

# Ball Detection for Robotic Soccer: A Real-Time RGB-D Approach

André Morais, Pedro Costa and José Lima

**Abstract** The robotic football competition has encouraged the participants to develop new ways of solving different problems in order to succeed in the competition. This article shows a different approach to the ball detection and recognition by the robot using a Kinect System. It has enhanced the capabilities of the depth camera in detecting and recognizing the ball during the football match. This is important because it is possible to avoid the noise that the RGB cameras are subject to for example lighting issues.

**Keywords** Image processing · Kalman filter · Kinect · Robotic football

## 1 Introduction

The evolution of technology in recent years has allowed the emergence of sensors that enable new and improved methods of interpreting the environment. This is relevant for many areas especially for robotics, as it allowed the creation and development of algorithms that perceive the world around us.

However, these algorithms are usually computationally demanding, making them unviable for robotic systems based on artificial vision that require real-time processes.

---

A. Morais · P. Costa

INESC TEC (formerly INESC PORTO) - Robotics and Intelligent Systems,  
Faculty of Engineering of University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal  
e-mail: {ee10276,pedrogc}@fe.up.pt

J. Lima

INESC TEC (formerly INESC PORTO) - Robotics and Intelligent Systems,  
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

J. Lima(✉)

Polytechnic Institute of Bragança, Campus Sta Apolónia, 5301-857 Bragança, Portugal  
e-mail: jllima@ipb.pt

Thus the development and creation of increasingly fast and efficient algorithms has been a challenge for the scientific community.

The aim of this work involves the development of a detection and object recognition system using the RGB-D sensors to perform the recognition of the ball in a robotic football game, that is computationally efficient and robust from outside interference such as lighting problems [9].

The motivation of this project is mainly to integrate all knowledge acquired in robotic football team FEUP (RoboSoccer-5DPO Team) and to get a good performance in a future robotic football competition, mainly because competitions have become more competitive and challenging which encourages students to seek new alternatives that have better results in order to improve their systems.

This article is divided into six sections, excluding this introduction section. The section 2 details the related work that has been found in the available literature. It is approached in the section 3 the detection system which aims to image segmentation and the identification of different image objects to be subsequently recognized. The football ball recognition system is specified in the section 4 where is specified the algorithm developed to recognize the different objects as a football ball. It is described in the section 5 the implementation and incorporation of the Kalman filter in the developed system. The conclusions and future developments are presented in the section 6.

## 2 Related Work

The detection and recognition of the objects using vision systems is a big part of the RoboCup competition. Through the years some rules have changed to improve the competition. This work comes in response to the latest change to the ball rules which is that the ball can be of any color. In the previous years the ball has always the same color and the teams use mainly the color information to localize the ball position. This new improvement to the competition presents a challenge to the participants because they cannot rely only in the color information.

The techniques used to detect and recognize objects using RGB-D sensors were analyzed because this systems are widely used for this type of applications [1], [2], [10].

In [3] a new perspective to the Hough Transformation is presented as a technique used in image processing in the detection of circular/lines features of the image. This article shows that this technique can be optimized using a set of descriptions and feature matching techniques following by a comparison with the data training that was done.

In [4] is detailed the SIFT method that enables the recognition of an object that is invariant to the object rotation, scale or projections in the 3D space. Also in [6]

an improvement of the SIFT method techniques increasing the algorithm processing time with similar results.

The techniques presented in [5] allow to measure the similarity of the object that has to be identified in the image from any other using histograms and descriptors acquired in a training phase.

In [7] the RANSAC algorithm is presented which finds in a set of data the presence of a mathematical model for example a plane.

### 3 Object Detection

This section presents the techniques used to make the detection of the ball. It is observed that during a football game, the ball may be present on the ground or in the air and these two cases have different properties. Thus, it was decided to follow this philosophy and detect the ball in two different cases. The following sections presents the procedures used to achieve effectively the detection of the football ball.

#### 3.1 Aerial Ball

**Clustering.** In order to separate the different objects in the depth image (see figure 1) it was chosen to separate the image into different clusters by applying the K-Means Clustering Technique based on OpenCV Libraries.

From figure 1 is possible to get the different clusters that are shown in figure 2 and also it shows that in the 2d sub-figure the ball is well defined and disconnected from other objects. The process time of the clustering and the recognition system is  $\sim 30$ ms that is quite fast, on an Intel i5 processor and using five clusters.

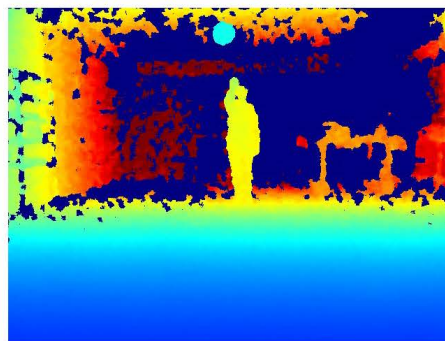
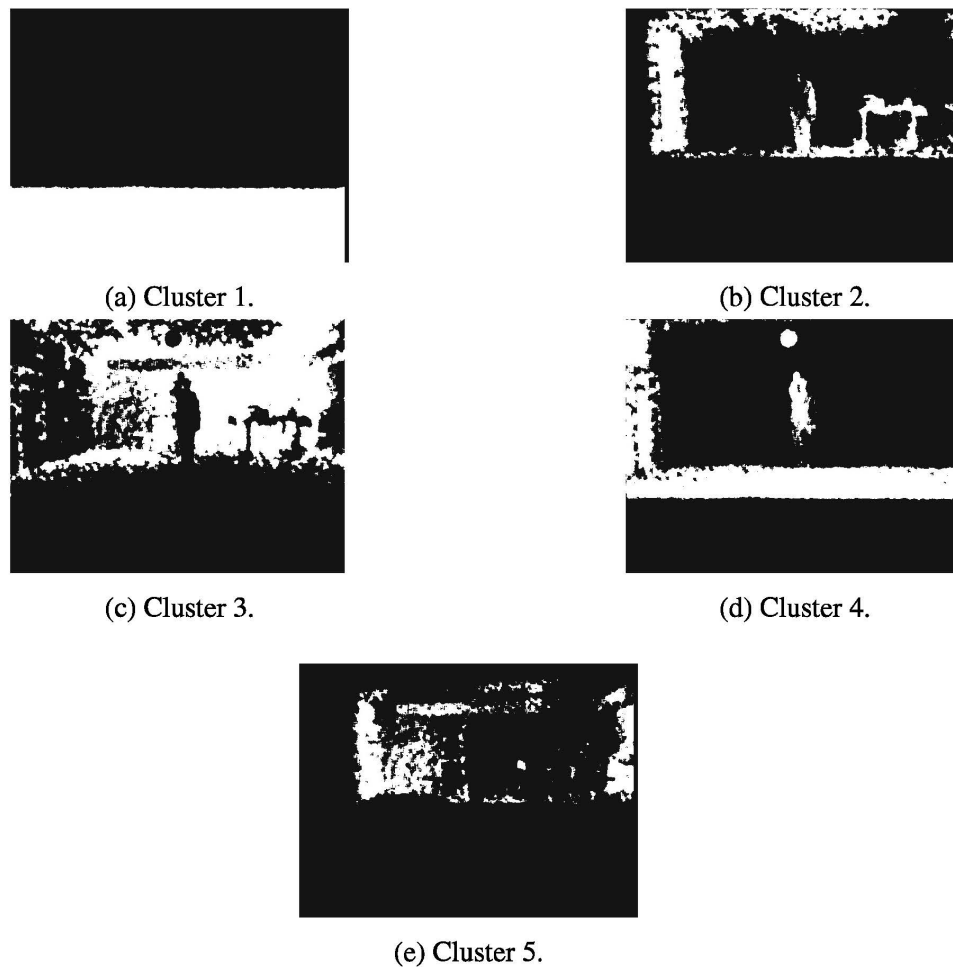


Fig. 1 Depth Image with an Aerial Ball present.



**Fig. 2** Clustering of the Depth Image in the figure 1.

### 3.2 Ground Ball

For the detection of the ball when it is in the ground it is not advantageous to use the technique used in the aerial ball detection without any pre-processing, because a large number of cluster was needed to archive similar results as in the aerial ball detection.

The increasing of the number of clusters will affect the computational time. This has to be avoided because it will reduce the reaction time by the goalkeeper, increasing the chances of suffer goal.

The technique used for this case will be detailed in the next sections.

**Field Segmentation.** It is known that the ground field is green, so this fact is used in order to find a set of points for calculate the ground plane equation.

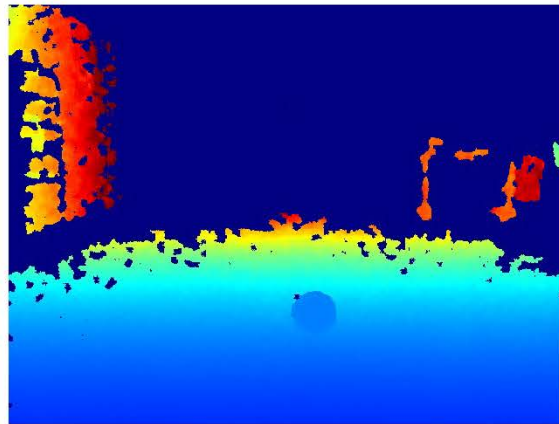


The use of a RGB image for the threshold of color is not the best option because of the high correlation between the different components, the image is highly sensitive to the scene illumination, etc [8]. To counter act this problem it is transformed the RGB image to the YUV color space. The YUV color space allows the separation of the luminance and the color in different components allowing a better threshold than the RGB image.

**Ground Plane.** After the threshold is selected a set of points acquiring the correspondent depth for each point based on a  $3 * 3$  mask.

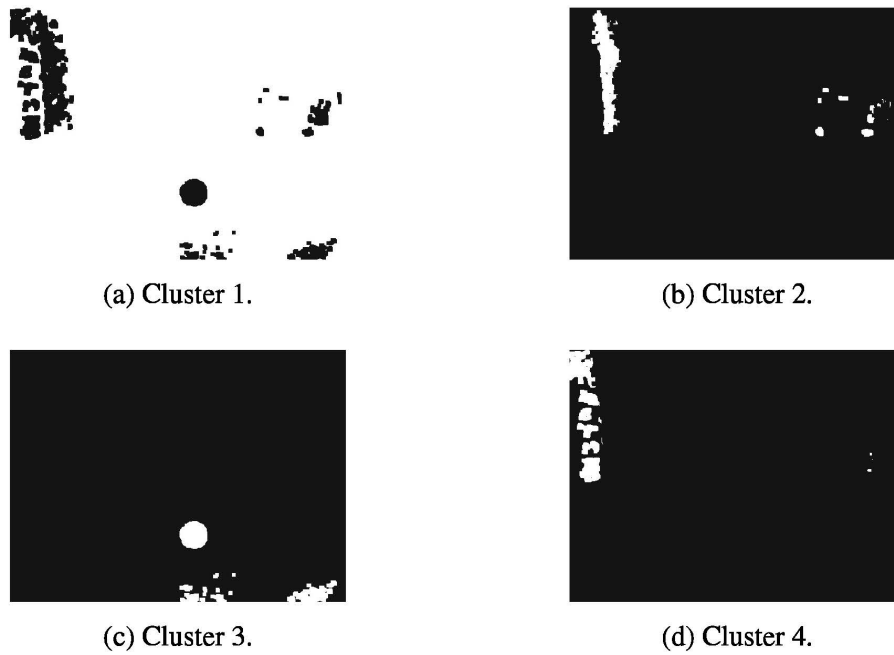
These points are used to calculate the ground plane parameters using the RANSAC algorithm on the beginning of the match. In order to remove the ground floor thus allowing to cluster the image and get similar results as in the aerial ball clustering.

**Results.** The next procedure is the same as used in the aerial ball which is the clustering of the example figure 3. In the figure 4 and 5 it is possible to see that after the football field segmentation the ball appears disconnected from other objects allowing to perform the recognition of the different image objects.



**Fig. 3** Depth Image with Ground Ball.

It was possible to see that the clustering of the depth image allows the detection of the ball in the image as an object disconnected from others objects with defined characteristics that are analyzed by the recognition system detailed in the following section. The process time of this task (including the recognition process) is slower than the aerial ball case taking a total time of  $\sim 80$ ms, on an Intel i5 processor, due the elimination of the ground planes takes a long time.



**Fig. 4** Clustering of the Depth Image in the figure 3.



**Fig. 5** Clustering of the Depth Image in the figure 3.

## 4 Object Recognition

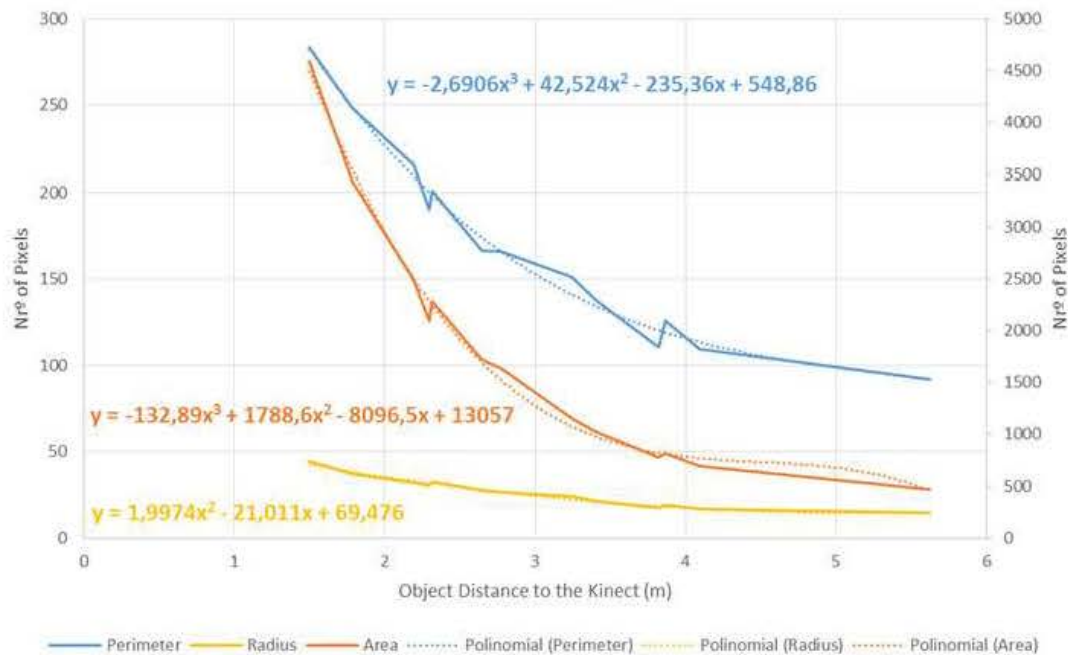
This section addresses the recognition process that is performed after the segmentation previously presented.

To develop an algorithm capable of identifying the presence of football ball is used the knowledge of the ball geometrical properties.

The features that yields better results to the conditions that are imposed to the system are: area, perimeter, circularity and the radius of the ball.

To create a dynamic recognition model was analyzed the variations of the considered object features through a series of test yielding the results present in the figure 6.

This results are approximated by a polynomial function that models the value of that feature.



**Fig. 6** Fitting curves to the different measured ball features, based on the distance tests to the Kinect.

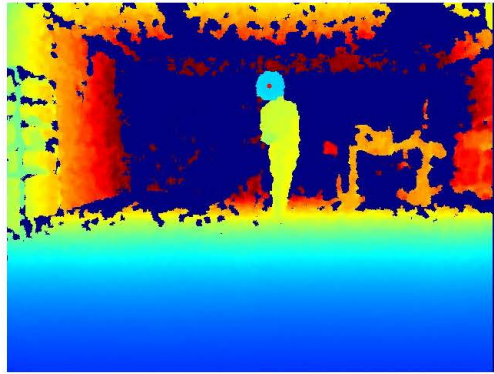
The developed ball recognition system initially removes the objects that exceed the maximum and minimum geometric characteristics considered. These limits were experimentally found taking into account the limitations of the equipment used, that is, the boundaries of the distance sensor that is able to return consistent results.

After this first step it is obtained, for the remaining objects, their distance to the camera using a 3x3 mask in order to avoid noise that may arise in the center of mass pixels of the object. Using the functions that were previously calculated that indicate the expected value for each characteristics. The objects that are out of the defined interval error around the expected value are discarded.

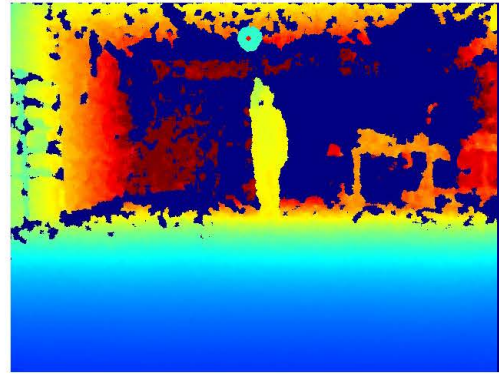
At the end it is considered that the object that passes all of these steps is considered to be a football ball.

## 4.1 Results

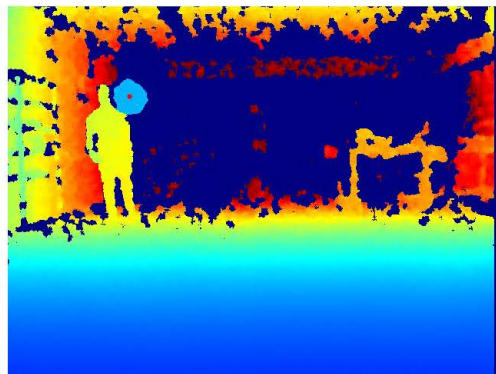
The results of this recognition system is presented in figures 7 and 8 where the football ball is shown with a red point.



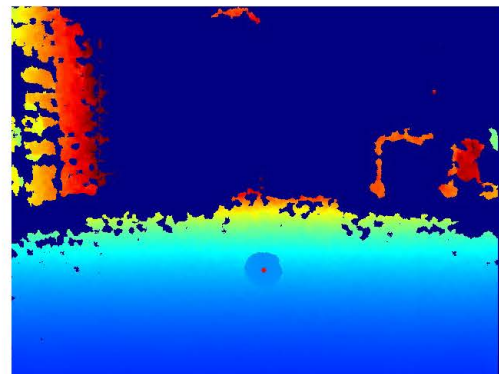
(a) Aerial Ball Result 1.



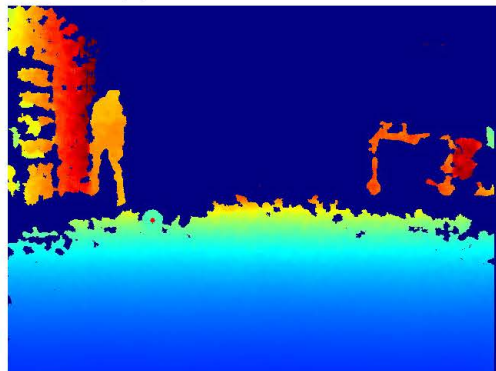
(b) Aerial Ball Result 2.

**Fig. 7** System Recognition Results.

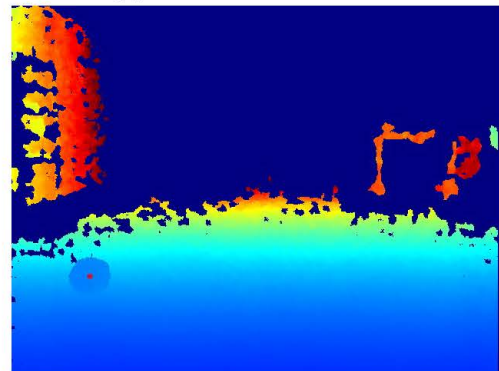
(a) Aerial Ball Result 3.



(b) Ground Ball Result 1.



(c) Ground Ball Result 2.



(d) Ground Ball Result 3.

**Fig. 8** System Recognition Results.



Finally it is possible to see that the recognition system yields good results. However it is not always possible to recognize the ball in every image frame because of the image acquisition rate or the natural occlusion that can occur, so to counter act this problem a Kalman Filter is implemented to predict the ball position when it is not possible to detect the presence of the ball as present in next section.

## 5 Kalman Filter

### 5.1 Aerial Ball Model

The model used for the case of the aerial ball is the model with constant acceleration. This model is used because being the ball in the air is knew that the gravity force is always present in the ball movement.

Equation 1 details the state of the ball used by the Kalman Filter. Where  $x$ ,  $y$  and  $z$  are the position of the ball in the world.

$$X = [x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z} \ \ddot{x} \ \ddot{y} \ \ddot{z}]^T \quad (1)$$

The calculation of the next state of the ball is done by using the equation 2 taking into account that there is no control system, the value  $B$  matrix is zero.

$$X_{k+1} = A \cdot X_k + B \cdot u_k \quad (2)$$

$$X_{k+1} = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{1}{2}\Delta t^2 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{1}{2}\Delta t^2 & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{1}{2}\Delta t^2 \\ 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot X_k \quad (3)$$

The calculation of the object's position is possible using the equation 4.

$$Y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot X \quad (4)$$

## 5.2 Ground Ball Model

For the next case the gravity force is present however, it doesn't make much interference in the movement. In this case the ball movement is detailed by the force applied by the kick executed by other robot.

It was considered that the ball velocity is equal through its movement, so it was used the constant velocity model to filter the ball position, as present in equation 5.

$$X = [x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z}]^T \quad (5)$$

The calculation of the next state of the ball is done by using the equation 6 taking into account that there is no control system, the value  $B$  matrix is zero and  $A$  is shown in 6.

$$X_{k+1} = A \cdot X_k + B \cdot u_k \quad (6)$$

$$X_{k+1} = \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}_k \quad (7)$$

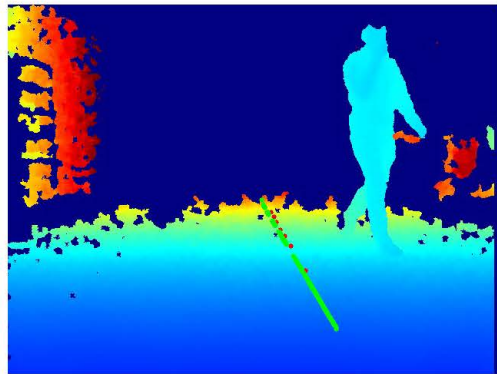
The calculation of the object's position is possible using the equation 8.

$$Y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \cdot X \quad (8)$$

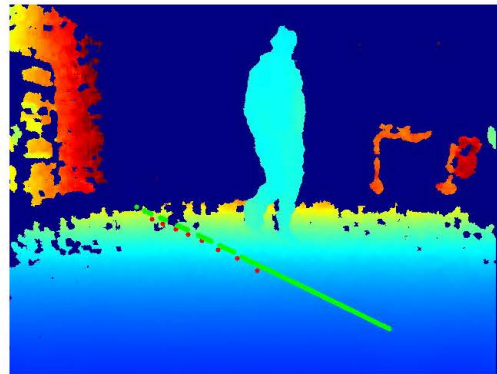
## 5.3 Results

The results of the Kalman Filter are showed in the figure 9, where the result of the recognition system are showed using a red point and the result of the Kalman Filter is showed using a green line.

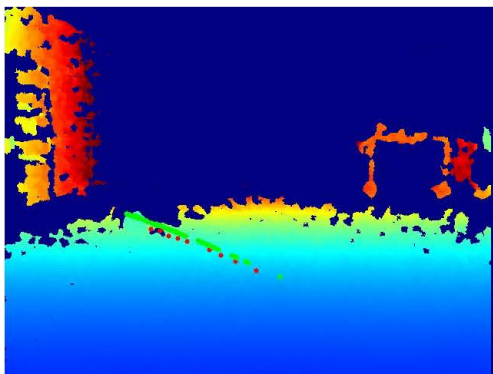
Table 1 shows the mean error of the actual position of the ball in comparison with the position returned by the Kalman filter. It is possible to see that the Kalman Filter gives a larger error in the aerial ball case than in the ground ball case. This fact is because the process and measure noise matrix are very challenging to determine. Where the number of pixels is the difference between the calculated center of the ball and the true center of the ball.



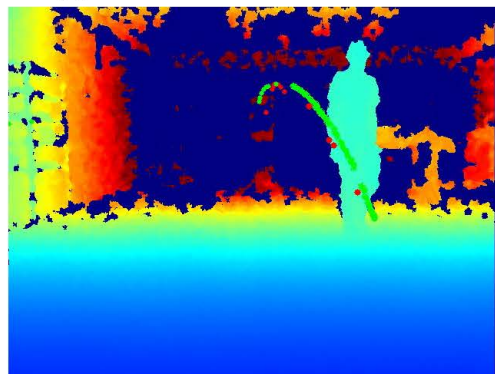
(a) Kalman Ground Ball Result 1.



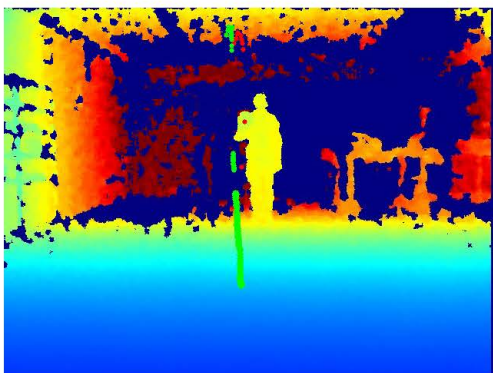
(b) Kalman Ground Ball Result 2.



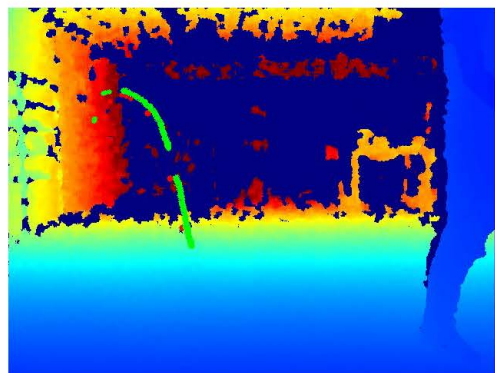
(c) Kalman Ground Ball Result 3.



(d) Kalman Aerial Ball Result 1.



(e) Kalman Aerial Ball Result 2.



(f) Kalman Aerial Ball Result 3.

**Fig. 9** Kalman Filter Results.

**Table 1** Mean error of the actual position of the running ball in comparison with the position returned by the Kalman filter.

	Number of Pixels
Kalman Aerial Ball 1	12.3241
Kalman Aerial Ball 2	15.0333
Kalman Aerial Ball 3	13.684
Kalman Ground Ball 1	3.2083
Kalman Ground Ball 2	3.6056
Kalman Ground Ball 3	6.2111

## 6 Conclusions

Summarily, it was observed that the separation of the detection and recognition of the football ball problem in two cases, aerial ball and ground ball, was beneficial because they present different problems which allowed the increasing of the algorithm speed.

The aerial ball case is quite interesting because it appears that while it is in the air it is disconnected from any image element which allows to quickly cluster the image based in the depth and make the ball recognition three times faster than for the ground ball case.

In the ground ball case it appears that the object is connected to the ground which difficults the cluster processing because it needs a lot of more cluster to provide the same results as in the aerial ball case.

So, it was necessary to eliminate the points that correspond to the ground plane. This process slows the algorithm  $\sim 40$ ms, on an Intel i5 processor, but this strategy does not become a bad choice.

The recognition system returns good results using that four features of the ball but it does not guarantee the recognition of the object in every frame due to the ball quick movement or occlusions.

To counter act this problem a Kalman filter was introduced to fill the gaps of the recognition algorithm. This filter can predict the position of the ball in an effective manner according to the inserted system model. It was observed that the biggest challenge is the modeling of the process noise and measurement matrix that can interfere significantly in its predictions.

This article illustrates one more time the utility of RGB-D sensors in robotic football. In the future it is expected that this algorithm is enhanced to detect and recognize the ball faster and more effectively.

For future developments it is proposed to improve the recognition algorithm that aims to circumvent the effects of the image depth by rapid movement of the ball. Also the method of choosing the aerial ball method or the ground ball method in a given time can be more studied.

The use of a depth sensor with better quality would be a plus because it was possible to get more accurate results avoiding some noise that comes in the picture.



**Acknowledgment** This work is financed by the ERDF – European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project FCOMP-01-0124-FEDER-037281.

## References

1. Marcel, J.: Object Detection and Recognition with Microsoft Kinect (2012)
2. Pheatt, C., Ballester, J.: Using the Xbox Kinect Sensor for Positional Data Acquisition, 1–14 (2011)
3. Tombari, F., Stefano, L.D.: Hough Voting for 3D Object Recognition under Occlusion and Clutter. *IPSN Transactions on Computer Vision and Applications*
4. Lowe, D.G.: Object recognition from local scale-invariant features. In: *IEEE International Conference on Computer Vision* (1999)
5. Rusu, R.B.: *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments* (2010)
6. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: *Computer Vision – ECCV 2006. Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, May 7–13, 2006. Lecture Notes in Computer Science*, vol. 3951, pp. 404–417 (2006)
7. Fischler, M.A., Bolles, R.C.: *Communications of the ACM* **24**, 381–395 (1981)
8. Qu, Z., Wang, J.: A color YUV image edge detection method based on histogram equalization transformation. In: *Natural Computation (ICNC)* (2010)
9. Budden, D., Fenn, S., Walker, J., Mendes, A.: A novel approach to ball detection for humanoid robot soccer. In: *AI 2012: Advances in Artificial Intelligence. LNCS*, vol. 7691, pp. 827–838 (2012)
10. Dias, P., Silva, J., Castro, R., Neves, A.J.R.: Detection of aerial balls using a kinect sensor. In: *RoboCup 2014: Robot World Cup XVIII. LNCS*, vol. 8992, pp. 537–548 (2015)