در تصویر اول، بازیکنان علامت‌گذاری شده بدون هم‌پوشانی و در تصویر دیگر هم‌پوشان با یکدیگر تشخیص داده شده‌اند. در شکل (4-10-ج)، بازیکنان شناسایی شده به همراه کلیشه آن‌ها در محدوده علامت‌گذاری شده برای شکل (4-10-الف)، به صورت بزرگ‌نمایی نشان داده شده است. شکل (4-10-د) نیز محدوده بازیکنان شناسایی شده در شکل (4-10-ب) را به همراه کلیشه آن نشان می‌دهد. همان‌طور که مشاهده می‌شود، 2 بازیکن به صورت هم‌پوشان تشخیص داده شده‌اند. شکل (4-10-و) نتیجه اعمال طبقه‌بندی پیکسل‌ها با استفاده از کلیشه بازیکنان در شکل (4-10-ج) را نشان می‌دهد. بازیکنان شناسایی شده پس از اعمال فیلتر میانه نیز در شکل (4-10-ی) نشان داده شده است.

4-3-5- شناسایی بازیکنان مبتنی بر کلیشه همبستگی

در حالتی که یک بازیکن با تابلوهای تبلیغاتی، آرم تلویزیونی یا بازیکن هم تیم خود هم‌پوشانی داشته باشد، روش مبتنی بر طبقه‌بندی پیکسل‌ها کارا نخواهند بود؛ زیرا امکان هم‌رنگ بودن پیکسل‌های بازیکن مورد نظر با دیگر اشیاء یاد شده و در نتیجه خطای طبقه‌بندی وجود دارد. روشی که در این حالت مورد استفاده قرار گرفته است، روش کلیشه همبستگی استاندارد می‌باشد. چون منطقه جستجو برای بازیکن مشخص و محدود است، انتظار می‌رود کلیشه همبستگی با خطای قابل قبول، بازیکن مورد نظر را با پیکسل‌های جدیدش شناسایی نماید. در این روش نیز همانند روش طبقه‌بندی پیکسل‌ها از به هنگام کردن کلیشه بازیکن صرف نظر می‌گردد؛ چرا که در غیر این صورت، خطاهای فریم‌های متوالی با یکدیگر جمع شده و در نهایت مشکل انحراف تدریجی پیش می‌آید.



(ب)



(الف)



(د)



(ج)



(ی)



(و)



شکل 4-10- شناسایی بازیکنان مبتنی بر طبقه‌بندی پیکسل‌ها: الف و ب) 2 تصویر از یک دنباله تصویر ج و د) نمای بزرگ از پیکسل‌های بازیکنان علامت‌گذاری شده در شکل‌های (الف) و (ب) به همراه کلیشه بازیکنان و) پیکسل‌های منتخب برای هر یک از بازیکنان شکل (د) که با طبقه‌بندی پیکسل‌ها به دست آمده است ی) بازیکنان شناسایی شده پس از اعمال فیلتر میانه

4-4- جمع‌بندی

در این فصل، ابتدا مشکلات موجود در زمینه ردیابی بازیکنان در تصاویر ویدیویی مسابقات فوتبال، شامل پویایی و پیچیدگی محیط، فعل و انفعالات دوربین، حرکت سریع و غیر قابل پیش‌بینی بازیکنان، دقت پایین تصاویر ویدیویی، غیر صلب بودن بازیکنان و مساله هم‌پوشانی مرور شد. سپس انواع روش‌های ردیابی معمول از قبیل ردیابی کلیشه همبستگی، ردیابی حرکت، ردیابی با استفاده از کلیشه فرم‌پذیر و ردیابی ساختاری مورد بحث و بررسی قرار گرفت و بر روی ضعف‌ها و

مشکلات این روش‌ها در مورد ردیابی بازیکنان فوتبال توضیحاتی ارائه گردید. در بخش سوم نیز روش پیشنهادی برای ردیابی بازیکنان یعنی ردیابی بازیکنان مبتنی بر مدل میدان معرفی گردید.

معمولاً در هنگام فیلم‌برداری از صحنه‌های گل، کارگردان تلویزیونی برای دنبال کردن توپ و بازیکنان مجبور به حرکت سریع دوربین (چرخش افقی، چرخش عمودی و بزرگ‌نمایی) می‌شود که این خود باعث تغییر در محدوده قابل رویت و در نتیجه تغییر در موقعیت اشیاء نسبت به تصویر می‌گردد. به عبارت دیگر، در صورتی که موقعیت اشیاء داخل زمین نسبت به سیستم مختصاتی مدل میدان ثابت باشد، ولی به دلیل تغییر دید به وجود آمده، موقعیت اشیاء داخل زمین نسبت به سیستم مختصاتی تصویر تغییر می‌یابد و هر چه فعل و انفعالات دوربین سریع‌تر باشد، این تغییرات بیشتر نمایان می‌شود.

با توجه به اطلاعات موجود نسبت به کارهای انجام شده در زمینه ردیابی بازیکنان ورزشی، موقعیت بازیکنان در هر فریم تنها از روی موقعیت آن‌ها در فریم قبلی و در برخی موارد با کمک روش‌های تخمین حرکت همچون فیلتر کالمن انجام گردیده است. چون در صحنه‌های گل حرکت بازیکنان، جدای از آن که با حرکت ناشناخته دوربین ترکیب می‌شود، از هیچ مدل خاصی تبعیت نمی‌کند و کاملاً ناشناخته و ناگهانی است، به نظر می‌رسد استفاده از روش‌های تخمین حرکت به شیوه‌های معمول برای این منظور کارایی مناسبی نداشته باشد. روشی که در این پایان‌نامه مورد استفاده قرار گرفته است، موقعیت واقعی بازیکن در هر فریم را نسبت به مختصات مدل میدان بازی به دست آورده و فقط در آن محدوده به دنبال بازیکن مورد نظر می‌پردازد که از لحاظ احتمال یافتن بازیکن در منطقه جستجوی مذکور از خطای کمتری در مقایسه با روش‌های پیشین برخوردار است.

فصل پنجم

ایجاد نمای دور

نمایش نمای دور، صحنه‌ای مانند آنچه را که یک تماشاچی حاضر در ورزشگاه می‌بیند، تداعی می‌کند. در این نمایش، بیننده قادر است علاوه بر جزئیات صحنه مانند بازیکنان، توپ و حرکت آن‌ها، نمای وسیع‌تری از آن صحنه که شامل موقعیت و حرکت اشیاء نسبت به یکدیگر و مدل میدان می‌باشد، را ببیند. به منظور ایجاد نمای دور از یک صحنه و در نتیجه برای کنار هم قرار دادن دنباله تصویر، لازم است دنباله تصویر با یک مدل مشخص که در این جا مدل میدان است، ترازبندی گردد و پس از آن دنباله تصویر به صورت مجتمع در کنار یکدیگر قرار گرفته و اشیاء بر روی آن حرکت داده شوند.

جهت ترازبندی دنباله تصویر با مدل میدان، تبدیل پرسپکتیو بر روی تقاطع نشانه‌های خطی محوطه جریمه انجام گرفته است. به همین منظور در بخش اول، مراحل در نظر گرفته شده برای شناسایی و ردیابی نشانه‌های خطی ارایه شده است. ابتدا روش پیشنهادی برای استخراج پیکسل‌های منتخب برای نشانه‌های خطی معرفی می‌شود. سپس الگوریتم تبدیل هاف برای شناسایی خطوط موجود در تصویر مورد استفاده قرار می‌گیرد. از آن جایی که شناسایی خطوط در تصویر با الگوریتم تبدیل هاف، بار محاسباتی بالایی دارد و ممکن است با خطا روبه‌رو شود، روشی مبتنی بر ترکیب الگوریتم "تبدیل هاف" با "فیلتر کالمن" برای ردیابی خطوط معرفی گردیده است.

در بخش دوم این فصل، مراحل لازم برای کنار هم قرار دادن دنباله تصویر برای ایجاد نمای دور که شامل ترازبندی، مجتمع‌سازی و نمایش پس‌مانده‌های مهم بر روی تصویر کنار هم قرار داده شده است، توضیح داده می‌شود. در این بخش، همچنین روش عمده برای کنار هم قرار دادن دنباله تصویر یعنی "کنار هم قرار دادن ایستا" و "کنار هم قرار دادن پویا" برای کاربردها و شرایط مختلف ارایه شده است.

5-1- شناسایی و ردیابی نشانه‌های خطی

همان‌طور که پیش از این اشاره شد، الگوریتم ارایه شده برای تعیین موقعیت بازیکنان و کنار هم قرار دادن دنباله تصویر و از این رو برای انجام تبدیل پرسپکتیو، به تقاطع نشانه‌های خطی نیازمند است. به دلایلی چون هم‌پوشانی به وسیله بازیکنان، رنگ پریدگی خطوط، کنده‌شدگی چمن در محل نشانه‌های خطی، نویز حاصل از تلفات فشرده‌سازی و تبدیلات ویدیویی، نشانه‌های خطی همواره به صورت واضح و کامل در تصویر مشاهده نمی‌شوند. بنابراین در مرحله اول، لازم است پیکسل‌های منتخب برای نشانه‌های خطی با دقت بالا استخراج گردند تا الگوریتم تبدیل هاف قادر به شناسایی خط مورد نظر باشد. همچنین

به منظور افزایش سرعت شناسایی خطوط و کاهش خطای احتمالی در الگوریتم تبدیل هاف، روشی مبتنی بر ترکیب الگوریتم "فیلتر کالمن" با "تبدیل هاف" برای ردیابی خطوط مورد استفاده قرار گرفته است.

در ابتدای این بخش، روش پیشنهادی برای استخراج پیکسل‌های منتخب برای نشانه‌های خطی معرفی شده است. سپس الگوریتم تبدیل هاف خطی که به منظور شناسایی خطوط از روی نقاط منتخب به کار می‌رود، تشریح شده است. زیربخش‌های بعدی، به توضیحاتی درباره الگوریتم "فیلتر کالمن خطی" و "فیلتر کالمن توسعه یافته" می‌پردازد. در انتهای بخش نیز روشی برای ردیابی خطوط که فیلتر کالمن را برای تخمین موقعیت نشانه‌های خطی در فضای هاف مورد استفاده قرار می‌دهد، معرفی شده است.

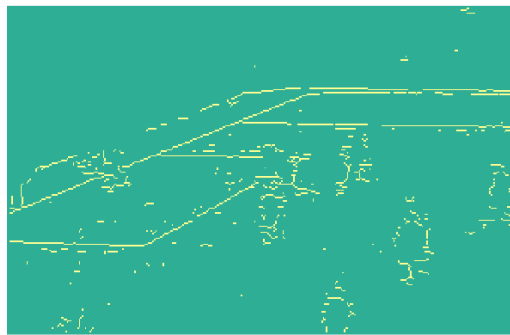
5-1-1- استخراج پیکسل‌های منتخب برای نشانه‌های خطی

با وجود آنکه نشانه‌های خطی در تصویر از سطح خاکستری پایینی برخوردار هستند، ولی این امکان وجود دارد که اشیاء دیگر نظیر تابلوهای تبلیغاتی، آرم‌های تلویزیونی، تماشاچیان و حتی بازیکنان دارای پیکسل‌های سفید رنگ باشند. علاوه بر این، ممکن است بخشهایی از نشانه‌های خطی به دلیل دقت پایین تصویر و شرایط زمین چمن دارای سطح خاکستری پایین بوده یا به دلیل هم‌پوشانی با بازیکنان قابل رویت نباشد. در نتیجه برای شناسایی نشانه‌های خطی باید روشی مورد استفاده قرار گیرد که با در نظر گرفتن شرایط یاد شده، قادر به استخراج دقیق و کامل پیکسل‌های نشانه‌های خطی باشد.

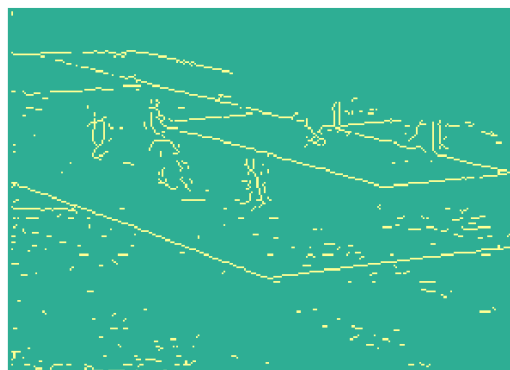
در روش پیشنهادی، به منظور استخراج پیکسل‌های منتخب برای نشانه‌های خطی، ابتدا فیلتر 2 بعدی h که در رابطه (5-1) تعریف شده است، بر روی تصویر سطح خاکستری اعمال می‌گردد. برای حذف نواحی متعلق به بازیکنان و تابلوهای تبلیغاتی، پیکسل‌هایی که در همسایگی 10×10 آن‌ها بیش از 75٪ پیکسل‌ها غیر چمن باشند (استخراج نواحی متعلق به چمن در فصل سوم آمده است)، از لیست پیکسل‌های منتخب حذف می‌شوند. علاوه بر این، برای حذف پیکسل‌هایی با سطح خاکستری و گرادیان پایین، پیکسل‌های منتخب که شرط رابطه (5-2) را برقرار کنند، حذف می‌گردند. شکل (5-1)، نتایج استخراج پیکسل‌های منتخب برای نشانه‌های خطی را برای تصاویر شکل (3-6) نشان می‌دهد.

$$h = \begin{pmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{pmatrix} \quad (1-5)$$

$$\begin{cases} I_A(x, y) < A_{peak} + \frac{A_t}{2} & \text{for all } A \in \{R, G, B\}, \\ GL(x, y) < GL_t \end{cases} \quad (2-5)$$



(الف)



(ب)



(ج)

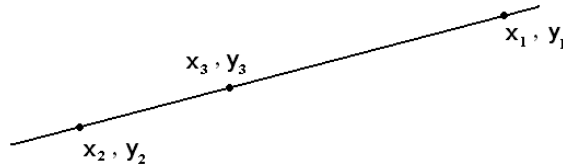
شکل 5-1- نتایج استخراج پیکسل‌های منتخب برای نشانه‌های خطی برای تصاویر شکل (3-6)

5-1-2- تبدیل هاف خطی

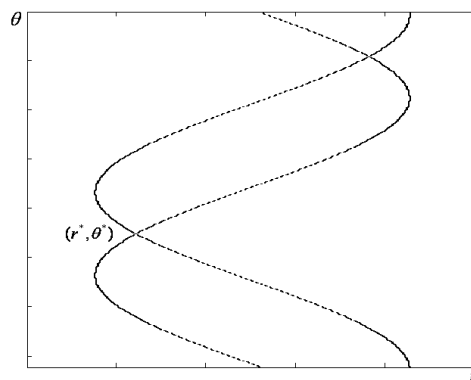
الگوریتم تبدیل هاف¹ خطی به ازای هر نقطه واقع بر روی یک خط راست، تمامی خطوط گذرنده از آن را در فضای پارامترها ترسیم می‌کند. معادل نقطه مشترکی که از برخورد خطوط در فضای پارامترها حاصل می‌شود، در فضای اصلی، معادله خط راست مورد نظر را نشان می‌دهد [45]. به عنوان مثال، در صورتی که 2 نقطه (x_1, y_1) و (x_2, y_2) همانند شکل (5-2)

¹ Hough Transform Algorithm

بر روی یک خط راست واقع شده باشند، معادل آن نقاط در فضای $r-\theta$ همانند شکل (3-5) به ترتیب موج‌های سینوسی $x_1 \cos(\theta) + y_1 \sin(\theta) = r$ و $x_2 \cos(\theta) + y_2 \sin(\theta) = r$ خواهد بود.

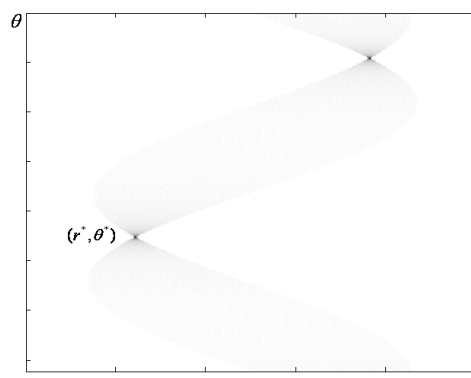


شکل 2-5- خط راست به همراه 3 نقطه مشخص از آن



شکل 3-5- معادل نقاط (x_1, y_1) و (x_2, y_2) در شکل (2-5) در فضای $r-\theta$: هر 2 موج سینوسی از نقطه (r^*, θ^*) عبور می‌کنند.

همان‌طور که در شکل (3-5) مشاهده می‌شود، در نقطه (r^*, θ^*) 2 موج سینوسی یاد شده یکدیگر را قطع می‌کنند. هر نقطه دیگری مانند (x_3, y_3) نیز که بر روی خط مورد نظر واقع شده باشد، در فضای $r-\theta$ از نقطه (r^*, θ^*) عبور می‌کند. در نتیجه با مشخص شدن نقطه (r^*, θ^*) ، معادله خط مورد نظر به صورت $x_1 \cos(\theta^*) + y_1 \sin(\theta^*) = r^*$ حاصل می‌شود. شکل (4-5) معادل تمامی نقاط واقع بر روی خط راست شکل (2-5) را در فضای $r-\theta$ نشان می‌دهد. همان‌گونه که مشاهده می‌شود تمامی موج‌های سینوسی از نقطه (r^*, θ^*) عبور می‌کنند.



شکل 4-5- معادل نقاط واقع بر روی خط راست شکل (2-5) در فضای $r-\theta$: تمامی موج‌های سینوسی از نقطه (r^*, θ^*) عبور می‌کنند.

برای پیاده‌سازی الگوریتم تبدیل هاف، یک آرایه انباره در فضای پارامترها در نظر گرفته شده و به ازای هر نقطه در تصویر، تمامی سلول‌های آرایه را که موج‌های سینوسی معادل در این فضا از آن می‌گذرند، یک واحد افزایش می‌یابد. سرانجام، سلولی که در فضای اصلی بیشترین مقدار را دارد، معادله خط راست موجود در تصویر را نشان می‌دهد.

5-1-3- فیلتر کالمن

فیلتر کالمن^۱ در سال 1960 توسط کالمن معرفی گردید [46]. فیلتر کالمن به صورت بازگشتی روش حداقل مربعات^۲ را پیاده‌سازی می‌کند. این فیلتر با حداقل کردن واریانس خطای تخمین، سعی در برآورد خروجی سیستم می‌نماید که برای رسیدن به هدف تخمین‌های گذشته، حال و حتی آینده را مورد استفاده قرار می‌دهد. به بیان ریاضی، عملکرد فیلتر کالمن در یک سیستم خطی عبارت است از تخمین حالت، $s \in \mathbb{R}^n$ ، از یک فرآیند تحت کنترل با رابطه تفاضلی خطی زیر [46]:

$$s_{k+1} = A_k s_k + B u_k + w_k \quad (3-5)$$

مقدار خروجی اندازه گرفته شده در این سیستم، $z \in \mathbb{R}^m$ ، برابر است با:

$$z_k = H_k s_k + v_k \quad (4-5)$$

متغیرهای تصادفی w_k و v_k به ترتیب نویز فرآیند و نویز اندازه‌گیری می‌باشند. در فیلتر کالمن، فرض بر آن است که نویزها سفید و مستقل از یکدیگر بوده و توزیع آن‌ها نرمال باشد؛ به عبارت دیگر:

$$p(w) \sim N(0, Q) \quad (5-5)$$

$$p(v) \sim N(0, R) \quad (6-5)$$

در رابطه‌های بالا A_k ماتریسی با ابعاد $n \times n$ بوده و ارتباط بین حالت در مرحله k با حالت در مرحله $k+1$ را در هنگامی که از نویز فرآیند و ورودی کنترل کننده، u_k ، صرف نظر می‌گردد، نشان می‌دهد. ماتریس B با ابعاد $n \times l$ نیز به ورودی کنترل کننده حالت، $u \in \mathbb{R}^l$ ، مربوط می‌شود. ماتریس H با ابعاد $m \times n$ در رابطه (3-5)، حالت را به مقدار اندازه‌گیری شده z_k مرتبط می‌سازد.

در صورتی که $\hat{s}_k^- \in \mathbb{R}^n$ تخمین حالت پیشین^۳ برای مرحله k ام با داشتن فرآیند مرحله پیش از k و $\hat{s}_k \in \mathbb{R}^n$ تخمین حالت پسین^۴ برای مرحله k با دانستن مقدار اندازه z_k باشد، خطای تخمین پیشین و پسین به ترتیب به این صورت تعریف می‌گردند:

$$\bar{e}_k \equiv s_k - \hat{s}_k^- \quad (7-5)$$

$$e_k \equiv s_k - \hat{s}_k \quad (8-5)$$

بنابراین کواریانس خطای تخمین پیشین و کواریانس خطای تخمین پسین به ترتیب عبارتند از:

$$P_k^- = E[\bar{e}_k \bar{e}_k^T] \quad (9-5)$$

$$P_k = E[e_k e_k^T] \quad (10-5)$$

¹ Kalman Filter

² Least-Square Method

³ Periori State Stimat

⁴ Posteriori State Stimat

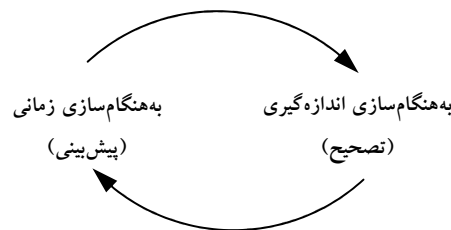
در فیلتر کالمن تخمین حالت پسین، $\hat{s}_k$ ، به صورت ترکیب خطی از تخمین حالت پیشین، $\hat{s}_k^-$ ، و تفاضل وزن دار بین مقدار واقعی z_k و مقدار پیش‌بینی شده $H_k \hat{s}_k^-$ در نظر گرفته می‌شود:

$$\hat{s}_k = \hat{s}_k^- + K(z_k - H_k \hat{s}_k^-) \quad (11-5)$$

که در آن ماتریس K با ابعاد $n \times m$ ، ضریب کالمن¹ نامیده می‌شود و باید به گونه‌ای انتخاب شود که کواریانس خطای پسین را به حداقل برساند. بنابراین می‌توان نوشت:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (12-5)$$

فیلتر کالمن، فرآیند یک سیستم را با استفاده از یک کنترل پس‌خور تخمین می‌زند؛ به این صورت که فیلتر، حالت فرآیند را تخمین زده و سپس پس‌خوری به شکل مقادیر خروجی که ممکن است همراه با نویز باشد، به دست می‌آیند. بنابراین رابطه‌های فیلتر کالمن به 2 دسته تقسیم می‌شوند: رابطه‌های به‌هنگام‌سازی زمانی² و رابطه‌های به‌هنگام‌سازی اندازه‌گیری³.
رابطه‌های به‌هنگام‌سازی زمانی، حالت فعلی و کواریانس خطای تخمین را به منظور محاسبه تخمین پیشین برای حالت بعدی مورد استفاده قرار می‌دهد (عملکرد پیش‌خور). از طرف دیگر، رابطه‌های به‌هنگام‌سازی اندازه‌گیری، عمل پس‌خور را انجام می‌دهند؛ به عبارت دیگر، مقادیر اندازه‌گیری شده جدید همراه با تخمین پیشین برای محاسبه تخمین پسین به کار گرفته می‌شوند. بنابراین می‌توان گفت رابطه‌های به‌هنگام‌سازی زمانی به صورت یک پیش‌بینی کننده و رابطه‌های به‌هنگام‌سازی اندازه‌گیری به شکل یک تصحیح کننده ظاهر می‌شوند. از این رو، چرخه فیلتر کالمن را می‌توان همانند شکل (5-5) در نظر گرفت.



شکل 5-5- چرخه فیلتر کالمن در نمای کلی

رابطه‌های به‌هنگام‌سازی زمانی و به‌هنگام‌سازی اندازه‌گیری برای فیلتر کالمن خطی گسسته به ترتیب در جدول (5-1) و (5-2) آمده است. در جدول (5-1)، A_k و B از رابطه (5-3)، Q_k از رابطه (5-5) و P_k^- و P_k به ترتیب از رابطه‌های (5-9) و (5-10) آورده شده است. در به‌هنگام‌سازی اندازه‌گیری برای مرحله k ام، ابتدا ضریب کالمن، K_k ، از روی رابطه (5-15) که مشابه رابطه (5-12) است، محاسبه می‌شود. قدم بعدی به دست آوردن مقدار خروجی سیستم، z_k ، و محاسبه تخمین پسین حالت توسط رابطه (5-16) است. قدم آخر نیز محاسبه تخمین پسین کواریانس با استفاده از رابطه (5-15) می‌باشد. پس از انجام به‌هنگام‌سازی اندازه‌گیری، تخمین پسین از حالت قبلی برای محاسبه تخمین پیشین در حالت جدید به کار گرفته می‌شود و به این ترتیب چرخه تکرار می‌گردد. شکل (5-6) نمای کلی از عملکرد فیلتر کالمن خطی را نشان می‌دهد.

¹ Kalman Gain

² Time Update

³ Measurement Update

جدول 5-1- رابطه‌های به‌هنگام‌سازی زمانی برای فیلتر کالمن خطی گسسته

$$\hat{s}_{k+1}^- = A_k \hat{s}_k + B u_k \quad (13-5)$$

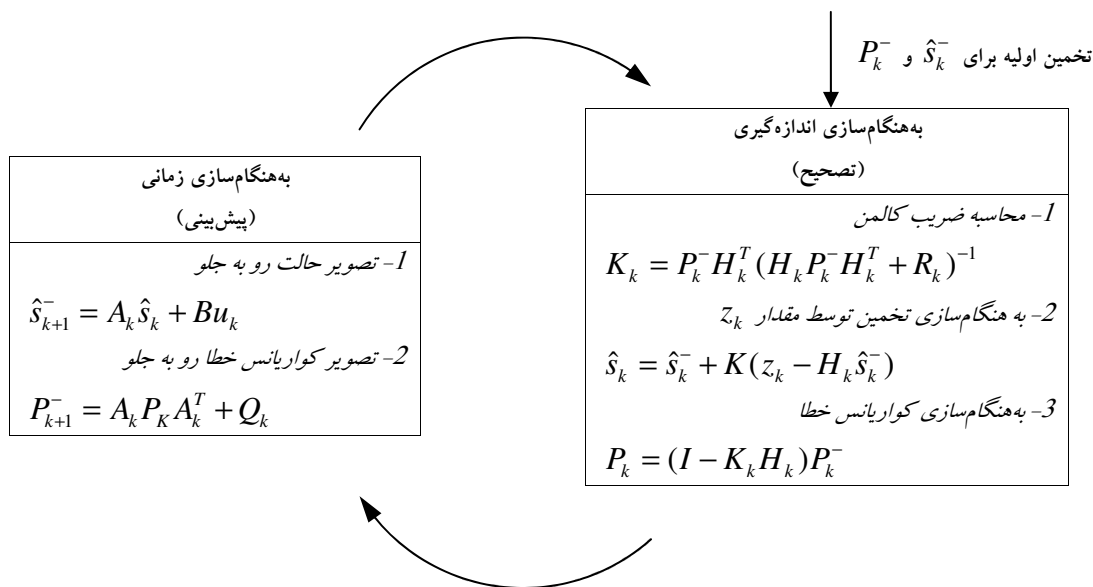
$$P_{k+1}^- = A_k P_k A_k^T + Q_k \quad (14-5)$$

جدول 5-2- رابطه‌های به‌هنگام‌سازی اندازه‌گیری برای فیلتر کالمن خطی گسسته

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (15-5)$$

$$\hat{s}_k = \hat{s}_k^- + K (z_k - H_k \hat{s}_k^-) \quad (16-5)$$

$$P_k = (I - K H_k) P_k^- \quad (17-5)$$



شکل 5-6- نمای کلی از عملکرد فیلتر کالمن خطی

5-1-4- فیلتر کالمن توسعه یافته

در صورتی که رابطه‌های فرآیند یا اندازه‌گیری به صورت غیر خطی باشند، لازم است فیلتر کالمن توسعه یافته که در ادامه به آن پرداخته می‌شود، مورد استفاده قرار گیرد. در چنین حالتی، می‌توان همانند سری‌های تیلور با به‌کارگیری مشتقات جزئی از توابع فرآیند و اندازه‌گیری، به صورت خطی تخمین را حول تخمین فعلی به دست آورد. فرض می‌شود حالت فرآیند یک سیستم غیر خطی به صورت:

$$s_{k+1} = f(s_k, u_k, w_k) \quad (18-5)$$

و با مقدار اندازه:

$$z_k = h(s_k, v_k) \quad (19-5)$$

باشد که در آن متغیرهای تصادفی w_k و v_k همانند قبل، به ترتیب نویز فرآیند و نویز اندازه گیری هستند. تابع غیر خطی $f(\cdot)$ در رابطه (5-18)، رابطه بین حالت در مرحله k ام و حالت در مرحله $(k+1)$ ام را نشان می دهد. تابع غیر خطی $h(\cdot)$ در رابطه اندازه گیری (5-19) نیز حالت s_k را به مقدار اندازه z_k مرتبط می سازد. بردارهای حالت و اندازه گیری در حالتی که از نویز فرآیند و نویز اندازه گیری صرف نظر شود، به ترتیب به صورت زیر نشان داده می شوند:

$$\tilde{s}_{k+1} = f(\hat{s}_k, u_k, 0) \quad (20-5)$$

$$\tilde{z}_k = h(\hat{s}_k, 0) \quad (21-5)$$

که $\hat{s}_k$ تخمین پسین حالت از مرحله قبل برای مرحله k ام را نشان می دهد. به منظور تخمین فرآیند با رابطه های غیر خطی حالت و اندازه گیری، رابطه های (5-20) و (5-21) به صورت خطی تخمین زده می شوند:

$$s_{k+1} \approx \tilde{s}_{k+1} + A(s_k - \hat{s}_k) + Ww_k \quad (22-5)$$

$$z_k \approx \tilde{z}_k + H(s_k - \hat{s}_k) + Vv_k \quad (23-5)$$

که

- s_{k+1} و z_k به ترتیب بردارهای واقعی حالت و اندازه گیری هستند.
- $\tilde{s}_{k+1}$ و $\tilde{z}_k$ به ترتیب تخمین حالت و اندازه گیری از روی رابطه های (5-20) و (5-21) هستند.
- $\hat{s}_k$ تخمین پسین از حالت در مرحله k است.
- w_k و v_k به ترتیب متغیرهای تصادفی برای نمایش نویز فرآیند و اندازه گیری می باشند.
- A ماتریس ژاکوبی از مشتق جزئی تابع $f(\cdot)$ نسبت به s است:

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_k, u_k, 0) \quad (24-5)$$

- W ماتریس ژاکوبی از مشتق جزئی تابع $f(\cdot)$ نسبت به w است:

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\hat{x}_k, u_k, 0) \quad (25-5)$$

- H ماتریس ژاکوبی از مشتق جزئی تابع $h(\cdot)$ نسبت به s است:

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_k, 0) \quad (26-5)$$

- V ماتریس ژاکوبی از مشتق جزئی تابع $h(\cdot)$ نسبت به v است:

$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{x}_k, 0) \quad (27-5)$$

که در رابطه های (5-24) الی (5-27) به منظور سادگی، از نمایش اندیس k برای مرحله زمانی خودداری شده است. مجموعه کامل رابطه های به هنگام سازی زمانی و به هنگام سازی اندازه گیری برای فیلتر کالمن توسعه یافته، به ترتیب در جدول های (5-3) و (5-4) نشان داده شده است. همان طور که مشاهده می شود، در رابطه های به هنگام سازی در جدول (5-3)

همانند جدول (5-1)، تخمین‌های حالت و کواریانس از مرحله k به مرحله $(k+1)$ ام تصویر می‌گردد. در این جدول، تابع $f(\cdot)$ از رابطه (5-18)، ماتریس‌های ژاکوبی فرآیند در مرحله k ام، A_k و W_k ، از رابطه‌های (5-24) و (5-25)، و کواریانس نویز فرآیند در مرحله k ام، Q_k ، از رابطه (5-5) به دست می‌آید.

در جدول (5-4)، رابطه‌های به‌هنگام‌سازی اندازه‌گیری با فرض تصحیح تخمین‌های حالت و کواریانس به وسیله مقدار اندازه z_k آمده است. در این جدول، تابع $h(\cdot)$ از رابطه (5-19)، ماتریس‌های ژاکوبی اندازه‌گیری در مرحله k ام، H_k و V_k ، از رابطه‌های (5-26) و (5-27)، و کواریانس نویز اندازه‌گیری در مرحله k ام، R_k ، از رابطه (5-6) به دست می‌آید. در شکل (5-7) نمای کلی از عملکرد فیلتر کالمن غیر خطی آورده شده است.

جدول 5-3- رابطه‌های به‌هنگام‌سازی زمانی برای فیلتر کالمن توسعه یافته

$$\tilde{s}_{k+1} = f(\hat{s}_k, u_k, 0) \quad (28-5)$$

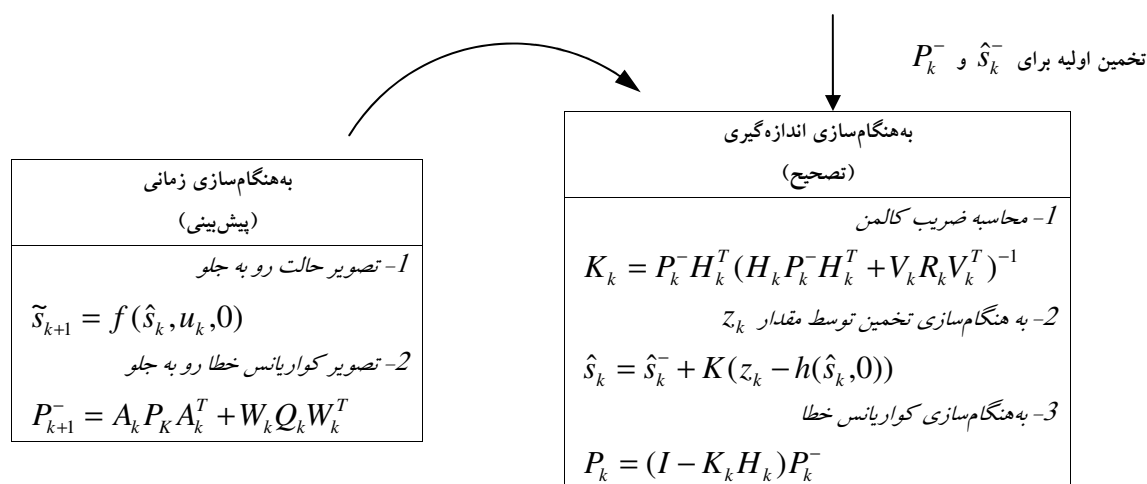
$$P_{k+1}^- = A_k P_k^- A_k^T + W_k Q_k W_k^T \quad (29-5)$$

جدول 5-4- رابطه‌های به‌هنگام‌سازی اندازه‌گیری برای فیلتر کالمن توسعه یافته

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \quad (30-5)$$

$$\hat{s}_k = \hat{s}_k^- + K(z_k - h(\hat{s}_k^-, 0)) \quad (31-5)$$

$$P_k = (I - K_k H_k) P_k^- \quad (32-5)$$



شکل 5-7- نمای کلی از عملکرد فیلتر کالمن توسعه یافته

5-1-5- ترکیب فیلتر کالمن با تبدیل هاف به منظور ردیابی نشانه‌های خطی

تبدیل هاف با جمع‌آوری مشاهدات درون تصویر در فضای پارامترها، ویژگی‌های مورد نظر تصویر را که معمولاً خطوط راست یا بیضی می‌باشند، پیدا می‌کند. مشکل تبدیل هاف در بار محاسباتی بالای آن است؛ زیرا همان‌گونه که در بخش 5-1-2 توضیح داده شد، در تبدیل هاف لازم است بر روی یک آرایه انباره که وظیفه جمع‌آوری مشاهدات تصویر در فضای پارامترها را دارد، محاسبات تکراری انجام شود. به طور دقیق، لازم است به ازای هر مشاهده در تصویر که در این تحقیق پیکسل‌های منتخب برای نشانه‌های خطی می‌باشد، مجموعه‌ای از سلول‌های انباره به ازای یک واحد افزایش یابند. در صورتی که ابعاد انباره هاف و ابعاد تصویر برابر $n \times n$ باشد، پیچیدگی محاسباتی تبدیل هاف برابر $O(n^3)$ خواهد بود [47].

همان‌گونه که پیش از این اشاره شد، به منظور نگاشت بین سیستم مختصاتی تصویر با سیستم مختصاتی مدل میدان، لازم است نشانه‌های خطی در دنباله‌ای از تصویر ردیابی گردند. در این حالت، فیلتر کالمن می‌تواند به صورت یک روش پیش‌بینی کننده ظاهر شود، چرا که فیلتر کالمن قادر است به صورت بازگشتی حالت یک ویژگی را تخمین بزند. فیلتر کالمن همچنین با پیش‌بینی و تخمین کواریانس خطا، قادر است دامنه جستجوی ویژگی مورد نظر را در فضای مربوطه کاهش دهد.

ترکیب تبدیل هاف با فیلتر کالمن چندین مزیت دارد؛ از جمله آنکه تبدیل هاف نسبت به نویز، شکستگی خطوط و هم‌پوشانی جزئی، انعطاف‌پذیری بالایی دارد و در نتیجه، قادر است مقدار خروجی مورد لزوم برای فیلتر کالمن را حتی در چنین حالت‌هایی فراهم کند. از طرف دیگر، فیلتر کالمن با تخمین حالت و کواریانس خطا می‌تواند جستجوی ویژگی تصویر را که در اینجا همان خطوط راست است، در فضای آرایه انباره‌ای هاف به فضای کوچکی محدود نماید. این مزیت کارایی محاسباتی تبدیل هاف را به شدت بالا می‌برد. همچنین فیلتر کالمن با در نظر گرفتن عدم قطعیت در تخمین حالت، قادر است تغییرات غیر قابل انتظار در ویژگی مورد نظر را که در این پژوهش حرکت ناگهانی دوربین است، بپوشاند.

میلز و همکارانش [47] فیلتر کالمن را برای ردیابی نقاط اوج در فضای هاف، به منظور ردیابی مجموعه خطوط راست در تصویر، مورد استفاده قرار دادند. در این روش مکان نقاط اوج در فضای پارامترها به عنوان مقدار خروجی برای فیلتر کالمن به کار گرفته شده است. فیلتر کالمن نیز پارامترهای حرکتی مجموعه خطوط مورد نظر را تخمین می‌زند.

در مدل میلز، هر یک از خطوط واقع در مجموعه n خطی، می‌توانند جهت‌های متفاوتی داشته باشند؛ ولی فرض می‌شود سرعت انتقالی و چرخشی برای تمامی آن‌ها در صفحه مسطح مدل میدان، ثابت باشد. به عبارت دیگر، فرض می‌شود موقعیت تمامی خطوط در بین 2 فریم متوالی، حول نقطه (x, y) با سرعت چرخشی ω و با سرعت انتقالی (u, v) تغییر می‌یابد. با این فرض، چون تمامی خطوط دارای پارامترهای حرکتی یکسان هستند، ردیابی خطوط به صورت مجموعه‌ای، و نه به صورت جداگانه انجام می‌شود که این امر دقت تخمین فیلتر کالمن را بالاتر می‌برد.

بردار حالت در فیلتر کالمن میلز، شامل مختصات قطبی خطوط، (ρ_i, θ_i) ، $1 \leq i \leq n$ ، و پارامترهای حرکتی یعنی مرکز چرخش، (x, y) ، سرعت چرخش، ω ، و سرعت انتقالی، (u, v) ، است. حالت‌ها در 2 فریم متوالی در صورت عدم وجود نویز به این صورت با یکدیگر ارتباط دارند:

$$\begin{aligned} s_{t+1} &= [x_{t+1} \quad y_{t+1} \quad u_{t+1} \quad v_{t+1} \quad \omega_{t+1} \quad \rho_{1,t+1} \quad \theta_{1,t+1} \quad \cdots \quad \rho_{n,t+1} \quad \theta_{n,t+1}] \\ &= [x_t + u_t \quad y_t + v_t \quad u_t \quad v_t \quad \omega_t \quad r_{1,t} \quad \theta_{1,t} + \omega_t \quad \cdots \quad r_{n,t} \quad \theta_{n,t} + \omega_t] = f(s_t) \end{aligned} \quad (33-5)$$

که در آن $r_{i,t}$ به صورت زیر تعریف می‌شود:

$$r_{i,t} = \rho_{i,t} - x_t \cos(\theta_{i,t}) - y_t \sin(\theta_{i,t}) + (x_t + u_t) \cos(\theta_{i,t} + \omega_t) + (y_t + v_t) \sin(\theta_{i,t} + \omega_t) \quad (34-5)$$

همچنین رابطه اندازه‌گیری در فیلتر کالمن از روی تبدیل هاف و به صورت زیر می‌باشد:

$$m_t = [\rho_{1,t}, \theta_{1,t}, \dots, \rho_{n,t}, \theta_{n,t}]^T \quad (35-5)$$

همان‌طور که ملاحظه می‌گردد، رابطه (34-5) غیر خطی است، به همین علت فیلتر کالمن توسعه یافته مورد استفاده قرار گرفته است.

به منظور در نظر گرفتن خطا در این مدل، 3 مورد زیر در نظر گرفته شده است:

• فرض این که پارامترهای حرکتی u ، v و w برای تمامی خطوط یکسان است، ممکن است همواره صحیح نباشد.

این خطا برای u ، v و w به ترتیب با $\mathcal{E}_u$ ، $\mathcal{E}_v$ و $\mathcal{E}_w$ نشان داده می‌شود و واریانس‌های آن‌ها به ترتیب σ_u^2 ، σ_v^2 و σ_w^2 می‌باشد. شایان ذکر است، این خطا بر روی تمامی خطوط به طور یکسان اثر می‌کند.

• این امکان وجود دارد که هر خط نسبت به مدل به طور مجزا انحراف داشته باشد. این حالت در نظر گرفتن حرکت

مجموعه خطوط به صورت صلبی شکل را نقض می‌کند. خطای ρ و θ برای i امین خط به ترتیب با $\mathcal{E}_{\rho_i}$ و $\mathcal{E}_{\theta_i}$ و واریانس‌های آن‌ها به ترتیب با $\sigma_{\rho_i}^2$ و $\sigma_{\theta_i}^2$ نشان داده می‌شود.

• مقدار اندازه‌های ρ و θ در رابطه (35-5) به علت کوانتیزاسیون انباره هاف دقیق نمی‌باشد. برای i امین خط این

خطا با e_{ρ_i} و e_{θ_i} و با واریانس‌های s_{ρ_i} و s_{θ_i} نشان داده می‌شود.

فرض می‌شود تمامی خطاهای یاد شده، مستقل، با میانگین صفر و توزیع آن به صورت گاوسی است. خطاهای $\mathcal{E}_u$ ،

$\mathcal{E}_v$ ، $\mathcal{E}_w$ ، $\mathcal{E}_{\rho_i}$ و $\mathcal{E}_{\theta_i}$ همگی بر روی انباره‌های u ، v ، w ، ρ_i و θ_i در رابطه‌های (33-5) و (34-5) و خطاهای e_{ρ_i} و

e_{θ_i} به ρ_i و θ_i در رابطه (35-5) اضافه می‌شوند. مجموعه خطاهای یاد شده، 2 بردار خطای فرآیند و اندازه‌گیری زیر را

برای مدل فیلتر کالمن توسعه یافته معرفی می‌کنند:

$$v_t = [\mathcal{E}_u, \mathcal{E}_v, \mathcal{E}_w, \mathcal{E}_{\rho_1}, \mathcal{E}_{\theta_1}, \dots, \mathcal{E}_{\rho_n}, \mathcal{E}_{\theta_n}]^T \quad (36-5)$$

$$w_t = [e_{\rho_1}, e_{\theta_1}, \dots, e_{\rho_n}, e_{\theta_n}]^T \quad (37-5)$$

با اعمال فیلتر کالمن توسعه یافته در هر فریم، موقعیت خطوط، (ρ_i, θ_i) ، تخمین زده می‌شوند. علاوه بر این،

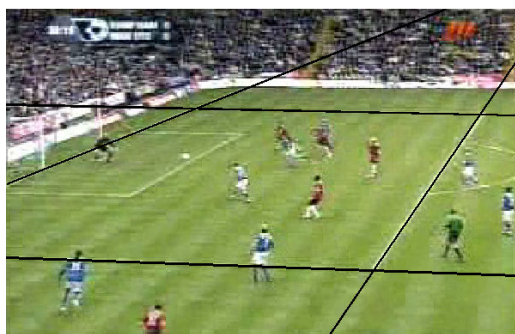
واریانس‌های σ_ρ^2 و σ_θ^2 نیز به دست می‌آیند. با به کارگیری اطلاعات به دست آمده از فیلتر کالمن، جستجو برای خطی در

فضای هاف به جستجوی آن خط در بازه $(\tilde{\rho} - k\sigma_\rho, \tilde{\rho} + k\sigma_\rho)$ در بعد ρ و $(\tilde{\theta} - k\sigma_\theta, \tilde{\theta} + k\sigma_\theta)$ در بعد θ

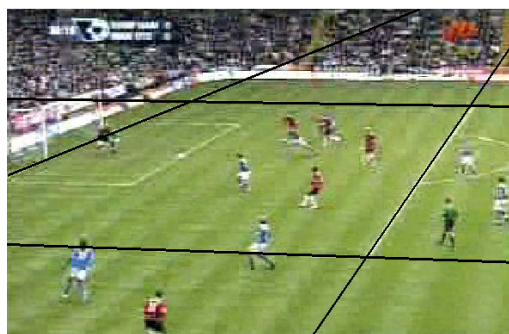
محدود می‌گردد که ضریب ثابت k ، برای کنترل بازه مورد استفاده قرار می‌گیرد. در این پایان‌نامه ضریب k برابر 3 انتخاب

شده است. شکل (8-5) نتایج شناسایی و ردیابی نشانه‌های خطی در یک دنباله تصویر را که با فعل و انفعالات سریع دوربین

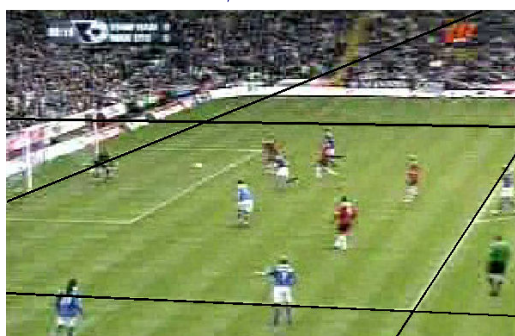
همراه است، نشان می‌دهد.



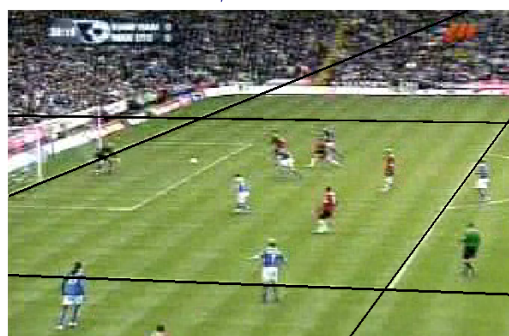
(ب) فریم 5



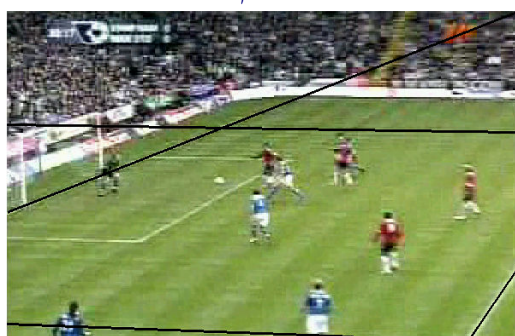
(الف) فریم 1



(د) فریم 14



(ج) فریم 9



(د) فریم 23



(و) فریم 19

شکل 5-8- نتایج شناسایی و ردیابی نشانه‌های خطی در یک دنباله تصویر

5-2- کنار هم قرار دادن دنباله تصویر به منظور ایجاد نمای دور

دنباله تصویر کنار هم قرار داده شده، یک دید دورنما از پس‌زمینه و حرکت اشیاء را به بیننده انتقال می‌دهد. کنار هم قرار دادن دنباله تصویر را می‌توان به 2 دسته ایستا و پویا طبقه‌بندی کرد که در ادامه به جزئیات بیشتری از آن پرداخته می‌شود.

5-2-1- کنار هم قرار دادن ایستا

دنباله تصویر کنار هم قرار داده شده ایستا می‌تواند برای نمایش رویدادهای ایستا، نمونه‌هایی از رخداد‌های مهم آن صحنه در هر لحظه را نشان دهد. برای تولید یک دنباله تصویر کنار هم قرار داده شده ایستا از یک صحنه، لازم است ابتدا پس‌زمینه

تمامی فریم‌ها در این دنباله به یک سیستم مختصات ثابت تراز شده، سپس تصاویر ترازبندی شده با استفاده از یک فیلتر زمانی^۱ مجتمع شوند و در پایان پس‌مانده‌های مهم برای هر فریم (داده‌های مهمی که در حین ایجاد تصویر کنار هم قرار داده شده حذف شده‌اند) که در این جا بازیکنان هستند، بر روی تصویر دور نما قرار داده شوند [48].

5-2-2- کنار هم قرار دادن پویا

با وجود آن که تصویر کنار هم قرار داده شده ایستا، شامل اطلاعات کلی از مجموعه‌ای از فریم‌ها می‌باشد، ولی قادر نیست به صورت جامع، جنبه پویایی دنباله تصویر ویدیویی را نشان دهد. علاوه بر این، در برخی موارد، به دلیل زیاد بودن تعداد بازیکنان در صحنه، تصویر کنار هم قرار داده شده ایستا پر ازدحام و شلوغ می‌شود، به طوری که تشخیص مسیر حرکت هر یک از بازیکنان دشوار می‌گردد.

در تصاویر کنار هم قرار داده شده پویا، محتویات هر تصویر دورنما با جدیدترین اطلاعات آخرین فریم به هنگام می‌شود. در این روش، پس از ایجاد تصویر دورنما از پس‌زمینه تمامی فریم‌ها، پس‌مانده‌های فریم‌های متوالی در هر لحظه بر روی تصویر پس‌زمینه نمایش داده می‌شوند. واضح است در حالت پویا بر خلاف حالت ایستا، قابلیت دستیابی تصادفی به فریم‌های مجزا از دست خواهد رفت [49].

5-2-3- کنار هم قرار دادن دنباله تصویر

به طور کلی، فرآیند ایجاد دنباله تصویر کنار هم قرار داده شده شامل 3 مرحله ترازبندی تصاویر در یک دنباله^۲، مجتمع سازی^۳ پس‌زمینه تصاویر در یک تصویر کنار هم قرار داده شده و نمایش پس‌مانده‌های مهم^۴ برای هر کدام از فریم‌ها بر روی تصویر پس‌زمینه است. در ادامه به جزییات هر یک از مراحل یاد شده پرداخته می‌شود.

5-2-3-الف- ترازبندی دنباله تصویر

ترازبندی ممکن است با مدل‌های حرکتی پارامتری^۲ بعدی مانند تبدیل آفین^۶ پارامتری (برای توضیحات بیشتر به بخش 2-6 مراجعه شود) یا تبدیل درجه دوم^۸ پارامتری (برای توضیحات بیشتر به بخش 4-3-1 مراجعه شود) انجام گیرد و یا اینکه مدل‌های حرکتی^۳ بعدی پیچیده به کار گرفته شود. با توجه به این که در مسابقه فوتبال، بازیکنان بر روی صفحه مسطح زمین فوتبال قرار دارند، مدل حرکتی^۲ بعدی برای ترازبندی دنباله تصویر با مدل میدان در نظر گرفته شده است.

در شکل (5-9) یک تصویر واقعی از مسابقه فوتبال و تصویر تراز شده آن بر روی مدل میدان با 2 روش تبدیل آفین^۶ پارامتری و تبدیل پرسپکتیو^۸ پارامتری نشان داده شده است. همان‌طور که مشاهده می‌شود، در شکل (5-9-ب) نشانه‌های خطی بر خلاف انتظار، کج شده‌اند. اعوجاجات لنز دوربین به خصوص در موارد بزرگ‌نمایی دوربین، باعث چنین رخدادی می‌شوند [41]. بنابراین با به کار بردن تعداد نقاط بیشتر، می‌توان خطای حاصل از اعوجاجات لنز دوربین را کاهش داد. به همین

¹ Temporal Filter

² Alignment

³ Integration

⁴ Significant Residual

علت، تبدیل پرسپکتیو که 4 نقطه متناظر بین تصویر و مدل میدان را به کار می‌برد، مورد استفاده قرار گرفته است (به شکل 5-9-ج) مراجعه شود).



(الف)



(ج)



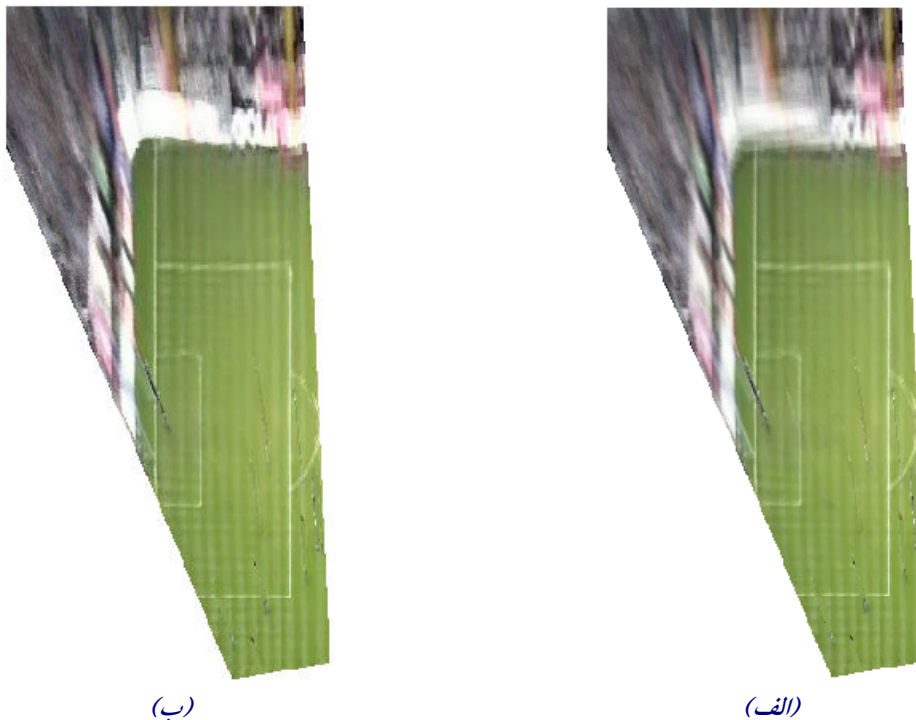
(ب)

شکل 5-9- ترازبندی تصویر به مدل میدان: (الف) تصویر واقعی (ب) تصویر تراز شده به مدل میدان با روش تبدیل آفین (ج) تصویر تراز شده به مدل میدان با روش تبدیل پرسپکتیو - نشانه‌های خطی در شکل (ب) بر خلاف شکل (ج) کج می‌باشند.

5-2-3-ب- مجتمع سازی دنباله تصویر ترازبندی شده

برای ایجاد تصویر کنار هم قرار داده شده، باید پس از ترازبندی فریم‌ها، مجتمع سازی انجام گیرد. البته برای داشتن پس زمینه دنباله تصویر، لازم است قبل از ترازبندی و مجتمع سازی، بازیکنان از دنباله تصویر حذف گردند. برای مجتمع سازی دنباله تصویر، می‌توان با یکی از 2 روش اعمال فیلتر میانگین گیری زمانی یا اعمال فیلتر میانه زمانی بر روی مقادیر روشنایی

تصاویر تراز شده اقدام نمود. شکل (5-10)، نتایج حاصل از کنار هم قرار دادن دنباله تصویر به 2 روش فیلتر میانگین گیری زمانی و فیلتر میانه زمانی را نشان می دهد.



شکل 5-10- تصویر کنار هم قرار داده شده پس زمینه از دنباله تصویر شکل (5-8) الف) به روش فیلتر میانگین گیری زمانی ب) به روش فیلتر میانه زمانی

5-2-3-ج- نمایش پس مانده های مهم بر روی تصویر کنار هم قرار داده شده

یک نمایش دنباله تصویر جامع، شامل دنباله تصویر کنار هم قرار داده شده از پس زمینه هر یک از فریم ها، پارامترهای تبدیل که تصویر کنار هم قرار داده شده را به هر فریم مجزا مربوط می سازد و نمایش پس مانده (بازیکنان شناسایی شده) برای هر یک از فریم ها بر روی تصویر پس زمینه می باشد. در حالت ایستا، بازیکنان به طور پی در پی بر روی تصویر پس زمینه قرار می گیرند. در صورتی که برای حالت پویا، بازیکنان برای هر فریم به طور مجزا نمایش داده می شوند؛ به عبارت دیگر بازیکنان در فریم قبلی جای خود را به بازیکنان در فریم جدید می دهند. در فصل بعدی، مثال های متعددی برای نمایش نمای دور در دنباله تصویر ویدیویی در 2 حالت ایستا و پویا آمده است.

5-3- جمع بندی

در تمامی کارهای انجام شده، برای ایجاد نمای دور از روی تصاویر ویدیویی مسابقه فوتبال، یا ابتدا تعدادی نقاط ویژگی به صورت دستی انتخاب شده و سپس از روی آن عمل تراز بندی تصاویر برای کنار هم قرار دادن آنها انجام گرفته است

[20]، یا اینکه روش حداقل کردن مربعات خطا¹ بدین منظور به کار گرفته شده است [3,5,19,27,28] که بار محاسباتی آن سنگین می‌باشد.

روشی که در این فصل ارائه شد، برای فریم اول با اعمال الگوریتم تبدیل هاف خطی بر روی نقاط منتخب برای نشانه‌های خطی، خطوط واقع در محوطه جریمه را شناسایی و سپس در فریم‌های بعدی با ترکیب الگوریتم "تبدیل هاف" و "فیلتر کالمن" به ردیابی آن‌ها می‌پردازد. پس از آن برای ایجاد نمای دور از صحنه مورد نظر، با استفاده از تبدیل پرسپکتیو بر روی نقاط تقاطع نشانه‌های خطی، پس‌زمینه دنباله تصویر را با مدل میدان ترازبندی و پس از مجتمع‌سازی این دنباله ترازبندی شده، بازیکنان را بر روی آن نمایش می‌دهد. این روش علاوه بر سرعت بیشتر و عدم نیاز به تعامل کاربر، دقت بالایی را به همراه دارد. علاوه بر این، چون در این روش نقاط تقاطع نشانه‌های خطی مورد استفاده قرار می‌گیرند، نقاط مجازی که در خارج از تصویر وجود دارند نیز به کار گرفته می‌شوند.

¹ Least-Square Error

فصل ششم

نتایج آزمایش‌ها

در این فصل نتایج آزمایش‌های انجام شده در 5 بخش زیر ارایه می‌گردد:

- شناسایی بازیکنان
- تعیین تیم بازیکنان
- شناسایی و ردیابی نشانه‌های خطی
- ردیابی بازیکنان
- ایجاد نمای دور

6-1- نتایج شناسایی بازیکنان

به منظور ارزیابی دقت شناسایی بازیکنان به روش ارایه شده (بخش 3-2)، سیستم بر روی 4 مسابقه در ورزشگاه‌های مختلف مورد آزمایش قرار گرفت که نتایج آن به روش فیلتر میانه به همراه عملگر مورفولوژی باز کردن و در 2 حالت اعمال اصلاحات کاهش خطا و بدون آن در جدول (6-1) آمده است. همان‌طور که ملاحظه می‌شود اعمال اصلاحات یاد شده، کاهش خطا را به میزان قابل توجه تضمین می‌نماید. لازم به ذکر است، در اغلب مواردی که سیستم با خطا روبه‌رو می‌شود، بازیکنان در تصویر به صورت مات ظاهر شده‌اند. بنابراین، با توجه به نتایج به دست آمده، سیستم قادر به شناسایی بازیکنان با خطای قابل قبول می‌باشد.

شکل (6-1) اولین فریم از یک دنباله تصویر نمونه را نشان می‌دهد. همان‌طور که پیش از این اشاره شد، لازم است بازیکنان در فریم اول شناسایی و سپس در فریم‌های متوالی ردیابی گردند. بازیکنان شناسایی شده در این فریم در شکل (6-2) نشان داده شده است. شایان ذکر است، بازیکن شماره 4 به دلیل مات بودن در تصویر (این بازیکن با نمای بزرگ در شکل (6-3) نشان داده شده است) به طور خودکار شناسایی نشده، بلکه به صورت دستی انتخاب شده است. همچنین بازیکنان 5 و 6 از طریق روش پیشنهادی به صورت هم‌پوشان شناسایی شده‌اند که به منظور تحلیل‌های بعدی به طور مجزا انتخاب شده‌اند.

جدول 6-1- دقت شناسایی بازیکنان به روش فیلتر میانه به همراه عملگر مورفولوژی باز کردن و در 2 حالت اعمال اصلاحات کاهش خطا و بدون اعمال آن بر روی تصاویر ویدیویی از 4 مسابقه در ورزشگاه‌های مختلف

تعداد بازیکنان شناسایی شده به طور صحیح	تعداد بازیکنان شناسایی نشده شناسایی شده به اشتباه	درصد بازیکنان شناسایی شده به طور صحیح	تعداد بازیکنان شناسایی شده به طور صحیح	تعداد بازیکنان شناسایی نشده شناسایی شده به طور صحیح	تعداد بازیکنان شناسایی شده به طور صحیح
تصاویر ویدیویی مسابقه اول (52 فریم)	با اعمال قوانین کاهش خطا	240	17	11	98/5
بدون اعمال قوانین کاهش خطا	248	9	37	84/3	
تصاویر ویدیویی مسابقه دوم (118 فریم)	با اعمال قوانین کاهش خطا	472	14	32	91/1
بدون اعمال قوانین کاهش خطا	479	7	70	86/1	
تصاویر ویدیویی مسابقه سوم (85 فریم)	با اعمال قوانین کاهش خطا	358	27	13	89/9
بدون اعمال قوانین کاهش خطا	370	15	81	79/3	
تصاویر ویدیویی مسابقه چهارم (206 فریم)	با اعمال قوانین کاهش خطا	678	14	21	95/0
بدون اعمال قوانین کاهش خطا	682	10	110	85/0	
کل تصاویر ویدیویی (461 فریم)	با اعمال قوانین کاهش خطا	1748	72	77	92/1
بدون اعمال قوانین کاهش خطا	1779	41	298	83/9	



شکل 6-1- اولین فریم از یک دنباله تصویر نمونه



شکل 6-2- بازیکنان شناسایی شده در اولین فریم از دنباله تصویر نمونه



شکل 6-3- بازیکن شماره 4 در تصویر شکل (6-2) با نمای بزرگ: بازیکن به صورت کامل مات در تصویر ظاهر شده است.

6-2- نتایج تعیین تیم بازیکنان

آزمایش‌هایی برای بررسی عملکرد سیستم مبتنی بر 3 روش "هیستوگرام مشترک"، "هیستوگرام باز تصویر شده" و "ماشین بردار پشتیبان"، و بر روی 4 مسابقه مختلف انجام گرفت. در این آزمایش‌ها، برای تعیین نمونه‌های هیستوگرام مولفه‌های رنگی هر تیم، حالت‌ها و موقعیت‌های مختلف بازیکنان در نظر گرفته شده است. همچنین در این آزمایش‌ها، تنها بازیکنانی مد نظر قرار گرفته‌اند که هم‌پوشانی نداشته باشند. نتایج به دست آمده در جدول (6-2) آورده شده است. همان‌طور که مشاهده می‌شود، بهترین نتایج با به کار بردن ماشین بردار پشتیبان حاصل شده است. شایان ذکر است، اغلب نتایج نادرست سیستم، به علت مات بودن بازیکن در تصویر یا مشابه بودن لباس بازیکنان به یکدیگر می‌باشد. با این وجود، همان‌گونه که مشاهده می‌شود، سیستم با دقت بالا به طبقه‌بندی تیم بازیکنان می‌پردازد.

جدول 6-2- دقت تعیین تیم متناظر برای بازیکنان بدون هم‌پوشان بر روی 4 مسابقه مختلف و مبتنی بر 3 روش متفاوت

ماشین بردار پشتیبان	هیستوگرام باز تصویر شده	هیستوگرام مشترک	
99/1	88/3	88/3	تصاویر ویدیویی مسابقه اول (52 فریم)
92/9	83/0	84/1	تصاویر ویدیویی مسابقه دوم (118 فریم)
94/0	88/6	90/2	تصاویر ویدیویی مسابقه سوم (85 فریم)
98/3	92/5	96/1	تصاویر ویدیویی مسابقه چهارم (206 فریم)
96/1	88/5	90/6	کل تصاویر ویدیویی (461 فریم)

6-3- نتایج شناسایی و ردیابی نشانه‌های خطی

همان‌طور که پیش از این اشاره شد (به شکل (4-6) مراجعه شود) اولین گام سیستم، یافتن پیکسل‌های منتخب برای نشانه‌های خطی می‌باشد. شکل (4-6) پیکسل‌هایی از تصویر شکل (1-6) را که در همسایگی 10×10 آن‌ها بیش از 75٪ پیکسل‌ها غیر چمن بوده یا متعلق به آرم‌های تلویزیونی باشد، نشان می‌دهد. این عمل باعث حذف پیکسل‌هایی که متعلق به بازیکنان، تابلوهای تبلیغاتی یا آرم‌های تلویزیونی است، می‌شود؛ زیرا در این نواحی، مجموعه‌ای از پیکسل‌های غیر چمن در کنار یکدیگر قرار گرفته‌اند.

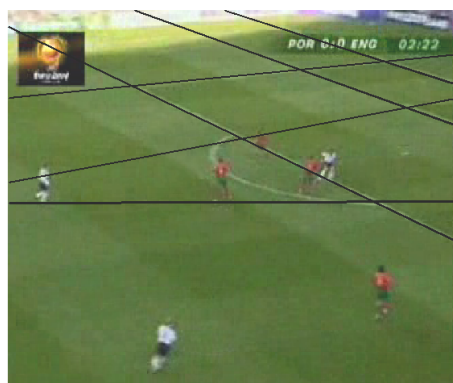


شکل 6-4- پیکسل‌هایی که در همسایگی 10×10 آن‌ها در تصویر شکل (6-1) بیش از 75٪ پیکسل‌ها غیر چمن بوده یا متعلق به آرم‌های تلویزیونی باشد

همان‌گونه که در شکل (6-4) مشاهده می‌شود، قسمت‌هایی از چمن نیز که روشن‌تر از سایر قسمت‌های دیگر است (در کناره‌های بالای تصویر)، به اشتباه به عنوان پیکسل‌های تابلو تبلیغاتی شناخته شده است. شکل (6-5) نیز پیکسل‌های منتخب برای نشانه‌های خطی پس از حذف نواحی مشخص شده در شکل (6-4) و اعمال شرایط (5-1) و (5-2) را نشان می‌دهد. نتایج الگوریتم تبدیل هاف بر روی پیکسل‌های منتخب نیز در شکل (6-6) نشان داده شده است که در پایان، به منظور طبقه‌بندی خطوط به کار بر عرضه می‌شود.

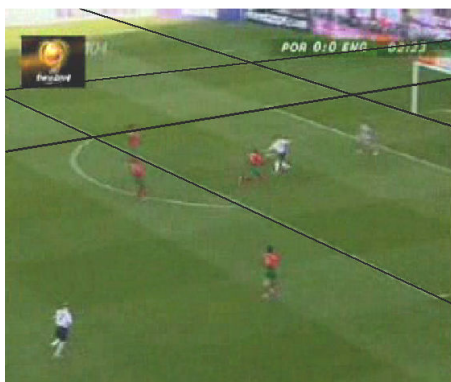


شکل 6-5- پیکسل‌های منتخب برای نشانه‌های خطی از تصویر شکل (6-1)

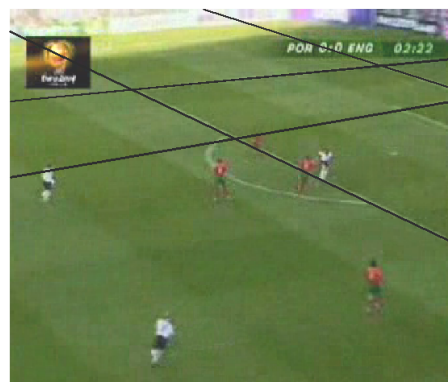


شکل 6-6- نشانه‌های خطی شناسایی شده با استفاده از الگوریتم تبدیل هاف بر روی پیکسل‌های منتخب در شکل (6-4)

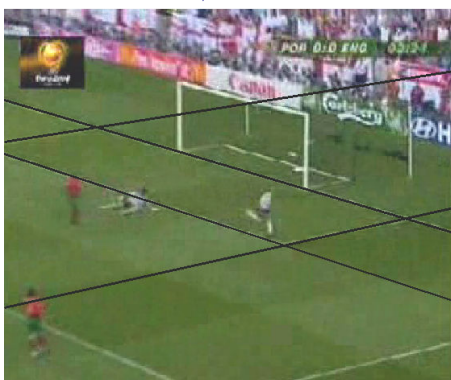
نتایج به دست آمده از ردیابی نشانه‌های خطی برای دنباله تصویر نمونه در شکل (6-7) نشان داده شده است. با توجه به این شکل مشاهده می‌شود، نشانه‌های خطی ردیابی شده در فریم‌های اول و 20ام با دیگر فریم‌ها متفاوت است. از آنجایی که ممکن است نشانه‌های خطی قابل رویت در فریم‌های متوالی یکسان نباشد، لازم است ردیابی نشانه‌های خطی در چنین حالتی متوقف شده و عمل شناسایی و ردیابی نشانه‌های خطی از نو آغاز گردد. علاوه بر این، نتایج شناسایی و ردیابی نشانه‌های خطی در یک دنباله تصویر نمونه دیگر در شکل (6-8) نشان داده شده است.



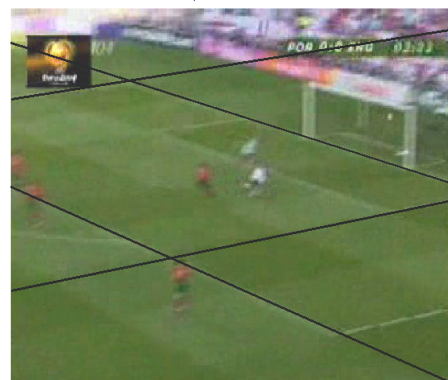
(ب) فریم 20



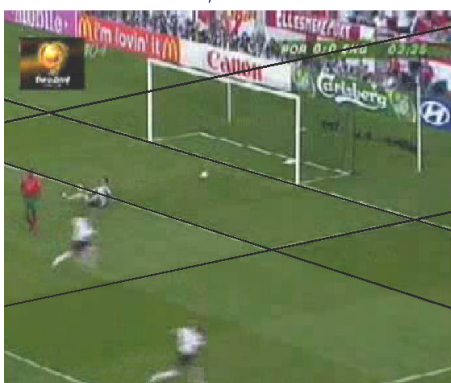
(الف) فریم 2



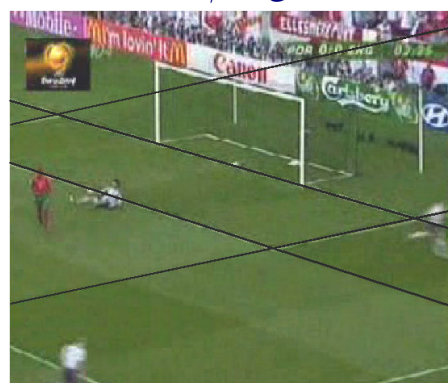
(د) فریم 60



(ج) فریم 40

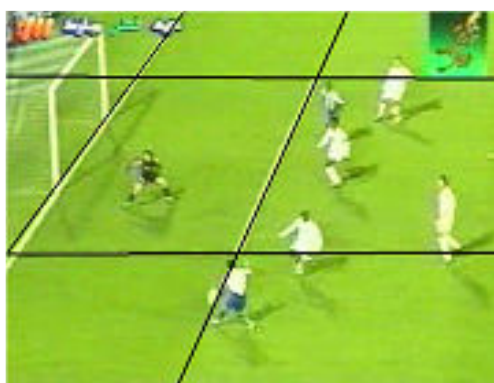


(ی) فریم 90

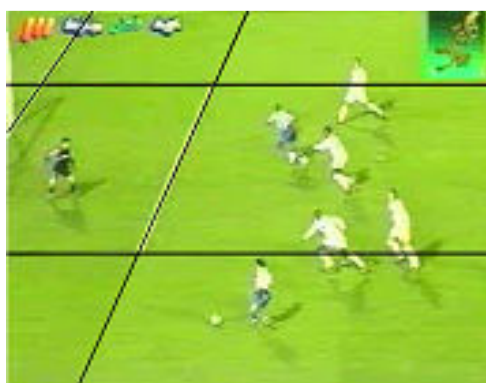


(و) فریم 80

شکل 6-7- نتایج ردیابی نشانه‌های خطی در دنباله تصویر نمونه برای 90 فریم



(ب) فریم 15



(الف) فریم 1



(د) فریم 45



(ج) فریم 30

شکل 6-8- نتایج شناسایی و ردیابی نشانه‌های خطی در یک دنباله تصویر نمونه

6-4- نتایج ردیابی بازیکنان

شکل (6-9) نتایج ردیابی بازیکنان برای فریم دوم از دنباله تصویر شکل (6-7) را نشان می‌دهد. در شکل (6-9) مشاهده می‌شود که بازیکن شماره 4، ناحیه کمتری نسبت به شکل (6-2) به خود اختصاص داده است؛ به عبارت دیگر، نتیجه شناسایی بازیکن در فریم دوم با دقت بیشتری انجام گرفته است. همچنین مشاهده می‌شود، بازیکنان شماره 5 و 6 به صورت هم‌پوشان شناسایی شده‌اند.



شکل 6-9- بازیکنان ردیابی شده در فریم دوم از دنباله تصویر شکل (6-7)

از آنجایی که بازیکنان 5 و 6 به تیم‌های متفاوت متعلق هستند، روش مبتنی بر طبقه‌بندی پیکسل‌ها برای شناسایی بازیکنان مورد استفاده قرار گرفته است. شکل (6-10) ناحیه مورد جستجو و نتیجه شناسایی هر یک از این بازیکنان را به روش مبتنی بر طبقه‌بندی پیکسل‌ها نشان می‌دهد.



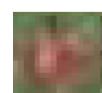
(ج)



(د)



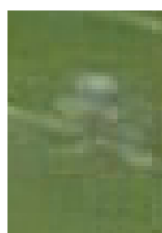
(الف)



(ب)

شکل 6-10- شناسایی بازیکنان در حالت هم‌پوشانی با تیم‌های مختلف (الف) محدوده جستجوی بازیکن شماره 5 از شکل (6-9) با نمای بزرگ (ب) محدوده جستجوی بازیکن شماره 6 (ج) پیکسل‌های شناخته شده برای بازیکن شماره 5 (د) پیکسل‌های شناخته شده برای بازیکن شماره 6

در شکل (6-11)، فریم 10ام از دنباله تصویر شکل (6-7) نشان داده شده است. در این فریم، یک بازیکن جدید به کناره‌های تصویر اضافه شده است که با نمای بزرگ نیز نمایش داده شده است. همان‌طور که مشاهده می‌شود، به علت مات بودن بازیکن، رنگ لباس بازیکن بسیار مشابه با رنگ زمین چمن است. در نتیجه الگوریتم شناسایی بازیکن که در کناره‌های تصویر به جستجوی بازیکنان جدید می‌پردازد، قادر به شناسایی بازیکن نمی‌باشد.



(ب)



(الف)

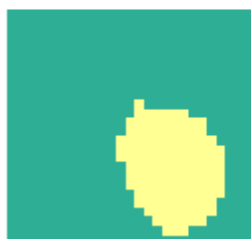
شکل 6-11- بازیکن اضافه شده به تصویر در حالت مات: (الف) فریم 10ام دنباله تصویر شکل (6-7) همراه با بازیکن علامت‌گذاری شده که در تصویر به صورت مات دیده می‌شود (ب) نمای بزرگ از بازیکن مات - به دلیل مات بودن بازیکن در تصویر، سیستم قادر به شناسایی بازیکن نیست.

در فریم 16ام دنباله تصویر شکل (6-7)، 2 بازیکن مشخص شده در شکل (6-12) به صورت هم‌پوشان شناسایی گردیده‌اند. با توجه به این که 2 بازیکن از یک تیم می‌باشند، روش تطبیق کلیشه همبستگی برای شناسایی بازیکن به کار گرفته

شده است. شایان ذکر است، کلیشه بازیکن نشان داده شده در شکل‌های (6-12-ج) و (6-12-د)، از فریم قبلی که به طور مجزا شناسایی شده، استخراج گردیده است. نتیجه اعمال تطبیق کلیشه همبستگی برای شناسایی بازیکن، در شکل (6-12-و) نشان داده شده است. نتایج به دست آمده از شناسایی و ردیابی بازیکنان در دنباله تصاویر شکل (6-7) و (6-8) به ترتیب در شکل‌های (6-13) و (6-14) نشان داده شده است.



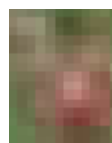
(الف)



(و)



(د)



(ج)



(ب)

شکل 6-12- شناسایی بازیکنان در حالت هم‌پوشانی با تیم هم‌نوع: الف) فریم 16ام دنباله تصویر شکل (6-7) همراه با بازیکنان علامت‌گذاری شده که به صورت هم‌پوشان شناسایی شده‌اند (ب) نمای بزرگ از بازیکنان هم‌پوشان (ج) الگوی مورد نظر بازیکن (د) پیکسل‌های بازیکن شکل (ج) و) پیکسل‌های شناسایی شده با به‌کارگیری روش تطبیق کلیشه همبستگی

در شکل (6-13) مشاهده می‌شود، دروازه‌بان در هیچ یک از فریم‌ها شناسایی نشده است؛ زیرا در هنگام وارد شدن بازیکن به تصویر به صورت مات بوده و الگوریتم شناسایی بازیکن در آن هنگام قادر به شناسایی آن نبوده است. از طرف دیگر، چون الگوریتم، بازیکنان جدید را فقط در کناره‌های تصویر دنبال می‌کند، دروازه‌بان پس از واضح شدن در تصویر (به عنوان مثال در فریم 80) شناسایی نمی‌گردد و در نتیجه سیستم در شناسایی و ردیابی دروازه‌بان با شکست مواجه می‌شود. با این وجود، در فریم 80، مشاهده می‌شود که بازیکن جدید در کناره پایین تصویر، به دلیل وضوح بیشتر، به درستی شناسایی شده است. از طرف دیگر، در فریم 40ام دنباله تصویر نمونه، مشاهده می‌شود با وجود آن که بازیکنان در تصویر مات هستند، ولی به درستی شناسایی شده‌اند؛ زیرا الگوریتم، محدوده جستجوی بازیکن را از روی تقاطع نشانه‌های خطی به درستی تشخیص داده و در نتیجه جستجوی بازیکنان در این محدوده با دقت بیشتری صورت گرفته است.



(ب) فریم 20



(الف) فریم 1



(د) فریم 60



(ج) فریم 40

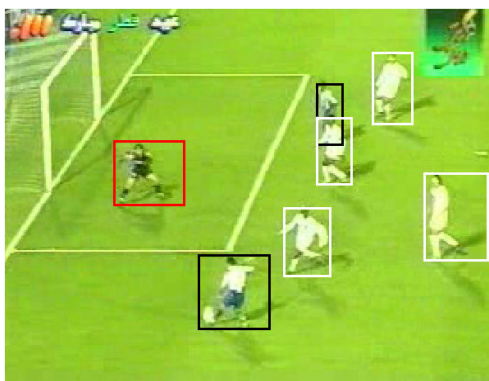


(ی) فریم 90

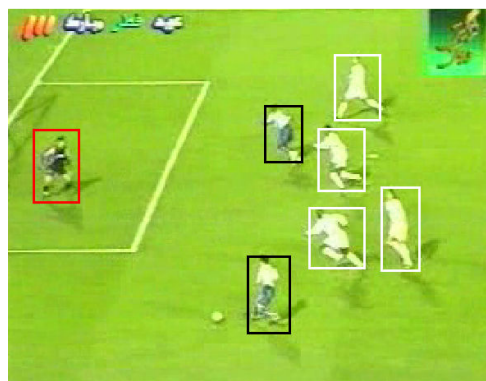


(و) فریم 80

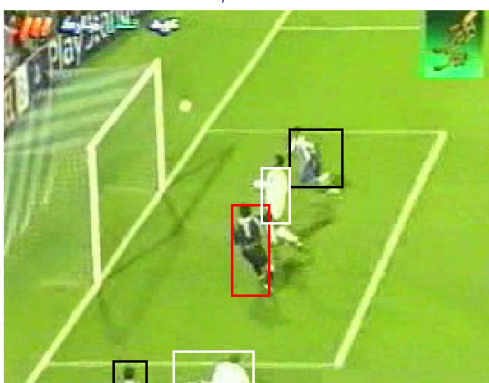
شکل 6-13- نتایج ردیابی بازیکنان در دنباله تصویر شکل (6-7) برای 90 فریم



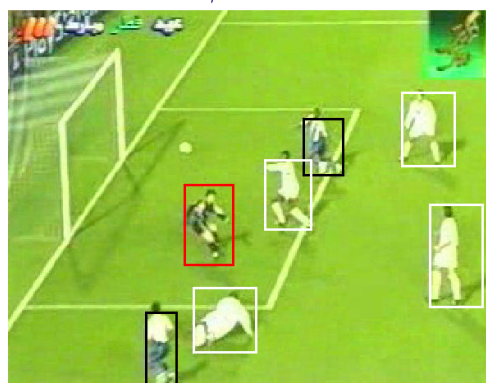
(ب) فریم 15



(الف) فریم 1



(د) فریم 45



(ج) فریم 30

شکل 6-14- نتایج شناسایی و ردیابی بازیکنان در دنباله تصویر نشان داده شده در شکل (6-8)

5-6- نتایج ایجاد نمای دور

تصویر حاصل از کنار هم قرار دادن دنباله تصویر شکل (6-7)، پس از حذف بازیکنان در دنباله تصویر نمونه، با اعمال فیلتر میانگین‌گیری زمانی و فیلتر میانه زمانی، در شکل (6-15) نشان داده شده است. همچنین نتیجه حاصل از قرار دادن بازیکنان حذف شده در این تصویر، در شکل (6-16) نشان داده شده است. همان‌طور که مشاهده می‌شود، در این تصویر (نمای دور ایستا)، دنباله تصویر به سیستم مختصاتی مدل میدان نگاشت شده است. به منظور نمایش بهتر می‌توان این تصویر را به هر سیستم مختصاتی دلخواه نگاشت کرد. به عنوان مثال، نتیجه نگاشت این تصویر به سیستم مختصاتی فریم اول در شکل (6-17) آمده است.



(ب)



(الف)

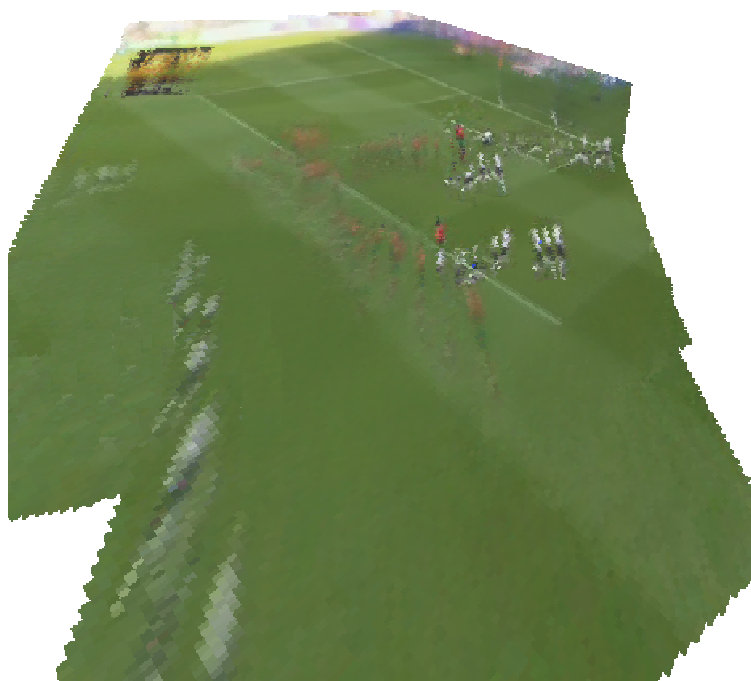
شکل 6-15- تصویر کنار هم قرار داده شده پس از حذف بازیکنان از دنباله تصویر شکل (6)-

(7: الف) با اعمال فیلتر میانگین‌گیری زمانی (ب) با اعمال فیلتر میانه زمانی



شکل 6-16- نتیجه حاصل از قرار دادن بازیکنان حذف شده در دنباله تصویر نمونه بر روی

شکل (6-15-ب)



شکل 6-17- نتیجه نگاشت شکل (6-16) به سیستم مختصاتی فریم اول

در شکل (6-18) نتایج نمای دور به صورت پویا، برای فریم اول و فریم 90ام نشان داده شده است. در این شکل مشاهده می‌شود، بازیکنان، به خصوص در فریم اول، به دلیل تفاوت نما با نمای حاصل از کنار هم قرار دادن دنباله تصویر بسیار کوچک دیده می‌شوند. با این وجود، با نمایش دنباله تصویر حاصل از کنار هم قرار دادن پویا، بیننده درک بهتری نسبت به موقعیت و مسیر هر یک از بازیکنان به دست می‌آورد، علاوه بر این که این نتایج می‌تواند برای تجزیه و تحلیل‌های خودکار مفید واقع شوند.

شکل (6-19) نمای دور ایستا از کنار هم قرار دادن 45 فریم متوالی شکل (6-8) را نشان می‌دهد. در شکل (6-19-الف)، نمای دور حاصل شده در مدل میدان و در شکل (6-19-ب)، نگاشت این نما در سیستم مختصاتی فریم اول نشان داده شده است. خطوط رسم شده در این شکل‌ها مسیر هر بازیکن را نشان می‌دهند و به منظور مشخص شدن جهت حرکت آن‌ها، بازیکنان در فریم‌های اولیه کم‌رنگ‌تر ظاهر شده‌اند. همچنین برای نمایش واضح‌تر بازیکنان و جلوگیری از شلوغ شدن تصویر، بازیکنان تنها در چندین فریم بر روی تصویر نمای دور قرار گرفته‌اند.

شکل (6-20) نیز یک دنباله تصویر دیگر را نشان می‌دهد که نمای دور ایجاد شده از این دنباله تصویر، در حالت ایستا و در سیستم مختصاتی فریم اول، در شکل (6-21) نشان داده شده است. در این تصویر نیز بازیکنان در فریم‌های اولیه کم‌رنگ‌تر ظاهر شده‌اند و مسیرشان به وسیله قطعه خطوط مشخص شده است.

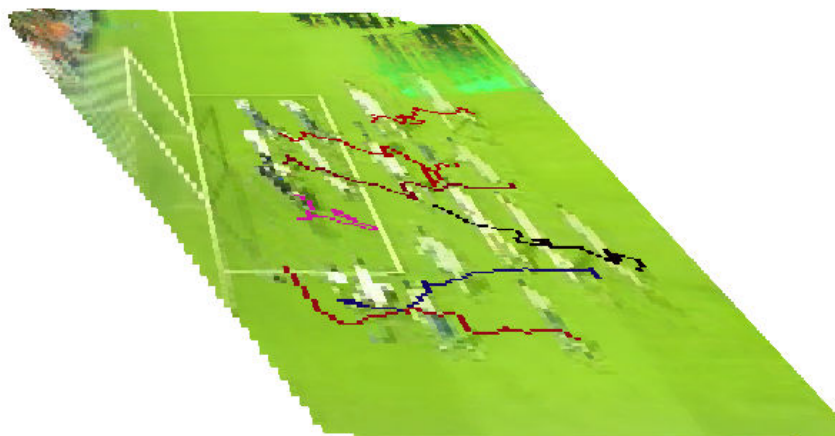


(ب) فریم 90



(الف) فریم 1

شکل 6-18- نتایج نمای دور به صورت پویا



(ب)



(الف)

شکل 6-19- نمای دور ایجاد شده از کنار هم قرار دادن دنباله تصویر شکل (6-8): (الف) نمای

دور ایستا در مدل میدان (ب) نمای دور ایستا در سیستم مختصاتی فریم اول



(ب) فریم 30



(الف) فریم 1

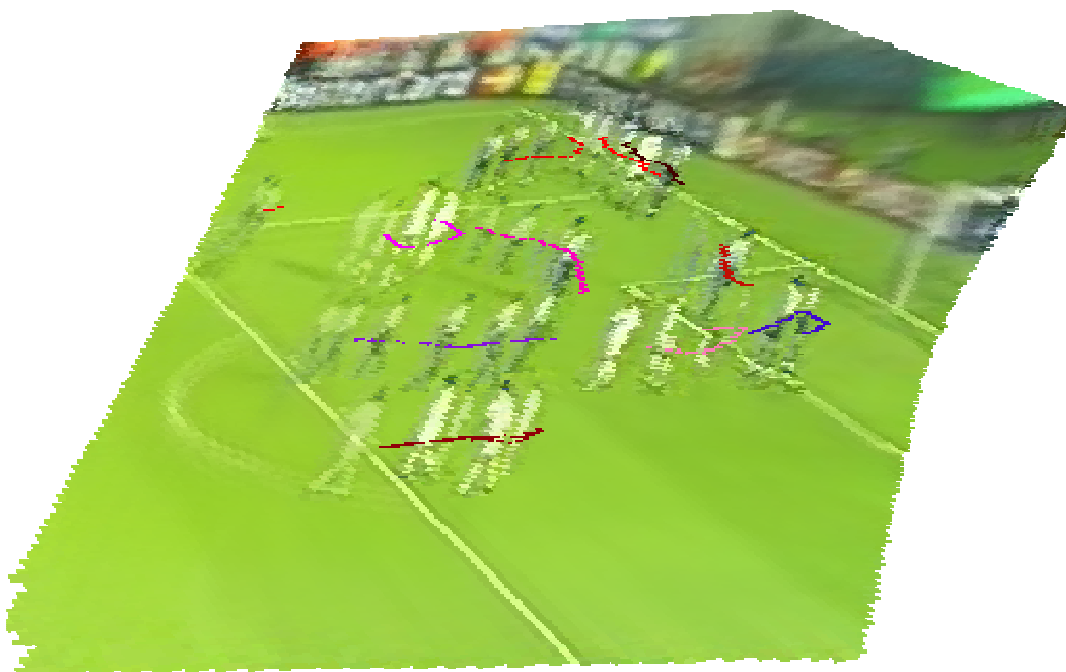


(د) فریم 100



(ج) فریم 60

شکل 6-20- فریم‌هایی از یک دنباله تصویر نمونه



شکل 6-21- نمای دور ایجاد شده از دنباله تصویر شکل (6-20) در حالت ایستا و در سیستم مختصاتی فریم اول

فصل هفتم

نتیجه گیری و پیشنهادات

در این پژوهش، سیستمی خودکار برای تجزیه و تحلیل صحنه‌های گل در تصاویر ویدیویی مسابقات فوتبال ارائه شده است. ابتدا چمن از دیگر اشیاء میدان بازی با استفاده از "هیستوگرام تصویر"، جدا شده است. سپس بازیکنان با اعمال فیلتر میانه و عملگر مورفولوژی "باز کردن" و با در نظر گرفتن اصلاحات برای کاهش خطا، شناسایی شده‌اند. پس از شناسایی بازیکنان، تیم بازیکنان بر اساس روش "ماشین بردار پشتیبان" مشخص شده است. ردیابی بازیکنان در فریم‌های متوالی با استفاده از موقعیت بازیکنان در مدل میدان انجام شده است که برای یافتن موقعیت بازیکنان در مدل میدان، نقاط تقاطع نشانه‌های خطی مورد استفاده قرار گرفته است. نشانه‌های خطی نیز با ترکیب الگوریتم "تبدیل هاف" و "فیلتر کالمن توسعه یافته" ردیابی شده‌اند و در پایان، با کنار هم قرار دادن دنباله تصویر، نمای دوری از صحنه گل ایجاد گردیده است.

در روش پیشنهادی برای استخراج زمین چمن، همانند اغلب روش‌های موجود، فرض غالب بودن آن در فضای تصویر به کار گرفته شده است. با این وجود، روش ارائه شده بر خلاف روش‌های دیگر، نواحی چمن را بدون وابستگی به تصاویر ویدیویی و شرایط زمین چمن به طور بهینه دست می‌آورد؛ به طوری که برای تمامی تصاویر مورد آزمایش، پارامترهای شناسایی چمن یکسان و قابل اعمال است.

به منظور شناسایی بازیکنان، روشی جدید مبتنی بر فیلتر میانه و عملگر مورفولوژی باز کردن ارائه شده است. این روش، پیکسل‌هایی را که در کنار یکدیگر مجتمع هستند به عنوان بازیکن شناسایی می‌کند. اعمال عملگر "مورفولوژی باز کردن" نیز در مواردی که بازیکنان در فاصله‌های نزدیک به هم قرار دارند، باعث تشخیص ناحیه‌های جداگانه می‌شود. علاوه بر این، همان‌طور که در فصل ششم نشان داده شد، با اعمال اصلاحات به منظور کاهش خطا، شناسایی بازیکنان با دقت بیشتری انجام می‌شود. البته در مواردی که بازیکن به صورت مات در تصویر ظاهر می‌شود، الگوریتم قادر به شناسایی بازیکن نخواهد بود؛ زیرا در این حالت، پیکسل‌های متعلق به بازیکن به عنوان پیکسل‌های چمن شناسایی می‌گردد.

تعیین تیم بازیکنان با استفاده از ماشین بردار پشتیبان نتایج بهتری نسبت به روشهای دیگر از جمله مقایسه توزیع هیستوگرام دارد. نحوه عملکرد روش‌هایی که توزیع هیستوگرام را به کار می‌برند، بدین شکل است که برای هر تیم، هیستوگرام بازیکنان، در حالت‌های مختلف به صورت یک هیستوگرام واحد ذخیره شده و آنگاه با مقایسه توزیع هیستوگرام بازیکن با هیستوگرام معرف هر تیم، برچسب‌گذاری بازیکن مجهول مشخص می‌شود. در صورتی که با استفاده از ماشین بردار پشتیبان، پارامترهای طبقه‌بندی از روی تمامی هیستوگرام‌های بازیکنان محاسبه می‌شود؛ به طوری که عمل دسته‌بندی داده‌ها که همان توزیع هیستوگرام است، به صورت بهینه انجام گیرد. همان‌طور که در قسمت نتایج حاصل از تعیین تیم بازیکنان در فصل ششم بیان

شد، در مواردی که بازیکن به صورت مات در تصویر ظاهر گردد یا لباس بازیکن در اثر دقت پایین، با لباس بازیکن تیم دیگر مشابه باشد، الگوریتم اشتباه عمل می‌کند. با این وجود، دقت این روش نسبت به سایر روش‌های دیگر نظیر مقایسه توزیع هیستوگرام بیشتر است.

همان‌گونه که از قبل بیان شد، الگوریتم ردیابی بازیکنان و ایجاد نمای دور بر اساس نشانه‌های خطی موجود در زمین عمل می‌کند. نشانه‌های خطی پس از شناسایی و طبقه‌بندی در اولین فریم، در فریم‌های متوالی با ترکیب الگوریتم "تبدیل هاف" و "فیلتر کالمن توسعه یافته" ردیابی شده‌اند. فیلتر کالمن، با تخمین حالت و کواریانس خطا، قادر است جستجوی خطوط راست در فضای آرایه انباره‌ای هاف را به فضای کوچکی محدود نماید که این امر باعث کارایی محاسباتی تبدیل هاف می‌گردد.

واضح است در صورتی که نشانه‌های خطی موجود در تصویر دقت پایینی داشته باشند، به صورتی که الگوریتم تبدیل هاف قادر به شناسایی خطوط نباشد، روش پیشنهادی برای ردیابی بازیکنان و ایجاد نمای دور کاملاً با شکست مواجه می‌شود. همچنین در حالتی که نشانه‌های خطی به تعداد کافی در تصویر مشاهده نشوند، نمی‌توان روش پیشنهادی را مورد استفاده قرار داد. در چنین شرایطی، استفاده از روش‌های حداقل کردن مربعات خطا و مانند آن، برای یافتن پارامترهای تبدیل بین فریم‌های دنباله تصویر می‌تواند مفید واقع شود.

چون در صحنه‌های گل، حرکت بازیکنان با حرکت کاملاً ناشناخته و ناگهانی دوربین ترکیب می‌شود و از هیچ مدل خاصی تبعیت نمی‌کند، استفاده از روش‌های تخمین حرکت به شیوه‌های معمول برای ردیابی آن‌ها کارایی مناسبی ندارد. با این وجود، موقعیت واقعی یک بازیکن یعنی موقعیت بازیکن نسبت به مدل میدان در شرایط مختلف تغییر چندانی ندارد. در نتیجه، روش به کار گرفته شده برای ردیابی بازیکنان، بر اساس موقعیت بازیکن نسبت به مدل میدان می‌باشد که بدین منظور تبدیل پرسپکتیو بر روی تقاطع نشانه‌های خطی مورد استفاده قرار گرفته است.

نتایج به دست آمده از ردیابی بازیکنان در شرایط مختلف مانند هم‌پوشانی، دقت پایین تصویر بازیکنان، حرکت سریع دوربین، و با تصاویر ویدیویی مختلف، حاکی از عملکرد مناسب سیستم می‌باشد. همان‌طور که در نتایج به دست آمده اشاره گردید، بازیکنان پس از شناسایی، حتی در صورتی که به دلیل حرکت سریع دوربین به صورت مات در فریم‌های بعدی ظاهر شوند، به درستی ردیابی می‌گردند. همچنین با به کار بردن روش مبتنی بر طبقه‌بندی پیکسل‌ها برای حالت هم‌پوشانی بازیکن با بازیکنان غیر هم تیم و به کار بردن روش تطبیق کلیشه برای حالت هم‌پوشانی بازیکنان هم تیم، ردیابی بازیکنان در صورت هم‌پوشانی با موفقیت انجام می‌گردد.

به منظور ایجاد نمای دور، تبدیل پرسپکتیو برای ترازبندی دنباله تصویر بر روی نقاط تقاطع نشانه‌های خطی اعمال شده است. پس از به دست آمدن دنباله تصویر تراز شده در مدل میدان، بازیکنان از این دنباله تصویر حذف گردیده و فیلتر زمانی برای ایجاد تصویر نمای دور پس‌زمینه اعمال گردیده است و در نهایت، بازیکنان در این تصویر به نمایش درآمده‌اند. در راستای ایجاد نمای دور، 2 نمای ایستا و پویا معرفی شده‌اند که نتایج آن در فصل ششم نشان داده شده است.

در سیستم‌های موجود، برای ترازبندی دنباله تصویر به مدل میدان، انتخاب دستی نقاط ویژگی یا روش کمترین مربعات خطا مورد استفاده قرار گرفته است. همان‌طور که بیان شد، در سیستم پیاده‌سازی شده، تقاطع نشانه‌های خطی برای ترازبندی دنباله تصویر به مدل میدان مورد استفاده قرار گرفته است. این روش، با سرعت بالا و بدون نیاز به تعامل کاربر، قادر به ایجاد نمای دور با دقت بالا می‌باشد.

به منظور ادامه فعالیت در این زمینه، موارد زیر پیشنهاد می‌گردد:

• شناسایی و ردیابی توپ

به منظور تجزیه و تحلیل کامل تصاویر ویدیویی مسابقه فوتبال، نیاز است علاوه بر ردیابی نشانه‌های خطی، توپ شناسایی، ردیابی و موقعیت آن در مدل میدان به دست آید. واضح است بدون داشتن مسیر توپ، درک صحنه گل کامل نمی‌باشد. روش‌هایی برای شناسایی و ردیابی توپ معرفی شده‌اند که در فصل اول به آن اشاره شده است. با این وجود، طبق نتایج گزارش شده، هیچ یک از روش‌های موجود با دقت بالا قادر به شناسایی توپ نمی‌باشند.

• تشخیص آفساید

واضح است که برای تشخیص آفساید باید موقعیت بازیکنان و توپ در دنباله تصویر نسبت به مدل میدان مشخص باشد. چون مسیر حرکت توپ، بر خلاف مسیر حرکت بازیکنان، بر روی فضای 3 بعدی است، تعیین موقعیت توپ به سادگی امکان‌پذیر نیست.

• رفع مشکل هم‌پوشانی در حالت‌های پیچیده

در برخی از صحنه‌های گل، بازیکنان در محوطه جریمه به گونه‌ای جمع می‌شوند که چندین بازیکن با یکدیگر هم‌پوشانی دارند. در صورتی که تعداد بازیکنان هم‌پوشان زیاد گردد، روش پیشنهادی برای رفع مشکل هم‌پوشانی با خطا مواجه می‌شود. همچنین در برخی موارد، یک بازیکن به صورت کامل در پشت بازیکن دیگر قرار می‌گیرد که این امر می‌تواند باعث بروز مشکلاتی در روش ارائه شده برای ردیابی بازیکنان گردد.

• تجزیه و تحلیل نوع حرکت بازیکنان

با داشتن نوع حرکت هر یک از بازیکنان که می‌تواند شامل سانتر، دریبل، دویدن، راه رفتن، تکل، شوت و موارد دیگر باشد، می‌توان تحلیل‌های بیشتری بر روی صحنه‌های گل انجام داد. به عنوان مثال، می‌توان نتایج آماری گل‌های زده شده با حرکات مختلف را به دست آورد. شایان ذکر است، برای رسیدن به این هدف، لازم است تصاویر از دقت بالایی برخوردار باشند.

• مرزبندی و طبقه‌بندی صحنه‌ها

یکی از مباحث مهمی که در مسابقه فوتبال نیز بسیار مورد مطالعه قرار گرفته است، مرزبندی و طبقه‌بندی صحنه‌ها به حالت‌های مختلف مانند صحنه گل، صحنه پنالتی، صحنه کرنر، صحنه اوت، صحنه نمای دور، صحنه نمای نزدیک و غیره می‌باشد. واضح است که انجام این مهم چقدر می‌تواند در تجزیه و تحلیل تصاویر ویدیویی مسابقه فوتبال نقش داشته باشد. به عنوان مثال، در این پایان‌نامه فرض شده است که صحنه‌های گل به صورت دستی مشخص می‌گردند؛ انجام این کار به صورت خودکار می‌تواند بسیار حایز اهمیت باشد.

• ایجاد تصاویر مجازی

با مشخص شدن بازیکنان و توپ به همراه موقعیت آن‌ها در مدل میدان، می‌توان دنباله تصاویر را به صورت مجازی تولید کرد. همان‌طور که از نتایج ایجاد نمای دور پیداست، بازیکنان با وضوح پایین در تصویر نمای دور پس‌زمینه قرار می‌گیرند. در صورتی که به صورت مجازی این عمل انجام گردد، تصاویر با وضوح بالاتری دیده می‌شوند.

• ایجاد زیرنویس ویدیویی

در صورت انجام تجزیه و تحلیل کامل تصاویر ویدیویی مسابقه فوتبال، می‌توان زیرنویس ویدیویی ایجاد کرد. به عنوان مثال، در صورت شناسایی خودکار اخطار کارت قرمز به یک بازیکن، زیرنویس ویدیویی با عنوان اخراج شدن بازیکن خطا کار می‌تواند به صورت خودکار انجام گیرد.