

## استخراج و بازسازی خودکار اطلاعات سه‌بعدی زمین فوتبال در تصاویر تک‌دوربینی کالیبره‌نشده

مرتضی نصیری<sup>۱</sup>، دانشجوی کارشناسی ارشد، علیرضا بهراد<sup>۲</sup>، دانشیار،

۱- دانشکده فنی و مهندسی - دانشگاه شاهد - تهران - ایران - mo.nasiri@shahed.ac.ir

۲- دانشکده فنی و مهندسی - دانشگاه شاهد - تهران - ایران - behrad@shahed.ac.ir

**چکیده:** تخمین مکان و موقعیت سه‌بعدی بازیکنان یکی از بحث برانگیزترین موضوعات در تحلیل ویدئویی بازی فوتبال است. ویدئوهای در دسترس از بازی فوتبال معمولاً با دوربین‌های کالیبره‌نشده تهیه می‌شوند؛ بنابراین کالیبراسیون دوربین و تخمین مکان آن، گام اول در محاسبه موقعیت سه‌بعدی بازیکن‌ها است. برای کالیبراسیون معمولاً به تعدادی نقطه از دنیای واقعی و متناظر تصویری آن نقاط نیاز داریم. بنابراین اولین مرحله در الگوریتم پیشنهادی، استخراج نقاط مشخص در زمین فوتبال به صورت خودکار است. سپس یک روش جدید برای کالیبراسیون دوربین و تخمین مکان آن ارائه می‌شود. این روش از هم‌صفحه بودن نقاط در زمین فوتبال و مکان‌های مشخص آن‌ها برای کالیبراسیون دوربین استفاده می‌کند. در این مقاله یک مدل فرم بسته برای محاسبه پارامترهای مختلف دوربین ارائه شده است. الگوریتم پیشنهادی با داده‌های واقعی و شبیه‌سازی شده تست شده و با الگوریتم‌های مشابه مقایسه گردید. نتایج آزمایشی نشان‌دهنده کارایی بالای الگوریتم پیشنهادی است.

**واژه‌های کلیدی:** کالیبراسیون خودکار دوربین، بازسازی صحنه‌های سه‌بعدی، تحلیل تصاویر بازی فوتبال.

### 3D Reconstruction and Pose Estimation in Monocular Uncalibrated Soccer Images

M. Nasiri<sup>1</sup>, A. Behrad<sup>2</sup>,

1, 2- Faculty of Engineering, Shahed University, Tehran, Iran.

**Abstract:** 3D pose estimation of players' location is one of the challenging issues in the soccer video analysis. The soccer videos are generally captured using uncalibrated cameras, therefore the first step to calculate 3D pose of the players is the camera calibration and estimating its location. For the purpose of calibration, it is necessary to extract some 3D points and their image correspondence. Therefore, in the first stage of the proposed algorithm, we automatically extract points and the required data for calibration. Then a new method for camera calibration is proposed based on the coplanarity of the soccer field. The method uses coplanarity of points in soccer plane and their known location for camera calibration. We derive a closed form model to calculate different camera parameters. We compare the results of the proposed method with those of existing methods and results show the efficiency of the proposed algorithm.

**Keywords:** Automatic camera calibration; 3D scene reconstruction; Soccer video analysis.

تاریخ ارسال مقاله: ۱۳۹۲/۱۰/۲۰

تاریخ اصلاح مقاله: ۱۳۹۲/۱۲/۴

تاریخ پذیرش مقاله: ۱۳۹۲/۱۲/۲۵

نام نویسنده مسئول: علیرضا بهراد

نشانی نویسنده مسئول: ایران - تهران - بزرگراه خلیج فارس - دانشگاه شاهد - دانشکده فنی مهندسی

## ۱- مقدمه

افزایش علاقه به بازی فوتبال در این روزها باعث شده است که این ورزش به یکی از محبوبترین بازی‌ها در سراسر دنیا تبدیل شود. تحلیل رایانه‌ای رخدادهای این بازی امری ضروری برای مربی‌ها و داوران و همچنین تماشاگران است تا به ارزیابی یک ویدئوی ضبط شده بپردازند. اخیراً در زمینه رایانه‌سازی و تحلیل ویدئویی بازی فوتبال روش‌هایی بر مبنای فناوری GPS ارائه شده است [۱ و ۲].

این روش‌ها دارای صحت بالایی هستند و به صورت بی‌درنگ<sup>۱</sup> کار می‌کنند، اما نیاز به تجهیزات گران‌قیمت و نسبتاً سنگین دارند که باید توسط ورزشکاران حمل گردند. این تجهیزات، ظریف و شکننده هستند و برای بازی‌های فوتبال رسمی قابل به‌کارگیری نیستند.

امروزه، سامانه‌های رایانه‌ای استفاده‌کننده از دوربین، ویدئو و الگوریتم‌های مبتنی بر بینایی ماشین متداول‌تر و پرکاربردتر شده‌اند. این روش‌ها دقت نسبتاً بالایی دارند؛ اما معمولاً نیاز به دوربین‌های کالیبره‌شده‌ای دارند که مکان آن‌ها هم معلوم است [۲۴-۳]. بخش اعظمی از منابع ویدئویی بازی فوتبال را تصاویری تشکیل می‌دهند که با استفاده از تک‌دوربین‌های متداول (غیراستریو) و کالیبره‌نشده تهیه شده‌اند. بنابراین تحلیل رایانه‌ای این ویدئوها بسیار مهم است. اولین مرحله در تحلیل ویدئوی تهیه‌شده با دوربین تک‌چشمی کالیبره‌نشده، کالیبراسیون دوربین و تخمین پارامترهای داخلی و خارجی دوربین است. چندین روش برای کالیبراسیون تصاویر ویدئویی فوتبال کالیبره‌نشده ارائه شده است که بیش‌تر آن‌ها بر مبنای هم‌صفحه بودن زمین فوتبال هستند [۲۵].

برخی روش‌ها از اطلاعات فریم‌های متوالی برای کالیبراسیون دوربین به صورت بی‌درنگ استفاده می‌کنند [۱۲]. برخی از روش‌های موجود برای این کار از مدل‌های زمین بازی<sup>۲</sup> [۱۴ و ۱۵] و برخی دیگر از نقاط دایره‌ای<sup>۳</sup> استفاده می‌کنند [۲۶]. استفاده از نقاط محوشونده<sup>۴</sup> نیز برای محاسبه پارامترهای خارجی دوربین، گزارش شده است [۱۳]. نقاط محوشونده، نقاطی از تصویر هستند که از تقاطع خطوط موازی در دنیای واقعی به دست می‌آیند. برخی از محققان از روش‌های بهینه‌سازی مختلف برای افزایش دقت کالیبراسیون دوربین در زمین فوتبال استفاده کرده‌اند [۲۰].

به منظور کالیبراسیون باید تعدادی نقطه متناظر در تصویر و دنیای واقعی استخراج شود. برای این منظور، برخی از روش‌ها از استخراج خطوط و منحنی‌ها با استفاده از تبدیل هاف<sup>۵</sup> بهره می‌برند [۲۷]. یکی از ویژگی‌های زمین فوتبال، هم‌صفحه بودن نقاط موجود در آن است. لذا از کالیبراسیون هم‌صفحه<sup>۶</sup> در الگوریتم پیشنهادی استفاده می‌شود.

کالیبراسیون هم‌صفحه دوربین فرایندی است که طی آن پارامترهای داخلی و خارجی دوربین از روی دسته‌ای از نقاط هم‌صفحه تصویر و نقاط متناظرشان در دنیای سه‌بعدی تعیین می‌گردد. از سوی دیگر کالیبراسیون غیر هم‌صفحه<sup>۷</sup> شامل نقاطی از دنیای سه‌بعدی

است که بر روی یک صفحه نیستند. با حل یک سری مسئله‌های بهینه‌سازی غیرخطی می‌توان برای هر دو نوع کالیبراسیون، پاسخ بهینه را یافت. اما تفاوت‌های ساختاری قابل ملاحظه‌ای بین رابطه‌بندی این دو نوع کالیبراسیون وجود دارد. در [۲۵] به این تفاوت‌های محاسباتی و الگوریتمی اشاره شده است. این مرجع یک دسته از الگوریتم‌های کارآمد را برای کالیبراسیون هم‌صفحه فراهم و گردآوری کرده است. کالیبراسیون هم‌صفحه، کاربردهای متعددی در سامانه‌های نظامی و تجاری دارد. آقای تسای یکی از پیشگامان کالیبراسیون دوربین، مخصوصاً کالیبراسیون هم‌صفحه است [۲۸]. ایشان روشی را طراحی کرده‌اند که مدل دوربین، همراه با اعوجاج دوربین در یک الگوریتم دو مرحله‌ای محاسبه می‌شود. در گام اول، برخی از پارامترهای دوربین به صورت فرم بسته به دست می‌آیند. پارامترهای دیگر مانند فاصله کانونی، مولفه عمق (سوم) بردار انتقال و اعوجاج شعاعی دوربین با استفاده از یک بهینه‌سازی غیرخطی و یک سری مراحل تکراری به دست می‌آیند. این روش برای محاسبه پارامترهای دوربین وقتی که صفحه نقاط بر جهت دید دوربین عمود است، ناموفق عمل می‌کند.

آقای سیریسانتیسامرید و همکارانش یک روش ساده برای تعیین پارامترهای دوربین با استفاده از کالیبراسیون هم‌صفحه پیشنهاد داده‌اند [۲۹]. در روش پیشنهادی ایشان، محاسبات پارامترهای کالیبراسیون به صورت مستقیم از روی تصویر رایانه‌ای پیکسلی انجام می‌شود و پارامترهای کالیبراسیون با استفاده از روش خطی به دست می‌آیند. البته این محاسبات، با صرف نظر از اعوجاج دوربین و این فرض که ضریب مقیاس یک است، انجام می‌شوند. این دو فرض در مورد بیش‌تر تصاویر ویدئویی فوتبال در نظر گرفته می‌شوند، اما در این روش، مختصات مرکز تصویر در هر دو جهت  $X$  و  $Y$  برابر فرض شده است ( $Cx = Cy$ ). این فرض باعث شده است که الگوریتم ایشان برای اغلب تصاویر ویدئویی فوتبال که مربعی نیستند ناکارآمد گردد.

در این مقاله یک روش جدید برای کالیبراسیون دوربین با استفاده از محدودیت هم‌صفحه بودن در تصاویر بازی فوتبال پیشنهاد داده شده است. ما فرض کردیم که محور نوری دوربین، از مرکز تصویر عبور می‌کند و با صرف نظر از اعوجاج دوربین، به یک روش فرم بسته برای محاسبه پارامترهای دوربین و مکان دوربین دست یافتیم. در الگوریتم پیشنهادی تصاویر می‌توانند مربعی یا مستطیلی بوده و محدودیتی در این زمینه قائل نشده‌ایم. همچنین به منظور استخراج نقاط در تصویر زمین فوتبال و نقاط متناظر آن‌ها در دنیای واقعی سه‌بعدی از یک روش جدید و خودکار استفاده می‌شود که عملکرد خوبی در مقایسه با روش‌های مشابه دارد.

ساختار این مقاله به صورت زیر است: به دنبال این بخش و در بخش بعدی الگوریتم استخراج خودکار نقاط و نقاط متناظر آن‌ها در فضای سه‌بعدی آورده شده است. در بخش‌های سوم و چهارم به ترتیب مدل دوربین و الگوریتم پیشنهادی برای کالیبراسیون دوربین بحث

$$x' = \frac{a_0 + a_1x + a_2y}{c_0 + c_1x + c_2y} \quad (1)$$

$$y' = \frac{b_0 + b_1x + b_2y}{c_0 + c_1x + c_2y} \quad (2)$$

در روابط فوق  $(x, y)$  و  $(x', y')$  به ترتیب مختصات نقاط زمین فوتبال در دو تصویر متفاوت از زمین فوتبال است. از سوی دیگر در صورتی که فرض کنیم  $(x, y)$  مختصات نقاط یک تصویر متعلق به دوربین عمود بر زمین باشد؛ در این حالت، بر اساس رابطه تبدیل تصویر و با توجه به ثابت بودن  $Z_c$  در این حالت، روابط (۳) و (۴) صدق می‌کند.

$$X = f(X_c / Z_c) = k_x X_c \quad (3)$$

$$Y = f(Y_c / Z_c) = k_y Y_c \quad (4)$$

با جایگذاری روابط (۳) و (۴) در رابطه‌های (۱) و (۲) روابط (۵) و (۶) به دست می‌آید:

$$x' = \frac{a'_0 + a'_1 X_c + a'_2 Y_c}{1 + c'_1 X_c + c'_2 Y_c} \quad (5)$$

$$y' = \frac{b'_0 + b'_1 X_c + b'_2 Y_c}{1 + c'_1 X_c + c'_2 Y_c} \quad (6)$$

در روابط فوق  $(x', y')$  مختصات نقاط در تصویر ورودی بوده و می‌تواند مربوط به تصویر دوربین با هر زاویه دلخواه باشد و  $(X_c, Y_c)$  مختصات متناظر آن‌ها در فضای سه‌بعدی دوربینی است که به صورت عمود به زمین فوتبال قرار دارد. برای تعیین مقادیر  $(X_c, Y_c)$  از اندازه استاندارد زمین که در شکل (۱) نشان داده شده است، استفاده می‌کنیم.



شکل (۱): اندازه‌های استاندارد زمین فوتبال

می‌شوند. در بخش پنجم نیز نتایج تجربی آورده شده و در انتها، نتیجه‌گیری انجام می‌شود.

## ۲- استخراج خودکار نقاط مهم در زمین فوتبال

اولین مرحله برای بازسازی اطلاعات سه‌بعدی زمین فوتبال، استخراج یک دسته نقاط در تصویر و متناظر سه‌بعدی آن‌ها می‌باشد. ما یک روش جدید برای استخراج یک دسته نقاط در تصویر زمین فوتبال ارائه کرده‌ایم که مختصات سه بعدی متناظر آن‌ها را برحسب اندازه‌های زمین فوتبال مشخص است.

برای این منظور در ابتدا باید به استخراج خطوط زمین فوتبال پرداخت تا به کمک قطع آن‌ها با هم بتوان نقاط مشخص را راحت‌تر یافت. اما وجود جایگاه تماشاگران باعث می‌شود که خطوط بسیاری که مزاحم و ناکارآمد هم هستند به همراه خطوط اصلی استخراج گردند. برای اجتناب از این موضوع در ابتدای کار، قسمت اصلی و مورد استفاده از تصویر که زمین بازی است را استخراج می‌کنیم. برای این کار از اجزای رنگی تصاویر استفاده می‌کنیم. معیاری که برای این کار از آن استفاده کردیم به صورت زیر تعریف می‌شود:

$$\text{Green} > \text{Red} \\ \text{Green} > \text{Blue}$$

به این ترتیب که برای همه پیکسل‌های تصویر، این شرط را بررسی کرده و در صورت برآورده شدن شرط برای هر پیکسل از تصویر، آن پیکسل به صورت منطقه سبز رنگ زمین فوتبال گزارش داده می‌شود. این معیار را با تعداد زیادی از تصاویر زمین فوتبال ارزیابی کردیم و نتایج بسیار خوب و مناسب بود.

اما به دلیل یکپارچه نبودن این منطقه، به کمک برجسب‌گذاری<sup>۸</sup> اجزای به هم پیوسته<sup>۹</sup> و یک رای‌گیری ساده، بزرگ‌ترین جز پیوسته را استخراج می‌کنیم و به عنوان منطقه بازی آن را از بقیه قسمت‌ها جدا می‌کنیم. در تصویر سطح خاکستری شده از ورودی آن قسمتی که باید حذف گردند را سیاه رنگ می‌کنیم.

از تصویر سطح خاکستری باقیمانده، با استفاده از تبدیل هاف استاندارد، خطوط موجود را استخراج می‌کنیم. سپس این خطوط را در دو دسته از نظر موازی بودن تقسیم می‌کنیم. این دو دسته خطوط باید در دنیای واقعی بر هم عمود باشند؛ ولی با توجه به عدم حفظ زاویه در تصاویر، از شیب و زاویه خطوط در فضای هاف برای دسته‌بندی خطوط در دو دسته موازی و عمود بر هم استفاده می‌شود. ما باید از میان این خطوط، خطوط مناسب را انتخاب کنیم.

انتخاب دقیق خطوط در این مرحله بسیار مهم می‌باشد، لذا ما از مدل هندسی بین دو تصویر زمین فوتبال برای این منظور استفاده کرده‌ایم. در دنیای واقعی خطوط زمین فوتبال حالت موازی و عمود برهم دارند. با توجه به هم‌صفحه بودن نقاط در زمین فوتبال، می‌توان ثابت کرد که رابطه بین دو تصویر از نقاط هم‌صفحه زمین فوتبال از مدل افکنشی<sup>۱۰</sup> به صورت روابط (۱) و (۲) پیروی می‌کند.

### ۳- مدل دوربین

در این مقاله از مدل دوربین حفره سوزنی<sup>۱۱</sup> برای کالیبراسیون استفاده شده است. مدل دوربین شامل پارامترهای داخلی و خارجی است. پارامترهای خارجی، شامل یک ماتریس سه در سه چرخش  $R$  که جهت دوربین را مشخص می کند و بردار انتقال  $t$  سه بعدی است که مکان سه بعدی دوربین را تعریف می کند. پارامترهای داخلی دوربین نیز عبارت از فاصله کانونی دوربین  $(f)$ ، مرکز نوری تصویر  $(C_x, C_y)$ ، نسبت منظر<sup>۱۲</sup> یا ضریب مقیاس تصویر  $(s)$  و اعوجاج عدسی می باشند. با در نظر گرفتن مدل حفره سوزنی برای دوربین روابط میان مکان نقاط سه بعدی در مختصات دوربین  $(X_c, Y_c, Z_c)$  و نقاط متناظرشان بر روی صفحه تصویر به صورت رابطه های  $(Y)$  و  $(A)$  خواهد بود.

$$x_u = -f (X_c / Z_c) \quad (7)$$

$$y_u = -f (Y_c / Z_c) \quad (8)$$

که  $(x_u, y_u)$  مختصات تصویری بدون اعوجاج و ایده آل است. این دو معادله به معادلات افکنش بعدنمایی<sup>۱۳</sup> مشهور است. با اعمال ماتریس چرخش و بردار انتقال به نقاط متناظر در مختصات جهانی دلخواه، داریم:

$$[X_w, Y_w, Z_w]^T = [R/t] [X_c, Y_c, Z_c]^T \quad (9)$$

در نهایت با صرف نظر از اعوجاج عدسی ها، رابطه نهایی میان نقاط بدون اعوجاج بر روی صفحه تصویر و متناظرشان بر روی فریم به صورت رابطه  $(10)$  خواهد بود.

$$\begin{aligned} (x_f - C_x) s_x &= x_u \\ (y_f - C_y) s_y &= y_u \end{aligned} \quad (10)$$

با ترکیب معادلات  $(Y)$  تا  $(10)$  و اعمال محدودیت هم صفحه بودن یعنی  $Z_w = 0$  داریم:

$$\begin{bmatrix} \alpha x_f \\ \alpha y_f \\ \alpha \end{bmatrix} = \begin{bmatrix} -f_x r_1 + C_x r_7 & -f_x r_2 + C_x r_8 & -f_x t_1 + C_x t_3 \\ -f_y r_4 + C_y r_7 & -f_y r_5 + C_y r_8 & -f_y t_2 + C_y t_3 \\ r_7 & r_8 & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \quad (11)$$

$$\begin{aligned} f_x &= \frac{f}{s_x} \\ f_y &= \frac{f}{s_y} \end{aligned} \quad (12)$$

که  $f_x$  و  $f_y$  فاصله کانونی موثر در جهت های  $x$  و  $y$  و  $\alpha$  یک ضریب غیر صفر به عنوان ضریب مقیاس است. پارامترهای  $r_1$  تا  $r_9$  و  $t_1$  تا  $t_3$

روابط  $(5)$  و  $(6)$  نشان دهنده صادق بودن رابطه افکنشی بین نقاط یک تصویر دلخواه از زمین فوتبال و مختصات متناظر آن ها در فضای سه بعدی دوربینی است. ما از این خاصیت به عنوان معیاری برای انتخاب درست خطوط استخراج شده در تصاویر زمین فوتبال و انتساب آن ها به خطوط واقعی زمین فوتبال استفاده کرده ایم. این کار به این دلیل است که در استخراج و تعیین خطوط زمین فوتبال ممکن است برخی از خطوط به دلیل زاویه دید دوربین یا عدم کنتراست مناسب تصویر ورودی تشخیص داده نشوند یا برخی خطوط اضافی تشخیص داده شوند. لذا تعیین این که هر خط استخراجی مطابق با کدام خط زمین در دنیای سه بعدی است، نیازمند الگوریتم مناسبی می باشد. برای این منظور ما از رابطه افکنشی موجود مختصات نقاط در تصویر ورودی یعنی  $(x', y')$  و مختصات نقاط زمین فوتبال در دنیای واقعی یعنی  $(X_c, Y_c)$  استفاده کرده ایم. این الگوریتم دارای مراحل زیر است:

۱. استخراج خطوط در تصویر ورودی زمین فوتبال
۲. تشخیص خطوط عمودی و افقی در تصویر ورودی زمین فوتبال
۳. انتخاب دو خط عمودی و افقی و نسبت دادن آن به دو خط عمودی و افقی زمین فوتبال در دنیای واقعی
۴. استخراج چهار نقطه تقاطع خطوط و محاسبه ضرایب مدل افکنشی در روابط  $(5)$  و  $(6)$
۵. تبدیل خطوط زمین فوتبال در دنیای واقعی با استفاده از مدل افکنشی محاسبه شده
۶. محاسبه میزان خطای بین خطوط تبدیل یافته و خطوط استخراج شده در زمین فوتبال
۷. تکرار مراحل ۲ تا ۶ برای انتخاب مدل بهینه
۸. استفاده از مدل بهینه و خطوط تصدیق کننده این مدل خطوط برای تعیین مختصات نقاط لازم برای کالیبراسیون

همان طور که در مراحل بالا ذکر شده است، معیار خطایی بر اساس ضرایب افکنش به دست آمده از تقاطع خطوط انتخاب شده تعریف می کنیم. برای تعریف این معیار در ابتدا فرض می کنیم که هر یک از این خط های تصویر ورودی یک خط مشخصی از زمین فوتبال هستند و با این فرض نقاط تقاطع خطوط را به دست می آوریم. بر اساس آن نقاط ضرایب تبدیل افکنش از دنیای واقعی به تصویر، استخراج می شوند. با اعمال مختصات استاندارد زمین فوتبال به این ضرایب یک سری نقاط در تصویر به دست می آید که با محاسبه فاصله اقلیدسی آن ها از نقاط اولیه تصویری که از تقاطع خطوط به دست آمده بودند برای هر دسته خطوط فرض شده، میزان خطا به دست آمده و ذخیره می شوند. در انتها دسته خطوط دارای کمترین خطا به عنوان بهترین دسته، گزارش می گردد. در ادامه کار با استفاده از این دسته نقاط برگزیده به کالیبراسیون دوربین می پردازیم.

به ترتیب درایه‌های ماتریس چرخش و بردار انتقال هستند که به صورت زیر بیان می‌شوند:

$$\mathbf{b} = \begin{bmatrix} (C_x r_7 - f r_1)/t_3 \\ (C_x r_8 - f r_2)/t_3 \\ (C_y r_7 - f r_4)/t_3 \\ (C_y r_8 - f_y r_5)/t_3 \\ r_7/t_3 \\ r_8/t_3 \\ (C_x t_3 - f t_1)/t_3 \\ (C_y t_3 - f t_2)/t_3 \end{bmatrix} \quad (19)$$

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (13)$$

$$\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (14)$$

#### ۴- معادلات فرم بسته برای کالیبراسیون دوربین

در روش پیشنهادی برای کالیبراسیون دوربین، معادلات فرم بسته‌ای استخراج شده است که با استفاده از آن‌ها به صورت مستقیم پارامترهای دوربین قابل محاسبه است. در این قسمت به توضیح معادلات فرم بسته برای کالیبراسیون دوربین و محاسبه پارامترهای دوربین می‌پردازیم. با فرض  $f_x=f_y=f$ ، تبدیل از مرجع مختصات جهانی به مرجع مختصات دوربین (رابطه ۱۱) را می‌توان به صورت دو معادله هم‌خطی<sup>۱۴</sup> (۱۵) و (۱۶) نوشت:

$$x_f = \frac{(-s f r_1 + C_x r_7) X_w + (-s f r_2 + C_x r_8) Y_w + (-s f t_1 + C_x t_3)}{r_7 X_w + r_8 Y_w + t_3} \quad (15)$$

$$y_f = \frac{(-f r_4 + C_y r_7) X_w + (-f r_5 + C_y r_8) Y_w + (-f t_2 + C_y t_3)}{r_7 X_w + r_8 Y_w + t_3} \quad (16)$$

با این فرض که محور نوری دوربین مرکز تصویر را قطع می‌کند، پارامترهای کالیبراسیون شامل ماتریس چرخش  $R$ ، بردار انتقال  $\mathbf{t}$  و فاصله کانونی  $f$  است. با داشتن  $n$  نقطه کنترلی به عنوان نقاط کالیبراسیون، معادلات (۱۵) و (۱۶) را می‌توان به صورت ماتریسی و به صورت رابطه (۱۷) نوشت:

$$\mathbf{A}\mathbf{b} = \mathbf{c} \quad (17)$$

که  $\mathbf{A}$ ،  $\mathbf{b}$  و  $\mathbf{c}$  به صورت روابط ۱۸ تا ۲۰ تعریف می‌شوند.

$$\mathbf{A} = \begin{bmatrix} X_{w_1} & Y_{w_1} & 0 & 0 & -X_{w_1} x_{f_1} & -Y_{w_1} x_{f_1} & 1 & 0 \\ 0 & 0 & X_{w_1} & -Y_{w_1} & -X_{w_1} y_{f_1} & -Y_{w_1} y_{f_1} & 0 & 1 \\ & & & & \vdots & \vdots & & \\ X_{w_n} & Y_{w_n} & 0 & 0 & -X_{w_n} x_{f_n} & -Y_{w_n} x_{f_n} & 1 & 0 \\ 0 & 0 & X_{w_n} & -Y_{w_n} & -X_{w_n} y_{f_n} & -Y_{w_n} y_{f_n} & 0 & 1 \end{bmatrix} \quad (18)$$

$$\mathbf{c} = \begin{bmatrix} x_{f_1} \\ y_{f_1} \\ \vdots \\ x_{f_n} \\ y_{f_n} \end{bmatrix} \quad (20)$$

روابط بالا یک دسته معادلات خطی برای بردار مجهول  $\mathbf{b}$  است که می‌توان آن را با استفاده از محدودیت کمینه میانگین مربعات<sup>۱۵</sup> (LMS) و روش شبه‌معکوس<sup>۱۶</sup> حل کرد. با محاسبه مقادیر بردار مجهولات  $\mathbf{b}$  به صورت  $\mathbf{b}^T = [b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6 \ b_7 \ b_8]$  و تعریف متغیرهای واسط به صورت  $a = b_4 b_5 - b_3 b_6$ ،  $e = b_1 b_4 - b_2 b_3$ ،  $c = a + d$ ،  $d = b_1 b_6 - b_2 b_5$  و  $k = \frac{C_y}{C_x}$  و  $\xi = \sqrt{\frac{b_5^2 + b_6^2}{a^2 + d^2}}$  پارامترهای دوربین را می‌توان با استفاده از روابط (۲۱) تا (۲۵) به دست آورد.

$$t_3 = \sqrt{\frac{1 - \xi^4 (e - c C_x - d C_x (k - 1))^2}{b_5^2 + b_6^2}} \quad (21)$$

$$f = \frac{t_3}{\xi} \quad (22)$$

$$r_1 = -\xi (b_1 - C_x b_5) \quad (23)$$

$$r_2 = -\xi (b_2 - C_x b_6)$$

$$r_3 = \xi t_3 a$$

$$r_4 = -\xi (b_3 - C_y b_5)$$

$$r_5 = -\xi (b_4 - C_y b_6)$$

$$r_6 = \xi t_3 d$$

$$r_7 = b_5 t_3$$

$$r_8 = b_6 t_3$$

$$r_9 = -\xi^2 (e - c C_x - d C_x (k - 1))$$

$$t_1 = \xi (C_x - b_7) \quad (24)$$

$$t_2 = \xi (C_y - b_8) \quad (25)$$

۱- میزان ارتفاع دوربین در برخی از تصاویر به صورت میانگین ۱۰/۵۴۵ متر تخمین زده شد. این ارتفاع برای دوربین ثبت وقایع فوتبال، پایین است. این ارتفاع پایین باعث می شود که خطوط به هم بسیار نزدیک دیده شوند و کار استخراج و انتخاب آن ها با مشکل روبرو شود (هر دو تصویر شکل (۸) ).

۲- کیفیت پایین تصاویر به دلیل حرکت سریع دوربین باعث شده است که خطوط با رنگ سبز زمینه ترکیب شود. هم چنین رنگ سفید کم رنگ این خطوط، هم عامل مزاحمی برای استخراج خطوط است (تصویر سمت راست شکل (۸) ).

۳- خطوطی در پشت دروازه مشخص شده است که این کار غیرمعمول است و بیش تر تزیینی است. این خطوط در هر دو تصویر شکل (۸) قابل مشاهده است. این خطوط به خصوص وقتی دوربین بزرگ نمایی دارد (همانند تصویر سمت چپ شکل (۸) ) بیش تر مشکل ساز می شوند؛ زیرا وجود این خطوط باعث می شود که به اشتباه به عنوان خطوط محوطه جریمه، شناسایی شوند.

۴- شکل شطرنجی زمین و تفاوت رنگی هر مربع باعث می شود که گاهی مرز بین این ناحیه ها به اشتباه به عنوان خطوط استخراج شود.

۵- ازدحام بازیکن ها در حوالی یک خط و عدم استخراج آن خط نیز در برخی از تصاویر باعث ایجاد خطا می شود.

۶- در برخی از تصاویر، تعداد خطوط لازم برای استخراج نقاط وجود ندارد.

الگوریتم کالیبراسیون پیشنهادی هم چنین با دو سری تصاویر مصنوعی شامل: (۱) ۳۰ تصویر مستطیلی و (۲) ۱۰ تصویر مربعی تست شد. برای هر دو این دسته داده ها، پارامترهای محاسبه شده توسط روش پیشنهادی با نتایج به دست آمده از روش تسای [۲۸] و روش آقای سیریسانتیسمرید و همکارانش [۲۹] مقایسه شد. جدول های (۱) و (۲) نتایج کالیبراسیون به دست آمده با این دو روش را نشان می دهند. این جدول ها به گزارش میانگین و واریانس خطای مطلق روش پیشنهادی و دو روش دیگر می پردازد. نتایج نشان داده شده در جدول، حاکی از این است که روش پیشنهادی دارای دقت مناسب تری نسبت به روش های دیگر می باشد.

در روش مرجع [۲۹] یک روش فرم بسته برای کالیبراسیون دوربین با استفاده از نقاط هم صفحه معرفی شده است ولی در این روش از فرض  $C_x = C_y$  استفاده شده است. این فرض برای تصاویر مربعی فرض درستی است ولی برای تصاویر واقعی که مستطیلی هستند، این فرض صحت خود را از دست داده و الگوریتم دارای خطای زیادی می شود. روش تسای یک الگوریتم چند مرحله ای برای کالیبراسیون می باشد. در این روش در طی مرحله اول از الگوریتم پارامترهای  $r_1$  تا  $r_9$  و  $t_1$  و  $t_2$  محاسبه شده و در مرحله دوم با فرض این که صفحه تصویر بر محور  $Z$  عمود نیست، پارامترهای  $f$  و  $t_3$

ماتریس چرخش محاسبه شده معمولاً متعامد نیست؛ بنابراین ما این ماتریس محاسبه شده را به کمک روش تجزیه مقادیر تکین  $SVD^{17}$  متعامد کرده و زوایای چرخش را محاسبه می کنیم.

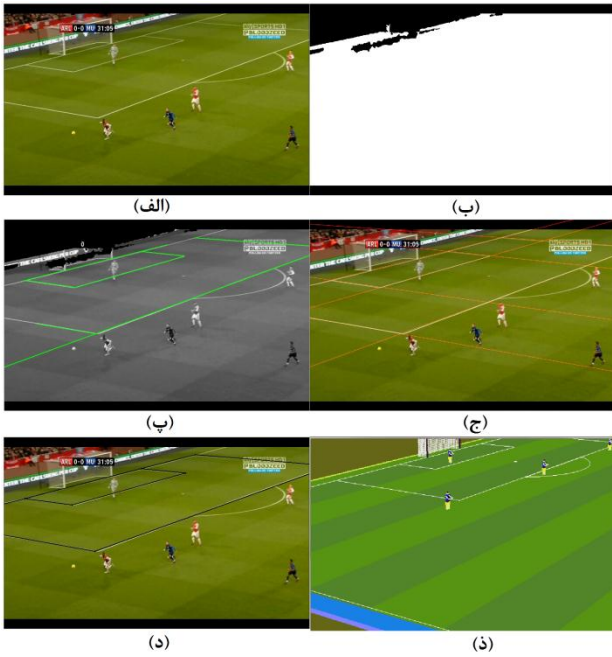
## ۵- نتایج تجربی

روش پیشنهادی به کمک برنامه MATLAB پیاده سازی شده و توسط داده های مصنوعی و واقعی مورد ارزیابی قرار گرفت. برای تهیه داده های مصنوعی یک برنامه ++VC نوشته شد که از کتابخانه OpenGL استفاده می کند. سپس به کمک برنامه داده های مصنوعی با پارامترهای داخلی و خارجی معلوم، تهیه کردیم.

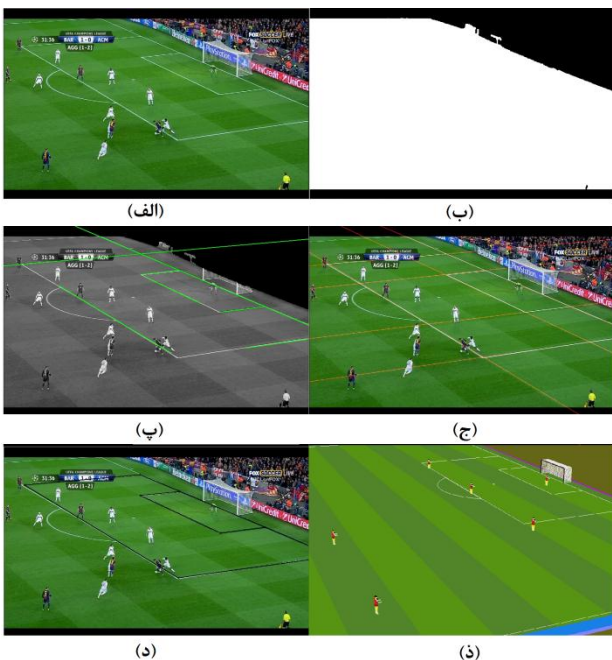
شکل های (۲) و (۳) نمونه ای از تصاویر واقعی و مصنوعی هستند که برای تست روش پیشنهادی مورد استفاده قرار گرفتند. در شکل های (۴) و (۵) مراحل الگوریتم پیشنهادی برای آشکارسازی خطوط در زمین فوتبال برای دو ناحیه مختلف تصویر نشان داده شده است. همان طور که در این شکل ها نشان داده شده است. ابتدا ناحیه چمن استخراج و سپس خطوط زمین با استفاده از تبدیل هاف مشخص می شوند. در این مرحله ممکن است برخی از خطوط شناسایی نشوند یا خطوط اضافی نیز آشکارسازی شوند. لذا با استفاده از مدل افکنشی بین زمین فوتبال و تصویر زمین، خطوط مشخص کننده زمین فوتبال استخراج و نقاط تقاطع مشخص می شوند.

برای تست عملکرد الگوریتم پیشنهادی، ۱۵۵ تصویر واقعی از صحنه ها و فایل های ویدئویی مختلف بازی فوتبال تهیه شد. سپس الگوریتم ناحیه بندی و استخراج خطوط به این تصاویر اعمال شد و پارامترهای دوربین محاسبه شد. از آنجا که مدل دوربین برای تصاویر واقعی نامعلوم هست؛ پارامترهای دوربین برای این تصاویر محاسبه کرده و با همان پارامترها، تصاویر مصنوعی با استفاده از نرم افزار پیاده سازی شده، ایجاد شد و نتایج از تصاویر واقعی به صورت چشمی مقایسه گردید. شکل (۶)، دو نمونه از تصاویر واقعی را به همراه تصاویر مصنوعی شبیه سازی شده متناظرشان را که پارامترهای آن توسط الگوریتم پیشنهادی به دست آمده اند نشان می دهد. هم چنین شکل (۷) نشان دهنده دو تصویر نوعی مصنوعی و تصاویر شبیه سازی شده آن ها به کمک پارامترهای به دست آمده از دوربین است.

الگوریتم ناحیه بندی و استخراج خطوط پیشنهادی بسته به نوع فایل ویدئویی و تصاویر آن، قابلیت ناحیه بندی و استخراج خطوط لازم با درصد موفقیت ۶۰ تا ۸۰ را دارد. در صورت استخراج موفقیت آمیز نقاط لازم برای الگوریتم کالیبراسیون پیشنهادی، این الگوریتم توانایی شبیه سازی مجدد تصویر ورودی با نرخ موفقیت ۱۰۰ درصد را دارد. براساس بررسی های صورت گرفته، عوامل متعددی باعث خطای الگوریتم پیشنهادی برای ناحیه بندی زمین و استخراج خطوط می شوند. شکل (۸) نمونه ای از تصاویر دارای خطا در استخراج خطوط را نشان می دهد. برخی از دلایل خطا در الگوریتم پیشنهادی برای ناحیه بندی زمین و استخراج خطوط به شرح زیر هستند:



شکل (۴): استخراج خودکار خطوط برای کالیبراسیون دوربین و بازسازی سه بعدی زمین فوتبال برای ناحیه جریمه سمت چپ، (الف) تصویر ورودی، (ب) ناحیه سبز استخراج شده، (پ) خطوط استخراج شده، (ج) تعیین محل برخورد خطوط، (د) تعیین خطوط نهایی، (ذ) تصویر بازسازی شده به کمک پارامترهای دوربین محاسبه شده و الگوریتم کالیبراسیون پیشنهادی.



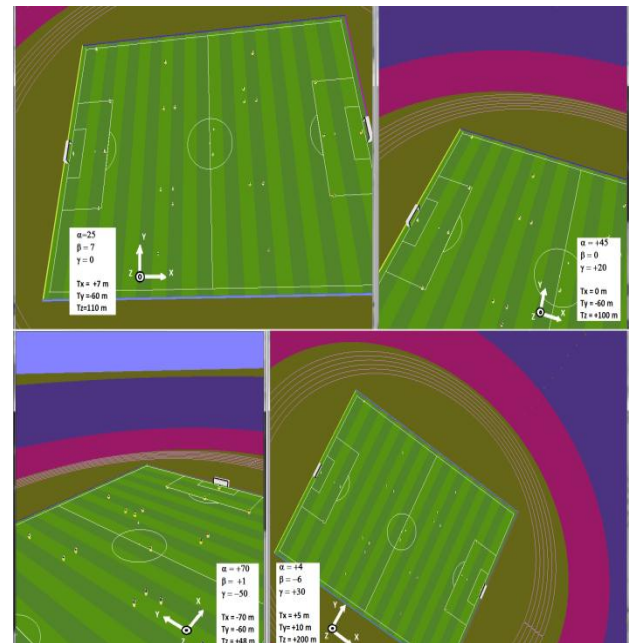
شکل (۵): استخراج خودکار خطوط برای کالیبراسیون دوربین و بازسازی سه بعدی زمین فوتبال برای ناحیه جریمه سمت راست، (الف) تصویر ورودی، (ب) ناحیه سبز استخراج شده، (پ) خطوط استخراج شده، (ج) تعیین محل برخورد خطوط، (د) تعیین خطوط نهایی، (ذ) تصویر بازسازی شده به کمک پارامترهای دوربین محاسبه شده و الگوریتم کالیبراسیون پیشنهادی.

تخمین زده می‌شوند. این فرض در تصاویری که دوربین به صورت عمود بر زمین فوتبال باشد، صدق نکرده و همان‌طور که در جدول‌های (۱) و (۲) نشان داده شده است، باعث ایجاد خطا در تخمین پارامتر  $t_3$  می‌شود.

روش پیشنهادی محدودیت‌های فوق را نداشته و می‌تواند برای کالیبراسیون دوربین در تصاویر بازی فوتبال و کاربردهای مربوطه مورد استفاده قرار گیرد.



شکل (۲): چهار مثال از تصاویر تست واقعی.



شکل (۳): چهار نمونه از تصاویر تست مصنوعی.

جدول (۱): نتایج کالیبراسیون تصاویر مستطیلی

ورایانس خطای مطلق (درجه)		میانگین خطای مطلق (درجه)		
روش پیشنهادی	تسای [۲۸]	روش پیشنهادی	تسای [۲۸]	روش
۰/۰۳	۰/۱	۰/۱۸	۰/۲۹	$\alpha$
۰/۸۶	۰/۸۹	۰/۳۱	۰/۴۶	$\beta$
۲/۲۶	۰/۸۹	۰/۵۳	۰/۳۹	$\gamma$
۰/۶۳	۲/۰۵	۰/۶۶	۱/۱۷	$t_1$
۱۰/۰۱	۲/۵۳	۱/۲	۱/۲۴	$t_2$
۱/۹۳	۹۹/۱۳	۰/۹۴	۶/۱۳	$t_3$
۲/۶۲	۱۷/۶	۰/۶۴	۱/۶۱	میانگین

جدول (۲): نتایج کالیبراسیون تصاویر مربعی

ورایانس خطای مطلق			میانگین خطای مطلق			
روش پیشنهادی	سیریسای نتیسامرید [۲۹]	تسای [۲۸]	روش پیشنهادی	سیریساینتیسامرید [۲۹]	تسای [۲۸]	روش
۰/۰۴	۱/۷۱	۰/۰۷	۰/۲۸	۱/۲۹	۰/۳۲	$\alpha$
۰/۰۹	۰/۰۹	۰/۰۳	۰/۳۵	۰/۳۵	۰/۱۲	$\beta$
۰/۲۷	۱/۲۱	۰/۳۴	۰/۴۴	۰/۹۹	۰/۳۸	$\gamma$
۱/۲۱	۱/۸	۱/۱۹	۱/۹۱	۱/۵	۱/۸۱	$t_1$
۱/۶۶	۱/۴	۲/۴۲	۱/۶	۱/۵	۱/۹	$t_2$
۰/۴۳	۰/۵۹	۷۷۲/۵	۰/۹۱	۰/۹۲	۱۴/۴۹	$t_3$
۰/۶۲	۱/۱۳	۱۲۹/۴	۰/۹۱	۱/۰۹	۳/۱۷	میانگین

## ۶- نتیجه گیری

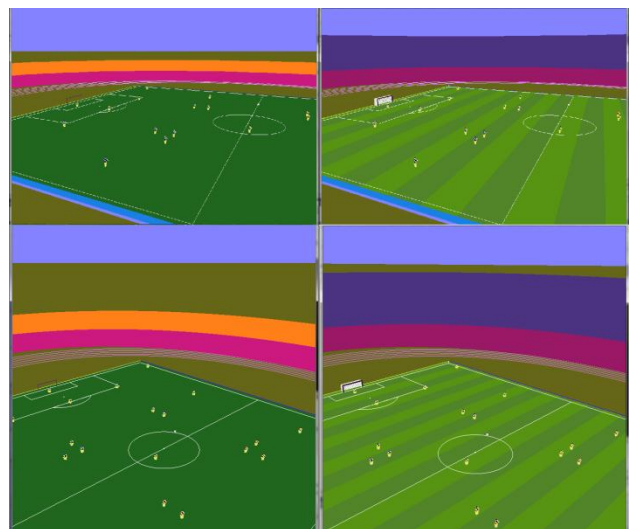
در این مقاله یک روش کارآمد برای استخراج خودکار نقاط مشخص در زمین فوتبال ارائه گردید. سپس به کمک این داده‌ها روش پیشنهادی جدیدی برای کالیبراسیون دوربین در تصاویر فوتبال معرفی شد. فرض شد که مرکز نوری دوربین را می‌توان با دقت بالایی با نقطه مرکزی تصویر تخمین زد. سپس با استفاده از این فرض، روشی که با استفاده از معادلات فرم بسته پارامترهای دوربین را محاسبه می‌کند، معرفی شد. الگوریتم پیشنهادی با تصاویر واقعی و مصنوعی ارزیابی شده و نتایج با نتایج روش‌های دیگر مقایسه گردید. نتایج عملی نشان دادند که پارامترهای محاسبه شده با الگوریتم پیشنهادی صحت بالایی نسبت به پارامترهای روش‌های آقای تسای [۲۸] و آقای سیریساینتیسامرید [۲۹] دارند و محدودیت‌های این دو روش را ندارند. الگوریتم پیشنهادی می‌تواند برای تحلیل بازی‌های فوتبال و آنالیز داورها مورد استفاده قرار گیرد.

## سپاس‌گزاری

در این قسمت لازم می‌دانیم از همکاری آقای مهندس علی آقاباقری که در جمع‌آوری داده‌های ارزیابی با ما همکاری داشتند؛ تشکر و قدردانی نماییم.



شکل (۶): دو مثال از تخمین موقعیت دوربین در تصاویر واقعی (تصاویر سمت چپ) و تصاویر مصنوعی تهیه‌شده با استفاده از پارامترهای دوربین محاسبه‌شده (تصاویر سمت راست).



شکل (۷): دو مثال از تخمین موقعیت دوربین در تصاویر مصنوعی (تصاویر سمت چپ) و تصاویر مصنوعی ساخته‌شده با استفاده از پارامترهای محاسبه‌شده دوربین (تصاویر سمت راست).



شکل (۸): دو نمونه از تصاویر ناموفق در استخراج خطوط مناسب (تصاویر بالایی). خطوط استخراج‌شده به اشتباه از تصاویر بالایی (تصاویر پایینی).



## مراجع

- [16] S. F. de Sousa Júnior, A. de A Araujo and D. Menotti, "An overview of automatic event detection in soccer matches," in 2011 IEEE Workshop on Applications of Computer Vision (WACV), 2011, pp. 31-38.
- [17] Q. Li and Y. Luo, "Automatic camera calibration for images of soccer match," in World Academy of Science, Engineering and Technology, 2005, pp. 170-173.
- [18] A. Ekin, A. M. Tekalp and R. Mehrotra, "Automatic soccer video analysis and summarization," IEEE Transactions on Image Processing, Vol. 12, pp. 796-807, 2003.
- [19] A. Hilton, J. Guillemaut, J. Kilner, O. Grau and G. Thomas, "3d-tv production from conventional cameras for sports broadcast," IEEE Transactions on Broadcasting, Vol. 57, pp. 462-476, 2011.
- [20] H. Kim and K. S. Hong, "Robust image mosaicing of soccer videos using self-calibration and line tracking," Pattern Analysis & Applications, Vol. 4, pp. 9-19, 2001.
- [21] T. Koyama, I. Kitahara and Y. Ohta, "Live mixed-reality 3d video in soccer stadium," in The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003, pp. 178-186.
- [22] D. Farin and J. Han, "Fast camera calibration for the analysis of sport sequences," in IEEE International Conference on Multimedia and Expo, 2005, pp. 482-485.
- [23] J. Puwein, R. Ziegler, L. Ballan and M. Pollefeys, "PTZ camera network calibration from moving people in sports broadcasts," in IEEE Workshop on Applications of Computer Vision (WACV), 2012, pp. 25-32.
- [24] G. A. Thomas, "Real-time camera pose estimation for augmenting sports scenes," in 3rd European Conference on Visual Media Production (CVMP 2006), 2006, pp. 10-19.
- [25] C. Chatterjee and V. P. Roychowdhury, "Algorithms for coplanar camera calibration," Machine Vision and Applications, Vol. 12, pp. 84-97, 2000.
- [26] X. Meng and Z. Hu, "A new easy camera calibration technique based on circular points," Pattern Recognition, Vol. 36, pp. 1155-1164, 2003.
- [27] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," Communications of the ACM, Vol. 15, pp. 11-15, 1972.
- [28] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," IEEE Journal of Robotics and Automation, Vol. 3, pp. 323-344, 1987.
- [29] K. Sirisantisamrid, T. Matsuura and K. Tirasesth, "A simple technique to determine calibration parameters for coplanar camera calibration," in IEEE Region 10 Conference TENCN 2004, 2004, pp. 677-680.
- [1] M. Nakazawa, "Analysis of match activities in high school soccer players using mobile GPS and VRT methods," in Conference of International Society of Biomechanics in Sports (ISBS), pp. 147-150, August 2005.
- [2] M. B. Randers, I. Mujika, A. Hewitt, J. Santisteban, R. Bischoff and R. Solano, "Application of four different football match analysis systems: A comparative study," Journal of Sports Sciences, Vol. 28, pp. 171-182, 2010.
- [3] R. Hamid, R. K. Kumar, M. Grundmann, K. Kim, I. Essa, and J. Hodgins, "Player localization using multiple static cameras for sports visualization," in 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, pp. 731-738.
- [4] V. Pallavi, J. Mukherjee, A. K. Majumdar and S. Sural, "Graph-based multiplayer detection and tracking in broadcast soccer videos," IEEE Transactions on Multimedia, Vol. 10, pp. 794-805, 2008.
- [5] M. Leo, T. D'Orazio, P. Spagnolo, P. L. Mazzeo and A. Distanto, "Multi-view player action recognition in soccer games," in Computer Vision/Computer Graphics Collaboration Techniques, ed: Springer, 2009, pp. 46-57.
- [6] S. H. Khatoonabadi and M. Rahmati, "Automatic soccer players tracking in goal scenes by camera motion elimination," Image and Vision Computing, Vol. 27, pp. 469-479, 2009.
- [7] N. Inamoto and H. Saito, "Virtual viewpoint replay for a soccer match by view interpolation from multiple cameras," IEEE Transactions on Multimedia, Vol. 9, pp. 1155-1166, 2007.
- [8] X. Desurmont, J. Hayet, J. Delaigle, J. Piater and B. Macq, "Trictrac video dataset: Public HDTV synthetic soccer video sequences with ground truth," in Workshop on Computer Vision Based Analysis in Sport Environments (CVBASE), 2006, pp. 92-100.
- [9] T. D'Orazio, M. Leo, P. Spagnolo, P. L. Mazzeo, N. Mosca, M. Nitti and A. Distanto, "An investigation into the feasibility of real-time soccer offside detection from a multiple camera system," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, pp. 1804-1818, 2009.
- [10] J. Kilner, J.-Y. Guillemaut and A. Hilton, "3d action matching with key-pose detection," in 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), 2009, pp. 1-8.
- [11] S. Hashimoto and S. Ozawa, "A system for automatic judgment of offsides in soccer games," in 2006 IEEE International Conference on Multimedia and Expo, 2006, pp. 1889-1892.
- [12] F. Szenberg, P. C. P. Carvalho and M. Gattass, "Automatic camera calibration for image sequences of a football match," in Advances in Pattern Recognition: ICAPR 2001, ed: Springer, 2001, pp. 303-312.
- [13] V. Babae-Kashany and H. R. Pourreza, "Camera pose estimation in soccer scenes based on vanishing points," in 2010 IEEE International Symposium on Haptic Audio-Visual Environments and Games (HAVE), 2010, pp. 1-6.
- [14] D. Farin, S. Krabbe and W. Effelsberg, "Robust camera calibration for sport videos using court models," in SPIE Proceedings, Vol. 5307, Storage and Retrieval Methods and Applications for Multimedia, 2003, pp. 80-91.
- [15] I. Jabri, "Camera calibration using court models for real-time augmenting soccer scenes," Multimedia Tools and Applications, Vol. 51, pp. 997-1011, 2011.

## زیر نویس ها

- 1 Real time
- 2 Court models
- 3 Circular points
- 4 Vanishing point
- 5 Hough transform
- 6 Coplanar
- 7 Noncoplanar
- 8 Labeling
- 9 Connected component
- 10 Projective

- 11 Pin hole
- 12 Aspect ratio
- 13 Perspective projection
- 14 Colinearity
- 15 Least mean square
- 16 Pseudo-inverse
- 17 Singular value decomposition