International Journal of Engineering & Technology, 7 (4.27) (2018) 6-10

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Athletes Tracking using Homography Method: a Preliminary Study

Khairul Imran Sainan[1], Mirwansah Mohamad[2], Zulkifli Mohamed[3*], Saiful Bahari Saari[4], Muhd Faiz Mat[5], Nurul Syuhadah Khusaini[6]

[1,3,6]*Sports Engineering Artificial Intelligence Centre, UiTM*
[2,4,5]*Fakulti Kejuruteraan Mekanikal, UiTM*
**Corresponding author E-mail: imransainan@salam.uitm.edu.my*

Abstract

Particle tracking has been used widely to track a single particle motion or trajectory in a medium. One of the applications of the particle tracking is in sports analysis. The tracking method can be divided into, the wearable device based system and the image based system. The wearable device based system utilizing global (GPS) and local (LPM) positioning system to track player movement. The image based system use video image processing to track the player movement and recorded is frames basis. However, the image processing method in a football match requires correction as it is normally recorded from the side of the field. Thus, to solve this problem a set of mathematical solution is needed to convert the image coordinate system (pixels) to the actual coordinate system (meters). The most commonly used is the homography method. The technique requires at least 4 reference points to transform the image coordinate into the actual coordinate system. In this project, a futsal game was recorded. The image coordinate of the player were marked in each frame with respect to the time. The image coordinate data were converted into the actual coordinate using homography matrix. Comparisons were made between the homography technique method and open-source available image warp processing method for validation. Based on the result, the homography coordinate transformation system produce a good agreement with actual player activity on the field.

Keywords: *Planar transformation, Homography matrix, Image perspective*

1. Introduction

The modern football data analysis includes many aspects such as the game speed, formation, tactical, set pieces, new rules and others. As the sport became more complex, football analyst have to find a way to decipher the dynamic code of football. Therefore, some analyses in the game are done with the help of mathematics. In football, tracking devices are used to record the x-y position of each player with respect to time. This system applies either manual or automated system. The abundant data collected needs mathematics for analysis. The companies who develop such system are ProZone, Amisco Pro and Tracab [1].

Position detection systems allows scientific investigations of player motion analysis in order to determine the extent of fatigue experience, variations of physical performances, and the effectiveness of training interventions [2]. The wearable based tracking system are routinely used compare to the image-based system [3]. However, the system were not permitted in official FIFA matches until recently because the system requires wearable active transponders on the player. The image based tracking system however, requires multiple camera in order to cover the whole playing field [4]. Through both tracking techniques, it has been possible to collect data for each player position on the field at every instant. The data collected is analyzed which included mathematical solution and such method is the application of adaptive calculus fractional calculus (FC) in order to improve the player tracking accuracy [5].

Direct Linear Transformation (DLT) is a technique to relate a point located in the object space/plane and the corresponding image point on the image plane of the camera. The DLT method converts the video images from broadcasting perspective into the actual coordinate according to the homography concept [6]. Research by Zhenxing, let $O = (X_w, Y_w)^T$ be the actual coordinates of a point on the field, and $o = (x,y)^t$ be its corresponding image coordinates on the video captured. The data is converted into the matrix form:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mu H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \tag{1}$$

Where μ is the scale factor and H is the homography 3 ×3 matrix [7]. The variable for this formula is the homography matrix, H value where it is defined by the position and angle of the camera. Different camera set up means a different set of homography matrix.

The trajectories of the player motion come from connecting points referring player particle at each selected frames. In Cartesian coordinate, the most common method is using trigonometry triangle to calculate distance and angle between two points. After the actual field coordinate has been obtained, then the data such as instantaneous speed, top speed, and distance covered, the angle between players can be calculated. The data collected reflect the performance of each player and their contribution to the match. Thus, it is more convenient for the coaches to monitor their player based on this data rather than direct observation of players on the sideline.

This project recorded data from a futsal match. The recorded data was then analyzed using homography technique to produce the actual motion of each player. The data was then compared with image warp technique from an open source software for validation. Only a single camera was used to achieve the intended objective.

2. Methodology

The project methodology presented focuses on the modelling used in the homography technique, the image warp technique used, camera setup, homography estimation and data analysis.

2.1. Homography Matrix Technique

Homography matrix, H is actually a transformation matrix between two planes or can be defined as planar transformation. The matrix is in a 3 × 3 matrices with 8 degrees of freedom and an arbitrary number. Let the image coordinate be $P(x_i,y_i)$ in a pixel unit and actual coordinate be $Q(x_r,y_r)$ in meter, m. Thus, the general equation is

$$P = HQ \tag{2}$$

and the equation is reversible;

$$Q = H^{-1}P \tag{3}$$

from the equation in (2), the inhomogeneous coordinate is

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} \tag{4}$$

in inhomogeneous coordinates where $x'_i = x_i/z_i$ and $y'_i = y_i/z_i$, therefore

$$x'_i = \frac{h_1x_r + h_2y_r + h_3z_r}{h_7x_r + h_8y_r + h_9z_r} \tag{5}$$

$$y'_i = \frac{h_4x_r + h_5y_r + h_6z_r}{h_7x_r + h_8y_r + h_9z_r} \tag{6}$$

without loss of generality, the $z_r = 1$ and $h_9 = 1$ as it is an arbitrary number:

$$x'_i(h_7x_r + h_8y_r + 1) = h_1x_r + h_2y_r + h_3 \tag{7}$$

$$y'_i(h_7x_r + h_8y_r + 1) = h_4x_r + h_5y_r + h_6 \tag{8}$$

these equations can be written in matrix form to represent one corresponding point:

$$\begin{bmatrix} x_{r1} & y_{r1} & 1 & 0 & 0 & 0 & -x_{r1}x_{i1} & -y_{r1}x_{i1} \\ 0 & 0 & 0 & x_{r1} & y_{r1} & 1 & -x_{r1}y_{i1} & -y_{r1}y_{i1} \end{bmatrix} h \tag{9a}$$

$$h = \begin{bmatrix} x_{i1} \\ y_{i1} \end{bmatrix} \tag{9b}$$

Then, the other 3 corresponding point matrix was positioned below the matrix above to make sure the matrix multiplication possible. In order to solve the value of h_1, h_2, h_3, h_4, h_5, h_6, h_7 and h_8, the 8 × 8 matrices were moved to the right-hand side to obtain the inverse matrix.

$$\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \end{bmatrix} = \begin{bmatrix} x_{r1} & y_{r1} & 1 & 0 & 0 & 0 & -x_{r1}x_{i1} & -y_{r1}x_{i1} \\ 0 & 0 & 0 & x_{r1} & y_{r1} & 1 & -x_{r1}y_{i1} & -y_{r1}y_{i1} \\ & & & & & \vdots & & \\ & & & & & \vdots & & \\ & & & & & \vdots & & \\ & & & & & \vdots & & \\ x_{r4} & y_{r4} & 1 & 0 & 0 & 0 & -x_{r4}x_{i4} & -y_{r4}x_{i4} \\ 0 & 0 & 0 & x_{r4} & y_{r4} & 1 & -x_{r4}y_{i4} & -y_{r4}y_{i4} \end{bmatrix}^{-1} \begin{bmatrix} x_{i1} \\ y_{i1} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ x_{i4} \\ y_{i4} \end{bmatrix} \tag{10}$$

Multiply the inversed matrix with the 8 × 1 matrix. Rearrange the value of h_1 to h_8 to get the homography matrix value ($h_9 = 1$ as an arbitrary number):

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \tag{11}$$

The actual coordinate was obtained simply by inversing the homography matrix and multiply by the image coordinate (in pixel unit) as follow:

$$\begin{bmatrix} x'_r \\ y'_r \\ z'_r \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{12}$$

The product of the matrix multiplication above is inhomogeneous coordinate, where to get the real value of the field coordinate is to make sure the value of z'_r is equal to 1 by divide all the value respect to the z'_r value. Each set of coordinate has a different value of z'_r. Thus, the final step is:

$$x_r = \frac{x'_r}{z'_r} \; ; \; y_r = \frac{y'_r}{z'_r} \tag{13}$$

This is one of the simplest forms in homography estimation. There are many ways to calculate the homography matrix and much more complex [6][8].

2.2. Image Warp Technique

This technique uses image shape transformation. The technique is provided by an open source software. The purpose of this method is to convert the irregular rectangular image form in the video into an actual shape of the court. After the desired shape of the court is obtained, then the image was measured to 38.4 m × 18 m dimension which is the futsal court actual dimension. Through this, all the pixel coordinate in the image was converted into the actual scale (in meter).

2.3. Camera Setup and Calibration

All data from the experiment come from a video recording. The football match was recorded on a futsal court. It was a 6-a-side (including the goalkeeper) match. Four players were selected to be analyzed. A calibration was done to improve data collection.

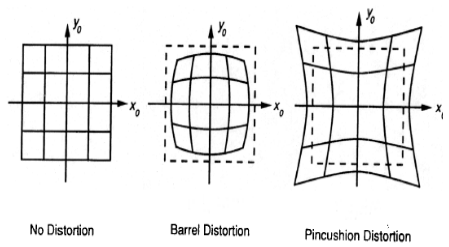

Fig. 1: Image distortion illustrations [9]

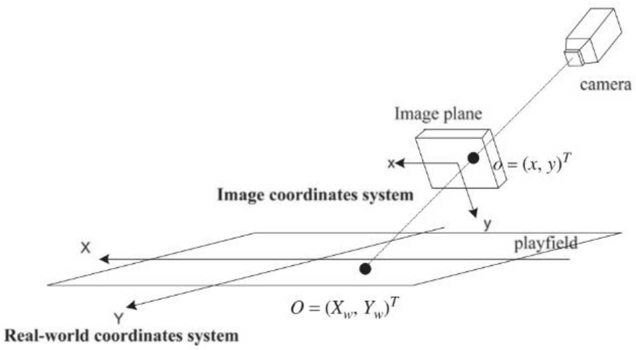

Fig. 2: Relationship between image coordinate system to real world coordinate system [7].

Radial distortion as illustrates in Figure 1 can cause huge error. It is still possible to do the planar transformation for this project however with a different set of a mathematical solution [10][11]. It is easy to identify if the image captured has distorted. Capture the image of the field line and use it the as reference line. Then, draw a straight line on the image and ensure the reference line is straight. For example, Figure 3 was taken from the futsal court. The figure clearly shows the sideline (white sideline) and the drawing line (red line) is parallel to each other. Thus, the camera used has no image distortion.

Fig. 3: Image distortion illustrations

The camera must be static for the whole calibration and recording process. A suitable tripod was used for this experiment.

Fig. 4: Camera set up

The camera is set up on the hills beside the futsal court approximately 4.7 meter above the futsal court.
Calibration needs to be done every time the camera is set up because it will produce a different set of homography value. Four reference coordinate is needed for the calibration and must be linked to the image coordinate [12]. On the futsal court, four markers were placed on the court. In the experiment, the nearest corner was blocked by the hills thus a point needs to be selected which was the half sideline point. Figure 5 illustrate the four points selected on the court.
Next step was to link the real world coordinate to the image coordinate with help of the open source software. The x-y axes was inserted into the video with the origin on the bottom left corner of the image. The four points corresponding to the points in Figure 5 was marked. The data produced is in pixels. The sideline was marked on the image coordinate with 5 marking point. The distance that connects the marked points were compared to the actual length of the line for validation purposes.

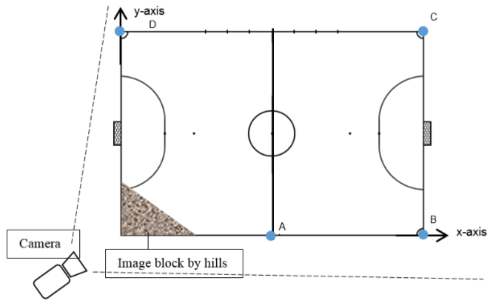

Fig. 5: Four points selected on the court

2.4. Homography Estimation

The center point of the court was taken as the initial point for the coordinate transformation. The first point taken in the center court point marked was 798 for x_i and 264.5 for y_i (unit in pixels). The point was tested in case there was an error in calibration. The converted coordinate calculation showed a value of 18.65 m for x_r and 8.75 m for y_r where the expected value should be 19.20 m for x_r and 9.00 m for y_r. The percentage error produced were 2.865 % and 2.778% for x and y position respectively.

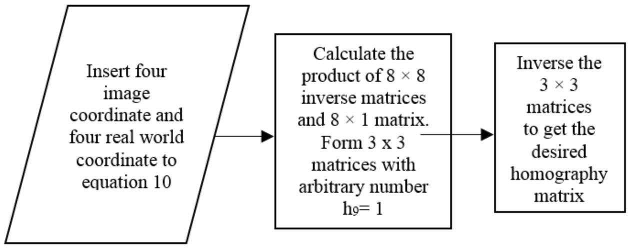

Fig. 6: Flowchart of the homography calculation

2.5. Data Analysis

The image coordinate was determined in each frames for calibration and player motion. Three important data were taken the x_i value, y_i value and the instantaneous time, t.
There were few assumptions made. The first assumption is that all particle movement must be lies on the same z-plane because the equation is limited on a 2-dimensional plane. Thus, all player considered to be on the ground surface for all time even if they were jumping or moved away from the ground surface. The second assumption is the court area is perfectly flat. Third, to mark the player on the image, the player foot is used because it lies on the surface. The player body is used as the centre of mass. The intersection between the lines is the marking point. The final assumption is the player considered to be inside the image at all the time of video recording.

Fig. 7: Player point marking method

The video was taken at 25 frames per second. Each step was taken at the interval of 0.40 seconds. The full-length video was about 391 seconds. The pixel quantity is based on the video resolution.

Fig. 8: Player4 motion tracked by the software

Image warp converts non-rectangle shape image into a rectangle shape image. Figure 9 shows the result of image warp for the whole court. The player image was distorted because of their body, not in the same plane with the court. Even the court image was distorted because of the uneven surface of the court. The higher the camera set up, the less distorted image is obtained especially the player body. After the entire coordinate has been transformed. Comparisons were made between both techniques used. The data were discussed on the player x-y position, speed, and the acceleration of the player.

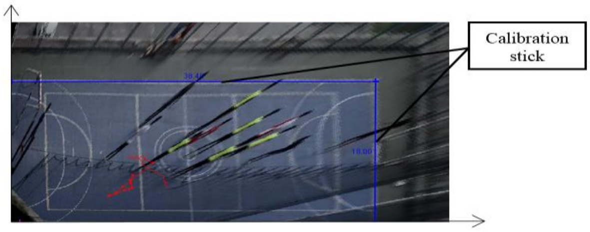

Fig. 9: Image warp technique

The four corresponding set of coordinates use in this experiment was:

Table 1: Four corresponding set of coordinates

Coordinates							
Real world coordinate (m)				Image coordinate (pixel)			
$x_{r,1}$	19.2	$y_{r,1}$	0	$x_{i,1}$	1085	$y_{i,1}$	214
$x_{r,2}$	38.4	$y_{r,2}$	0	$x_{i,2}$	1240	$y_{i,2}$	297.5
$x_{r,3}$	38.4	$y_{r,3}$	18	$x_{i,3}$	857.5	$y_{i,3}$	343.5
$x_{r,4}$	0	$y_{r,4}$	18	$x_{i,4}$	47	$y_{i,4}$	216

The homography matrix obtained for the coordinate transformation is

$$H^{-1} = \begin{bmatrix} 0.0082224265 & 0.01374 & -3.35536257 \\ -0.008433684 & 0.01566 & 5.800303815 \\ -0.0004673649 & -0.00181 & 0.881582612 \end{bmatrix} \quad (14)$$

3. Results

The results presented focuses on the differences of x-y position, speed and accelerations between all four players involve using the homography and the image warp techniques.

3.1. Data Validation

The court sidelines were measured using the homography method. The lines were marked with five points to illustrate the particle movement on that line. All the points were then converted into real world coordinate. The data was compared with the actual length. The length of the three lines were 19.96 m, 18.32 m, and 18.80 m. The difference produces were 3.958%, 4.583%, and 2.083 % respectively which shows good agreement.

3.2. Comparison between two Different Techniques

Figure 9 and Figure 10 shows the x-y position of four players on the court. Based on the results, there is not much difference between the homography technique and image warp technique. From the Figure 3-1, the percentage different of average x position were 0.958%, 1.295%, 0.579% and 0.225% meanwhile in y position shows 1.386%, 8.239%, 2.540% and 2.138% for Player1, Player2, Player3 and Player4 respectively. Player4 has less variance because the Player4 stays closest to the camera.

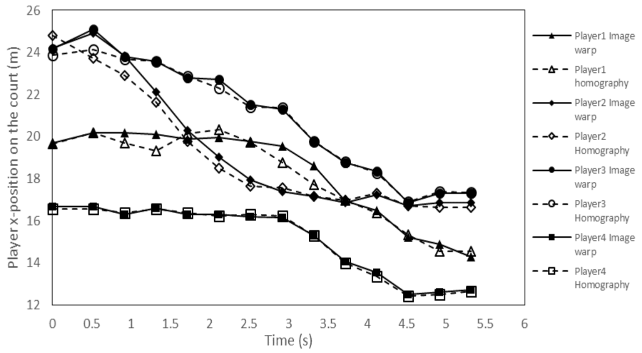

Fig. 9: Players x-position on the court between two different techniques at interval 0-5.32 s

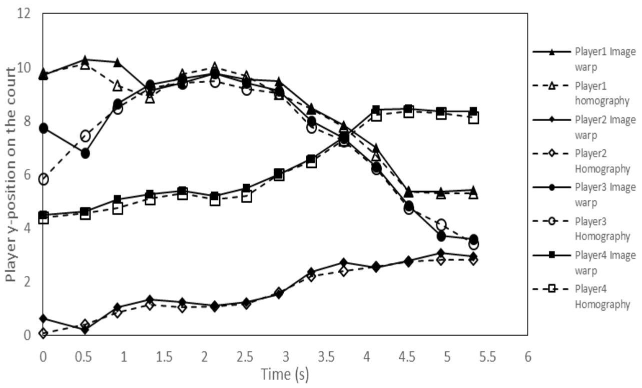

Fig. 10: Players y-position on court between two different techniques at interval 0-5.32 s

Figure 11 and Figure 12 shows the difference in speed and acceleration between the two techniques. The different in the maximum variance of speed value were 1.98, 1.08, 2.72 and 0.84 while the different in the maximum variance of acceleration value were 8.39, 4.09, 8.27 and 2.29 for Player1, Player2, Player3 and Player4 respectively. The image warp produce was distorted therefore the difference in value is expected.

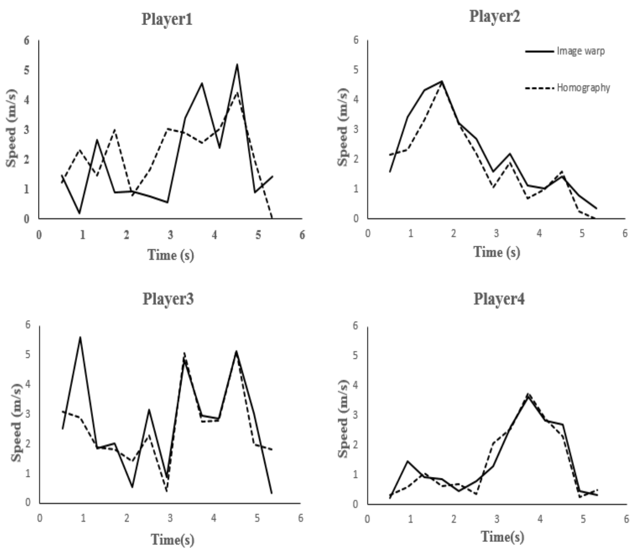

Fig. 11: Players speed against time between two different techniques at interval 0-5.32 s

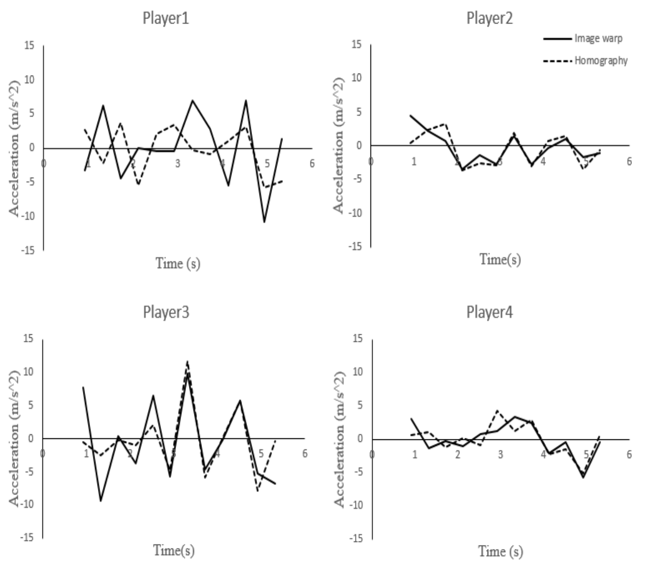

Fig. 12: Players acceleration against time between two different techniques at interval 0-5.32 s

3.3. Data of Players

Four players selected produced 979 sets of data each. From the Table 2, Player3 has made more than any player on the court covered 183.38 meters in total. Player3 was the fastest player follow closely by Player2. Compare to professional football player they can reach the speed of 27 km/h in a single sprint [13]. Player4 only manage to reach 13.57 km/h. Player3 shows the highest average speed meaning that player always in motion throughout the match without resting in one particular coordinate. That also explains why Player3 has the highest distance covered. The average x-y position shows the common area played by particular player.

Table 2: Data comparison of the four players

Particle	Player1	Player2	Player3	Player4
Distance covered (m)	111.60	127.56	183.38	162.55
Top speed (m/s)	4.24	4.60	5.12	3.77
Top speed (km/h)	15.26	16.56	18.43	13.57
Average speed (m/s)	2.16	1.85	2.56	1.38
Average x-y position (m)	(23.6, 10.4)	(16.9, 11.1)	(17.2, 7.5)	(14.3, 8.5)

4. Conclusion

The study carried out has been successfully achieved and satisfy the objectives of using homography matrix to convert image coordinate to actual coordinate and extracting data of player motion from the video recording. The result shows a similar pattern between the homography and image warp technique for the x-y coordinate results but varied slightly in speed and acceleration. This is comparable with published research [14]. However, the result can be improved further with better equipment and devices set up. In future, data that were recorded will be compared against wearable based system such as GPS system.

Acknowledgement

The authors would like to thank the Fakulti Kejuruteraan Mekanikal, Universiti Teknologi MARA (UiTM) to finance the project under Sports Engineering Artificial Intelligence Centre (SEA-IC) for supporting this manuscript is gratefully acknowledged.

References

[1] J. Gudmundsson and T. Wolle, “Football analysis using spatio-temporal tools”, *Comput. Environ. Urban Syst.* Vol. 47, (2014), pp. 16–27.

[2] C. Carling, J. Bloomfield, L. Nelsen, and T. Reilly, “The role of motion analysis in elite soccer”, *Sport. Med.* Vol. 38, No. 10, (2008) pp. 839–862.

[3] M. Siegle, T. Stevens, and M. Lames, “Design of an accuracy study for position detection in football”, *J. Sports Sci.* Vol. 31, (2012), pp. 1–7.

[4] P. Figueroa, N. Leite, R. M. L. Barros, I. Cohen and G. Medioni, “Tracking soccer players using the graph representation”, *Proceedings of the International Conference on Pattern Recognition 4*, (2004), pp. 787-790.

[5] M. S. Couceiro, F. M. Clemente, and F. M. L. Martins, “Analysis of football player’s motion in view of fractional calculus”, *Cent. Eur. J. Phys.* Vol. 11, No. 6, (2013), pp. 714–723.

[6] R. Hartley , A. Zisserman , Multiple View Geometry in Computer Vision, 2nd ed., Cambridge University Press, (2003), pp. 88-92.

[7] Z. Niu, X. Gao, and Q. Tian, “Tactic analysis based on real-world ball trajectory in soccer video”, *Pattern Recognit.* Vol. 45, No. 5, (2012), pp. 1937–1947.

[8] P. Chen and D. Suter, “Error analysis in homography estimation by first order approximation tools: A general technique”, *J. Math. Imaging Vis.* Vol. 33, No. 3, (2009) pp. 281–295.

[9] J. Mrovlje, and D. VranCic, "Distortion impact on a stereo distance", *in Proc. of the 10th International PhD Workshop on Systems and Control: young generation viewpoint,* (2009).

[10] M. Alemán-Flores, L. Alvarez, L. Gomez, P. Henriquez, and L. Mazorra, “Camera calibration in sport event scenarios”, *Pattern Recognit.* Vol. 47, No. 1, (2014), pp. 89–95.

[11] F. E. E. Devernay and O. Faugeras, “Straight lines have to be straight: automatic calibration and removal of distortion from scenes of structured enviroments”, *Mach. Vis. Appl.* Vol. 13, No 1, (2001), pp. 14–24.

[12] S. Y. Park and G. G. Park, “Active calibration of camera-projector systems based on planar homography”, *Proceedings of the International Conference on Pattern Recognition*, (2010), pp. 320-323.

[13] D. P. Wong, V. Pialoux, C. Hautier, K. Chamari, and A. Dellal, “Physical activity during a prolonged congested period in a top class European football team”, *Asian J Sports Med*. Vol. 5, No. 1, (2014), pp. 47–53.

[14] J. Mara,S. Morgan, K. Pumpa, and K. Thompson, “The Accuracy and Reliability of a New Optical Player Tracking System for Measuring Displacement of Soccer Players”, *International Journal of Computer Science in Sport*, Vol. 16, No. 3, (2017), pp. 175-184.