

Mobile Path and Spin 3D Tracking and Reconstruction

Federico Cristina¹, Sebastián H. Dapoto¹, Claudia Russo¹, Armando Degiusti¹,
and María José Abásolo²

¹ Instituto de Investigación en Informática LIDI
Facultad de Informática Universidad Nacional de La Plata
La Plata, Buenos Aires, Argentina
{fcristina, sdapoto, crusso, degiusti}@lidi.info.unlp.edu.ar

² Departamento de Matemática e Informática de la Universidad
de las Islas Baleares
Palma, Baleares, España
{mjabasolo}@uib.es

Abstract. The branch of Computer Science known as digital image processing includes different topics of Investigation and Development, as well as applications encompassing different stages that go from data acquisition, enhancement, and segmentation, up to the analysis, classification and interpretation of images [1]. Particularly, the reconstruction of 3D movements from 2D images (photos, filming) is a complex area, which becomes significant when it is a matter of obtaining real time responses. [1, 2, 3]. This project aim is developing a three-dimensional analysis system considering the processing of a soccer ball trajectory and rotational speed for its later computer-generated graphical modeling. The objective of the system is to improve the player's skills and the training methodology, and is framed within the research line of this Institute and within the area of signal and image processing.

Keywords: Image processing, objects tracking, path, spin.

1 Introduction

Path tracking and analysis is of particular importance in industrial environments, especially robotics. [4]

Focusing now on the subject matters of this project, we shall mention those robot applications called car-like robots. These vehicles have to move in unknown environments and avoid obstacles fast and effectively.

They are used, for instance, in games employing robots, which move in a small scenario taking the ball towards the rival goal. They have also been found in some tests with games of several sports, in which the ball is tracked in order to analyze shots [5, 6].

Path tracking requires a previous retrieval of the environment, be it manual or automatic, in order to obtain the information on the location and dimension of potential obstacles. On the other hand, if the generated path is to be tracked, we need to detect the specific object within the environment, so as to place it along the path and make the corresponding decisions. [5, 7]

The system aims at improving the players' free-shot technique by means of the immediate display of the results obtained after each attempt. It also allows qualifying each shot according to the achieved precision and speed, showing comparatives among the different attempts. Thanks to this, each player will have its history, which will let us make note of their evolution.

3D reconstruction of the moving object (the ball) is of utmost importance since not only can the observation point vary but also other effects of sports importance can be studied (for example, ball spin on its axis over a lineal path). This requires identifying the points of the mobile object and reconstructing the spin path simultaneously with the advance path [4, 8].

This aspect presents a particular complexity since the ball's translational movement should be properly isolated from that of rotation. We should also take into account the labeling criteria over the surface in order to obtain the proper results.

In order to carry out the system development, several steps had to be followed, including camera calibration, image filtering, point triangulation in space, path obtaining, and ball spin. The following section presents a theoretical explanation of each technique used and the developed implementation for its solution.

2 Camera Calibration

In order to obtain a proper triangulation of the object in space, we have to know first the characteristics of the camera to be used [8, 11, 12, 14].

There exist two sets of parameters that should be obtained by the camera calibration process. These are the intrinsic and extrinsic parameters.

2.1 Intrinsic Camera Calibration

Camera intrinsic calibration allows us to obtain its intrinsic parameters. These do not depend on the orientation and position of the camera within the 3D world; they are typical of this camera as optical device. The intrinsic parameters are the following:

- Principal point or position of image center with respect to the camera reference system. It is the intersection point between the camera optical axis and the image plane.
- Focal length: distance separating the image plane optical center.
- Scale factors for x and y axis, to convert pixels into units.
- Distortion coefficients. Geometric distortion affects the points in the image plane, as a result of a series of imperfections in the manufacture and assemblage of the lenses of the optical system. There are three types of distortion: radial, decentering, and prismatic.

2.2 Extrinsic Camera Calibration

Extrinsic calibration of a camera allows us to obtain its extrinsic parameters. These define the orientation and position of the camera in relation to a determined coordinate system, which is known as the world coordinate system. Three parameters

define the movement and another three parameters (three angles with respect to axis x and y) define orientation.

2.3 Calibration Methodology in the Developed System

A camera intrinsic and extrinsic calibration module has been developed. This has the following features: left and right camera calibration, and stereo camera.

2.3.1 Left and Right Camera Calibration

In order to obtain the intrinsic calibration, we have used a known-size calibration grid, as Figure 1 shows. A series of shots are obtained with different positions and the grid's intersections are detected.

Figure 2 shows how the left and right camera were calibrated. For this, it was necessary to inform the system about the size of the frames of the grid (in millimeters) and the quantity of internal points of intersections (both horizontal and vertical).

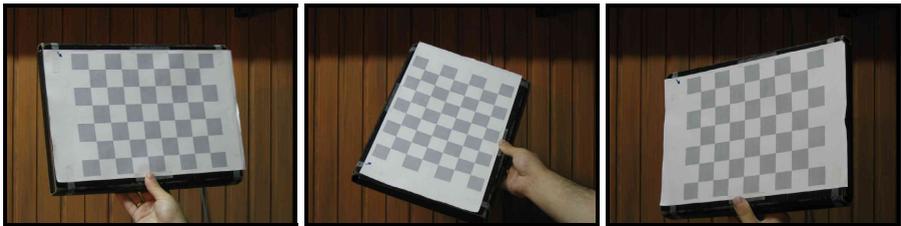


Fig. 1. Different shots of the Calibration Grid

| | | |
|--|--|---|
| Calibración de Cámara: Derecha Tamaño de los cuadrados (mm): 30 Cantidad de Intersecciones: 7 x 5 <input type="button" value="Seleccionar las imágenes y calibrar"/> | | Información Intrínseca Cámara Izquierda: Distancia Focal X: 793.960265805237 Distancia Focal Y: 793.960265805237 Punto Principal X: 365.627888461818 Punto Principal Y: 212.205391058194 Distorsion Param1: -0.250799716105071 Distorsion Param2: 0.37266010621031 Distorsion Param3: 0.00436802357410647 Distorsion Param4: -0.00641981160968308 |
| Calibración Stereo Tamaño de los cuadrados (mm): 30 Cantidad de Intersecciones: 7 x 5 <input type="button" value="Seleccionar las imágenes y calibrar"/> | | Información Intrínseca Cámara Derecha: Distancia Focal X: 763.85569902998 Distancia Focal Y: 763.85569902998 Punto Principal X: 353.534072207849 Punto Principal Y: 227.363459528849 Distorsion Param1: -0.167173996179779 Distorsion Param2: -0.0923596371782595 Distorsion Param3: 0.000507731311851995 Distorsion Param4: -0.00615210287931833 |

Fig. 2. Information obtained after Camera Calibration

2.3.2 Stereo Calibration

Once both cameras are calibrated, the following step entails the stereo or extrinsic calibration. For this, it is necessary to inform once again the features of the grid used. Once this is finished, a stereo image of the left camera and a stereo image of the right



Fig. 3. Images corresponding to Stereo Calibration

camera are selected. Both images should be correspondent to one another, as Figure 3 shows.

3 Filtering of the Object of Interest

Since a color filtering technique is used, the corresponding module is developed in order to detect without mistakes the red ball in each of the squares making up the videos to process. Such module allows setting up a configuration of the maximum and minimum allowable color thresholds, such as follows:

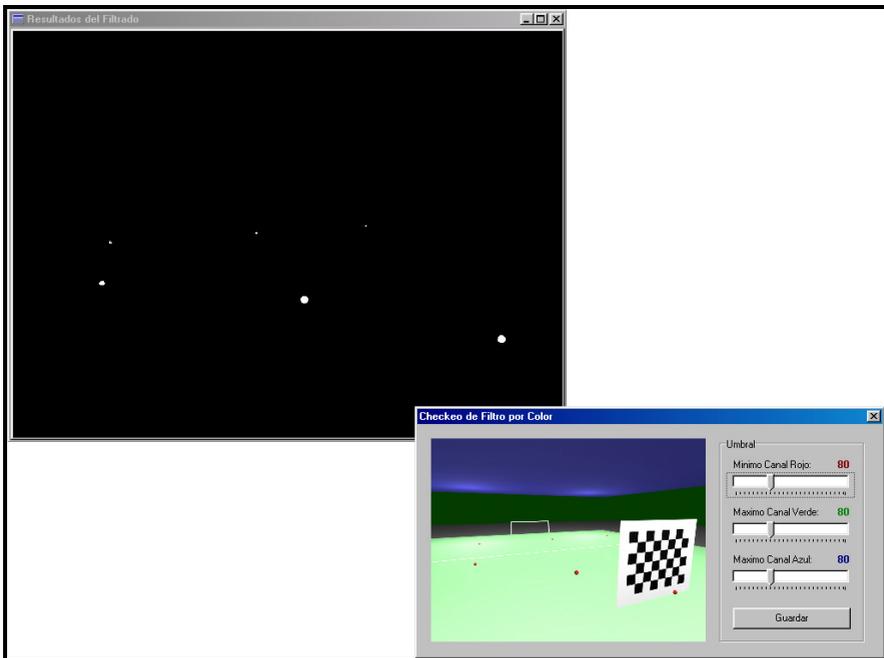


Fig. 4. Ball Color Filtering

- Red Channel Minimum: it represents the minimum allowable red intensity of the image. Those pixels with a red intensity lower than such value will not be taken into account.
- Green Channel Maximum: it represents the maximum allowable green intensity of the image. Those pixels with a green intensity higher than such value will not be taken into account.
- Blue Channel Maximum: it represents the maximum allowable blue intensity of the image. Those pixels with a blue intensity higher than such value will not be taken into account

4 Path Computation and Visualization

Once each stereo pair corresponding to each other and making up the video sequence are filtered, the ball triangulation is to be carried out. This technique is referred to as stereovision. [9, 16].

4.1 Stereovision

The correspondence problem is the most difficult to solve within stereovision. It means deciding which points of two images are the projection of the same real point. Once I_1 and I_2 are established to be the projection of the same point P in the two image planes R_1 and R_2 , it is possible to obtain the 3D coordinates of such point by triangulation.

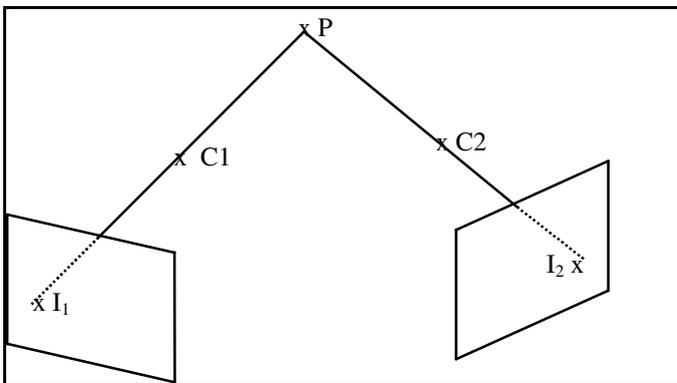


Fig. 5. Corresponding Points in Two Stereo Images

In principle, any point in the image plane R_2 could correspond to any point in R_1 . In order to solve this ambiguity, certain geometrical restrictions depending on the image system setup (camera positions) could be used. The most important is the epipolar restriction, which allows us converting a two-dimensional search into a one-dimensional.

4.2 Object-of-Interest Triangulation Methodology in the Developed System

The ball triangulation is carried out by the corresponding video processing. The processing time will mainly depend on the quantity of squares per second of the videos (the higher the quantity of squares per second, the longer will be the processing time) and the resolution of the images making up the video (the higher the resolution, the larger the volume of information to be processed, and thus the longer the processing time).

Once the video sequence is processed, enough information is obtained to render the following results:

1. Numerical Information:
 - a. Initial ball acceleration.
 - b. Point-to-point ball speeds along the path.
 - c. Maximum ball height.
 - d. Average shot speed.
 - e. Ball location in space.
2. Visual Information:
 - a. Shot starting point.
 - b. Impact point in the barrier or arrival at the goal line.
 - c. Real path description.
 - d. Estimated path description
 - e. Comparison between the real and the estimated path.

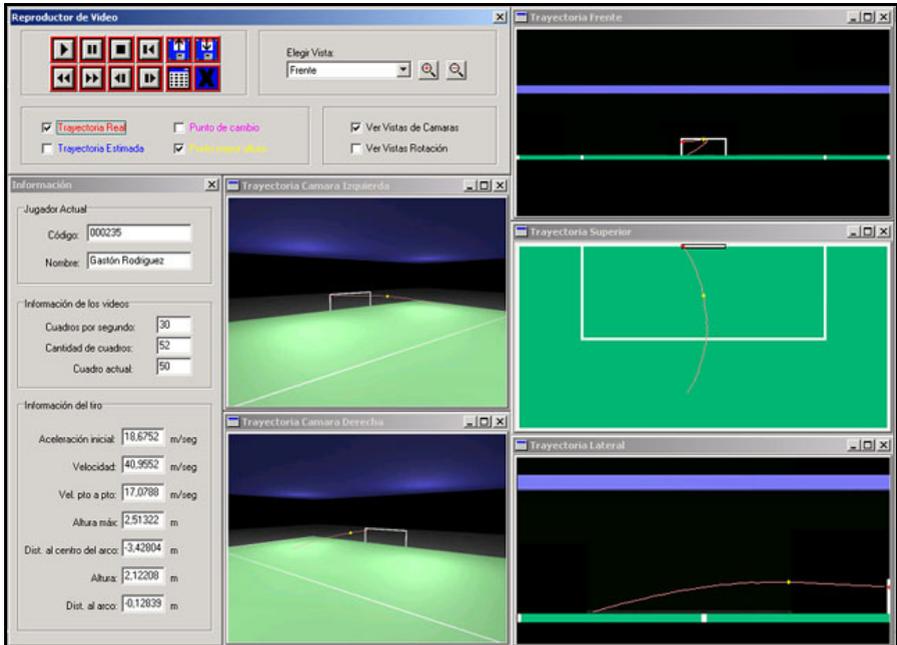


Fig. 6. Shot Playback Environment

The shot playback environment allows functionalities similar to those of a video player with different views (front, upper, lateral) as well as shot numerical data, such as figure 6 shows.

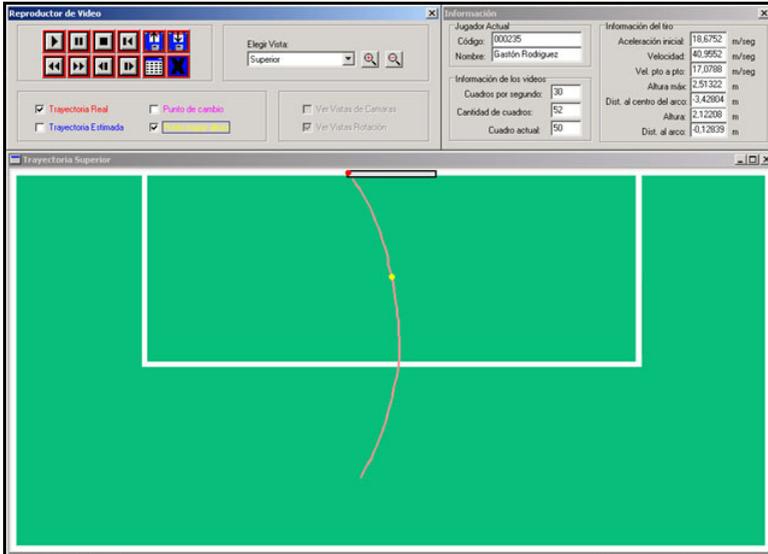


Fig. 7a. Zooming in of the upper view in the Shot Playback environment

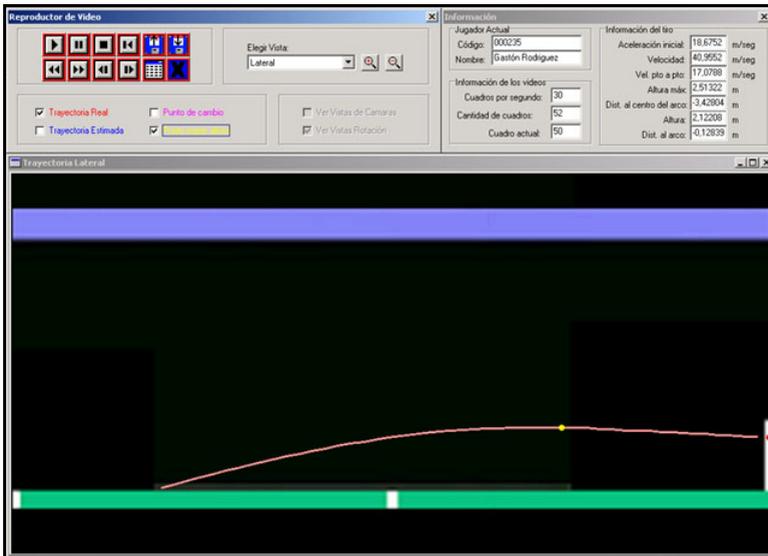


Fig. 7b. Zooming in of the lateral view in the Shot Playback environment

The shot playback environment also presents the possibilities of storing the shots, for a later load and playback of them if necessary, without the need of re-processing. It also provides the possibility of zooming in one of the views for a greater detail of the path traced by the ball, such as figure 7 shows.

5 Spin Computation

In order to determine the ball spin speed, it is necessary to analyze the movement made by the ball between one frame and the next one. In these cases, movement estimation or optical flow techniques are used, which allows us to detect the movements generated along the video sequence. [18, 19].

5.1 Movement Estimation

It is a process by means of which the object (or pixels) movement between two images is measured. Figure 8 shows that the moving image can be expressed in a three-dimensional space resulting from moving two consecutive images through the time axis.

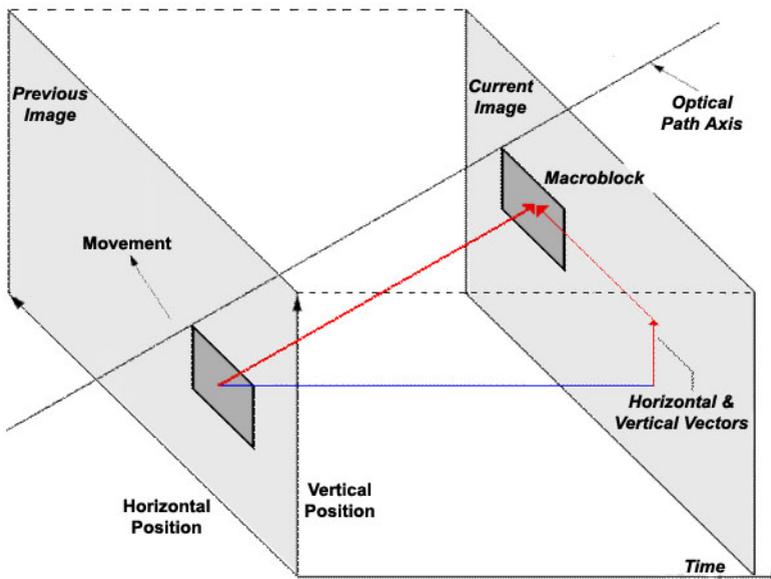


Fig. 8. Objects moving in a Three-Dimensional Space

In the case in which the object is static, its movement is only seen over the time axis. However, when an object is moving, it moves over the optical path axis (horizontal and vertical axis in time) which is not parallel to the time axis.

5.2 Object-of-Interest Spin Computation Methodology in the Developed System

The main problem to solve in this case entailed determining the correct labeling over the ball surface in order to know the movement the ball is making at all times.

Several alternatives [10] have been analyzed and we found that the best option – according to the project’s requirements – is to make marks with a distinctive color. In this case, yellow marks are used to distinguish them from the ball, the goal, and the field. The distribution consists of two marks per axial axis, such as figure 9 shows.



Fig. 9. Display of marks in the ball

Thus, there exists the problem of identifying each mark individually. This means, given an image and a mark, determining the new location of such mark in the following image.

Techniques of both movement estimation (block matching) and optical flow have been tested. We eventually decided to develop an ad-hoc method according to the posed requirements.

The procedure has three main stages: yellow marks filtering, their labeling, and label tracking.

Yellow marks filtering was carried out by means of light intensity level thresholding in RGB channels.

The labeling of marks consists in assigning an identifier (label) to each of the marks obtained after the filtering.

The objective of label tracking is to determine the position of each label in the previous and subsequent images of the referential one. Knowing the distance the marks have traveled between a frame and the following of the video sequence and its capture speed (FPS), we are able to determine the ball spin speed.

In order to obtain a higher precision when estimating the rotation, the procedure searches the frame along the video in which the ball shows a defined mark in the center. The three previously mentioned stages are carried out over the central mark.

The case this mark does not have any clear movement over the capture view point, or no defined central mark is found along the video may arise. That is why perimeter marks are taken into account.

The algorithm assures the existence of the selected mark along the frames to be used for the spin computation.

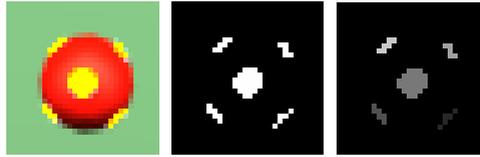


Fig. 10. Filtering and Labeling

Figure 10 shows the mark filtering and labeling process over the optimum image found along the video sequence.

A camera has been exclusively placed in order to carry out the ball spin computation at the time this comes closer to the goal.

6 Used Resources

6.1 Software

In order to ease the initial tests, a virtual model was made with the 3D Studio program. Thanks to this model, we were able to determine the technical requirements to be fulfilled by the cameras and their optimal position.

Then a ToolBox for MatLab [15] was used in order to understand the camera calibration process, the involved variables (focal length, aberrations, etc), and carry out real calibration tests.

An optimum language was then determined in order to carry out the application. The OpenCV library (“Open Computer Vision”) [17] was selected as adequate together with the C++Builder development environment.

It is important to mention that certain changes to the libraries making up OpenCV were made in order to adapt it to the project’s requirements. Due the considerably long distance between the cameras and the calibration board, the empiric bound for minimal allowed perimeter for squares and the minimal distance between image etalon points were changed to more suitable values.

6.2 Hardware

6.2.1 Cameras

After a detailed analysis, the optimum setup of the cameras making up the stereo system was determined. The cameras should have the following characteristics:

- 640x480 pixel resolution
- 100 Frame Per Second (FPS)
- 1/1000 of Shutter Speed

Using images of 640x480 pixels, we can attain a reasonable precision, taking into account that the greatest distance between the camera and the ball will be of approximately 20 meters.

Capturing 100 frames per second, the ball traces a path no bigger than 20 centimeters between a frame and the other.

With a shutter speed of 1/1000, the problem of capturing images in which the ball appears in the form of a tail will be avoided.

We opted to use two stereo systems instead of one in order to avoid occlusion problems and obtain a higher precision in the triangulation computations.

6.2.2 Capturing System

The digital capture system should allow large quantities of information per second. Using four cameras for the path computation and one camera for the spin computation, we will need a storage speed of approximately 450 Megabytes per second.

7 Achieved Results

As regards the path computation, a precision with a 10 cm maximum error at a maximum distance was obtained, taking into account the specified hardware - being able to minimize such error if we counted with images of higher resolution.

Each shot processing time is closely related to two main factors: quantity of frames per second (FPS) of the video and its resolution. The higher the FPS and/or resolution, the larger will be the required processing time. The tests carried out rendered the following results: it takes 20 seconds approximately to process the 5 videos without compression of 300 frames (3 seconds).

8 Current Research Lines

There exists an attempt to improve the shot playback environment and take it to a 3D playback environment. This aims at providing a visualization of the shots from any point in the space, thus allowing placing virtual cameras in endless locations for a better interpretation of results. [13]

9 Conclusions

An analysis and computer graphic modeling system was developed providing detailed information on the path and a ball spin computation, thus being able to obtain results allowing the technical improvement of players and their training methodology.

In order to develop it, several techniques, tools, and computer vision methods have been used. A customized spin computation method was also developed. In addition, a shots capturing, processing, and visualization environment was developed.

A future expansion of the system will be the three-dimensional representation of the path from any point of view, which will allow the analysis from strategic points of view, such as the goalkeeper's vision, the shooter player' vision, etc.

References

- [1] Digital Video Processing – A. Murat Tekalp – Prentice Hall - 1995
- [2] Visión por Computador, Fundamentos y Métodos – Arturo de la Escalera – Prentice Hall - 2001
- [3] Machine Vision – Ramesh Jain, Rangachar Kasturi, Brian G. Schunck – Mac Graw-Hill Internacional Edition - 1995
- [4] Real-time 3D Soccer Ball Tracking from Multiple Cameras. Jinchang Ren, James Orwell, Graeme Jones, Ming Xu. Digital Imaging Research Centre, Kingston University. Surrey, KT1 2EE, U. K. - 2004
- [5] Estimating the Position of a Football from Multiple Image Sequences. Jinchang Ren, James Orwell and Graeme A. Jones. Digital Imaging Research Center, Kingston University. Penrhyn Road, Kingston upon Thames, Surrey, KT1 2EE, UK - 2003
- [6] Fútbol de Robots Uruguayo para Torneos. Alvaro Castromán & Ernesto Copello. Tesis de Grado de la Carrera de Ingeniería en Computación. Facultad de Ingeniería - Universidad de la República - 2004
- [7] Tracking Players and a Ball in Video Image Sequence for Retrieving Scenes in Soccer Games. Akihito Yamada, Yoshiaki Shirai and Jun Miura. Dept. of Computer-Controlled Mechanical Systems, Osaka University. 2-1, Yamadaoka, Suita, Osaka 565-0871, JAPAN - 2000
- [8] Introductory Techniques for 3D Computer Vision – Emanuele Trucco, Alessandro Verri - Prentice Hall - 1998
- [9] An Effective and Fast Soccer Ball Detection and Tracking Method. Xiao-Feng Tong, Han-Qing Lu, Qing-Shan Liu. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences - 2003
- [10] An image recognition system for the measurement of soccer ball spin characteristics. Paul Neilson¹, Roy Jones, David Kerr¹ and Chris Sumpter. Institute of physics publishing - LE11 3EH, UK - 2004
- [11] Estereoscopía y Calibración de Cámaras. María Jose Abásolo - 2003
- [12] A compact algorithm for rectification of stereo pairs. Andrea Fusiello, Emanuele Trucco, Alessandro Verri. Machine Vision and Applications - 2000
- [13] 3D Polygon Rendering Pipeline. Greg Humphreys. University of Virginia CS 445 - 2003
- [14] A Flexible New Technique for Camera Calibration. Zhengyou Zhang. Technical Report MSR-TR-98-71 - 2002
- [15] Camera Calibration Toolbox for Matlab. Last access: Jul 2005. www.vision.caltech.edu/bouguetj/calib_doc/index.html
- [16] Seguimiento tridimensional usando dos cámaras. Pablo Barrera González, José María Cañas Plaza - 2004
- [17] OpenCV - Open Source Computer Vision Library Community. Last access: Feb 2006. <http://groups.yahoo.com/group/OpenCV>
- [18] Search Algorithms for Block-Matching in Motion Estimation. Deepak Turaga, Mohamed Alkanhal. Last access: Feb 2006. http://www.ece.cmu.edu/~ee899/project/deepak_mid.htm
- [19] Optical Flow. David Marshall. Last access: Oct 2005. http://www.cs.cf.ac.uk/Dave/Vision_lecture/node45.html