

UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR



ANÁLISIS DE ACTIVIDAD EN VÍDEOS DE BALONCESTO

Rubén García García
Tutor: Rafael Martín Nieto
Ponente: José María Martínez Sánchez

-PROYECTO FIN DE CARRERA-

Ingeniería de Telecomunicación

Escuela Politécnica Superior
Universidad Autónoma de Madrid
Julio 2016

ANÁLISIS DE ACTIVIDAD EN VÍDEOS DE BALONCESTO

Rubén García García

Tutor: Rafael Martín Nieto

Ponente: José María Martínez Sánchez



Video Processing and Understanding Lab

Departamento de Tecnología Electrónica y de las Comunicaciones

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Julio 2016

**Trabajo parcialmente financiado por el gobierno español bajo el proyecto
TEC2014-53176-R (HAVideo)**



Resumen.

El gran abanico de posibilidades de trabajo que ofrecen los sistemas de análisis de contenido en vídeos deportivos hizo que se desarrollase en el grupo de procesamiento de vídeo (VPU-Lab) de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid una aplicación interactiva de detección y seguimiento de jugadores. Este prototipo consistía en la mejora de usabilidad e interacción con el usuario final de un sistema ya realizado previamente en el cual se conseguían buenos resultados tanto de detección como de seguimiento.

Dicha aplicación se había creado tanto para deportes individuales como colectivos, pero en este último caso sólo trabajaba con vídeos de fútbol. Debido a esto, surgió una nueva línea de investigación que consistía en la adaptación de este prototipo a otro deporte colectivo como es el baloncesto.

Por tanto, en este proyecto se ha trabajado sobre esa línea, solucionando la problemática presente y creando un primer prototipo sobre el cual seguir investigando y desarrollando. Se trabajó en la adaptación de la programación del sistema para que se pudiese ejecutar con un número distinto de vídeos entrantes.

A continuación se modificó la interfaz gráfica para que pudiese representar de una forma óptima los nuevos resultados obtenidos, manteniendo toda la funcionalidad que tenía en el anterior prototipo.

Seguidamente, se investigó la mejora de los resultados obtenidos en estos nuevos vídeos con el sistema que estaba implantando. Se trató de obtener los mejores resultados posibles tanto de detección como de seguimiento ajustando diferentes parámetros.

Por último, puesto que la interfaz es la plataforma visible de este proyecto, se ha trabajado en mejorarla dotándola de un nuevo módulo de creación de la homografía que hace que se simplifique aún más el uso de esta interfaz por el usuario.

Palabras clave.

Detección y seguimiento de personas, fusión de trayectorias, interfaz gráfica de usuario, vídeos deportivos, extracción de fondo.

Abstract.

The large range of functional possibilities that content analysis systems offer in sports videos has prompted the development, in Escuela Politécnica Superior of the Universidad Autónoma de Madrid, of an interactive detection and tracking application for players. This prototype consisted of a usability and end-user interaction improvement of a previously known system in which both detection and tracking results were satisfactory.

Such application had been created both for individual and collective sports, but in this last case, it only worked for soccer videos. Due to this, a new line of investigation arose, that consisted in the adaptation of this prototype to another sport such as basketball.

Therefore, this Project has worked among the same lines, solving the existing issues and creating a similar prototype on which to continue investigation and further development. The system programming was adapted, in order to be able to be executed with a different number of inputted videos.

Following this, the graphic interface has been modified, to be able to optimally represent the newly obtained results, maintaining the full functionality of the previous prototype.

Furthermore, an improvement of the obtained results in these new videos was researched, with the system that was being implanted. The goal was to obtain the best possible results in both detection and tracking by adjusting different parameters.

Lastly, due to the interface being the visible part of this project, there have been updates made to improve it, by including a new homography creation module, that simplifies the use by the end user even more.

Keywords

People detection and tracking, trajectories fusion, graphical user interface, sport videos, background subtraction.

Agradecimientos.

En primer lugar, quiero dar las gracias a Chema por darme la oportunidad de realizar este proyecto y a mi tutor Rafa por haberme dado la motivación necesaria para acabarlo. Por su ayuda constante y, a pesar de la distancia, siempre disponible para consultas y solucionar las dudas que fuesen surgiendo con gran paciencia y buen humor.

También, a cada uno de los componentes del VPU Lab que de una forma u otra me han ayudado a realizar este trabajo, especialmente Pencho. Se agradece que en este laboratorio exista un buen ambiente de trabajo y gran disposición de ayudar en todo momento.

Dedicar un espacio de agradecimiento a mis compañeros y amigos de la carrera, que me han acompañado en esta larga etapa de mi vida. Son muchas horas de apoyo y estrés compartido y, a pesar de eso, conseguir siempre pasar buenos momentos.

A todos y cada uno de mis amigos, que han estado ahí siempre que se ha necesitado, regalando risas y buenos momentos, y que han colaborado en meterme presión para que acabase ya el proyecto.

Mención especial debe tener mi familia, que es la que ha tenido que aguantar tanto momentos buenos como malos a lo largo de estos años. Fundamentalmente a mis padres y hermana, que son los que me han ayudado y confiado en mí en todo momento, y a mis abuelos, que han vivido el principio de esta etapa y seguro se sentirán muy orgullosos de que finalmente lo haya conseguido.

A todos muchas gracias.

Rubén García García

Julio 2016.

Índice general

Resumen	v
Abstract	VII
Agradecimientos	IX
1. Introducción.	1
1.1. Motivación.	1
1.2. Objetivos.	2
1.3. Estructura de la memoria.	2
2. Estado del arte.	3
2.1. Introducción al análisis de vídeos deportivos	3
2.2. Técnicas de análisis de vídeos deportivos	5
2.2.1. Extracción de fondo	6
2.2.2. Análisis de frente y detección de <i>blobs</i>	7
2.2.3. Identificación de equipo	8
2.2.4. Seguimiento del jugador y fusión de trayectorias	9
2.3. <i>Datasets</i>	10
2.3.1. CVLAB	10
2.3.2. CVBASE 06	11
2.3.3. APIDIS	12
2.4. Productos comerciales	14
2.4.1. SHOT TRACKER TEAM	14
2.4.2. LONGOMATCH	15
2.4.3. VIDSWAP	15
2.4.4. KROSSOVER	16
2.4.5. KINOVEA	16
2.5. Sistemas existentes en publicaciones de investigación	16
2.5.1. ROBUST CAMERA CALIBRATION AND PLAYER TRACKING IN BROADCAST VIDEO	16
2.5.2. AUTOMATIC PLAYER POSITION DETECTION IN BASKETBALL GAMES	17
2.5.3. A TEMPLATE-BASED MULTI-PLAYER ACTION RECOGNITION OF THE BASKETBALL GAME	18

2.5.4.	JOINT MULTI-PERSON DETECTION AND TRACKING FROM OVERLAPPING CAMERAS	18
2.5.5.	ROBUST REAL-TIME TRACKING OF MULTIPLE OBJECTS BY VOLUMETRIC MASS DENSITIES	19
2.5.6.	REAL-TIME POSITION ESTIMATION AND TRACKING OF A BASKETBALL	20
2.5.7.	BASKETBALL PLAYER TRACKING AND AUTOMATED ANALYSIS	20
3.	Sistema base.	23
3.1.	<i>Background Substraction</i>	24
3.2.	Post-procesamiento	24
3.3.	<i>Blob Extractor</i>	25
3.4.	<i>Blob Tracking</i>	25
4.	Aplicación de baloncesto	27
4.1.	Adaptación a videos de baloncesto	28
4.1.1.	Fusión de cámaras	28
4.1.2.	Modificación de la interfaz gráfica	30
4.2.	Modificación de parámetros	34
4.3.	Ventana de creación de la homografía	37
5.	Conclusiones y trabajo futuro.	45
5.1.	Conclusiones.	45
5.2.	Trabajo futuro.	46
	Bibliografía	48
	A. Presupuesto	53
	B. Pliego de condiciones	55

Índice de figuras

2.1. Diagrama de bloques del sistema global	6
2.2. Imágenes obtenidas de http://cvlab.epfl.ch/data/pom#basketball	11
2.3. Imágenes obtenidas de http://vision.fe.uni-lj.si/cvbase06/downloads.html	12
2.4. Imágenes de las cámaras de APIDIS	13
2.5. Mapa de la situación de las cámaras de APIDIS	14
3.1. Diagrama de bloques del sistema base. Fuente [1]	23
3.2. Diagrama de bloques de obtención de la trayectoria. Fuente [2]	26
4.1. De izquierda a derecha, imágenes de las cámaras 1,2 y 6 de APIDIS	28
4.2. Lista de asociación. Fuente [1]	29
4.3. Vista de la ventana principal	30
4.4. Vista de la ventana principal en ejecución	31
4.5. Vista de la ventana de corrección de trayectorias	32
4.6. Vista de la ventana de corrección en uso	33
4.7. Vista de la ventana de <i>Frames</i>	34
4.8. Imágenes de la máscara con VARNOISE=8 (izquierda) y VARNOI- SE=16 (derecha)	35
4.9. Imágenes de la máscara con ALPHA=0.0 (izquierda) y ALPHA=0.2 (derecha)	35
4.10. <i>Frame</i> del vídeo de salida con 2 (izquierda) y 4 (derecha) iteraciones de la función de erosión	36
4.11. <i>Frame</i> del vídeo de salida con un tamaño de ventana de 15x30 (izquier- da) y de 10x20 (derecha)	36
4.12. Ventana de la aplicación con la nueva opción de crear la homografía	37
4.13. Ventana de inicio de la creación de la homografía	38
4.14. Ventana con las imágenes cargadas	39
4.15. Ventana con los puntos elegidos marcados	40
4.16. Cuadros editables de las coordenadas elegidas por el usuario	40
4.17. Ventana con el botón de siguiente imagen remarcado	41
4.18. Ventana con una nueva imagen de otra cámara	41
4.19. Ventana con el botón de guardar la homografía remarcado	42
4.20. Mensaje de confirmación	42

Capítulo 1

Introducción.

1.1. Motivación.

En la actualidad, los vídeos deportivos han cobrado un gran interés tanto a nivel comercial como a nivel profesional. La alta demanda que el público provoca hace que las investigaciones y trabajos relacionados con este material hayan crecido y se continúe avanzando en la evolución relativa a este tema. Este hecho se demuestra en que cada vez las tecnologías usadas para el análisis de actividad de los deportistas van mejorando más y muchos profesionales de este campo se aprovechan de ello para aumentar las capacidades de los deportistas.

Uno de los campos que más destaca dentro del estudio tecnológico en el deporte es el análisis de actividad. Gracias a ello se pueden obtener numerosos datos pertenecientes al juego como puede ser la trayectoria seguida por un jugador, la velocidad del trayecto, la distancia recorrida, zona de influencia . . .

Todos estos datos son muy útiles en el baloncesto, que es un deporte muy seguido en diversos países del mundo, sobre todo el seguimiento de trayectorias, puesto que los movimientos y las jugadas tácticas que implican desplazamiento del jugador son una constante en este deporte. Por ello, muchos entrenadores y demás profesionales de este campo, recurren cada vez más a la tecnología y al análisis de vídeo para obtener mejores resultados.

Como respuesta a esta demanda, se va a realizar una adaptación del prototipo que se implantó y desarrolló en [1] y que se mejoró en [2]. La aplicación que se tiene realiza la localización y el seguimiento a partir de vídeos multicámara fija de fútbol, pudiendo corregir posibles errores en la detección y fusión de trayectorias de una forma sencilla mediante una interfaz que cualquier usuario puede manejar. El problema de este algoritmo, pese a conseguir muy buenos resultados, es que está centrado en el

fútbol, por lo que no se puede aplicar a otras disciplinas deportivas, como el caso del baloncesto.

1.2. Objetivos.

Dadas las motivaciones mencionadas anteriormente, el objetivo de este proyecto es encontrar la solución a esos problemas y realizar un primer desarrollo de la aplicación adaptada a baloncesto.

Para ello, se debe estudiar el trabajo desde el que se parte y realizar los cambios necesarios para que el funcionamiento de la aplicación sea adecuado en otro escenario completamente distinto como es el que ofrecen los vídeos de baloncesto. Estos cambios deben ser tanto a nivel interno (a nivel de desarrollador, modificando el código) como a nivel externo (modificaciones visuales de la aplicación).

Aparte de este objetivo principal, se debe trabajar también en la mejora de la interfaz, facilitando todavía más la interacción con el usuario final.

Cabe destacar que estos cambios van a realizarse orientados al caso del baloncesto, creando un primer prototipo sobre el cual ir trabajando para obtener mejores resultados.

1.3. Estructura de la memoria.

La memoria del proyecto consta de los siguientes capítulos:

- Capítulo 1: Introducción y motivación del proyecto.
- Capítulo 2: Estudio detallado del estado del arte de análisis de vídeos deportivos, en especial baloncesto, y de las aplicaciones comerciales e investigaciones realizadas existentes.
- Capítulo 3: Descripción del sistema base del que se parte.
- Capítulo 4: Trabajo desarrollado para la adaptación y mejora de la aplicación existente.
- Capítulo 5: Conclusiones y trabajo futuro
- Referencias y anexos.

Capítulo 2

Estado del arte.

En este capítulo se realiza un estudio del estado del arte del análisis de vídeos deportivos. Para ello, en primer lugar, se presenta el análisis de este tipo de vídeos y su aplicación e importancia en la actualidad (sección 2.1). Seguidamente, se han analizado las distintas técnicas presentes en el estado del arte y sus características (sección 2.2). Posteriormente, se han resumido los *datasets* públicos existentes relativos al baloncesto, entre los que se encuentra el que se ha elegido para este trabajo, (sección 2.3), así como diversos productos comerciales con sus características y utilidades (sección 2.4). Por último, se han investigado algunas de las publicaciones existentes que tratan el mismo tema que se trabaja en este proyecto (sección 2.5). Nótese que algunas secciones de este capítulo utilizan como base el trabajo realizado en [2].

2.1. Introducción al análisis de vídeos deportivos

Hoy en día los vídeos relacionados con el deporte han adquirido una mayor importancia, ya que ha aumentado la demanda de este contenido audiovisual. Debido a ello ha surgido mucho material sobre el que trabajar a petición de los usuarios y consumidores de este tipo de vídeos.

Uno de los deportes que más aumento de acogida ha tenido estos últimos años es el baloncesto, convirtiéndose en el segundo deporte más seguido en muchos países o incluso el primero.

Las líneas de investigación y trabajo de estos vídeos deportivos pueden tener como meta principal tanto sistemas de indexación y recuperación (que se basan en las consultas semánticas de alto nivel) como sistemas de realidad aumentada (para mejorar y completar la información visual que tienen los usuarios al consumir este tipo de vídeos).

El resultado de estas investigaciones puede tener diversas aplicaciones según lo que busque el usuario final (el público en general se interesa por jugadas espectaculares como mates o triples, mientras que los entrenadores se preocupan por el movimiento de los jugadores en las distintas jugadas del partido o su rendimiento). Por este motivo surgen aplicaciones del procesamiento de vídeo tan variadas como las siguientes :

- Resúmenes [3]: versión corta de los vídeos deportivos en la cual aparecen los momentos más relevantes para el consumidor (como mates, tapones o triples) de la versión larga.
- Analizar tácticas y rendimiento [4, 5]: sacar conclusiones con respecto al rendimiento de los jugadores y la efectividad de los movimientos y jugadas elegidas durante los partidos, tanto en ataque como en defensa, por lo que se convierte en una herramienta de gran utilidad para el cuerpo técnico en este caso.
- Reconstrucción 3D [6]: ayuda al consumidor a ver determinados lances del juego con mayor claridad, llegando incluso a poder elegir el punto de vista desde el que se quiera observar la jugada.
- Adaptación a pequeños dispositivos [7]: en la actualidad, un porcentaje importante de las reproducciones de estos vídeos se dan en dispositivos como *smartphones* o *tablets*.
- Ayuda arbitral: asiste y sirve de apoyo para decisiones importantes o difíciles que tenga que tomar el equipo arbitral, llegando a poder sustituirlos en algún momento.

Para trabajar en estos objetivos hay que tener en cuenta que existen dos tipos de vídeos :

- *Broadcast* editado [8, 9, 10, 11]: vídeos editados usados para la difusión de los mismos. Incluyen repeticiones, cambios del punto de vista... siendo el tipo de vídeo más usado.
- Multicámara: sin editar, se dividen en dos tipos dependiendo de la cámara.
 - ◊ Fija [12, 13, 14]: cámara inmóvil de principio a fin del vídeo, por lo que no es necesario ningún operario de cámara.
 - ◊ Móvil [15]: manejadas por operadores que siguen jugadores y varían el zoom según sea necesario. Se utilizan para la emisión *broadcast* previa edición y mezcla de estas imágenes obtenidas.

Con todo ello, se pueden realizar distintos enfoques y métodos de diseño al trabajar en los algoritmos de detección que se quieren desarrollar, como la clasificación que se usa en [16] :

- Características bajo nivel vs. relacionadas con el objeto: Las de bajo nivel se obtienen directamente del vídeo con el que se trabaja, mientras que las relacionadas con el objeto (posición de la pelota, del jugador. . .) se obtienen mediante algoritmos complejos.
- Monocanal vs. Multicanal: La información se puede obtener mediante un canal (vídeo) o mediante varios (en este caso vídeo y audio).
- Genérico vs. Específico: Usualmente se implementan para un determinado juego, pero lo ideal es que sea válido para múltiples juegos (ya tengan características semejantes o no).
- Modelo de características vs. Patrón de contexto: A veces las relaciones entre características y eventos son directas, pero otras veces los algoritmos no pueden modelar las relaciones de forma intuitiva, por lo que se recurre al uso de patrones de contexto para establecer ese enlace mediante la información de contexto.
- Basado en aprendizaje vs. No basado en aprendizaje: En el primer caso, las relaciones entre las características y los eventos se realizan mediante el estudio de estadísticas y optimización de información. Esta es por la que se suele optar si las relaciones son difíciles de establecer intuitivamente.
- Vídeo *broadcast* vs. no *broadcast*: La mayoría de algoritmos se orientan hacia vídeos *broadcast*, ya que es el tipo más usado en difusiones deportivas y estos algoritmos son un complemento a estos vídeos. Complementando a todo lo anterior, los algoritmos deben tener en cuenta también detalles tales como la estructura del juego (por ejemplo las partes de un partido), la calibración de la cámara (para determinar la correlación entre imagen y coordenada real necesaria a la hora de reconstrucciones 3D o de juntar diferentes imágenes tomadas al mismo tiempo de posiciones distintas) o la detección y seguimiento de las distintas estructuras involucradas en el juego (jugadores, pelota. . .).

2.2. Técnicas de análisis de vídeos deportivos

El estudio del sistema de detección y seguimiento de los jugadores en vídeos deportivos para su posterior representación en un campo de juego se puede dividir en distintas etapas como las se representan en el esquema de la figura 2.1.

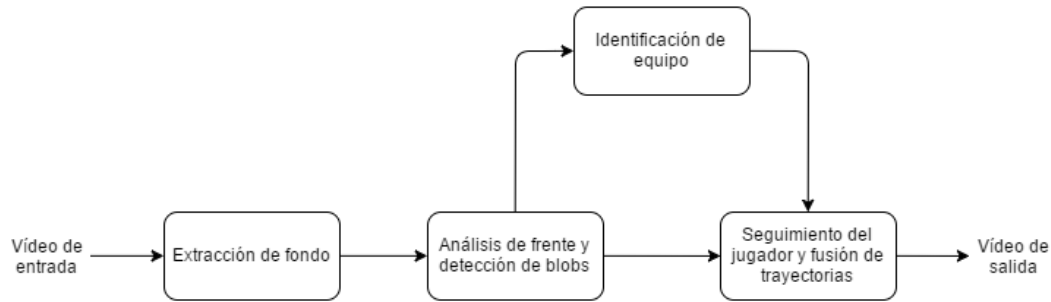


Figura 2.1: Diagrama de bloques del sistema global

A continuación se procede a describir cada uno de estos módulos.

2.2.1. Extracción de fondo

La extracción de fondo es la primera tarea del procesamiento de vídeo que consiste en eliminar la imagen “*estática*” y diferenciar aquellos objetos que serán de interés para el análisis. Esto se puede realizar a partir de una imagen referencia en la que esté el terreno de juego vacío, o a partir de un fragmento del vídeo a analizar. Su complejidad y capacidad de éxito depende en muchos casos de elementos externos tales como la iluminación, presencia de sombras y reflejos (muy notorias en los vídeos de baloncesto), cambios del fondo de la imagen. . .

En [17], la información de la pista se obtiene a través de una máscara binaria creada mediante el estudio del color dominante basado en el modelo *Gaussian mixture model* (GMM), que analiza el color mayoritario obteniendo los histogramas periódicamente a lo largo de una gran cantidad de *frames*. Una vez obtenida esta máscara previa con la digitalización de la imagen, se rellenan los huecos generados por los jugadores que se encuentran en la pista.

La probabilidad de que sea un píxel perteneciente al fondo en [18] se define como la diferencia entre el *foreground* segmentado y una máscara artificial creada a partir de las detecciones obtenidas en las cámaras a trabajar.

También se puede aprovechar este módulo para obtener información relativa al juego, como en el caso del [19]. La segmentación del fondo se obtiene a partir del color, comparando histogramas de *frames* consecutivos en el vídeo a trabajar, y de predicción de movimiento (con lo que puedes saber qué tipo de jugada se está produciendo dependiendo de la velocidad que tengan los vectores del fondo de la pista).

Otro método es el usado en [20], en el cual se obtiene la diferencia entre una media de varios *frames* anteriores que se van almacenando con el *frame* consecutivo a analizar que sigue a este grupo. Esta diferencia se umbraliza, consiguiendo identificar los píxeles

que pertenecen al fondo de la imagen y los que se encuentran en movimiento.

En [21], la segmentación de la pista se realiza mediante un detector *Máximo A Posteriori* (MAP) que se establece con el estudio de los primeros 10 *frames*, obteniendo de éstos la media de los valores RGB que tienen los píxeles del campo. Después este detector umbraliza y binariza la imagen obtenida originando una máscara ruidosa que con una serie de operaciones morfológicas consigue rellenar los huecos que se habían creado por la presencia de los jugadores en la pista, consiguiendo así una máscara binaria clara que elimina todos los elementos externos a la pista. Esto puede generar el problema de que jugadores que se encuentren cercanos al borde desaparezcan de la imagen a analizar, por lo que se aplica una dilatación a la máscara que se había obtenido para ampliarla y tener un margen de seguridad extendiendo el campo a analizar.

2.2.2. Análisis de frente y detección de *blobs*

Una vez se diferencia el fondo del resto de la imagen, se procede a extraer los *blobs* que se obtienen a partir de las componentes conexas de frente. Estos *blobs* en principio se corresponderán con los jugadores que se quieren identificar, por lo que uno de los objetivos principales es que la determinación del número de *blobs* sea consecuente con el juego a analizar. El ajuste de este *blob* con el contorno del jugador debe ser lo mayor posible para así poder diferenciar jugadores que estén muy próximos o incluso se encuentren superpuestos.

Algunas investigaciones usan las máscaras que se crean para la extracción de fondo, como es el caso de [17], generando una máscara de jugadores. Partiendo de la imagen binarizada obtenida por el estudio del color dominante usando el modelo GMM (mencionado anteriormente) se conectan las regiones de píxeles que no tienen el color dominante mediante diversas operaciones morfológicas que eliminan también objetos pequeños y ruido de la máscara.

Otras recurren a métodos basados en la reconstrucción 3D que se origina de juntar las imágenes obtenidas desde distintas cámaras colocadas en diversas posiciones. En [18] se estudian los elementos verticales que se han originado de esta reconstrucción, eliminando aquellos *blobs* que tienen un tamaño menor que el de un ser humano medio. Con el resto de posibles *blobs* se analiza el espacio que ocupa y se calcula el número de personas que se encuentran en ese mismo grupo sabiendo lo que ocupa un ser humano de media.

En [21], una vez se multiplica la imagen por la máscara obtenida para la eliminación de fondo, se realiza un estudio YCbCr mediante el cual se binariza la imagen con los colores previamente establecidos del equipo a estudiar y se aplica una operación

morfológica de cierre. Así, se amplía el tamaño de las detecciones con un rectángulo de 25x20 y se quedan con las 10 zonas conectadas más grandes (ya que no siempre se detecta únicamente a los jugadores).

En [22] se recurre al modelo de identificación *Deformable Part Model* (DPM) que, a grandes rasgos, localiza al jugador por completo y después separa las distintas partes (en este caso 6 partes y 3 ratios de aspecto) de éste para ajustar lo máximo posible, solucionando problemas tales como la oclusión entre jugadores.

2.2.3. Identificación de equipo

Esta información se obtiene a partir del *blob* detectado anteriormente, generalmente analizando la región o regiones que engloba la detección realizada. Según la información del color del uniforme se podrá diferenciar de qué equipo se trata, si es árbitro o ninguno de los dos casos. Además, en baloncesto, la diferencia entre colores de equipos es muy clara puesto que en la mayoría de las ligas se obliga a que un equipo vista de oscuro y otro de color claro, por lo que los histogramas van a ser muy distantes entre ellos.

El histograma de color RGB usado en [22] sirve para clasificar los *blobs* en equipo A, equipo B y otros, por lo que lo utiliza para eliminar falsos positivos también. Se entrena un clasificador de regresión logística en el que la imagen se mapea en zonas representadas por el histograma RGB obtenido.

La separación de equipos en [23] se obtiene a partir de diferenciar el tronco del resto del cuerpo (hecho posible ya que se identifican numerosos puntos de interés del cuerpo del jugador). Una vez se determina el área a estudio del jugador, se convierte al espacio de color HSV, discriminando los píxeles que tengan tonalidad semejante a la piel por el valor de H. Teniendo el dato que se ha mencionado anteriormente de que los equipos visten uno de equipación clara y otro oscura, la diferenciación de los equipos se obtiene por la componente de saturación S (alto nivel de saturación indica color oscuro).

En [17] se hace uso de un algoritmo basado en *CamShift* mediante el cual de generan ventanas de tamaño variable sobre la máscara para localizar a los jugadores y ajustarla lo máximo posible. Dentro de esta ventana se estudia la distribución de color y se obtiene la información necesaria para poder clasificar al *blob* identificado dentro de esa ventana.

2.2.4. Seguimiento del jugador y fusión de trayectorias

En este módulo se tiene que determinar la posición del jugador a analizar a lo largo de varios *frames* consecutivos, fusionando estas coordenadas para poder determinar el movimiento del jugador. Una vez se obtiene esta fusión, se dibuja la trayectoria que ha seguido el jugador bajo estudio bien para representarla en una interfaz o para obtener más resultados a partir de ella. A continuación se enumeran distintos modos o algoritmos que se han usado en el estado del arte.

Mediante el uso del algoritmo de tracking basado en *CamShift*, [17] genera la distribución de probabilidad de color del recuadro que engloba al jugador a estudiar, clasificándolo en un equipo u otro. Una vez obtenida, se localiza el máximo de esta distribución y se marca con un punto y el gradiente que marca la dirección predecible que va a tener el jugador en los *frames* sucesivos.

En el nuevo *frame* se busca la región delimitada de un jugador que contenga el punto de máximo de la distribución de color marcado anteriormente, y si se da este hecho, se añade el nuevo centro del recuadro del *CamShift* que engloba al jugador a la trayectoria que sigue éste.

Para concluir con el *tracker*, se establece un umbral de los valores de probabilidad dentro de la ventana de seguimiento mediante el cual, si el valor es menor, se da por finalizado el *tracker* puesto que significará que el jugador ha abandonado la imagen.

En [22] el *tracking* se establece a partir de las detecciones obtenidas y las predicciones que se generan. Se asocia el *match* del *frame* a estudiar con una trayectoria ya existente e iniciada mediante un sistema de dos estados en el cual se determina si pertenece o no a esta trayectoria por la distancia euclídea que separa el centro de la detección y el centro de la predicción realizada.

La predicción de detección se basa en que la posición del jugador con respecto al tiempo en un periodo corto va a seguir una trayectoria aproximadamente lineal, con lo que el punto futuro va a poder ser predicho sabiendo la información de un número determinado de *frames* y la trayectoria que se ha seguido en éstos por el jugador a estudio. También se hace uso de un filtro de *Kalman* para poder rellenar posibles huecos que deje la no detección del jugador en futuros *frames*, por lo que se soluciona así posibles problemas de oclusión y pérdidas de la trayectoria.

La creación de nuevas trayectorias se da cuando el punto de la localización del jugador no se corresponde con ningún *tracker* existente y la finalización de una trayectoria se origina al abandonar el jugador los límites de la imagen a estudio o si pasa un tiempo umbral establecido en el que no se ha vuelto a detectar ningún *blob* en la trayectoria predicha.

Otros usan información para casos determinados, como es el caso de [20], en el que

se trabaja en el seguimiento de la pelota. Se representa en una gráfica las coordenadas x e y por separado de todos los candidatos existentes a balón con respecto al número de *frame* en el que se han detectado.

Una vez se tiene esta información, se trabaja sobre ella calculando la distancia que existe entre estos puntos y los que se predicen de forma matemática. Al ser el caso concreto de una pelota de baloncesto, se puede anticipar que la trayectoria de la pelota va a ser lineal en el eje x y en el eje y va a seguir una curva cuadrática debido a la gravedad. Sabiendo esto, se puede dibujar la trayectoria predicha sobre los puntos obtenidos anteriormente, quedándonos con la trayectoria resultante de los puntos que más se aproxime a la que por la teoría debería tener. Esto da lugar a que se pueda seguir a la pelota aunque se haya tapado en algún momento del vídeo por algún jugador o algún otro elemento externo.

2.3. *Datasets*

Escoger correctamente la base de datos con la que vas a trabajar es muy importante para obtener buenos resultados que se ciñan a las especificaciones que tiene el proyecto. Dado que este trabajo parte de uno anterior, la elección se basa en que se tiene que ajustar también a las características de su predecesor (sistema multicámara sincronizado grabado con cámara fija). Debido a esto se trabajará con el dataset API-DIS, que es el que mejor encaja con el proyecto a realizar. A continuación se describen distintos *datasets* de baloncesto, entre ellos el elegido para trabajar con él.

2.3.1. CVLAB

Este laboratorio de la Universidad de Lausanne ofrece en su página web el acceso a secuencias de vídeo que ayudan al desarrollo y testeo de sus programas, con libre descarga. Se trata de 4 vídeos obtenidos por 4 cámaras DV a 25 fps de un entrenamiento de un equipo local.



Figura 2.2: Imágenes obtenidas de <http://cvlab.epfl.ch/data/pom#basketball>

2.3.2. CVBASE 06

Los organizadores de CVBASE 2006 proporcionan un *dataset* disponible en su página web para descargar en el que se encuentran dos vídeos sincronizados tomados desde arriba de un partido de baloncesto. Están disponibles en dos modos de calidad distintos:

- 368x288, 25 fps, DivX 5 @1580 Kbit/s (5 minutos)
- 720x576, 25 fps, DivX 5 @4000 Kbit/s (2 minutos)

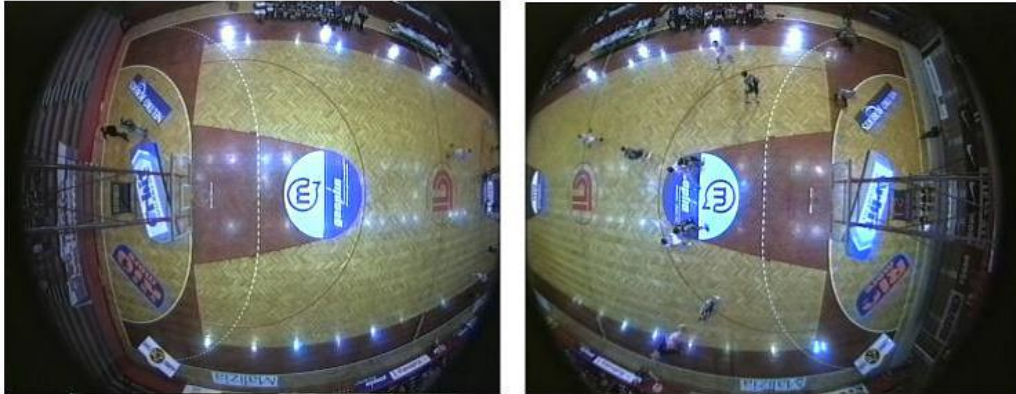


Figura 2.3: Imágenes obtenidas de <http://vision.fe.uni-lj.si/cvbase06/downloads.html>

2.3.3. APIDIS

Este *dataset* fue grabado en Abril de 2008 en Namur (Bélgica), en el transcurso de la final femenina de la liga de baloncesto de Bélgica. Se compone de lo siguiente:

- Siete cámaras de color de 2 *Mpixels* situadas tanto alrededor de la cancha como encima de la misma.
- *Time stamp* en cada *frame*. Todas las cámaras capturan por un único servidor a unos 22 fps.
- Anotaciones manuales de eventos relevantes de todo el partido, tales como tiros, faltas...
- Anotaciones manuales durante un minuto de los objetos identificados durante ese tiempo (posiciones de jugadores, árbitros, balón...).
- Información para calibración, como imágenes de referencia o medidas de la pista de baloncesto.

A continuación se ven las imágenes de ejemplo de los vídeos y un mapa con la situación de las cámaras en la pista.

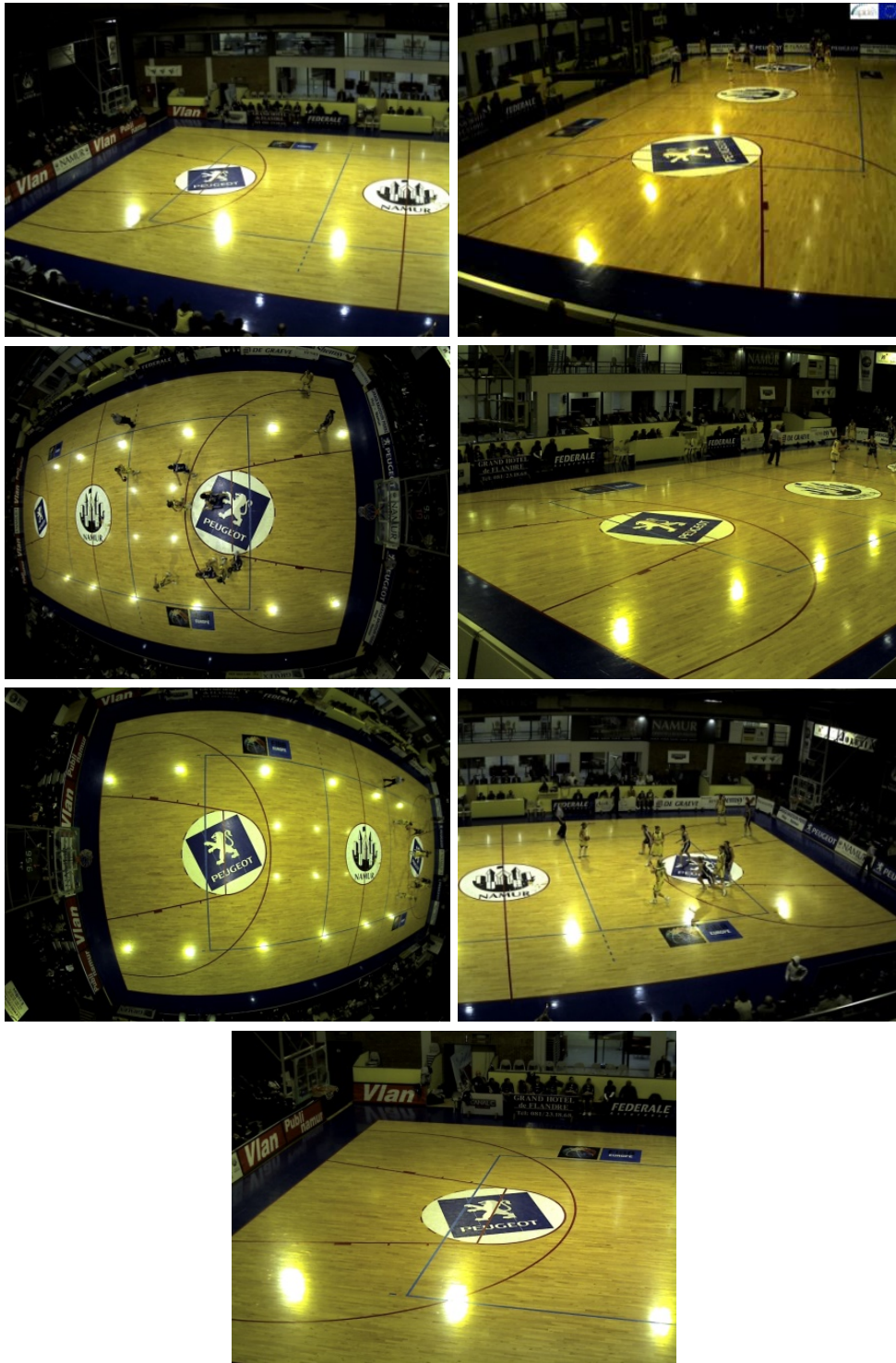


Figura 2.4: Imágenes de las cámaras de APIDIS

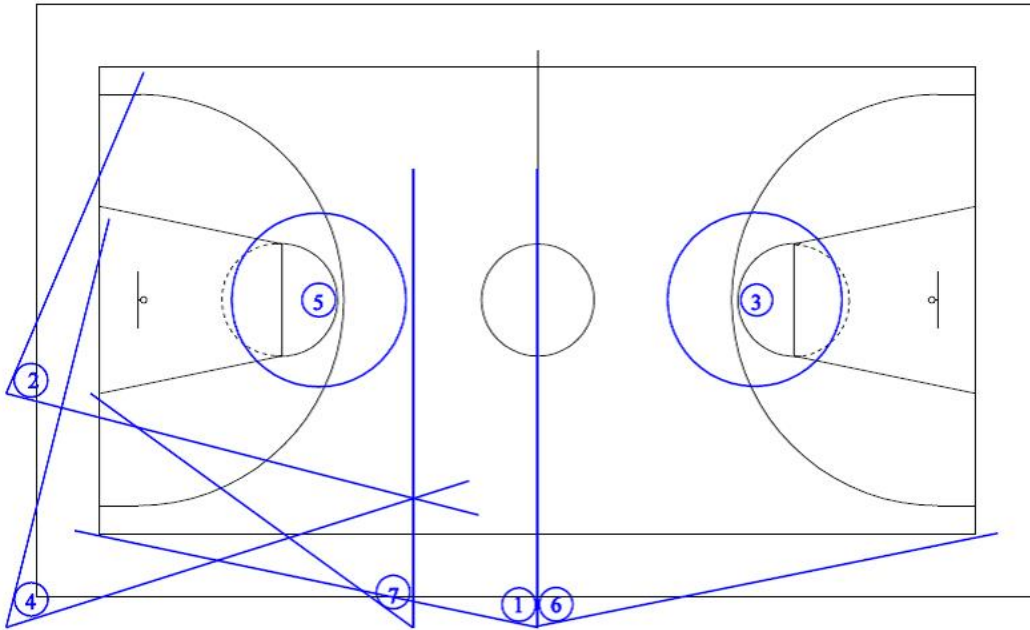


Figura 2.5: Mapa de la situación de las cámaras de APIDIS

La elección de este *dataset* reside en la similitud con el dataset del sistema del que partimos, puesto que se compone también de cámaras laterales que abarcan gran parte del campo o casi su totalidad. Otro punto que favorece su selección es que tiene la posibilidad de realizar evaluaciones en un futuro debido a la vista superior que se tiene de dos cámaras situadas en el techo y en las cuales se puede observar la totalidad de la pista. Aparte, ofrece un minuto de vídeo con los objetos anotados manualmente que facilitan futuras pruebas que se hagan.

2.4. Productos comerciales

En este apartado se describirán algunas de las aplicaciones relacionadas con el estudio y análisis de vídeos de baloncesto (entre otros deportes) que se han comercializado y cuyas especificaciones se encuentran en páginas web donde puedes hacer uso de ellas. Aunque la mayoría sea de pago, algunas ofrecen la posibilidad de probar el producto de manera gratuita.

2.4.1. SHOT TRACKER TEAM

Este programa, creado y desarrollado con la colaboración de *Spalding*, captura las estadísticas de tiro y posición de todos los jugadores en tiempo real. Para ello,

se instalan sensores localizados por toda la pista y, junto con la información que se obtiene de los sensores que se ajustan a las zapatillas de los jugadores, se determina la situación de éstos en el campo.

Los sensores de las zapatillas también sirven para poder diferenciar a los jugadores, obteniendo así los datos personalizados de cada uno de ellos (como porcentajes de acierto, movimientos realizados, lugares de acierto en la cancha...).

Spalding ha colaborado con la creación de un balón que a su vez actúa como sensor, pero manteniendo el aspecto tradicional del balón clásico, mediante el cual se recoge su posición en la pista en todo momento y saber si un tiro ha sido fallido o no.

Esta aplicación es muy novedosa y ya la usan tanto jugadores profesionales, como por ejemplo Klay Thompson, como entrenadores de élite de las ligas estadounidenses.

2.4.2. LONGOMATCH

Creada en 2008 y adquirida en 2014 por FLUENDO para su comercialización, esta herramienta ofrece una gran variedad de funcionalidades, comenzando por la forma en la que puedes importar el archivo a la plataforma.

Se puede, a diferencia del resto, grabar en directo, analizando y obteniendo los resultados durante el transcurso del partido, o también existe la posibilidad de importar hasta 8 vídeos simultáneos ya realizados anteriormente.

En cuanto al análisis, se puede personalizar todo ajustándote a tus necesidades, etiquetando las jugadas o lances del partido para posteriores filtrados e incluso dibujar sobre las imágenes. Los resultados del análisis se pueden exportar a hojas Excel, obtener estadísticas avanzadas y crear *playlist* por tipos de jugada de varios partidos por ejemplo, exportándolo a un archivo de vídeo con el cual se facilita el compartirlo.

2.4.3. VIDSWAP

Se encarga de diseccionar cualquier vídeo deportivo que cargues en unas 6-12 horas. Especificando el deporte a analizar, tienes el vídeo clasificado por timeline con los eventos más destacables del partido, como faltas, tiros, rebotes...

A partir de esto, en la web puedes trabajar sobre estos datos creando vídeos de *highlights*, analizando datos de jugadores específicos o incluso pintar sobre estos mismos vídeos, pudiendo compartirlos con el resto de miembros de la comunidad o mandándolos por *email*.

2.4.4. KROSSOVER

Al igual que VIDSWAP, esta aplicación divide los vídeos que cargues, en este caso por posesiones, y etiqueta todos y cada uno de los eventos importantes del partido.

La plataforma para ver y crear los vídeos es más intuitiva a nivel usuario y ofrece más opciones de filtrado. También se pueden observar diversas estadísticas en forma de números y porcentajes, o incluso en mapas del campo (lugar de tiros errados o acertados).

En cuanto a compartir los vídeos diseccionados, te da la opción de crear una especie de grupo para compartir entre todos los miembros diferentes informaciones y datos.

2.4.5. KINOVEA

Kinovea es un reproductor de vídeo usado por diferentes campos profesionales de diversos deportes. Con él puedes editar vídeos, compararlos, sincronizarlos, ajustar la imagen, realizar *zooms* o incluso dibujar trayectorias.

Su principal diferencia con el resto es que es completamente gratis y de código abierto.

2.5. Sistemas existentes en publicaciones de investigación

En este apartado se enumera y resume una serie de papers existentes que guardan relación y tratan el mismo tema (o similar) que el que se trabaja en este proyecto:

2.5.1. ROBUST CAMERA CALIBRATION AND PLAYER TRACKING IN BROADCAST VIDEO

En este *paper* [17] se trata de conseguir un algoritmo mediante el cual se obtenga un seguimiento de jugadores de baloncesto en vídeos *broadcast*, por lo que se va a observar las técnicas usadas, aunque en el caso del proyecto en el que se va a trabajar no se trate con *broadcast* vídeos.

La información sobre la pista se obtiene primero creando una máscara binaria mediante detección de color dominante y rellenando los huecos generados por líneas y jugadores. A partir de esto, se obtienen los límites de la pista.

Una vez se tienen los bordes principales, los 4 puntos que limitan la zona de debajo de la canasta se calculan por cruces de las líneas detectadas cerca y paralelamente de las principales. De los puntos de corte, eligen los que se ajustan mejor al modelo de

pista ya establecido, al área que limitan y a la geometría de la pista. En los siguientes *frames* se busca el punto cerca de donde estaba en el anterior *frame*.

Para los jugadores se usa un algoritmo basado en *CamShift*, por el cual se generan ventanas de tamaño variable sobre la máscara que se crea anteriormente, localizando así a los jugadores. En esa ventana se calcula la distribución de color para ir ajustando el centro de la ventana en el centro del jugador y también sirve para diferenciar el equipo al que pertenece. Así, la situación del jugador se da por la equipación, por lo tanto es necesario un ajuste para pasar la localización de este a sus pies.

Los resultados son bastante buenos, con altos porcentajes de acierto, pero genera problemas cuando hay oclusión de jugadores o se juntan dos o más de un mismo equipo, que lo detecta como uno.

2.5.2. AUTOMATIC PLAYER POSITION DETECTION IN BASKETBALL GAMES

Este sistema [23] trata la detección de posición de jugadores de vídeos de baloncesto, aunque estos vídeos sobre los que se trabaja también sean *broadcast*.

Esta detección se basa en el reconocimiento de 14 puntos del cuerpo (tobillos, rodillas, codos, cadera, cabeza...) y en la obtención de los puntos entre medias de los que se han localizado anteriormente (26 en total). A continuación, se convierten estos puntos en áreas con el mismo tamaño y se aplica el algoritmo HOG para que la detección sea robusta a cambios de luz y sombras.

La eliminación de falsos positivos se lleva a cabo después de la detección de bordes, que se realiza con el algoritmo *Canny* para la detección de bordes (obtiene las mayores líneas de píxeles existentes tanto en horizontal como en vertical).

Una vez se tienen los límites de la cancha, toda aquella detección cuyos tobillos estén fuera del área de la pista se desecha. En cuanto a los árbitros, al tener las diferentes partes del cuerpo localizadas, se compara el histograma de color de las piernas con el que ya se sabe que tienen todos los árbitros.

Diferenciar equipos es relativamente sencillo en este algoritmo, puesto que únicamente se realiza el estudio del espacio comprendido por los puntos que limitan la zona de la camiseta (el tronco). Una vez se tiene esto, se eliminan los píxeles de la piel y la pista y se obtiene el histograma de color para saber si es un color claro u oscuro.

Para determinar la posición del jugador se parte de la imagen obtenida por el algoritmo de *Canny* y se determinan las líneas que limitan la pintura. Así, se tienen los puntos de referencia de la cancha y, mediante una matriz de transformación, se puede establecer con cierta exactitud la posición del jugador en la representación de la pista con la que se trabaja. El punto exacto es el punto medio entre los dos puntos

correspondientes a los tobillos.

2.5.3. A TEMPLATE-BASED MULTI-PLAYER ACTION RECOGNITION OF THE BASKETBALL GAME

Se presenta un método en [24] por el cual se pueden reconocer jugadas de un equipo comparándolo con un “*playbook*” previamente diseñado.

Comienza con la división de juego en *large-scale* y *small-scale*. *Large-scale* es la partición de fases del partido, que en este caso es ataque, defensa y tiempo muerto. *Small-scale* se enfoca a los elementos relacionados de cada jugador como la posición de partida, el movimiento por la pista o hacer “*pantallas*” o bloqueos.

La diferenciación de fases del partido se realiza calculando el centro de gravedad del equipo, el punto medio en la pista, junto a su dirección de movimiento y a partir de ahí determinar en qué situación del juego se encuentran. También se tiene en cuenta la variable tiempo y la probabilidad de pasar de una fase a otra.

En cuanto a la *small-scale*, se ha creado un modelo de pista con diferentes zonas muy determinadas que han dado gran resultado en las pruebas. Después se comparan la posición real y los movimientos del jugador con el modelo de jugada a través de distancias euclídeas. Se detectan también bloqueos observando distancias entre jugadores y velocidad de ejecución del movimiento.

Con todo esto se crean “*agendas*” que recogen las acciones de cada jugador y juntando todas ellas se compara con el “*playbook*” y se permite un 10 % de diferencia para considerar que se trata de la misma jugada o, por el contrario, no figura en la base de datos.

2.5.4. JOINT MULTI-PERSON DETECTION AND TRACKING FROM OVERLAPPING CAMERAS

Este artículo [18] consta de una gran enumeración de bibliografía relacionada con el tema de la detección en sistemas multicámara y diferentes algoritmos que se usan para trabajar con ello.

El método que se especifica en este documento se basa en la reconstrucción 3D juntando las imágenes de las distintas cámaras que se tienen en cada base de datos a trabajar. Una vez se realiza esta proyección 3D, se estudian los elementos verticales que han surgido, eliminando aquellos que son más estrechos que un hombre. Los demás elementos tendrán que ser analizados para determinar cuántas personas (u objetos) están presentes en esa “*mancha*”. Para ello, se tiene el espacio que ocupa una persona de media y se calcula el número de personas que hay en el grupo.

Las trayectorias se determinan con grafos entre dos variables (*tracks* y medidas) que, dependiendo de cómo se conecten entre ellas, determinan una trayectoria nueva, que continúa o que ha acabado. A partir de esto, los seguimientos de las trayectorias se realizan con diferentes cálculos de probabilidades y la combinación optimizada de ellas.

A continuación, en el *paper* se presentan diferentes pruebas realizadas en las que se compara el método usado con otros presentes en la bibliografía antes mencionada. Se observa en las tablas que mejora considerablemente lo ya presente en el estado del arte en diferentes entornos, tanto en exteriores como en interiores.

Los resultados obtenidos con la base de datos APIDIS, que es lo que interesa en este proyecto, son aceptables en comparación con los obtenidos con los otros métodos salvo por un detalle, y es que no detecta de forma inmediata nuevas incorporaciones a la imagen, sino que aparece la persona localizada al cabo de un tiempo transcurrido.

2.5.5. ROBUST REAL-TIME TRACKING OF MULTIPLE OBJECTS BY VOLUMETRIC MASS DENSITIES

El sistema presentado en este *paper* [25] se basa en el análisis de sistemas multi-cámara fijas, como es el caso que se trata en este trabajo.

Inspirado por el uso de escenas 3D, se centran simplemente en el concepto de volumen ocupado. Hallando el centro de masas de estos volúmenes detectados mediante la densidad de masa obtenida de la reconstrucción 3D, se determinan las coordenadas de partida de los objetos detectados y se mapea usando Voronoi para reducir en lo posible el problema de posibles oclusiones.

Aparece el concepto de (2+1)D, que es la coordenada en cartesianas con una estimación en el eje z, lo que reduce considerablemente el coste computacional que se tiene en la búsqueda 3D. También se plantea la idea de regiones de entrada (donde se van a iniciar y cancelar trayectorias) en las imágenes para reducir considerablemente la aparición de “fantasmas” en mitad de la imagen.

Los resultados que se obtienen con este método son muy buenos comparados con los ya existentes en el estado del arte como el *tracker K-Shortest Paths* KSP . El problema es que en los estudios con la base de datos de APIDIS se ha observado que hay problemas a la hora de diferenciar jugadores, ya que únicamente se basa en la geometría de los objetos detectados y en baloncesto las formas de los jugadores son muy parecidas, dando lugar a errores, pero los análisis de MOTP y MOPA resultan ser muy destacables.

2.5.6. REAL-TIME POSITION ESTIMATION AND TRACKING OF A BASKETBALL

En este *paper* [20] se propone un algoritmo de detección y seguimiento de un balón, tema inédito hasta ahora en análisis de vídeo de baloncesto y cuyo estado del arte no es muy extenso debido a su complejidad.

Primero, de los *frames* que entran, se detectan los objetos que se mueven. Para ello se sustrae el fondo a través de la diferencia entre *frames* consecutivos (su principal problema, que es la posible movilidad de la cámara, en el caso de este proyecto no existe), se hacen las operaciones de apertura y cierre y después dilatación para acabar este proceso con el algoritmo detector de bordes *Canny edge*.

Una vez tenemos esto, se filtran los posibles candidatos según “*size*”, “*shape*” y “*compactness*” (área del objeto entre área del rectángulo más pequeño que lo abarca).

Con los candidatos que se tienen, se procede al análisis de trayectoria. Se representan los centroides en dos gráficas del eje X y del eje Y, y se unen aquellos que en *frames* contiguos cumplan con el umbral de la función de predicción expuesta en el *paper*. A continuación se analizan los movimientos obtenidos en las gráficas y se hace que concuerde con el movimiento observado en la imagen 2D, teniendo en cuenta que dure más de un número determinado de *frames* y que cumpla el umbral de la ecuación de predicción.

Por último, para evitar posibles oclusiones, se usa la ecuación antes mencionada de predicción y se sigue esa trayectoria si no existe ningún candidato cerca de esa posición, hasta que el número de *frames* en el que no se detecte supere un umbral.

Los resultados son dignos de destacar, obteniendo un 100 % de exactitud en la trayectoria del balón.

2.5.7. BASKETBALL PLAYER TRACKING AND AUTOMATED ANALYSIS

Se explica brevemente varios aspectos y técnicas del reconocimiento de jugadores usadas en el algoritmo estudiado en este *paper* [21].

Comienza por la detección de la pista, que se realiza mediante el color RGB medio medido en los 10 primeros *frames*. A esta máscara binarizada obtenida se le aplican operaciones morfológicas y se la dilata para ampliar los bordes y detectar a los jugadores próximos a estos.

La detección de jugadores se realiza a través de un detector MAP en el que se ha determinado el color de las equipaciones. A continuación, se analizan las posiciones de estos posibles jugadores en *frames* consecutivos y, mediante un umbral, se determina

si las posiciones consecutivas determinan una trayectoria o si ha sido un falso positivo al principio.

Por último, se realiza la representación de la posición y trayectoria mediante una matriz de transformación 3x3 (asumiendo que la posición de la cámara es fija, como en nuestro caso).

Los resultados son buenos en cuanto a porcentajes de detección se refiere, pero el problema de las oclusiones entre dos jugadores está presente y no se aporta posible solución.

Capítulo 3

Sistema base.

En este capítulo se va a describir el sistema implementado en [2], que a su vez es el sistema adaptado del trabajo [1]. Dado que este proyecto parte de ambos PFCs antes mencionados, debe quedar claro el funcionamiento de los sistemas de detección y seguimiento previamente diseñados, con sus ventajas y sus limitaciones.

Por ello se va a explicar a grandes rasgos el sistema en el que se basa este trabajo, ya que las explicaciones exhaustivas se encuentran en ambos documentos, tanto en [2] como en [1].

La aplicación de partida sigue el diagrama descrito por la figura 3.1 en la que podemos observar que se recoge la secuencia (vídeo) de entrada, se produce la segmentación de fondo y, una vez se tienen los *blobs*, se procede al *tracking* de estos *blobs* detectados.

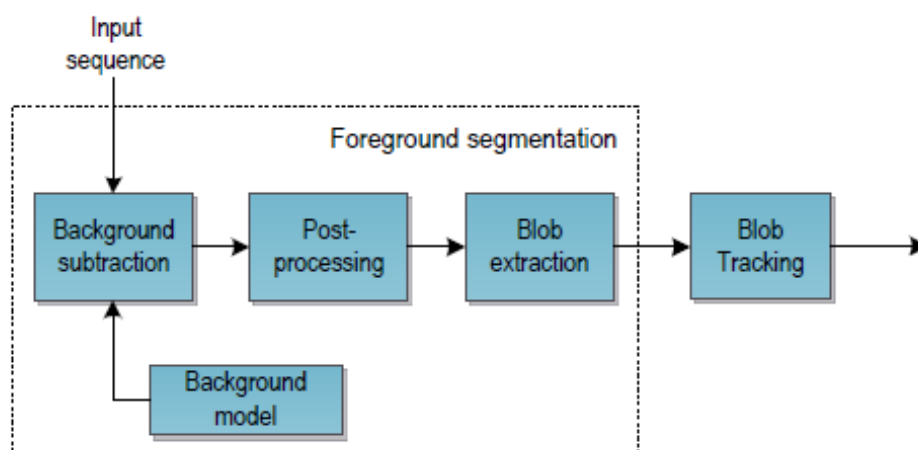


Figura 3.1: Diagrama de bloques del sistema base. Fuente [1]

Este trabajo tiene como punto de partida un sistema de análisis de vídeo diseñado para la detección de objetos robados y abandonados en el *Video Processing and Understanding Lab* (VPU Lab) cuya información y diseño se encuentra en [26] y [27] con más detalle.

A continuación se explicará brevemente los diferentes módulos que se encuentran en el sistema base desde el que se ha partido para el inicio de este proyecto.

3.1. *Background Substraction*

El primer paso según se recoge un *frame* entrante en el sistema es la creación de una máscara binaria en la cual se puedan diferenciar los píxeles pertenecientes al fondo y los que no (frente), bien sean estáticos o móviles.

Para obtener esta máscara, se ha de introducir una imagen obtenida en la cámara en la cual únicamente aparezca el fondo, en este caso que sólo se vea la pista vacía de jugadores. Con este modelo de fondo se obtienen los píxeles que no pertenecen al fondo y se obtiene la máscara binaria con la que se trabajará a posteriori para conseguir diferenciar a los jugadores.

3.2. Post-procesamiento

Una vez se tiene la máscara binaria, se debe trabajar con ella para eliminar elementos que puedan causar falsos positivos, como pueden ser sombras y reflejos, o ruido presente en la máscara.

Para ello se tienen una serie de variables, las cuales se modificarán más adelante (en el siguiente capítulo 4.2), con las que se puede controlar y eliminar ciertos elementos que enturbien la máscara. Estas variables son:

- **VARNOISE** → Establece el valor umbral para determinar la sensibilidad del sistema con respecto al ruido. Un valor alto se corresponde con una menor sensibilidad y viceversa, un valor pequeño indica una mayor sensibilidad a la variación de ruido.
- **ALPHA** → Determina la velocidad a la cual un objeto de la imagen se absorbe y pasa a considerarse parte del fondo.
- **WINDOW_Q** → Tamaño de la ventana que se establece para la diferencia en la escala de grises entre el *frame* entrante y los píxeles del fondo.
- **TH_MOTION** → Parámetro usado para el cálculo de la función *Gamma* con el que se contrarrestan posibles movimientos de la cámara, que en nuestro caso van

a ser o inexistentes o muy leves ya que son cámaras fijas dentro de pabellones en los que no van a afectar elementos externos que muevan la cámara como el viento. A continuación, se procede a aplicar operaciones morfológicas a la imagen tales como erosión y dilatación de la máscara de fondo obtenida para poder diferenciar mejor *blobs* que se encuentren demasiado juntos en la imagen y para eliminar también posibles zonas pequeñas ocasionadas por el ruido en la máscara.

3.3. *Blob Extractor*

A partir de esta máscara binaria que se ha obtenido y se ha trabajado con ella para tener los menores errores posibles, se procede a extraer los *blobs* resultantes de la imagen. Estos *blobs* serán los resultantes de obtener de la imagen analizada las componentes conexas de la máscara de frente.

3.4. *Blob Tracking*

En esta etapa se asocia un identificador para cada *blob* extraído en la secuencia de *frames* y se comienza a realizar el *tracking* de los mismos. La figura 3.2 muestra el esquema seguido para obtener la trayectoria de los jugadores.

Se recoge toda la información necesaria para poder seguir a los *blobs* y dibujar las trayectorias. A partir de los vídeos y máscaras recibidas se generan los datos de seguimiento de cada una de las cámaras. Junto con los puntos de referencia que se tienen del archivo de homografías que se determina al principio de la aplicación, se obtiene la vista superior de seguimiento de cada cámara.

Una vez se tienen estos datos se procede a filtrarlos (número mínimo de *frames* para el *blob*, anchura y altura del *blob*, región del campo...) y a fusionarlos a continuación con el resto de las cámaras.

La etapa de fusonado se explica con detalle en [2].

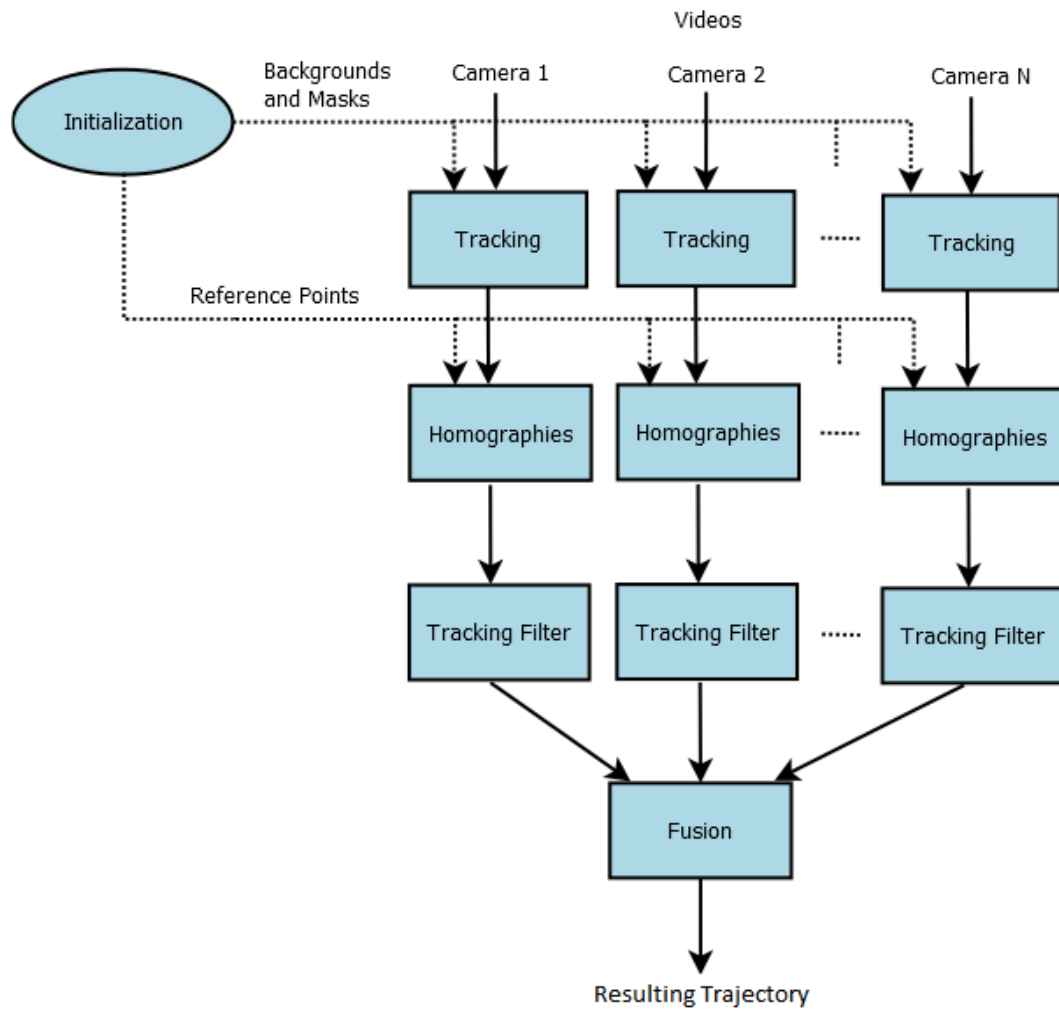


Figura 3.2: Diagrama de bloques de obtención de la trayectoria. Fuente [2]

Capítulo 4

Aplicación de baloncesto

En este capítulo se van a describir los ajustes y mejoras realizadas a la aplicación que se desarrolló en [2] para crear esta primera aproximación de un sistema de detección y seguimiento en vídeos deportivos de baloncesto.

Se parte del sistema definido en el anterior capítulo, y basado en [2] y [1], en los cuales se implementó una aplicación en la que se pueden cargar vídeos de deportes de equipo (fútbol en ese caso) para que, mediante una interfaz gráfica de usuario, se puedan representar los movimientos de los jugadores dibujados por el campo. En ella se pueden corregir trayectorias de una manera más interactiva para que sea fácil a nivel usuario.

A continuación se detallan las adaptaciones que se han tenido que realizar para que este sistema funcione de la forma más correcta posible, dentro de las limitaciones que tiene el sistema del que se parte en el caso de trabajar con vídeos de baloncesto. Estas limitaciones se deben a que las técnicas usadas en el anterior sistema no se adaptan a las características que tienen los vídeos de baloncesto, como por ejemplo la eliminación de sombras y reflejos en la substracción de fondo o que las operaciones de erosión y dilatación sufren ya que en este deporte hay numerosas oclusiones y cruces de trayectorias.

Posteriormente se describirá una interfaz nueva añadida a la aplicación existente en la que se puede crear cualquier archivo de homografía (ver sección 4.3) seleccionando los puntos que el usuario crea que son los óptimos para crear la equivalencia entre la imagen del campo del vídeo y la imagen en la que se representan las trayectorias.

4.1. Adaptación a vídeos de baloncesto

Este proyecto se basa en uno anterior en el que se trabajaba con vídeos de fútbol, por lo que se requiere una adaptación previa de toda la aplicación.

El cambio más significativo es que se pasa de trabajar con 6 cámaras a necesitar únicamente 3 para realizar el seguimiento de actividad. En el caso del fútbol se necesitaba la totalidad de las cámaras para abarcar toda la visión del campo de una forma más fiable, puesto que la pista de juego es mucho mayor que en el deporte que se está tratando en este proyecto.

Se ha decidido que, de las 7 cámaras disponibles en el dataset APIDIS con el que se trabaja, seleccionando 3 se puede abarcar toda la pista. Además esta configuración se elige para que se pueda adaptar el sistema a cualquier cancha de baloncesto. Si se escogiese por ejemplo la vista cenital, el sistema no sería aplicable para la mayoría de las pistas de baloncesto. Dos de los puntos de vista seleccionados enfocan la misma mitad del terreno de juego y la tercera está dirigida hacia la otra mitad restante, como podemos ver en la figura 4.1.



Figura 4.1: De izquierda a derecha, imágenes de las cámaras 1,2 y 6 de APIDIS

Se pasa, por tanto, de un sistema en el que hay que fusionar las cámaras de dos en dos al estar enfrentadas y después asociar estas parejas entre sí, a un sistema en el que las asociaciones de las vistas de las cámaras para la fusión de las trayectorias cambian completamente.

En consecuencia, los cambios a realizar son tanto al nivel de la fusión de las trayectorias como a nivel de interfaz, puesto que está diseñada para el caso de tener 6 cámaras.

4.1.1. Fusión de cámaras

En este prototipo, la fusión a bajo nivel sigue siendo la misma que se define y explica en [2], puesto que esta fusión se produce con las coordenadas reales en metros de la pista. Esto provoca que no afecte el hecho de que las cámaras, como es el caso en el que se está trabajando, no se encuentren enfrentadas.

	1	2
1	5	15
2	20	0
3	0	0
4	21	0
5	4	19
6	0	0

Figura 4.2: Lista de asociación. Fuente [1]

A un nivel más superficial, el tener la mitad de cámaras con las que trabajar sí que afecta. En este caso, la fusión se produce entre la cámara 1 y la cámara 2 que enfocan la misma parte del terreno de juego, y el resultado de esta superposición se junta con lo obtenido únicamente en la cámara 3.

Para la pareja de cámaras se realiza una lista de asociación en la cual aparecen en cada fila qué *blobs* de la otra vista debe ser fusionado con el *blob* que tenga el número de la fila. Por ejemplo, como vemos en la figura 4.2, si en la fila 5 de esta lista aparecen los números 4 y 19, el *blob* 5 obtenido en la primera cámara se debe fusionar con los *blobs* 4 y 19 de la otra cámara a fusionar, o si no hay ningún número como en la fila 3, no existe ninguna fusión a realizar. Las listas de asociación se explican con mayor detalle en [1].

El determinar si un *blob* debe fusionarse con otro o no depende de si un valor, que se ha denominado *score*, es menor que un umbral que se ha impuesto. Este valor numérico se define como la distancia cuadrática media por *frame* que separa los *blobs* a estudio. La fórmula que define este valor es la siguiente:

$$Score(blob_I, blob_J) = \begin{cases} \sum_{f=n}^m \left((x_{f,i} - x_{f,j})^2 + (y_{f,i} - y_{f,j})^2 \right) & , \text{ambos blobs en frame} \\ \infty & , \text{en caso contrario} \end{cases}$$

Una vez se tiene la fusión de las cámaras 1 y 2, al resultado obtenido hay que fusionarle la cámara 3. Dado que en este prototipo los datos que se recogen para combinar son únicamente de una cámara, hay que adaptar el prototipo original a este caso particular. No existe una lista de asociación para la cámara 3, por lo que la fusión total se va a realizar con los datos de la fusión resultante de 1 y 2 y los datos que se han recogido de la cámara 3.

Se repite el módulo de fusión descrito anteriormente y se obtiene la trayectoria general que dibujan los jugadores a lo largo de la pista con los datos recogidos por las tres cámaras que abarcan la totalidad del campo.

4.1.2. Modificación de la interfaz gráfica

Otra adaptación necesaria para este prototipo debido al cambio del número de fuentes desde las que se recoge la información de los jugadores, es la que se ha realizado a la interfaz gráfica que maneja el usuario de la aplicación.



Figura 4.3: Vista de la ventana principal

Como se puede observar en la figura 4.3, aparte del cambio obvio de la imagen del campo, se pasa de 6 a 3 cámaras, lo que provoca un cambio pronunciado a nivel

estético en la ventana principal. La imagen del campo puede ser cualquier imagen que desee el usuario de la aplicación. En los sucesivos ejemplos que se muestran en las figuras, se puede comprobar que no tiene por qué guardar una relación total ya que el campo es diferente al que se observa en el vídeo (la zona comprendida entre el tiro libre y la canasta en un caso es cuadrada y en otro tiene una forma triangular).



Figura 4.4: Vista de la ventana principal en ejecución

Ahora se aprovecha mucho más espacio, por lo que los vídeos se ven con un tamaño

más grande que en este caso se necesita ya que los espacios en la pista de baloncesto son más pequeños que en un campo de fútbol.

En cuanto a la siguiente ventana, la de modificar las trayectorias que se observa en la figura 4.5, los cambios son también funcionales, ya que, como se ha dicho anteriormente, la cámara 3 no se fusiona con ninguna otra, por lo que no va a ser necesario corregir ninguna fusión errónea que se produzca al no haber fusión en sí.

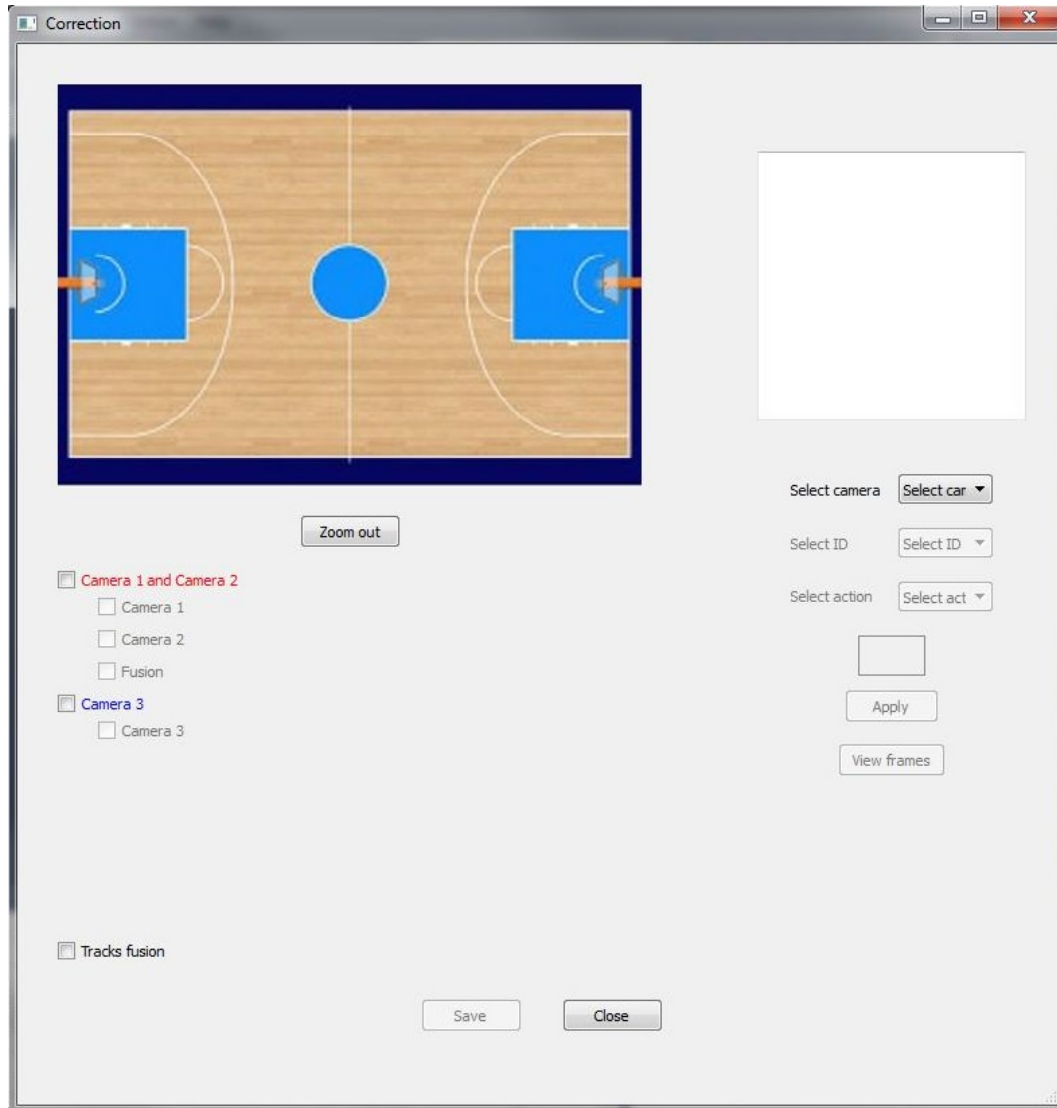


Figura 4.5: Vista de la ventana de corrección de trayectorias

Debido a este hecho, como se ve en la figura 4.6, el panel con el listado de las fusiones de los *blobs* se encuentra vacío, pero se encuentra habilitado por si el usuario quiere hacer uso de la ventana de *frames* que se observa en la figura 4.7 para observar

con más detalle el vídeo y algún *id* de ejemplo.

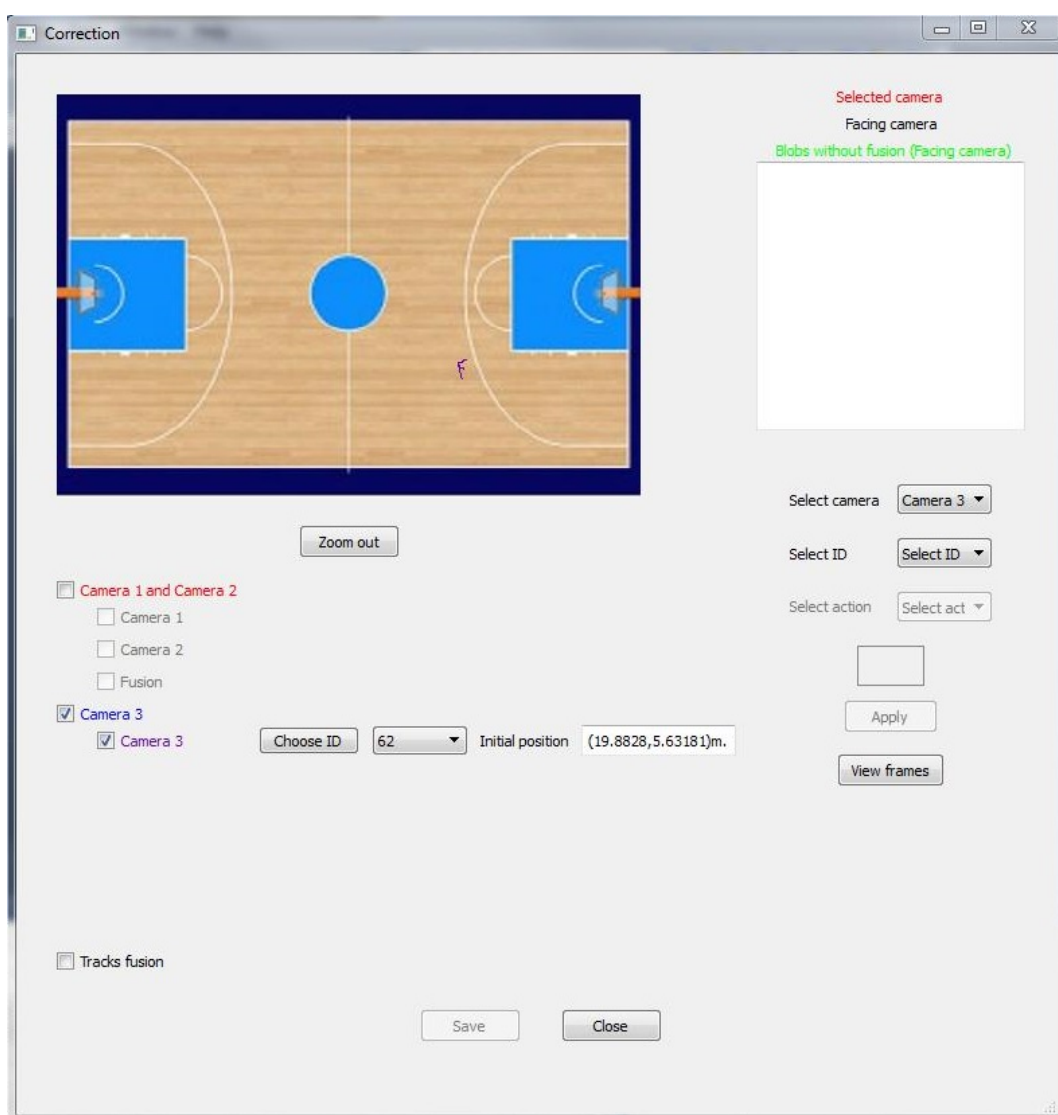


Figura 4.6: Vista de la ventana de corrección en uso

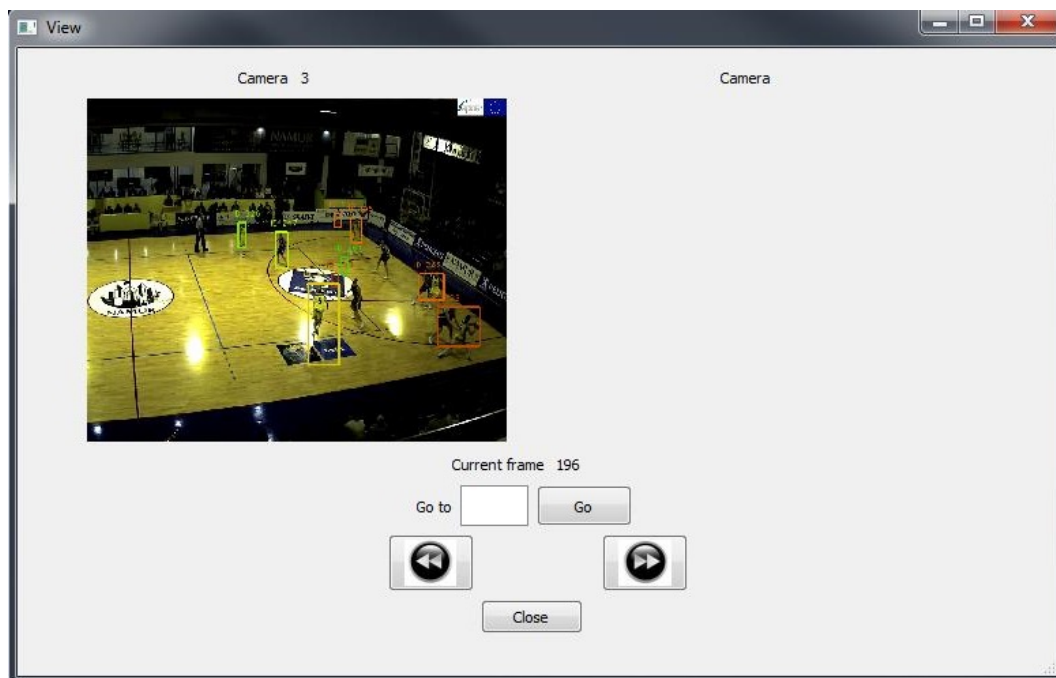


Figura 4.7: Vista de la ventana de *Frames*

4.2. Modificación de parámetros

Dado que las características de los vídeos de baloncesto son diferentes a los de fútbol (cámara más cercana a la pista, jugadores más juntos con mayor número de cruces y oclusiones, iluminación de pabellón *indoor*, suelo de parquet que produce sombras y reflejos...), existe la necesidad de tener que ajustar los parámetros editables que se tienen en las operaciones de extracción de fondo y obtención de *blobs*.

Estos cambios se han decidido a base de pruebas, puesto que crear un barrido de parámetros sería muy costoso a nivel computacional y no es posible al no disponer de anotaciones manuales con las que evaluar los resultados de forma automática. Se han ajustado lo máximo posible para tener los mejores resultados subjetivos que se pueden obtener, con las limitaciones que nos impone el sistema desde el cual partimos, ya que está pensado para otro tipo de situaciones.

A continuación se definen estos parámetros y se muestra cómo afecta cada una de las modificaciones a la máscara con la que se trabaja.

- **VARNOISE** → Establece el valor umbral para determinar la sensibilidad del sistema con respecto al ruido. Un valor alto se corresponde con una menor sensibilidad y viceversa, un valor pequeño indica una mayor sensibilidad a la

variación de ruido. En el caso desde el que se parte, el valor se encuentra establecido en 8, que al observar la máscara representada en la figura se ve que convendría establecer un valor más alto para que el sistema no sea tan sensible al ruido, por lo que se establece un valor de 16.

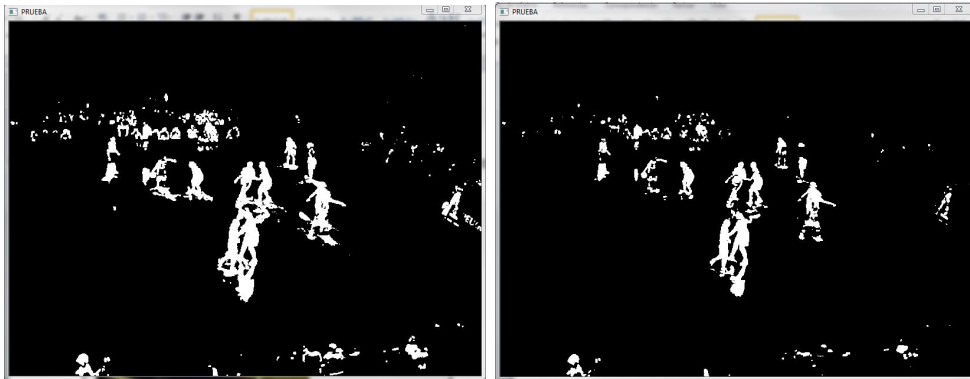


Figura 4.8: Imágenes de la máscara con $\text{VARNNOISE}=8$ (izquierda) y $\text{VARNNOISE}=16$ (derecha)

- **ALPHA** → Determina la velocidad a la cual un objeto de la imagen se absorbe y pasa a considerarse parte del fondo. Se ha observado, a base de pruebas, que los resultados se mejoran considerablemente si aumentamos ligeramente este parámetro, pasando de tener un valor de 0 (no se absorben objetos) a un valor de 0.2. De esta forma, es posible eliminar detecciones indeseadas en los alrededores de la pista.

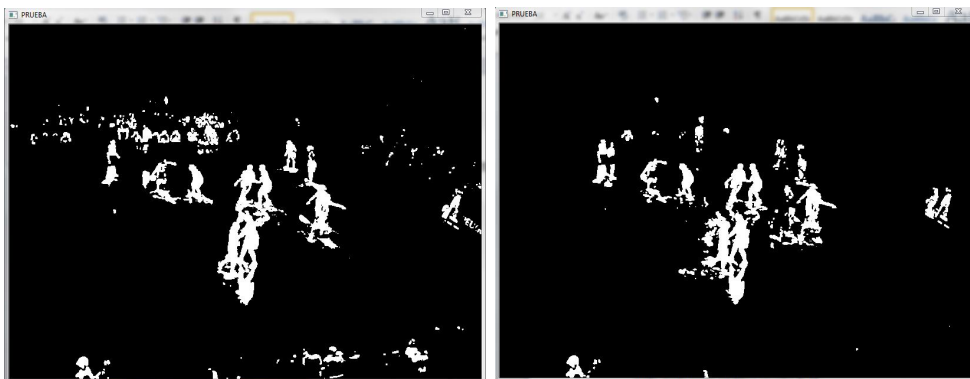


Figura 4.9: Imágenes de la máscara con $\text{ALPHA}=0.0$ (izquierda) y $\text{ALPHA}=0.2$ (derecha)

- **EROSIÓN DE LA MÁSCARA** → Esta operación morfológica hace que los *blobs* que se puedan deducir de la máscara que se ha obtenido, se separen más y de

esta forma poder determinar si hay más de un jugador involucrado en una región conexas que se haya detectado. En el caso que nos incumbe, se necesita aumentar el número de iteraciones que se realizan de la función de erosión (cvErode) que se aplica a la máscara puesto que los jugadores se encuentran muy juntos en este deporte en concreto. Debido a esto, se han añadido dos iteraciones de la función de erosión más para la imagen de la máscara.

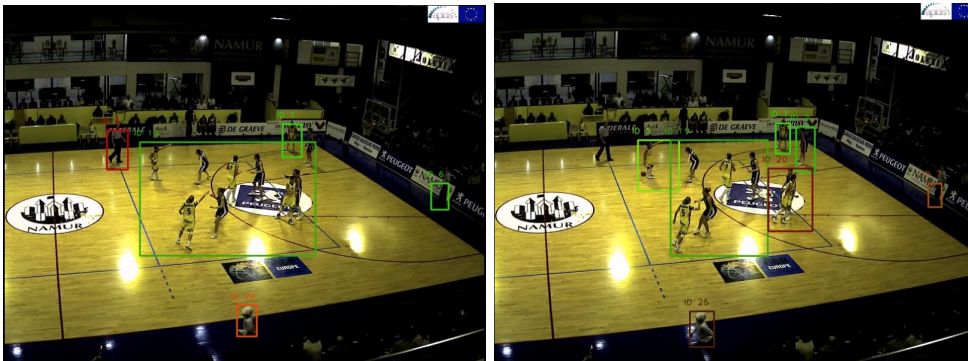


Figura 4.10: *Frame* del vídeo de salida con 2 (izquierda) y 4 (derecha) iteraciones de la función de erosión

- **TAMAÑO DE LA VENTANA** → Es otro parámetro que se ha modificado para poder detectar a los jugadores de la mejor forma posible. Se ha procedido a establecer una ventana menor que enmarque a los jugadores. De esta forma, el recuadro que identifica los *blobs* se ajusta más a la silueta de los jugadores y ayuda también a separar posibles detecciones de varios jugadores cercanos en la imagen. El tamaño establecido antes era de 15x30 y ahora es de 10x20.

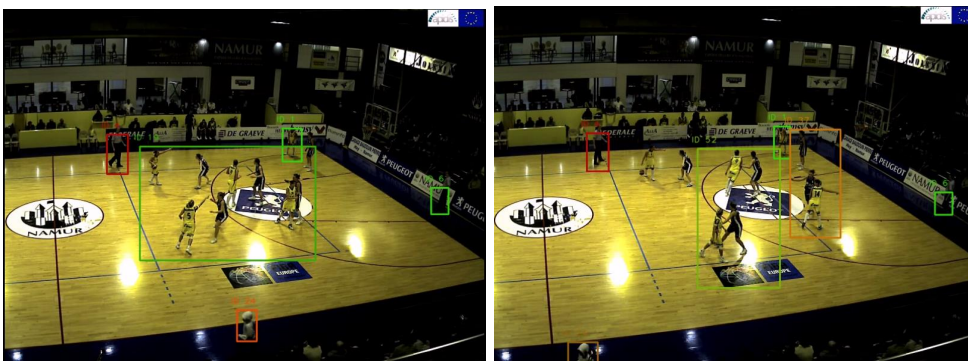


Figura 4.11: *Frame* del vídeo de salida con un tamaño de ventana de 15x30 (izquierda) y de 10x20 (derecha)

4.3. Ventana de creación de la homografía

En este apartado se describe el funcionamiento de la aportación más relevante a la aplicación de la que se parte: la creación de una nueva opción en el menú principal (figura 4.12) en la cual se puede crear un archivo de homografías desde una interfaz sencilla para usar en la ejecución del programa principal.



Figura 4.12: Ventana de la aplicación con la nueva opción de crear la homografía

La homografía es una parte muy importante en todo análisis de actividad y seguimiento de vídeos. Este hecho se demuestra en que todas las referencias que se han explicado en el Estado del Arte [2] dedican un porcentaje relativamente alto a describir las técnicas que han usado para calcular estas homografías.

Su importancia reside en que es la parte del procesado que se encarga de situar los objetos detectados en las coordenadas correctas del lugar en el que se encuentran, aunque la cámara sea lateral y no haya correspondencia.

El cálculo puede ser costoso para el usuario de esta aplicación, por lo que se ha desarrollado esta interfaz en la que su cálculo se obtiene de una forma muy sencilla.

Al seleccionar la opción de crear homografía, se abre una nueva ventana (figura 4.13) en la cual se observan dos partes diferenciadas: parte derecha, que se va a corresponder con la parte real de la homografía (puntos de las coordenadas metros), y parte izquierda, que se corresponde con la parte de las coordenadas de las imágenes (puntos de las coordenadas en píxeles).

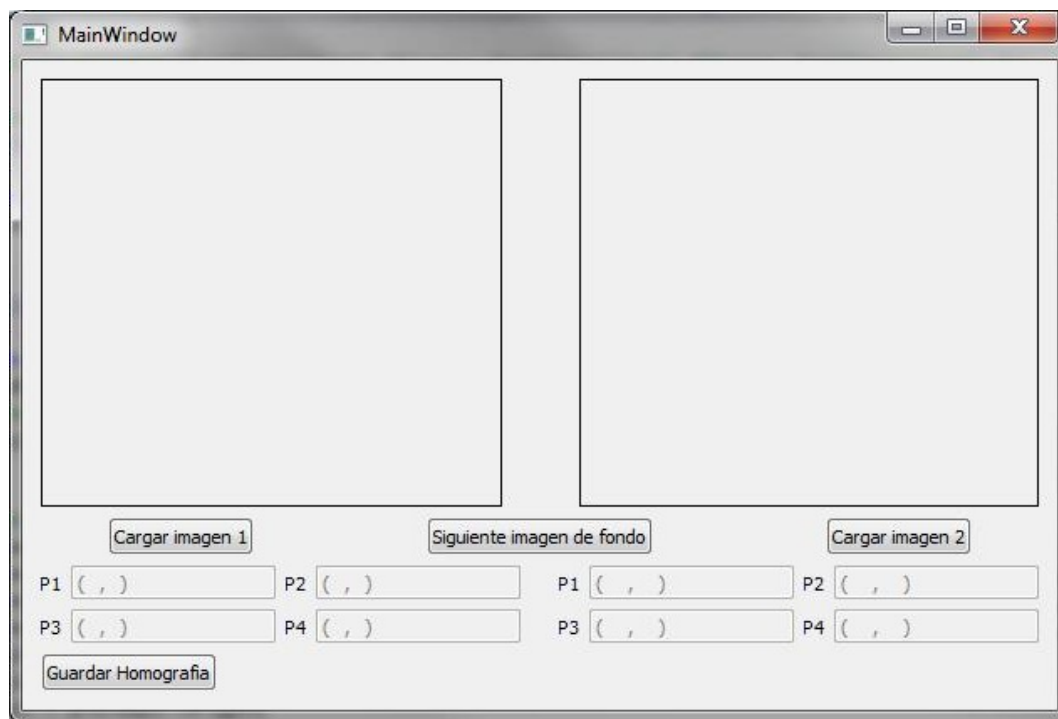


Figura 4.13: Ventana de inicio de la creación de la homografía

Una vez se tiene la ventana a disposición del usuario, se procede a cargar las imágenes que el usuario crea oportunas para crear correctamente la homografía. En el lado derecho se debe poner la imagen del vídeo con el que se va a trabajar en la aplicación y en la derecha una imagen que represente el campo del deporte con el que

se va a trabajar (en este caso una pista de baloncesto).

Las únicas premisas para la elección de la imagen de referencia son que el dibujo del campo sea igual al que se encuentra en la cámara (en la figura 4.14 se observa un cambio en la imagen de la pista usada con respecto a la de la aplicación principal de la figura 4.3, puesto que en la que aparece aquí el dibujo de la zona es triangular, al igual que en la pista del vídeo) y que los márgenes del campo estén muy próximos a los márgenes de la imagen (ya que se toma como punto de origen el extremo superior izquierdo de la imagen cargada). Si no es así, la elección de una imagen que no cumple estas especificaciones puede derivar en fallos de precisión.

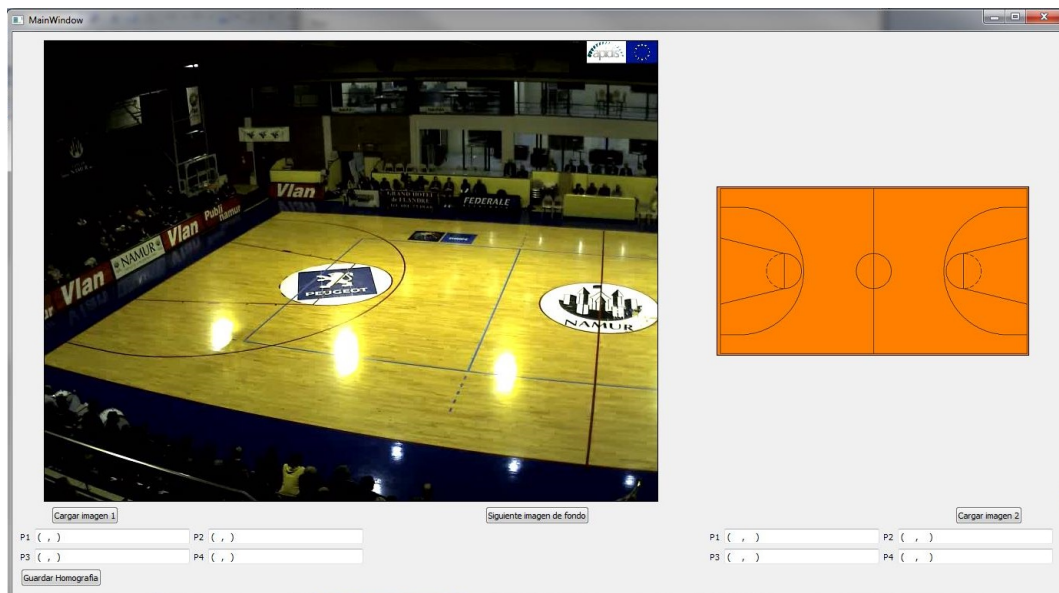


Figura 4.14: Ventana con las imágenes cargadas

Como se muestra en la figura 4.15, la interfaz muestra ambas figuras para que ahora el usuario pueda trabajar sobre ellas. Se debe seleccionar los puntos que se consideren referencia (cruce de líneas limítrofes de la pista, esquinas de fondo, cortes de líneas destacadas del campo...) tanto en una imagen como en la otra, y se queda marcado con un cuadrado rojo el lugar donde se ha pinchado con el ratón. El orden de los puntos es importante, puesto que el primer punto de la figura de la izquierda se corresponderá con el primer punto de la derecha, y así con el resto de los 4 puntos, pero es indiferente si se seleccionan los 4 a la vez en una imagen y después en la otra o si se hace de forma alternativa.

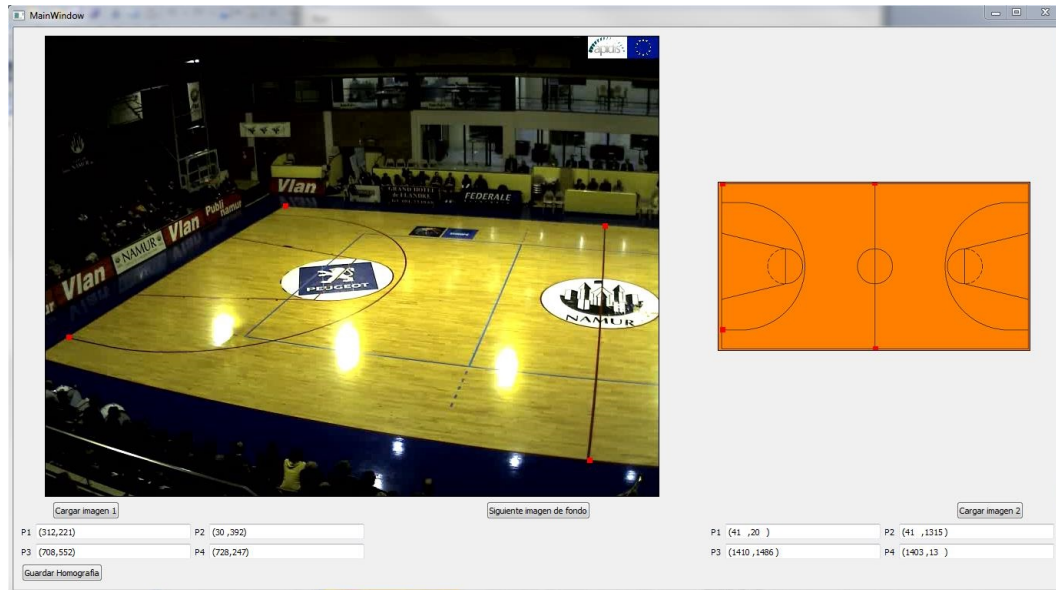


Figura 4.15: Ventana con los puntos elegidos marcados

Las coordenadas de los puntos que se han seleccionado aparecen en el margen inferior de su correspondiente imagen. Estos recuadros son editables (como se demuestra en la figura 4.16), por lo que el usuario podrá hacer ajustes en estas coordenadas que se verán reflejados tanto en el punto de la imagen como en el posterior archivo de homografía. Nótese que las coordenadas de la izquierda se encuentran en píxeles, y las coordenadas de la derecha se encuentran en centímetros, para facilitar los cálculos de la correspondencia entre la imagen de la pista y las coordenadas reales de ésta.

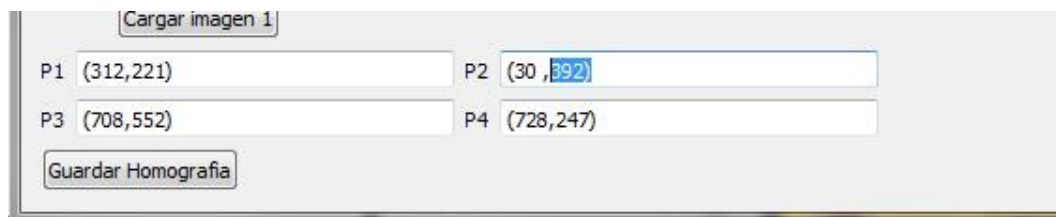


Figura 4.16: Cuadros editables de las coordenadas elegidas por el usuario

Cuando ya se ha terminado con la elección de los puntos, se pasa a seleccionar las correspondencias homográficas de la siguiente cámara. Para ello se debe pinchar en el botón de cargar siguiente imagen que se muestra en la figura 4.17. Una vez se ha cargado ésta, se repite el mismo proceso que se ha explicado anteriormente, pero esta vez con una imagen de otra cámara y la imagen del campo que se tenía anteriormente (figura 4.18).

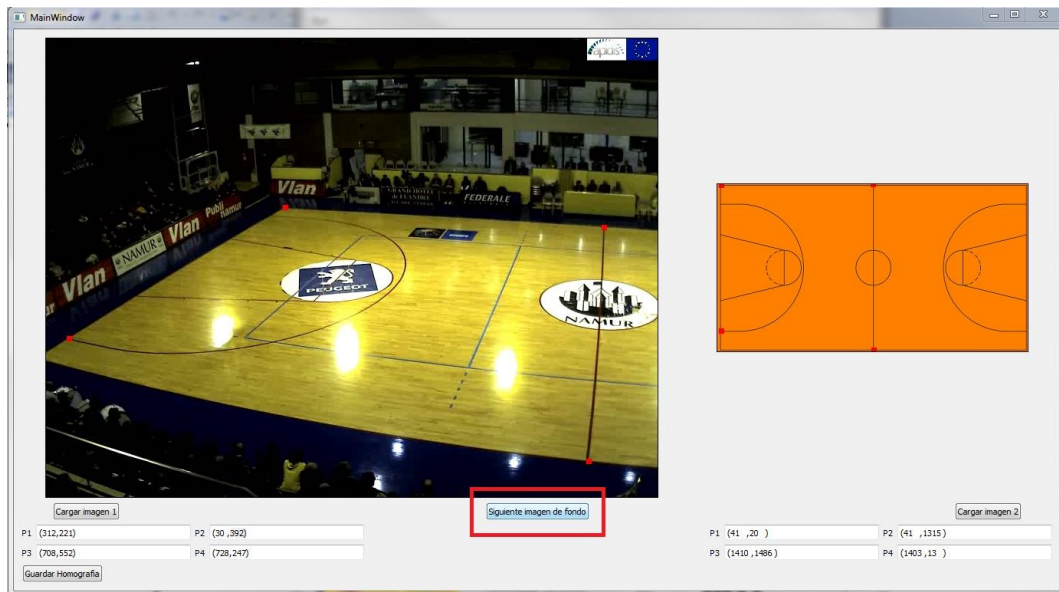


Figura 4.17: Ventana con el botón de siguiente imagen remarcado

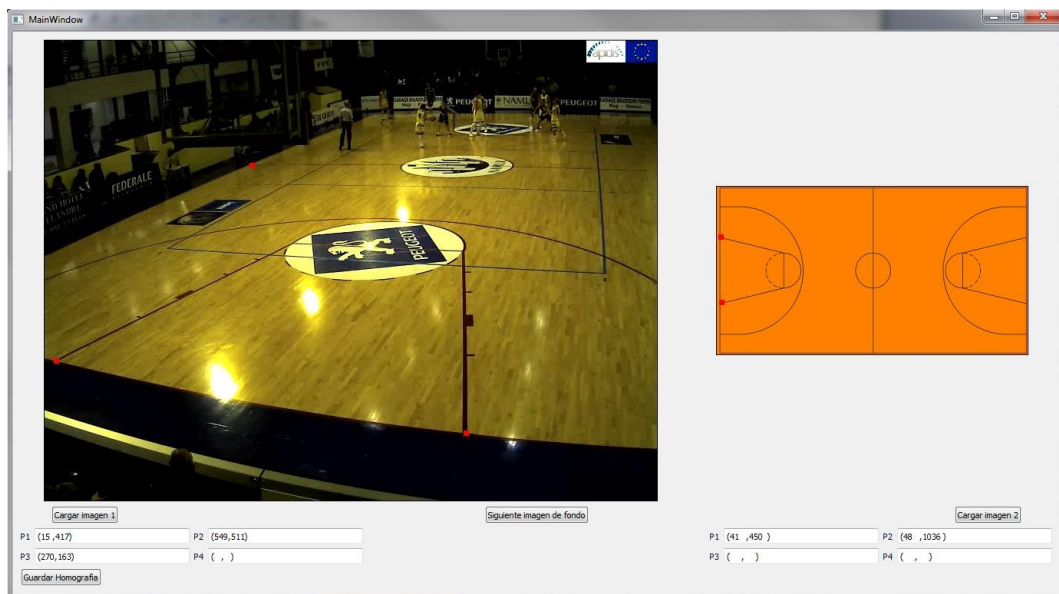


Figura 4.18: Ventana con una nueva imagen de otra cámara

Estos pasos se pueden repetir tantas veces como cámaras haya en el dataset que se va a estudiar después en la aplicación.

Una vez finalizado el proceso, se tiene que generar y guardar el archivo (tipo .txt) que se va a usar en la siguiente parte de la aplicación. Para ello, se pincha en el botón de guardar que se muestra en la figura 4.19 y aparecerá un mensaje de confirmación

del estilo al de la figura 4.20 en el cual se muestra el número de cámaras para el que se ha calculado la homografía. A continuación se abre el explorador de carpetas para que el usuario pueda decidir el lugar y el nombre con el que se va a guardar este archivo de homografía.

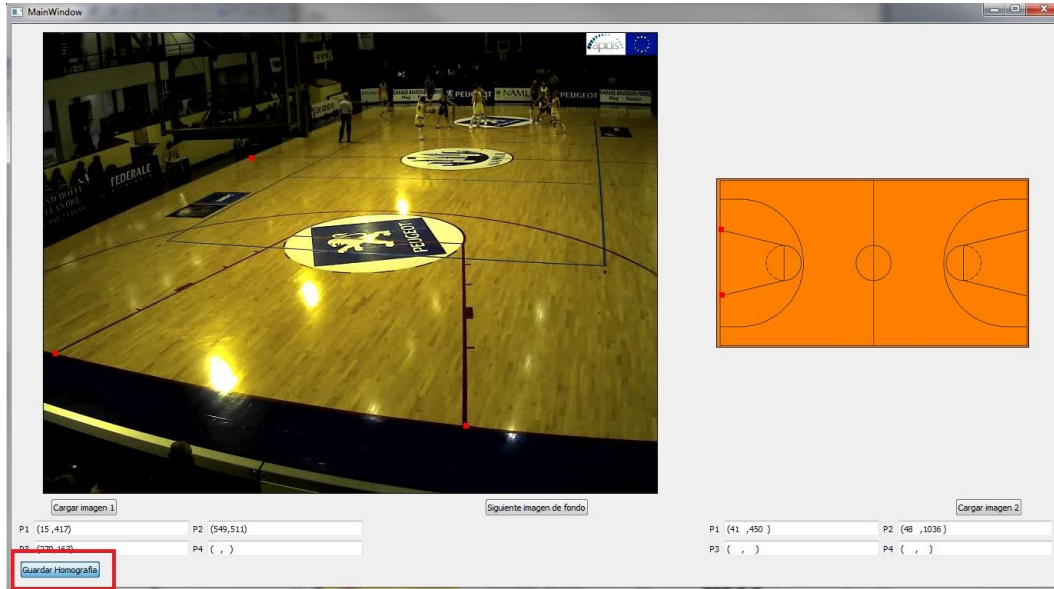


Figura 4.19: Ventana con el botón de guardar la homografía remarcado

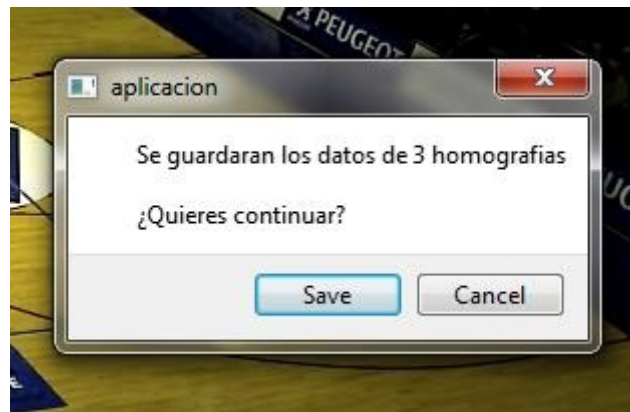


Figura 4.20: Mensaje de confirmación

Finalizado este proceso, se cierra la ventana y se vuelve a la ventana principal de la aplicación desde la cual ya se puede cargar el nuevo archivo de homografías que se ha generado. Este documento será un documento de texto en el que estarán escritos varios bloques (en este caso 3, que es el número de cámaras utilizadas) compuestos por las 4 coordenadas en píxeles en líneas distintas y las 4 coordenadas correspondientes

en metros. Los puntos se encuentran separados por un espacio y los decimales, que se encuentran aproximados a 2 cifras, separados por un punto .

Para que los puntos elegidos sean los adecuados y puedan dar buenos resultados, éstos deben seguir unas pautas determinadas. Estas pautas son las siguientes:

- Separación más grande posible, para poder abarcar toda la imagen y la proporción sea correcta.
- Nunca 3 puntos deben encontrarse alineados, puesto que desaprovechamos uno de ellos (la proporción va a ser lineal, por lo que tener un tercer punto resulta redundante).
- Deben ser unos puntos característicos que se puedan reconocer fácilmente tanto en las imágenes de la cámara como en el campo real. Si se obedecen estas premisas, los resultados de localización de los objetos serán muy próximos a las coordenadas reales que tiene ese objeto.

Capítulo 5

Conclusiones y trabajo futuro.

5.1. Conclusiones.

El objetivo de este proyecto era adaptar el funcionamiento de una aplicación existente de análisis de actividad de vídeos deportivos a vídeos de baloncesto. El sistema vigente trabajaba tanto con deportes individuales como colectivos, pero en el caso de los deportes colectivos estaba enfocado a trabajar con vídeos de fútbol, cuyas características diferían de las características de los deportes de baloncesto, lo que no permitía su uso para este deporte.

Puesto que el baloncesto es otro deporte con mucho seguimiento por parte del público y que el escenario de trabajo es muy diferente al ya usado de fútbol (pista de interior, con diferente iluminación que provoca sombras y reflejos, mayor oclusión de jugadores. . .), surgió este proyecto para realizar un primer prototipo de aplicación de análisis de actividad en vídeos de baloncesto.

Este proyecto ha conseguido cumplir los objetivos inicialmente planteados:

- Tanto el sistema como la interfaz se adaptan al nuevo dataset y al nuevo número de cámaras que se consideran necesarios para obtener toda la información de la pista.
- Se ha ajustado y obtenido los mejores resultados posibles en la detección y seguimiento de los jugadores usando los módulos existentes.

Aparte de esto, se debe continuar trabajando en la mejora de la interfaz gráfica para facilitar el manejo de la aplicación a un usuario sin previos conocimientos.

Una vez especificado el marco en el que se ha trabajado, los pasos que se han seguido para la realización del proyecto son los siguientes:

En primer lugar, se ha realizado la adaptación total del sistema de partida, tanto del código como de las ventanas de la aplicación. Se ha conseguido que el sistema trabaje con un número de cámaras inferior e impar partiendo de que el existente usaba una fusión de cámaras que necesitaba de un número par de vídeos de entrada.

Seguidamente, se han adaptado ciertos parámetros tanto en la extracción de fondo como en la obtención de los *blobs* para poder mejorar los resultados que se tenían con el sistema original. Para una primera aproximación, y dadas las limitaciones presentes en los módulos usados en la aplicación de partida en el nuevo escenario de trabajo, los resultados son aceptables, teniendo un largo margen de mejora.

Por último, se ha añadido una nueva interfaz integrada en la aplicación original en la cual se puede generar el archivo de homografías de forma intuitiva para que un usuario sin conocimientos de procesado de vídeos pueda realizar este archivo de una forma sencilla.

Por tanto, se puede concluir que se han cumplido los objetivos establecidos al principio del proyecto para un primer prototipo de la aplicación enfocada al baloncesto en la cual seguir trabajando en un futuro para continuar la mejora del mismo.

5.2. Trabajo futuro.

Una vez se ha concluido con este proyecto, surgen diversos caminos sobre los que se pueden realizar trabajos futuros. A continuación se enumeran algunas ideas que podrán ser consideradas posibles para trabajar sobre ellas.

Por un lado, a nivel de prototipo las mejoras pueden ir dirigidas a modificar y/o realizar pruebas con diferentes módulos a los usados en este proyecto. Principalmente, el módulo que debe ser el punto de partida es el extractor de fondo, que se tiene que adaptar a las condiciones presentes en los vídeos de baloncesto (eliminación de reflejos y sombras). Otro módulo a investigar sería el extractor de *blobs* para mejorar los resultados que incluyen oclusiones y jugadores muy próximos. También sería interesante integrar un detector de equipos, mediante el análisis de color de los *blobs* de los jugadores, que facilitaría la fusión de trayectorias reduciendo los errores.

Por otro lado, las mejoras también pueden ir enfocadas a la interfaz y la interacción con el usuario. Una de las ideas que surgen puede ser que el usuario pueda elegir el número de *frames* sobre el que trabajar para que las correcciones se realicen cuando el cliente estime oportuno (con mayor o menor frecuencia, a su elección). Otra idea sería realizar una interfaz en la cual poder ejecutar una u otra aplicación a elección del usuario dependiendo del deporte sobre el que quiera realizar el análisis de actividad correspondiente.

Finalmente, otra línea de trabajo futuro para trabajar sobre el sistema desarrollado sería generar anotaciones de los jugadores mediante la vista superior, para poder obtener una evaluación objetiva de los resultados obtenidos por el sistema, de forma que las futuras mejoras se puedan cuantificar usando esta evaluación.

Bibliografía

- [1] R. Martín, “Detección y seguimiento en vídeo deportivos multicámara,” *Proyecto Fin de Carrera, Escuela Politécnica Superior, Universidad Autónoma de Madrid*, Septiembre 2012.
- [2] A. Mora, “Análisis de actividad en vídeos deportivos multicámara,” *Proyecto Fin de Carrera, Escuela Politécnica Superior, Universidad Autónoma de Madrid*, Marzo 2016.
- [3] Y. Rui, A. Gupta, and A. Acero, “Automatically extracting highlights for tv baseball programs,” in *Proceedings of the eighth ACM international conference on Multimedia*, pp. 105–115, ACM, 2000.
- [4] T. Taki, J.-i. Hasegawa, and T. Fukumura, “Development of motion analysis system for quantitative evaluation of teamwork in soccer games,” in *Image Processing, 1996. Proceedings., International Conference on*, vol. 3, pp. 815–818, IEEE, 1996.
- [5] X. Yu, H. W. Leong, J.-H. Lim, Q. Tian, and Z. Jiang, “Team possession analysis for broadcast soccer video based on ball trajectory,” in *Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, vol. 3, pp. 1811–1815, IEEE, 2003.
- [6] N. Inamoto and H. Saito, “Free viewpoint video synthesis and presentation from multiple sporting videos,” *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, vol. 90, no. 2, pp. 40–49, 2007.
- [7] K. Wan, X. Yan, and C. Xu, “Automatic mobile sports highlights,” in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, pp. 4–pp, IEEE, 2005.
- [8] Y. Huang, J. Llach, and S. Bhagavathy, “Players and ball detection in soccer videos based on color segmentation and shape analysis,” in *Multimedia Content Analysis and Mining*, pp. 416–425, Springer, 2007.
- [9] D. Liang, Y. Liu, Q. Huang, and W. Gao, “A scheme for ball detection and tracking in broadcast soccer video,” in *Advances in Multimedia Information Processing-PCM 2005*, pp. 864–875, Springer, 2005.
- [10] X.-F. Tong, H.-Q. Lu, and Q.-S. Liu, “An effective and fast soccer ball detection and tracking method,” in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4, pp. 795–798, IEEE, 2004.

- [11] P. Figueroa, N. Leite, R. M. Barros, I. Cohen, and G. Medioni, "Tracking soccer players using the graph representation," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4, pp. 787–790, IEEE, 2004.
- [12] C. Poppe, S. D. Bruyne, S. Verstockt, and R. de Walle, "Multi-camera analysis of soccer sequences," in *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, pp. 26–31, IEEE, 2010.
- [13] M. Xu, J. Orwell, and G. Jones, "Tracking football players with multiple cameras," in *Image Processing, 2004. ICIP'04. 2004 International Conference on*, vol. 5, pp. 2909–2912, IEEE, 2004.
- [14] Y. Seo, S. Choi, H. Kim, and K.-S. Hong, "Where are the ball and players? soccer game analysis with color-based tracking and image mosaick," in *Image Analysis and Processing*, pp. 196–203, Springer, 1997.
- [15] G. Kayumbi, P. L. Mazzeo, P. Spagnolo, M. Taj, and A. Cavallaro, "Distributed visual sensing for virtual top-view trajectory generation in football videos," in *Proceedings of the 2008 international conference on Content-based image and video retrieval*, pp. 535–542, ACM, 2008.
- [16] X. Yu and D. Farin, "Current and emerging topics in sports video processing.," in *ICME*, pp. 526–529, 2005.
- [17] M.-C. Hu, M.-H. Chang, J.-L. Wu, and L. Chi, "Robust camera calibration and player tracking in broadcast basketball video," *IEEE Transactions on Multimedia*, vol. 13, no. 2, pp. 266–279, 2011.
- [18] M. C. Liem and D. M. Gavrila, "Joint multi-person detection and tracking from overlapping cameras," *Computer Vision and Image Understanding*, vol. 128, pp. 36–50, 2014.
- [19] S. Liu, M. Xu, H. Yi, L.-T. Chia, and D. Rajan, "Multimodal semantic analysis and annotation for basketball video," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, pp. 1–13, 2006.
- [20] B. Chakraborty and S. Meher, "Real-time position estimation and tracking of a basketball," in *Signal Processing, Computing and Control (ISPCC), 2012 IEEE International Conference on*, pp. 1–6, IEEE, 2012.
- [21] S. Parsons and J. Rogers, "Basketball player tracking and automated analysis,"
- [22] W.-L. Lu, J.-A. Ting, J. J. Little, and K. P. Murphy, "Learning to track and identify players from broadcast sports videos," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 7, pp. 1704–1716, 2013.
- [23] Z. Ivankovic, M. Rackovic, and M. Ivkovic, "Automatic player position detection in basketball games," *Multimedia Tools and Applications*, vol. 72, no. 3, pp. 2741–2767, 2014.

- [24] M. Perše, M. Kristan, J. Perš, and S. Kovačič, “A template-based multi-player action recognition of the basketball game,” in *In: Janez Pers, Derek R. Magee (eds.), Proceedings of the ECCV Workshop on Computer Vision Based Analysis in Sport Environments, Graz, Austria*, pp. 71–82, 2006.
- [25] H. Possegger, S. Sternig, T. Mauthner, P. M. Roth, and H. Bischof, “Robust real-time tracking of multiple objects by volumetric mass densities,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2395–2402, 2013.
- [26] L. Campos, “Contribuciones a la detección de objetos robados y abandonados en secuencias de vídeo-seguridad,” *Proyecto Fin de Carrera, Escuela Politécnica Superior, Universidad Autónoma de Madrid*, Julio 2011.
- [27] J. C. SanMiguel and J. M. Martínez, “A semantic-based probabilistic approach for real-time video event recognition,” *Computer Vision and Image Understanding*, vol. 116, no. 9, pp. 937–952, 2012.

Apéndice A

Presupuesto

1. Ejecución Material

- Compra de ordenador personal (Software incluido)2.000€
- Alquiler de impresora láser durante 6 meses260 €
- Material de oficina 150 €
- Total de ejecución material 2.400 €

2. Gastos generales

- 16 % sobre Ejecución Material 352 €

3. Beneficio Industrial

- 6 % sobre Ejecución Material 132 €

4. Honorarios Proyecto

- 1800 horas a 15 € / hora27.000 €

5. Material fungible

- Gastos de impresión.....280 €
- Encuadernación 200 €

6. Subtotal del presupuesto

- Subtotal Presupuesto.....32.774 €

7. I.V.A. aplicable

- 21 % Subtotal Presupuesto6.882,5 €

8. Total presupuesto

- Total Presupuesto39.656,5 €

Madrid, julio 2016

El Ingeniero Jefe de Proyecto

Fdo.: Rubén García García

Ingeniero de Telecomunicación

Apéndice B

Pliego de condiciones

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de *análisis de actividad en vídeos de baloncesto*. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho entorno. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

Condiciones generales

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.
2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.
3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.
4. La obra se realizará bajo la dirección técnica de un Ingeniero de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.
6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.
7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.
8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.
9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.
10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.
11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado

en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partidaalzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.
13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.
14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.
15. La garantía definitiva será del 4 % del presupuesto y la provisional del 2 %.
16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.
17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.
18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.
19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.
20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean

oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.
22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.
23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad "Presupuesto de Ejecución de Contrata" y anteriormente llamado "Presupuesto de Ejecución Material" que hoy designa otro concepto.

Condiciones particulares

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.
2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.

3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.
4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.
5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.
6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.
7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.
8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.
9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.
10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.
11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.
12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.