PRACTICE 4: A MATTER OF HEIGHT

ADVANCED TOPICS IN COMPUTER VISION

Marc Zoel Arias Mitjà
Yeray Navarro Soler

November 10, 2024

To understand the process followed to obtain these results, we encourage the reader to take a look at the MATLAB code, which is designed to be executed once in order to plot the results discussed on this document.

The aim of this practice is to conclude wether the controversial play that took place in minute 28:40 of LaLiga game Real Madrid - FC Barcelona was goal or not. The match ended 3-2, in which the white team is said to have been favored by the referee.

Even with cutting-edge technology such as VAR (Video Assistant Referee), which costs around 4 million euros per season, it was not possible to determine whether the ball was in or out of the goal. Thus, the students of *Advanced Topics in Computer Vision* are encouraged to act. Based on the techniques studied in this course, we will now carry out a study in order to clarify all doubts and draw conclusions about what happened, exposing the lies of the press that, week after week, tries to deceive the fans.

The solution will go through the use of the DLT algorithm, from which we will obtain a homography matrix H that will help us to transfer the points of the given images to a plane of the goal seen from the front.

APPLIED METHODOLOGY

First of all, we have overlaid a pure red circle (255, 0, 0) that delimits the perimeter of the ball in some of the provided images and others taken from the Internet.

(a) Image 1 (b) Image 2

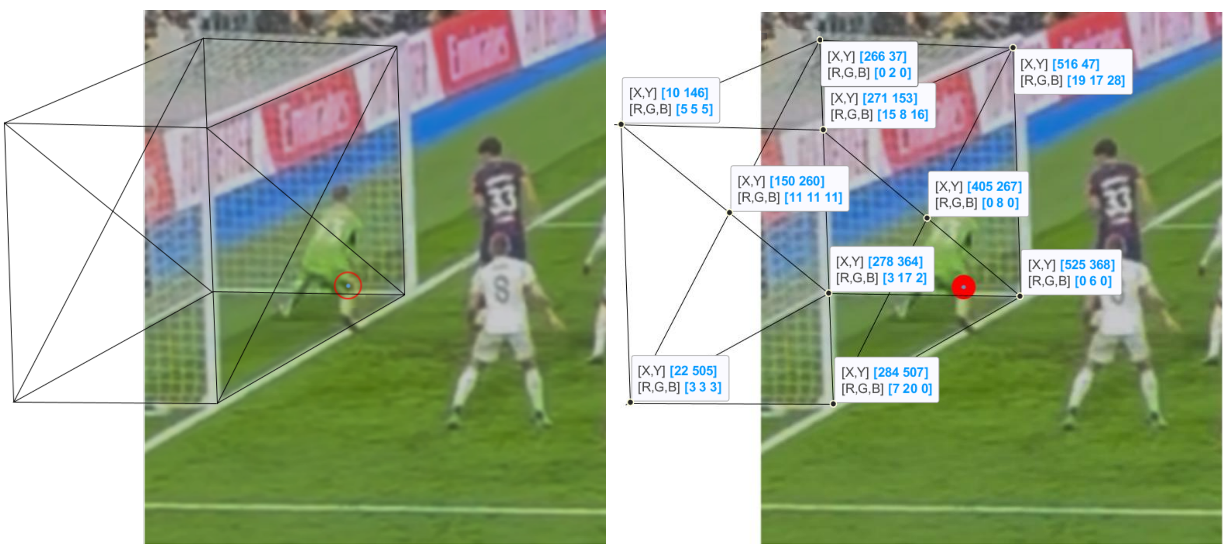
Figure 1: Overlaying a pure red circle on the ball

A total of 4 images have been considered in order to perform the study (see *Annex* at the end of the document), which are not necessarily captured at the same instant of time. From the images provided, just the first 3 have been used, whereas the 4^{th} and 5^{th} have been discarded. The 4^{th} one has been rejected because there are not enough visible corners of the goal, whereas the 5^{th} has been rejected because the image is distorted since the goal lines is seen as a curve. Next, it is necessary to mark the edges that delimit the rectangular prism of the goal. Then, we will be able to label the pixel coordinates of the corners of the goal.

(a) Goal edges (b) Pixel coordinates

Figure 2: Visible goal. Locating the corners of the goal

Initially, only the 5 points contained in the goal plane have been used. But later, with a method tried that will be explained afterwards, also the 5 points of the back of the goal -where the net is- are also needed.



(a) Goal edges (b) Pixel coordinates

Figure 3: Partially visible goal. Locating the corners of the goal

Since the available images do not always show the entire goal, we have had to approximate the non-visible part as follows: projections of the lines that can be seen are made, and the intersections between these are considered to be the corners of the goal (see *Figure* 3).

The following step consists on estimating the coefficients of the homography matrix H from the 3D-2D correspondences using the DLT algorithm. For numerical stability reasons, we will take the origin of the world coordinate system at the center of the goal plane.

To do so, the homogeneous linear equation system -obtained by the DLT definition- is solved using the singular value decomposition of the matrix of coefficients of the system. Since we are considering a transformation with 2 dimensions, the homography matrix will be 2x2. Also, only 5 correspondences are needed since there are only 9 parameters to compute.

Note that this process has to be repeated for each of the images considered, obtaining a total of 4 different homography matrices. One of the obtained can be seen below.

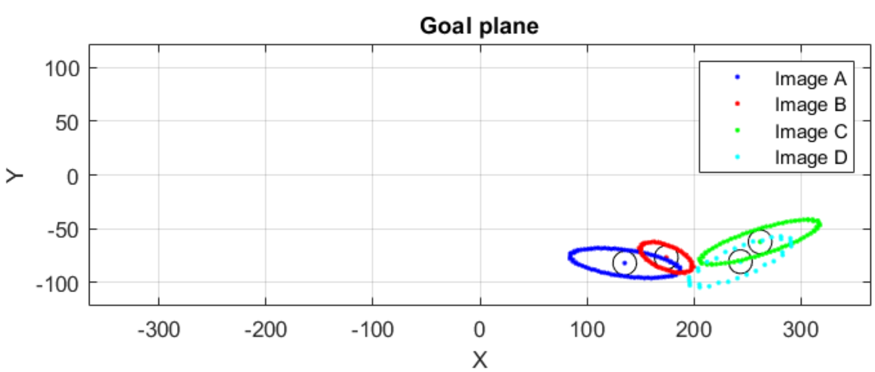
$$H_2 = \begin{bmatrix} 0.0008 & -0.0001 & 0.8163 \\ 0.0003 & -0.0015 & 0.5776 \\ 0 & 0 & 0.0014 \end{bmatrix}$$

Once the homography is computed, the ball perimeter is detected, by using the pure red circle added to the images. The pixels of the perimeter are then transformed by multiplying their coordinates by the inverse of the homography matrix (see *Equation* 1). Recall that these values need to be neormalized by the third component of the vector, then, these values are the ones projected into the goal plane.

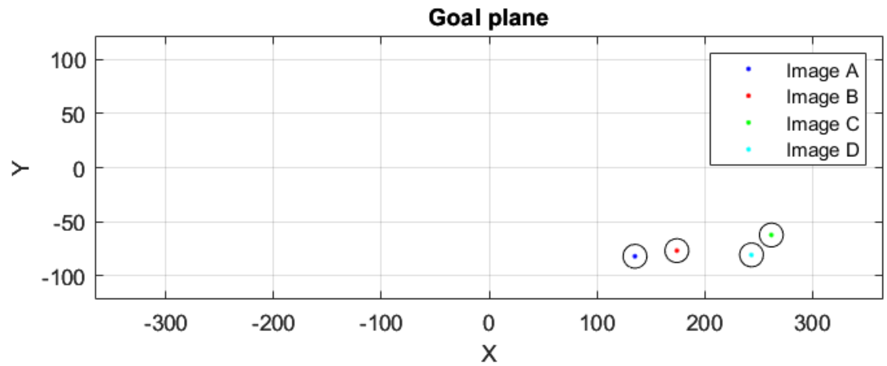
$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = H^{-1} \cdot \begin{bmatrix} x_{pixel} \\ y_{pixel} \\ 1 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} x_{world\ coord} \\ y_{world\ coord} \\ 1 \end{bmatrix} = \begin{bmatrix} A/C \\ B/C \\ 1 \end{bmatrix} \tag{1}$$

After having the projected perimeter, the center of the ball is computed as the mean of all these points. Also, in the goal plane a circle is drawn with the radius of the football ball.

The homography of the ball in the goal plane for each of the four perspectives looks as follows:



(a) Homografied balls



(b) Homografied centres

Figure 4: Goal plane

As can be seen in the figures above, it can be noticed how the position of the camera, and the perspective seen, affects the direction of the ellipsoids axis. Images A and B are from the same side of the court, whereas images C and D are from the opposite side.

In order to be able to determine the minimum and maximum height that the ball must have to be inside the goal, there would be necessary to have images taken facing the goal, which are the ones we have, and images taken from behind the goal, which we don't have any.

Also, observe that the center of the ball for the different perspectives are not coincident. Why can this happen? It can be caused by the different time instants when the images were taken, but is this the case?

For instance, take the differences between the ball center of image A and C, its separation is of 128 cm (see *Figure* 5). If we considered a time lapse of 1 second between the capture of the images, this would mean that the ball is moving at approximately 1.28 m/s, which is equivalent to 4.6 km/h. If a time lapse of 0.5 seconds is considered, the ball speed will be 9.2 km/h. These velocities are really low, considering that a shot with maximum power could arrive to 85 km/h. Even though, the action analyzed was not a direct shot to the goal, the velocity of the ball should be higher.

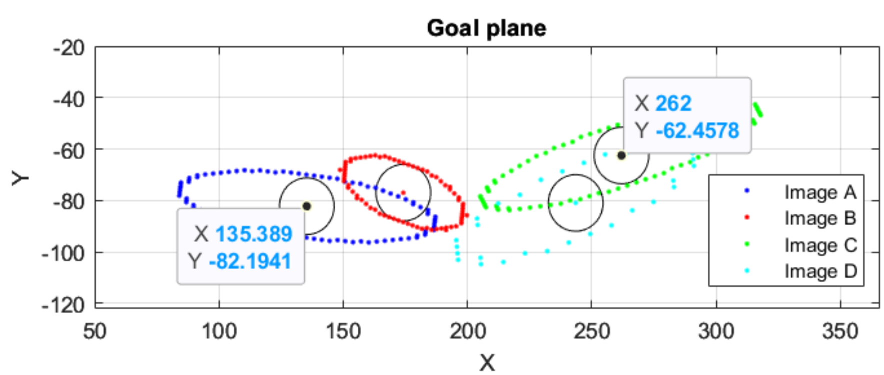


Figure 5: Zoom of the goal plane projections

Being said that, we could consider that the images are taken basically at the same time instant. So, in order to determine the position of the ball, and whether it was a goal or not, the next reasoning has been followed.

FURTHER ANALYSIS

The homography of the ball in a plane should lead to the same solution even if perspectives change -we are using four different perspectives-. This means that, if the considered plane is the actual plane of the ball, the four transformed center coordinates should be overlaid, since we are considering that the images are taken at the exact same instant.

For example, in *Figure 4*, the different perspectives of the ball do not match on the same coordinates of the goal plane. That means that this is not the actual plane on which the ball is located.

Talking with some other colleagues, we came up with the following method. In order to determine the real plane in which the center of the ball is located, parallel planes to the goal plane in both directions, that is, in the direction inward and outward of the goal will be placed. The idea is to compute the homography for many different planes -represented on *Figure 6*- in order to find the plane which the centres of the ball overlay. This plane will correspond to the actual plane of the ball.

(a) Goal planes (b) No goal planes

Figure 6: Imaginary planes representation

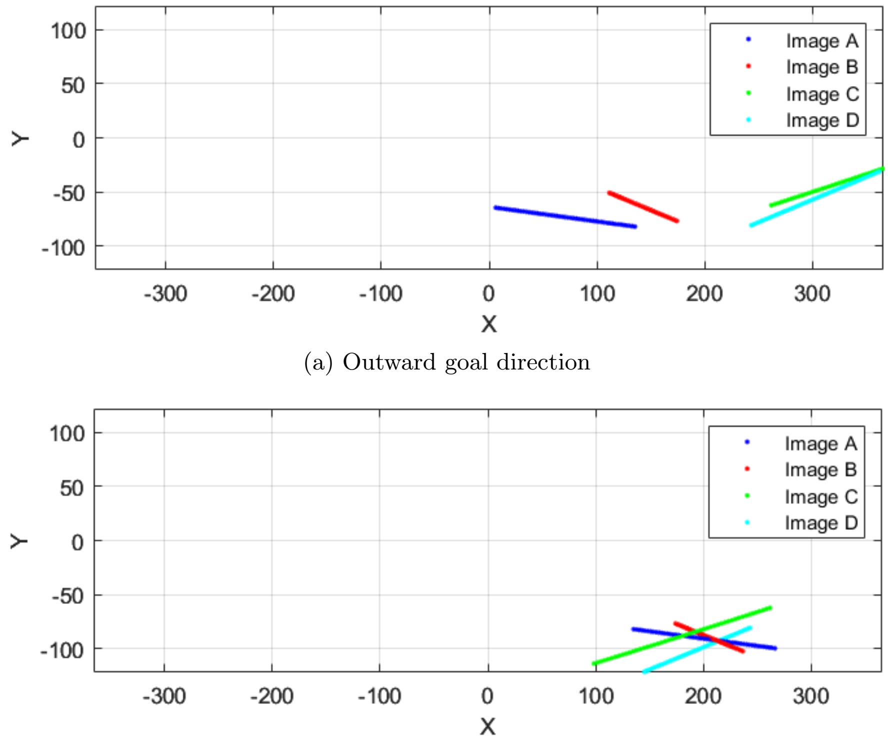
Note that due to the inaccuracies of determining the ball position in the image, and selecting the corner of the goals, some error is assumed, and the centers of the ball should, at least be as close as possible, in the actual plane where the center of the ball is located, i.e. we need to find the plane with the minimum distance between centres.

First of all, the 10 points -corners and centers- of the goals are defined as explained before. Then, the distance between the goal line and the back of the net has been divided into 300 slices. Each slice will represent a plane, and each plane will have its own corner and center points. These points has been computed by doing a linear interpolation between the original 10 points acquired from the image:

$$point_{new\ plane} = point_{goal\ plane} \pm \frac{i}{300} \cdot (point_{goal\ plane} - point_{net\ plane})$$

Where i represents the current plane that is being studied. Also, the sign will be negative when the inward planes are being considered, whereas the sign is positive when the outward planes are.

After computing the homography of the ball for each imaginary plane -green/red- on each of the four considered images, the following results have been obtained.



(a) Outward goal direction

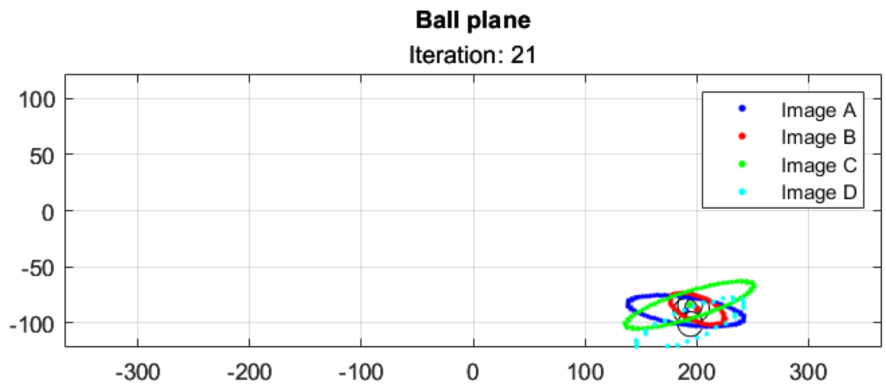
(b) Inward goal direction

Figure 7: Homography evolution

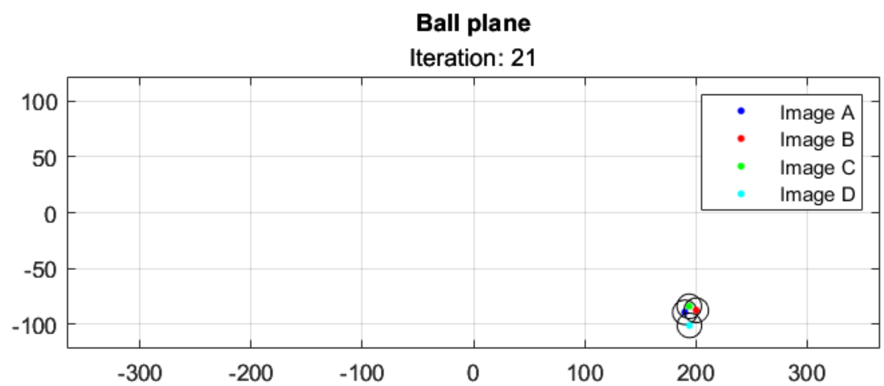
On the image above, only the ball centers have been represented, in order to see the evolution of the projections of the ball for each image. *Figure 7a* represents the outwards direction of the goal, where it can be seen that the centers become more distant as the plane moves away from the goal plane. *Figure 7b* represents the direction toward the goal, where it can be seen that the centers get closer as the plane gets closer to the net plane.

Since we are looking for the plane where the centers overlay, we can say that this plane will be inside the goal, but will it be enough to be a goal? There are at least 2 intersections between the centers evolution lines, one for images A, B, C and the other for images A, B, D. Nonetheless, the centers must belong all to the same plane.

In order to determine which of all the new planes is the actual ball plane, the distances between the centre projections have been computed during the iterations. These distance has been computed as the sum of distances between all the pair of centers (1-2, 1-3, 1-4, 2-3, 2-4, 3-4) . It has been obtained that the minimum distance between centers is obtained at the 21^{st} plane, with a distance sum of 68.84 cm. The homography of the ball in all the perspectives in this plane is the following.



(a) Homografied balls



(b) Homografied centres

Figure 8: Ball plane

As it can be seen, more clearly in the second image, is that all the centers are really close to each other. The one with the most distance to the others is the one from image D.

Once the real plane of the ball is determined, its depth with respect to the goal plane can be computed. To do so, the following expression has been used:

$$z_{plane} = -i \cdot \frac{Goal\ depth}{300} = -12.81$$

Where $i = 21$. The goal depth has been computed by taking the proportion of pixels -in image B- between the height and depth of the goal in pixels. Knowing that its height is 244 cm, the depth of the goal has been considered to be 183 cm.

If the center of the ball is located at 12.5 cm towards the net, and the radius of the ball is considered to be 11.5 cm, then all the ball is inside the goal, so it is a goal.

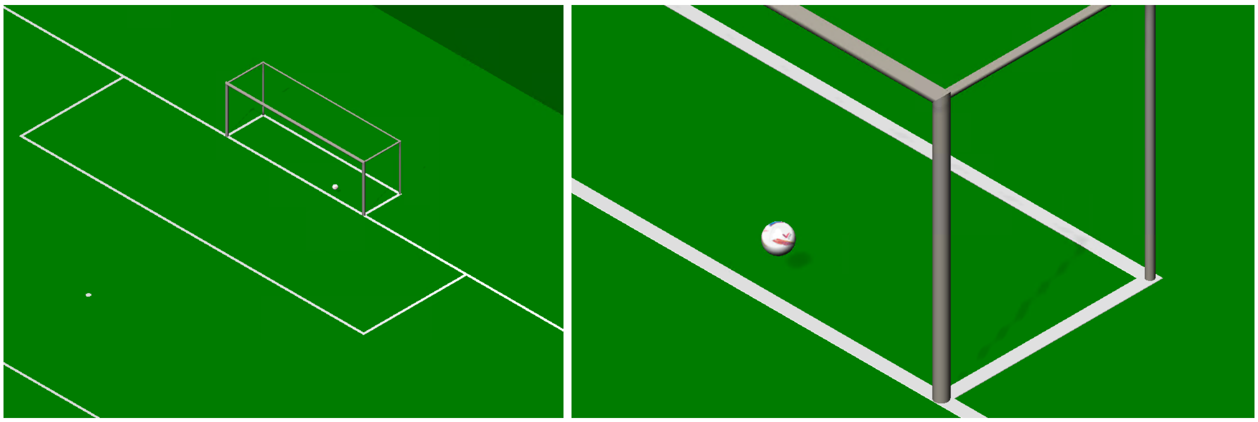
Finally, knowing the plane where the center of the ball is, and their projections of the different images, the coordinates of the ball center can be fully computed. To do so, the mean of the transformed centers has been considered to be the real position of the ball, which in this case is:

$$[x \quad y \quad z] = [194.96 \quad -87.17 \quad -12.81]$$

EXTRA BONUS: 3D POSITION OF THE BALL

Once we have estimated the position of the ball, we are going to define it in a simulated 3D scenario in order to visually perceive the real location of the ball without occlusions. We used a computer-aided design (CAD) software for the 3D representation.

Knowing from previous sections that the ball is located on the position $[x \;\; y \;\; z] = [194.96 \;\; -87.17 \;\; -12.81]$ cm, its isometric representation looks as follows.



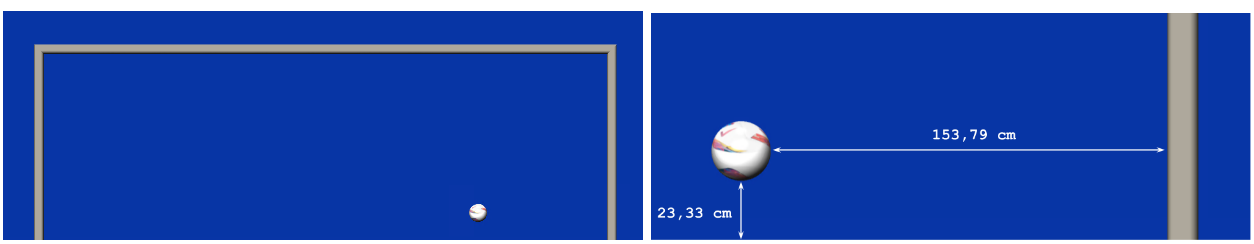
(a) Isometric

(b) Zoomed isometric

Figure 9: 3D scenario definition

We are going to observe in detail the main views that are commonly used to fully describe the geometry of an object.

First of all, the elevation view, which shows the goal as seen from the front. It is as if you were looking at the object from the front at eye level.



(a) Standard

(b) Zoom

Figure 10: Elevation view

As it can be observed in *Figure 10*, the ball is 153.79 cm away from the right post of the goal. It is also at a height of 23.33 cm from the floor. Evidently, this measures only indicate from which position would the ball enter the goal -in case it does-, but it does not provide any information of the depth Z axis.

Therefore, the plan view is analyzed. It shows the object as seen from above, i.e. it is as if you were looking at the object from the ceiling.

(a) Standard

(b) Zoom

Figure 11: Plan view

It can be observed how the ball is really close to the crossbar of the goal, even making difficult to visualize it. The image is zoomed in to appreciate its details. Notice that this image will determine if the ball is inside the goal or not.

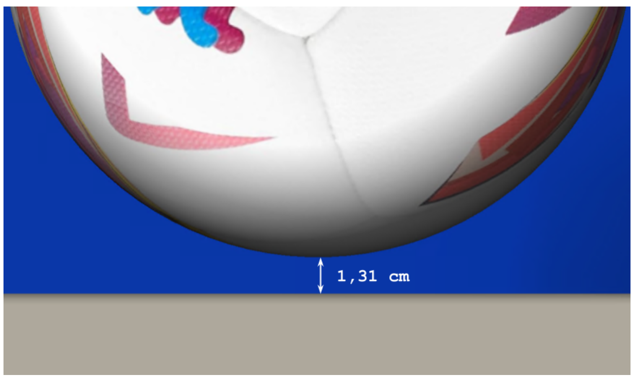
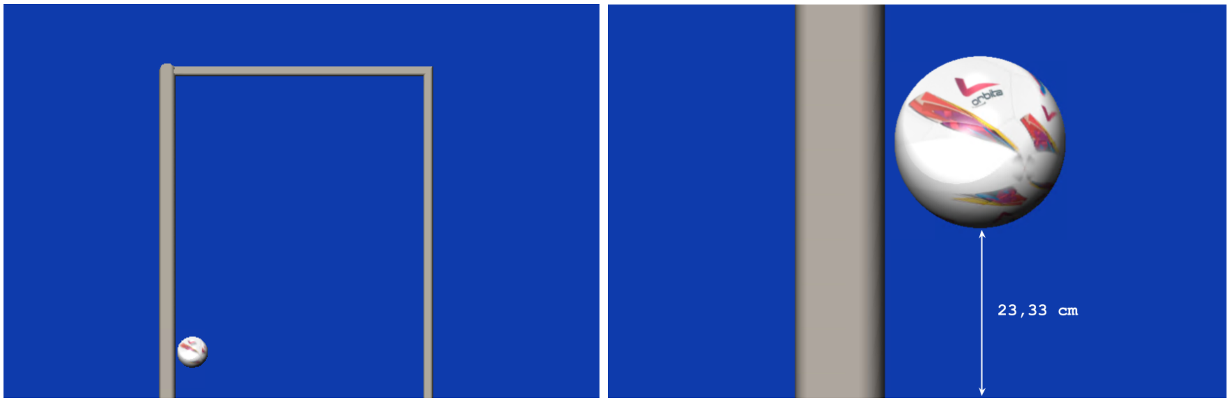


Figure 12: Plan view. Extra zoom

The ball is inside the goal as it has exceeded the goal plane by 1.31 cm, which is a really tight distance making this situation hard to take a decision in the same moment.

Even knowing that it was goal, let's observe the profile view which shows the object as seen from the right side.



(a) Standard

(b) Zoom

Figure 13: Profile view

As before, the ball is so close to the posts so we need to zoom in the image.

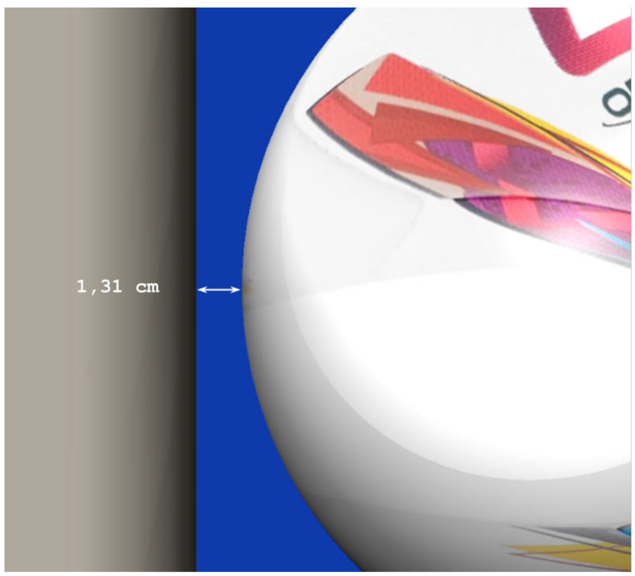


Figure 14: Profile view. Extra zoom

Having positioned the ball in a simulated 3D scenario with as precision as possible according to the performed homographies, we can conclude that the ball entered the goal plane in minute 28:40 of LaLiga game Real Madrid - FC Barcelona. Even though the ball is inside the goal, it is understandable that the referee was unable to decide wether it was goal or not at the same instant of time. However, it is not understandable that VAR technology, which costs around 4 million euros per season, is not able to decide in this hard situation.

This has not been the only case. A few weeks later, in the minute 103 of the Real Madrid - FC Bayern Munich of the UEFA Champions League semifinal match and with Real Madrid ahead (2-1), the referee signaled offside for the Germans before the ball entered the Real Madrid goal, making it impossible for the video referee to review the action.

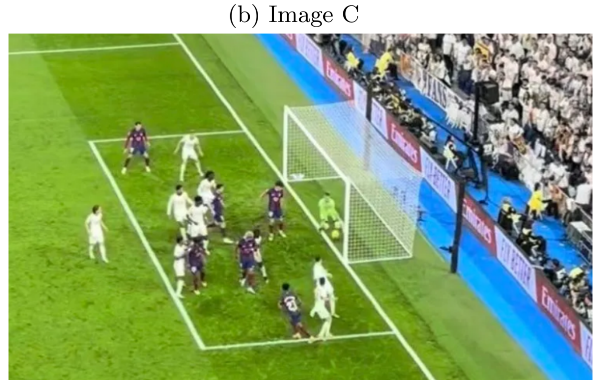
Fortunately, the students of *Advanced Topics in Computer Vision* are capable of solving this type of critical situations.

Annex

(a) Image A

(b) Image C

(c) Image B



(d) Image D

Figure 15: Considered images with different perspectives