

Seguimiento y reconstrucción 3D en tiempo real de la trayectoria y rotación de un móvil

Federico Cristina, Sebastián H. Dapoto, Claudia Russo
{fcristina, sdapoto, crusso} @lidi.info.unlp.edu.ar

Instituto de Investigación en Informática LIDI
Facultad de Informática. Universidad Nacional de La Plata.
La Plata, Buenos Aires, Argentina.

CN Sports Argentina Especialización Futbolística S.A.
Plata, Buenos Aires, Argentina.

*III Workshop de Computación Gráfica, Imágenes y Visualización
XI Congreso Argentino de Ciencias de la Computación CACIC 2005*

Resumen

La rama de la Informática conocida como procesamiento de imágenes digitales comprende diferentes temas de Investigación y Desarrollo, así como aplicaciones que involucran las distintas etapas que van desde la adquisición de los datos de entrada, el mejorado, segmentación, hasta el análisis, clasificación e interpretación de imágenes. [1]

En particular la reconstrucción de movimientos 3D a partir de imágenes 2D (fotografías, filmaciones) es un área compleja, que se vuelve particularmente significativa cuando se trata de obtener respuestas en tiempo real. [1][2][3]

Este proyecto tiene por objetivo el desarrollo de un sistema de análisis tridimensional que contempla el procesamiento de la trayectoria y el cálculo de rotación de una pelota de fútbol para su posterior modelización gráfica por computadora. El objetivo del sistema es mejorar la técnica del jugador y la metodología de entrenamiento, y se encuentra enmarcado dentro de la línea de investigación del Instituto en el área de tratamiento de señales e imágenes.

Palabras claves: Procesamiento de imágenes, seguimiento de objetos, trayectoria, rotación

Abstract

The branch of the Computer Science known as digital image processing includes different topics of Investigation and Development, as well as applications that involve different stages that go from the acquisition of data, enhancement, 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's objective is the development of a three-dimensional analysis system that contemplates the processing of the trajectory and rotational speed of a soccer ball for its later computer generated graphical modeling.

The aim of the system is to improve the player's skills and the methodology of training, and is framed inside the line of investigation of the Institute in the processing of signals and images area.

Keywords: Image processing, objects tracking, trajectory, rotation

1 INTRODUCCIÓN

El tema de seguimiento y análisis de trayectorias tiene particular importancia en ambientes industriales, en particular en robótica. [4]

Orientándonos hacia los temas de interés de este proyecto, debemos mencionar las aplicaciones de robots denominados car-like robots. Estos vehículos tienen que moverse por entornos desconocidos y evitar obstáculos de forma rápida y efectiva.

Se utilizan por ejemplo en juegos que usan robots, los cuales se movilizan en un pequeño escenario llevando una pelota hacia el arco rival. También se han encontrado algunas pruebas con videos de partidos de diversos deportes, en las que se realiza un seguimiento de la pelota para un posterior análisis de los tiros. [5][6]

El seguimiento de trayectoria requiere de la previa captura del ambiente, ya sea de forma manual o automática, para obtener la información referente a la ubicación y dimensión de los posibles obstáculos. Por otra parte, si se desea seguir la trayectoria generada, se requiere detectar el objeto específico dentro del medio, para así poder ubicarlo a lo largo del trayecto y poder tomar las acciones correspondientes. [5][7]

El sistema tiene como meta mejorar la técnica de tiros libres de los jugadores mediante la exposición inmediata de los resultados obtenidos luego de cada intento. Permite además calificar cada tiro de acuerdo a la precisión y velocidad lograda; brindando comparativas entre varios intentos. Gracias a esto se tendrá un historial de cada jugador con el que se podrá notar la evolución de los mismos.

Resulta particularmente importante la reconstrucción 3D del objeto en movimiento (la pelota), ya que no sólo se puede variar el punto de observación sino que se pueden estudiar otros efectos de importancia deportiva (por ejemplo la rotación sobre su eje de la pelota dentro de un trayecto lineal). Esto requiere identificar puntos propios del objeto móvil y reconstruir la trayectoria de giro, simultáneamente con la trayectoria de avance. [4][8]

Este aspecto posee una complejidad particular; ya que debe aislarse adecuadamente el movimiento de traslación que sufre la pelota del de rotación. Se debe tener en cuenta además el criterio de etiquetado sobre la superficie de la misma con el fin de lograr resultados correctos.

Para llevar a cabo el desarrollo del sistema; fue necesario seguir una serie de pasos que abarcó la calibración de cámaras, el filtrado de imágenes, la triangulación de puntos en el espacio y obtención de la trayectoria y rotación de la pelota. A continuación se realiza una explicación teórica de cada técnica utilizada y la implementación desarrollada para su solución.

2 CALIBRACIÓN DE CÁMARAS

Con el fin de lograr una correcta triangulación del objeto de interés en el espacio, se requiere primeramente conocer las características de las cámaras a utilizar [8][11][12][14].

Existen dos conjuntos de parámetros que deben ser obtenidos mediante el proceso de calibración de cámaras. Éstos son los parámetros intrínsecos y extrínsecos.

2.1 Calibración Intrínseca de cámaras

La calibración intrínseca de una cámara permite obtener los parámetros intrínsecos de la misma. Éstos son los que no dependen de la orientación y la posición de la cámara en el mundo 3D, sino que son propios de la misma como dispositivo óptico. Los parámetros intrínsecos son los siguientes:

- Punto principal o posición del centro de la imagen, respecto del sistema de referencia de la cámara. Es el punto de intersección entre el eje óptico de la cámara y el plano de la imagen.
- Distancia focal: distancia que separa el centro óptico del plano de la imagen.
- Factores de escala para los ejes x e y, para convertir píxeles en unidades métricas.
- Coeficientes de Distorsión. La distorsión geométrica afecta a los puntos en el plano de la imagen; como resultado de una serie de imperfecciones en la fabricación y el montaje de las lentes que forman el sistema óptico. Hay tres tipos de distorsión: radial, descentral y prismática.

2.2 Calibración Extrínseca de cámaras

La calibración extrínseca de una cámara permite obtener los parámetros extrínsecos de la misma. Éstos definen la orientación y la posición de la cámara, respecto a un sistema de coordenadas determinado, al que se conoce como sistema de coordenadas del mundo. Tres parámetros definen el desplazamiento y otros tres parámetros (tres ángulos respecto a los ejes x, y, z) definen la orientación.

Los sistemas de referencia de las cámaras izquierda y derecha están relacionados a través de los parámetros extrínsecos. Estos definen una transformación rígida en el espacio 3D definida por un vector de traslación una matriz de rotación. Los puntos en los cuales la recta que une los centros de proyección de las cámaras corta a los planos de proyección se llaman epipolos. El plano epipolar, interseca cada imagen en una línea llamada línea epipolar. Se establece entonces una relación entre puntos de la imagen izquierda y rectas de la imagen derecha y viceversa. Por lo tanto se puede restringir la búsqueda de las correspondencias a lo largo de la línea epipolar, reduciendo la tarea a un problema de una dimensión.

2.3 Metodología de Calibración en el Sistema Desarrollado

Se desarrolló un módulo de calibración intrínseca y extrínseca de cámaras. El mismo cuenta con las siguiente funcionalidad: calibración de la cámara izquierda y derecha; y calibración estéreo.

2.3.1 Calibración de Cámaras Izquierda y Derecha

Con el fin de lograr la calibración intrínseca se utilizó una grilla de calibración de tamaño conocido, como muestra la figura 1. Se realiza una serie de capturas de la misma en distintas posiciones y se detectan las intersecciones de ésta.

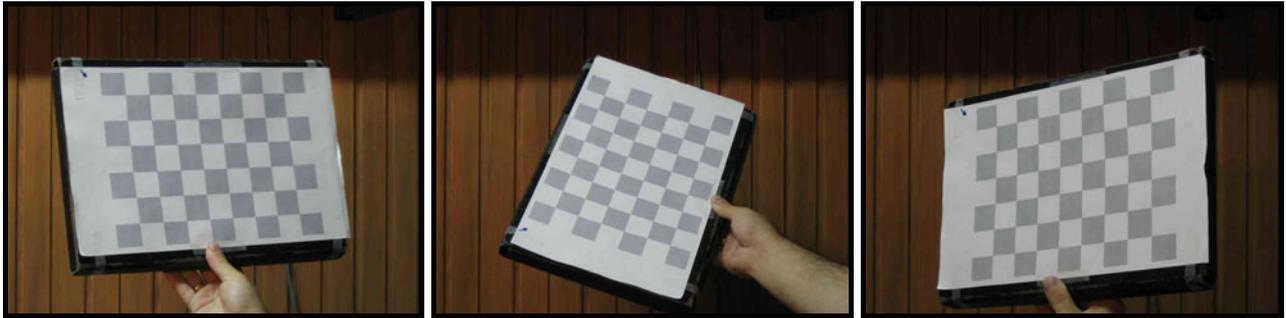


Figura 1: Diferentes Capturas de la Grilla de Calibración

En la figura 2 se puede observar cómo fueron calibradas tanto la cámara izquierda como la derecha. Para esto fue necesario informarle al sistema el tamaño de los cuadros de la grilla (en milímetros) y la cantidad de puntos de intersección internos que contiene la misma (tanto horizontal como verticalmente).

<p>Calibración de Cámara: Derecha</p> <p>Tamaño de los cuadros (mm): 30</p> <p>Cantidad de Intersecciones: 7 5</p> <p>Seleccionar las imágenes y calibrar</p>	<p>Información Intrínseca Cámara Izquierda:</p> <p>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</p>
<p>Calibración Stereo</p> <p>Tamaño de los cuadros (mm): 30</p> <p>Cantidad de Intersecciones: 7 5</p> <p>Seleccionar las imágenes y calibrar</p>	<p>Información Intrínseca Cámara Derecha:</p> <p>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</p>

Figura 2: Información Obtenida Luego de la Calibración de Cámaras

2.3.2 Calibración Estéreo

Una vez que ambas cámaras están calibradas, el siguiente paso es realizar la calibración estéreo o extrínseca. Para esto se debe nuevamente informar las características de la grilla utilizada. Una vez hecho esto se pasa a seleccionar una imagen estéreo de la cámara izquierda y una imagen estéreo de la cámara derecha. Dichas imágenes deben ser correspondientes entre si; tal como se puede observar en la figura 3.



Figura 3: Imágenes Correspondientes para Calibración Estéreo

3 FILTRADO DEL OBJETO DE INTERÉS

Debido a que se utilizó una técnica de filtrado por color, se desarrolló el módulo correspondiente con el fin de detectar de manera inequívoca la pelota roja en cada uno de los cuadros que componen los videos a procesar. Dicho módulo permite realizar una configuración del umbral de colores máximos y mínimos permitidos, tal como se detalla a continuación:

- Mínimo Canal Rojo: los píxeles que cuenten en su conformación con una intensidad de rojo por debajo de dicho valor no serán tomados en cuenta.
- Máximo Canal Verde: los píxeles que cuenten en su conformación con una intensidad de verde por encima de dicho valor no serán tomados en cuenta.
- Máximo Canal Azul: los píxeles que cuenten en su conformación con una intensidad de azul por encima de dicho valor no serán tomados en cuenta.

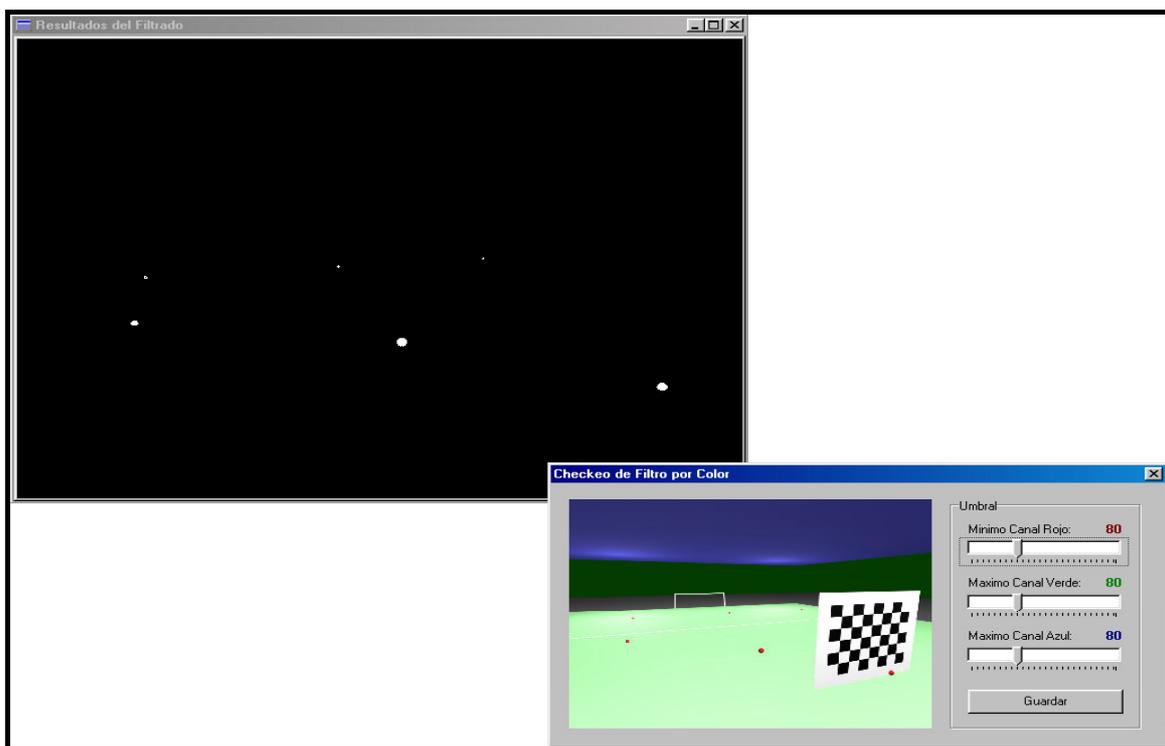


Figura 4: Filtrado por Color de la Pelota

4 CÁLCULO Y VISUALIZACIÓN DE LA TRAYECTORIA

Una vez que se han filtrado cada par de imágenes estéreo correspondientes entre si que conforman la secuencia de video, se deberá llevar a cabo la triangulación de la pelota. Esta técnica se denomina estereovisión. [9][16].

4.1 Estereovisión

El problema de la correspondencia es el de más difícil resolución dentro de la estereovisión. Se trata de decidir qué puntos de dos imágenes son la proyección de un mismo punto real. Una vez

establecido que I_1 e I_2 son la proyección de un mismo punto P en los dos planos imagen R_1 y R_2 , es posible ya obtener las coordenadas de dicho punto, mediante triangulación.

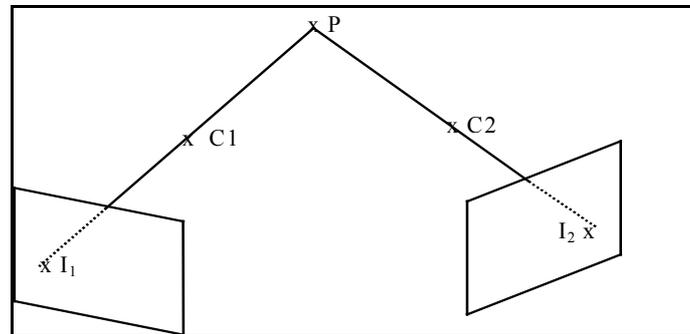


Figura 5: Puntos Correspondientes en Dos Imágenes Estéreo

En principio cualquier punto en el plano imagen R_2 podría corresponder a cualquier punto en R_1 . Para resolver esta ambigüedad se pueden emplear ciertas restricciones geométricas debidas a la configuración del sistema de imagen (posición de las cámaras). La más importante es la restricción epipolar, que nos permite convertir una búsqueda bidimensional en unidimensional.

4.2 Metodología de Triangulación del Objeto de Interés en el Sistema Desarrollado

La triangulación de la pelota se realiza mediante el procesamiento de los videos correspondientes. El tiempo de procesamiento dependerá principalmente de la cantidad de cuadros por segundo con que cuentan los videos (a mayor cantidad de cuadros por segundo, mayor tiempo de procesamiento) y la resolución de las imágenes que conforman el video (a mayor resolución mayor volumen de información a ser procesada, y por ende mayor tiempo de procesamiento).

Una vez procesada la secuencia de video, se obtiene la información suficiente para poder brindar los siguientes resultados:

1. Información numérica:
 - a. Aceleración inicial de la pelota.
 - b. Velocidades punto a punto de la pelota a lo largo de la trayectoria.
 - c. Altura máxima que toma la pelota.
 - d. Velocidad promedio del tiro.
 - e. Ubicación de la pelota en el espacio.

2. Información visual:
 - a. Punto de partida del tiro.
 - b. Punto de impacto en la barrera o de llegada a la línea de fondo.
 - c. Descripción de la trayectoria real.
 - d. Descripción de la trayectoria estimada.
 - e. Comparativa entre la trayectoria real y la estimada.

El entorno de reproducción de tiros permite funcionalidades similares a las de un reproductor de video acompañado de diferentes vistas (frente, superior, lateral) así como de los datos numéricos del tiro, tal como se muestra en la figura 6.

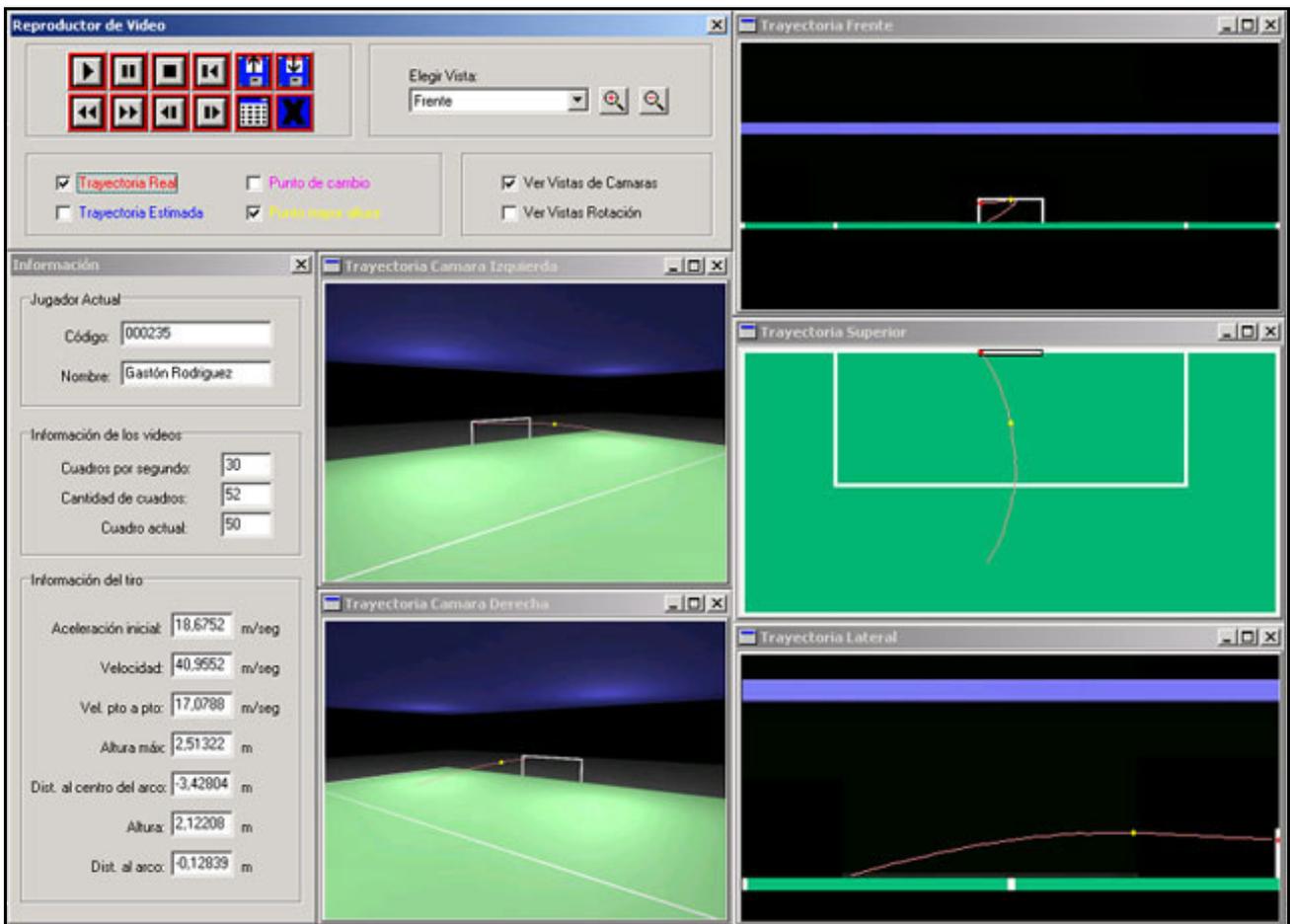
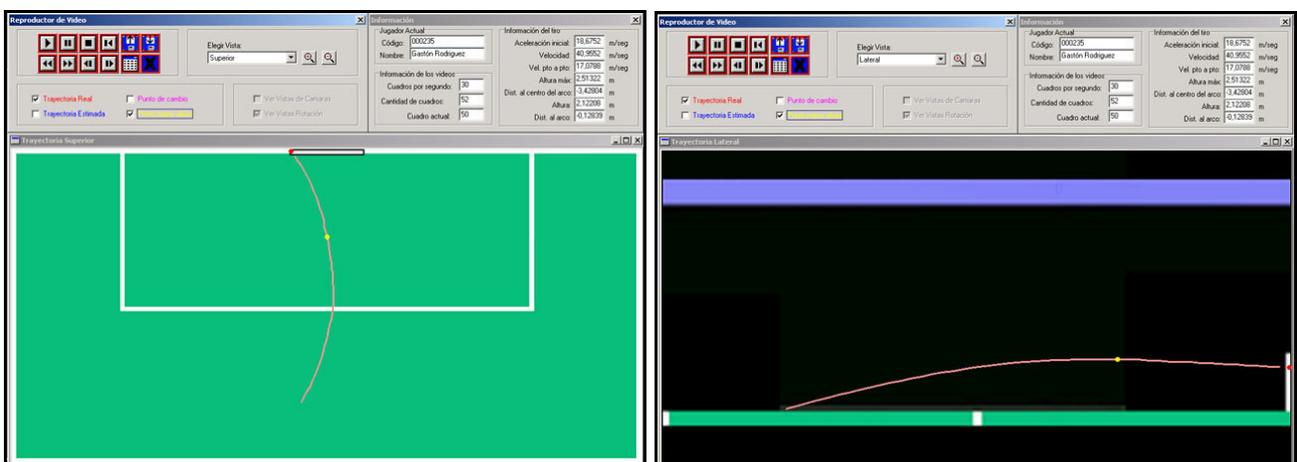


Figura 6: Entorno de Reproducción de Tiros

El entorno de reproducción de tiros presenta además la posibilidad de almacenar los tiros, para una posterior carga y reproducción del mismo cuantas veces sea necesario, sin necesidad de reprocesamiento. Brinda también la posibilidad de realizar un acercamiento sobre alguna de las vistas para un mayor detalle de la trayectoria que ha realizado la pelota, tal como lo muestra la figura 7.



(a)

(b)

Figura 7: Acercamiento de la vista (a) Superior y (b) Lateral en el Entorno de Reproducción de Tiros

5 CÁLCULO DE ROTACIÓN

Con el fin de determinar la velocidad de rotación de la pelota, se requiere analizar el movimiento que la misma ha sufrido entre un cuadro y el siguiente. Se utilizan en estos casos técnicas de estimación de movimiento o flujo óptico, las cuales permiten detectar los movimientos que se van generando a lo largo de la secuencia de video [18][19].

5.1 Estimación de Movimiento

Es un proceso mediante el cual se mide el movimiento de los objetos (o píxeles) de una imagen a otra. La figura 8 muestra que la imagen en movimiento puede ser expresada en un espacio de tres dimensiones que resulta de mover a lo largo del eje del tiempo dos imágenes consecutivas.

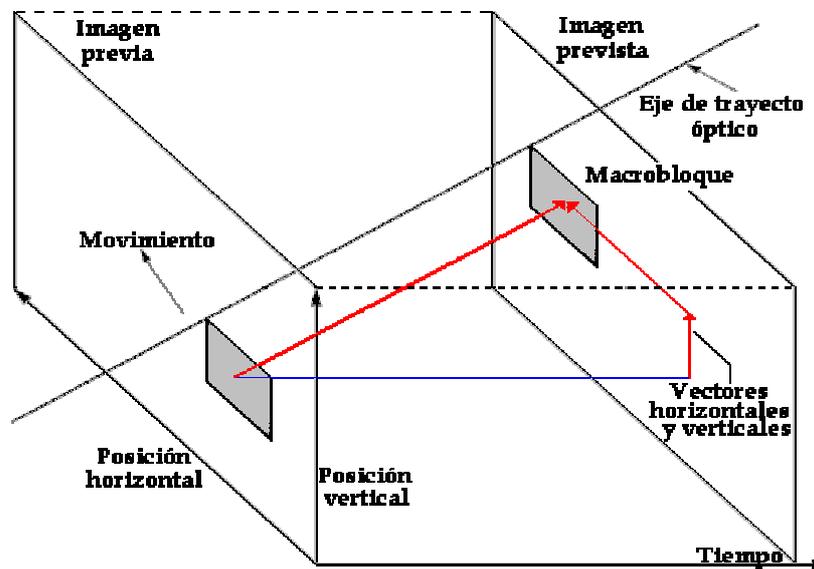


Figura 8: Objetos Viajando en un Espacio de Tres Dimensiones

En el caso de un objeto que permanezca estático, el movimiento de este solo se ve en el eje del tiempo. Sin embargo, cuando un objeto está en movimiento, éste se mueve en el eje de trayecto óptico (eje horizontal y vertical en el tiempo) que no es paralelo al eje del tiempo.

5.2 Metodología de Cálculo de Rotación del Objeto de Interés en el Sistema Desarrollado

El principal problema a resolver en este caso fue determinar un correcto etiquetado sobre la superficie de la pelota con el fin de poder conocer en todo momento el movimiento que ésta realiza.

Se han analizado varias alternativas [10] y se encontró que la mejor opción de acuerdo a los requerimientos del proyecto es la de realizar marcas con un color distintivo. En este caso se utilizaron marcas amarillas para distinguirlas de la pelota, del arco y del campo de juego. La distribución es de dos marcas por eje axial, tal como se puede apreciar en la figura 9.

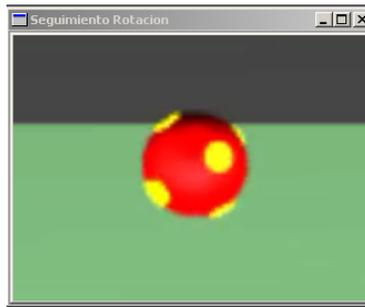


Figura 9: Disposición de las marcas en la pelota

Existe entonces el problema de identificar individualmente cada marca. Esto es, dada una imagen y una marca, determinar la nueva ubicación de dicha marca en la siguiente imagen.

Se han probado técnicas tanto de estimación de movimiento (block matching) como de flujo óptico. Finalmente se decidió desarrollar un método ad-hoc acorde a los requerimientos planteados.

El procedimiento consta de tres pasos principales: filtrado de marcas amarillas, etiquetado de las mismas, seguimiento de las etiquetas.

El filtrado de marcas amarillas se realizó mediante una umbralización por niveles de intensidad lumínica en los canales RGB.

El etiquetado de marcas consiste en asignarle un identificador (etiqueta) a cada una de las marcas obtenidas luego del filtrado.

El objetivo del seguimiento de las etiquetas es el de determinar la posición de cada etiqueta en las imágenes anterior y siguiente a la referencial. Conociendo la distancia que las marcas han realizado entre un cuadro y el siguiente de la secuencia de video y la velocidad de captura de la misma (FPS) se puede determinar la velocidad de rotación de la pelota.

Para lograr una mayor precisión en la estimación de la rotación, el procedimiento busca el cuadro a lo largo del video en el que la pelota muestre una marca definida en el centro de la misma. Se realizan luego los tres pasos explicados anteriormente sobre la marca central.

Se puede presentar el caso en el que esta marca no presente movimiento aparente sobre el punto de vista de la captura, o bien que a lo largo del video no se detecte una marca central definida. Es por esto que también se toman en cuenta las marcas perimetrales.

En la figura 10 se puede observar el proceso de filtrado y etiquetado de marcas sobre la imagen óptima encontrada a lo largo de la secuencia de video.

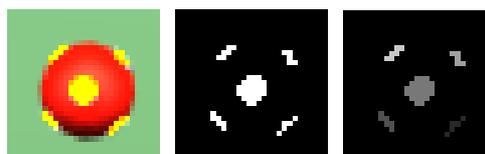


Figura 10: Filtrado y etiquetado

Se ha colocado una cámara exclusivamente para realizar el cálculo de rotación de la pelota al momento de que ésta se aproxime al arco.

6 RECURSOS UTILIZADOS

6.1 Software

Para facilitar las pruebas iniciales; se realizó una maqueta (modelo virtual) mediante el programa 3D Studio. Gracias a este modelo se pudieron determinar los requerimientos técnicos que deben cumplir las cámaras y su óptima posición.

Posteriormente se utilizó un ToolBox para MatLab [15] con el fin de comprender el proceso de calibración de cámaras, las variables que involucra (distancia focal, aberraciones, etc) y realizar pruebas de calibración reales.

Se determinó luego el lenguaje óptimo para realizar la aplicación. Se seleccionó como adecuada la librería OpenCV (“Open Computer Vision”: librería Open Source sobre visión por computador) [17] junto al entorno de desarrollo C++ Builder.

Cabe mencionar que fueron realizados ciertos cambios a la librerías que componen OpenCV con el fin de adaptar la misma a los requerimientos del proyecto.

6.2 Hardware

6.2.1 Cámaras

Luego de un minucioso análisis se determinó la óptima configuración de cámaras que componen el sistema estéreo; las cuales deberán presentar las siguientes características:

- Resolución de 640x480 píxeles
- 100 Cuadros Por Segundo (FPS)
- 1/1000 de Velocidad de Obturación

Utilizando imágenes de 640x480 píxeles se puede llegar a lograr una precisión razonable; teniendo en cuenta que la mayor distancia entre la cámara y la pelota será de aproximadamente 20 metros.

Capturando 100 cuadros por segundo, la pelota recorre una trayectoria no mayor a 25 centímetros entre cuadro y cuadro.

Mediante una velocidad de obturación de 1/1000 se evitará el problema de capturar imágenes en las que la pelota se muestre en forma de estela.

Se optó por utilizar dos sistemas estéreo en lugar de uno con el fin de evitar problemas de oclusión y lograr además mayor precisión en los cálculos de triangulación.

6.2.2 Sistema de Captura

El sistema de captura digital debe permitir almacenar grandes cantidades de información por segundo. Al utilizar cuatro cámaras para el cálculo de la trayectoria y una cámara para el cálculo de rotación se requerirá una velocidad de almacenamiento aproximadamente 450 Megabytes por segundo.

7 RESULTADOS OBTENIDOS

En lo que respecta al cálculo de trayectoria, se ha logrado una precisión con un error máximo de 10cm a la máxima distancia, teniendo en cuenta el hardware especificado; pudiendo minimizar dicho error de contar con imágenes de mayor resolución.

El tiempo de procesamiento de cada tiro se encuentra estrechamente relacionado con dos factores principales: la cantidad de cuadros por segundo (FPS) con los que cuenta el video y su resolución. A mayor FPS y/o resolución se requerirá mayor tiempo de procesamiento. Las pruebas realizadas arrojaron que para procesar los 5 videos sin compresión de 300 frames (3 segundos) se requieren 44 segundos aproximadamente.

8 LÍNEAS ACTUALES DE INVESTIGACIÓN

Se piensa mejorar el entorno de reproducción de tiros y llevarlo a un entorno de reproducción 3D. El fin de esto es brindar una visualización de los tiros desde cualquier punto en el espacio, permitiendo así colocar cámaras virtuales desde un sin fin de ubicaciones para una mejor interpretación de los resultados. [13]

9 CONCLUSIONES

Se ha logrado desarrollar un sistema de análisis y modelización gráfica por computadora que brinda información detallada de la trayectoria y cálculo de rotación de una pelota de fútbol; logrando de esta forma obtener resultados que permiten mejorar la técnica de los jugadores y la metodología de entrenamiento de los mismos.

Para realizar el desarrollo del mismo, se han analizado y utilizado variadas técnicas, herramientas y métodos de visión por computadora. Se diseñó a su vez un método de cálculo de rotación a medida. Se desarrolló además un entorno de captura, procesamiento y visualización de tiros.

Una futura ampliación del sistema será la representación tridimensional de la trayectoria desde cualquier punto de vista, lo que permitirá el análisis desde puntos de vista estratégicos tales como la visión del arquero, del jugador ejecutante, etc.

10 REFERENCIAS

[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] Estereoscopia 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.
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.
<http://groups.yahoo.com/group/OpenCV>
- [18] Search Algorithms for Block-Matching in Motion Estimation. Deepak Turaga, Mohamed Alkanhal. http://www.ece.cmu.edu/~ee899/project/deepak_mid.htm
- [19] Optical Flow. David Marshall. http://www.cs.cf.ac.uk/Dave/Vision_lecture/node45.html