

Generación de mapas de calor de un partido de básquetbol a partir del procesamiento de video

Jimena Bourlot, Gerónimo Eberle, Eric Priemer, Enzo Ferrante, César Martínez,
Enrique M. Albornoz

Instituto de investigación en señales, sistemas e inteligencia computacional, sinc(i)
UNL-CONICET, Ciudad Universitaria, Ruta Nac. N° 168, km 472.4, (3000) Santa Fe
jimebourlot@gmail.com, geroo_ebeerle_11@hotmail.com, ericpriemer@yahoo.com
{eferrante,cmartinez, emalbornoz}@sinc.unl.edu.ar

Resumen.

El seguimiento de los jugadores en partidos de básquetbol para la obtención de estadísticas de los mismos representa actualmente una nueva fuente de información muy útil a la hora de la preparación de los partidos de un equipo en la temporada. En este trabajo presentamos un método para obtener el mapa de calor de dos equipos mediante una combinación de una red neuronal profunda, a fin de obtener a los jugadores dentro de la cancha, y técnicas de procesamiento digital de imágenes para convertir cada frame del video en un imagen de vista aérea (superior y perpendicular a la cancha), para luego mostrar la posición de los jugadores en un instante de tiempo especificado. Como caso de estudio, aplicamos el método para analizar un partido de los Juegos Asiáticos masculinos 2010 entre India y Afganistán. El sistema describe consistentemente la distribución de los jugadores a lo largo del partido.

Palabras claves: mapas de calor, procesamiento de video, básquetbol, redes neuronales profundas.

1 Introducción

Dentro de los distintos deportes en equipo, una buena manera de interpretar la forma de juego de cada uno es por medio de la distribución de sus jugadores en la cancha, que usualmente se observa por medios de mapa de calor que permiten visualizar las zonas donde hubo más presencias que en otras. A partir de esto, surge la motivación de poder generar mapas de calor de un partido de básquetbol televisado. Actualmente, la forma de realizar este trabajo es con costosos sensores que mediante GPS o acelerómetros miden la trayectoria del jugador en todo el partido. Estos sensores se

colocan en distintas partes del cuerpo para realizar el seguimiento [11], siendo algunos molestos para el atleta en su actividad. Algunos ejemplos son los sensores que se ubican como fajas en ciertas zonas como la rodilla [13], chips que se adhieren al cuerpo del deportista [12], o incluso insertos en protectores bucales [7]. En la Figura 1 se pueden ver imágenes de algunos de estos dispositivos. En este trabajo se pretende realizar una tarea de registro de ubicaciones utilizando únicamente la transmisión en video. Esto podría permitir que cualquier equipo, que no disponga de los recursos económicos para adquirir otra tecnología, pueda hacerlo de manera accesible, y además, es un sistema claramente no invasivo ya que no se debe “vestir” ningún sensor durante la actividad.



Figura 1. Sensores utilizados para realizar el seguimiento de personas en distintas disciplinas.

Habitualmente, las cámaras de televisión no realizan un paneo completo del campo de juego, sino que se capturan distintos sectores a partir de una rotación de las mismas. Es por esto que para realizar un mapa de calor se requiere como primer etapa la registración de los distintos frames del video, y así es posible generar una imagen panorámica que permita visualizar la cancha completa. Luego, se debe transformar el resultado en una vista perpendicular aérea, donde la visualización de los mapas de calor de los jugadores sea más sencilla.

La detección de objetos en imágenes es un campo muy incipiente, sobre todo con la llegada de modelos robustos de redes neuronales profundas [10]. Estos modelos permiten identificar objetos dentro de imágenes lo cual nos resultará útil para encontrar los jugadores y así obtener los datos que buscamos.

A continuación se describen los objetivos planteados en este trabajo, luego se introduce la metodología y los métodos utilizados y finalmente, se presentan los resultados y conclusiones.

2 Objetivos

El objetivo general es desarrollar un sistema que permita la generación de un mapa de calor de la distribución de los jugadores en un partido de básquetbol. En primer lugar se pretende definir la localización y características esperadas para el registro del video. Posteriormente, estudiar e implementar métodos de registración de imágenes para obtener una imagen de cancha entera. Luego, se procederá a estudiar e implementar rutinas para la identificación de jugadores en la cancha. Finalmente, se generarán mapas de calor a partir de las localizaciones de los jugadores en el video analizado.

3 Metodología

Se diseñó el sistema en base a una serie de procesamientos en bloques sucesivos. La Figura 2 muestra un esquema general del sistema.

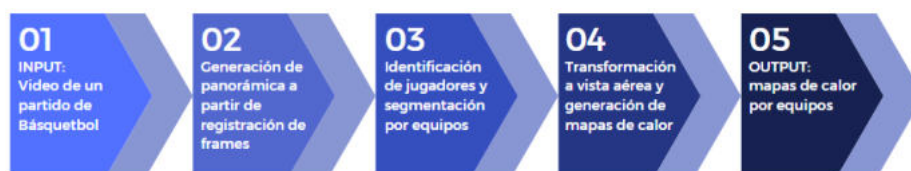


Figura 2. Diagrama de bloques del sistema.

La entrada es un video que tiene una resolución de 600x400 píxeles, tomados a 30 frames por segundo. El primer paso es registrar estos frames para obtener una imagen panorámica de la cancha completa. Al realizar la registración, se obtiene la transformación que se realiza a cada frame del video y esto es útil para trasladar la posición de los jugadores del frame original a la vista panorámica en la cancha a cada instante de tiempo. Para una mejor interpretabilidad, se realiza otra transformación que lleva estos resultados a una vista aérea.

3.1 Generación de Panorámica a partir de la registración de frames

Para generar la vista de la cancha sobre la cual se mostrarán los resultados, en primer lugar se obtiene una vista panorámica a partir del video de entrada [2]. Para esto, se utilizó un algoritmo de registración de imágenes que consta en una instancia de identificación de descriptores y puntos claves mediante el *ORB detector* [5] y luego se realiza una búsqueda de correlación entre estos puntos utilizando el *BF Matcher* [5]. A partir de la obtención de los puntos relevantes entre dos frames sucesivos, es posible obtener la transformación que desplaza el frame de la vista original a la vista panorámica, utilizando un modelo de estimación lineal *RANSAC* [4]. Todos los

métodos utilizados se encuentran implementados en la librería *OpenCV* [8]. El resultado final de la registración de todos los frames del video puede verse en la Figura 3. La información calculada en este punto será relevante posteriormente para hallar la posición de los jugadores en esta vista.

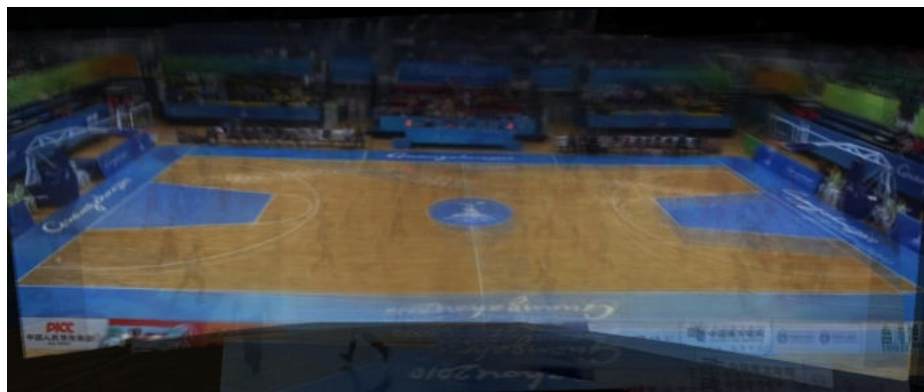


Figura 3. Vista panorámica generada por la registración de los frames del video.

3.2 Identificación de jugadores y segmentación por equipos

En esta etapa se utiliza el algoritmo de detección de objetos *YOLO* [10]. El mismo es un algoritmo basado en redes neuronales profundas entrenada para segmentar una gran variedad de objetos, entre ellos animales, pelotas, personas, entre otros. Esta red es una de las más usadas ya que mediante una sola iteración de la imagen puede realizar una segmentación muy precisa de la misma, lo cual hace que se pueda implementar tanto offline como en tiempo real. Este algoritmo fue entrenado con la librería *Tensorflow* en lenguaje Python [1]. Los parámetros necesarios fueron modificados de forma tal que se obtienen sólo las personas presentes en cada frame. Luego, se segmenta la cancha para separar jugadores del público, obteniendo así los jugadores de cualquiera de los equipos, y/o los árbitros. Finalmente, se realiza un post-procesamiento que permite identificar a qué equipo pertenece cada jugador, y además descartar a los árbitros. Para tal fin, se utilizan segmentaciones utilizando distintos modelos de color [9]. El resultado de la segmentación puede verse en la Figura 4. En la misma se encuentran encuadrados con distintos colores los jugadores identificados de cada uno de los equipos, y en rojo los árbitros.

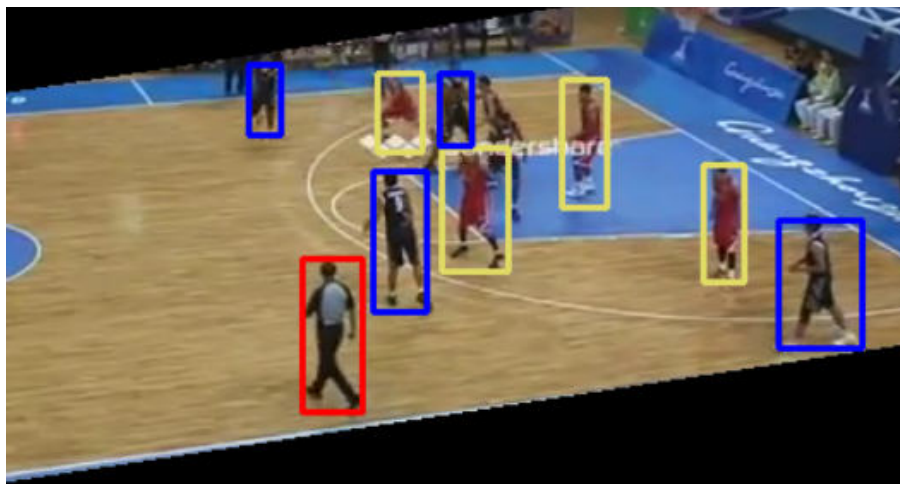


Figura 4. Resultado de la identificación de jugadores (recuadros azules y amarillos) y árbitros (recuadros rojos).

3.3 Transformación a vista aérea y formación de mapas de calor

Se pretende mostrar el mapa de calor en una vista aérea (superior y perpendicular a la cancha), ya que de esta forma se pueden ver los lados fuertes de cada jugador, por dónde realizan las jugadas principales los equipos, etc. Para ello, se realiza una transformación a la vista panorámica de la cancha. Esto se logra ubicando los puntos extremos de la cancha en la vista panorámica, y los puntos de destino que respetan las proporciones estándar de una cancha de básquetbol. Se utiliza para este proceso una matriz de transformación de perspectiva [15]. El resultado es una imagen aérea de la cancha sobre la cual se visualizarán los mapas de calor del partido, tal como puede verse en la Figura 5.



Figura 5. Vista aérea de la cancha.

Dado que la transformación se realiza sobre una imagen panorámica, construida a partir del promediado de un conjunto de reconstrucciones parciales obtenidas por la registración de los distintos frames del video, al realizar la transformación a la vista superior se vuelve notable que los jugadores no fueron eliminados por completo de la cancha. Si bien, se obtiene un resultado que se aproxima a una vista superior de una cancha de Básquetbol, tiene algunos pequeños errores que podrían mejorarse con algunas técnicas de reducción del ruido (jugadores que aún aparecen parcialmente), y mejorando el sistema de detección de los bordes de la cancha. Es por esto que se puede optar por una representación esquemática de la cancha, como la que se observa en la Figura 6, y así se logra una visualización más limpia de la información relevante, es decir, los mapas de calor del partido. La primera opción podría ser de utilidad para realizar un procesado de una transmisión televisiva, mientras la segunda tendría más utilidad para un entrenador que analice el partido.



Figura 6. Representación esquemática de una cancha de básquetbol.

Una vez obtenidas las coordenadas que representan la ubicación de cada uno de los jugadores en la cancha, se procede a realizar las dos transformaciones de forma sucesiva.

La cancha se divide en celdas no solapadas de 50 cm. de lado [3]. Con esta discretización se define una matriz para cada equipo, cuyos valores contendrán la cantidad de jugadores detectados dentro de la celda correspondiente. Posteriormente, con esta información se calculan los mapas de calor de cada uno de los equipos utilizando la librería *seaborn* [14] que es utilizada para la visualización de datos estadísticos. Esta es de código abierto, está basada en *matplotlib* [6] e implementada en *Python*, y cuenta con varios modelos que permiten realizar mapas de calor con distintas variantes.

4 Resultados y Conclusiones

En este trabajo se ha desarrollado un método que permite, a partir de un video de transmisión televisiva, obtener mapas de calor de cada uno de los equipos que juegan un partido de básquetbol. Una salida del sistema puede verse en la Figura 7. En la misma se observa el mapa de calor de uno de los equipos presentes en el video analizado. Las zonas más cálidas (rojas) representan aquellas donde el equipo tuvo más presencia de jugadores a lo largo del partido. Las referencias en los ejes de la izquierda e inferior están indicando la discretización en celdas mencionadas previamente.

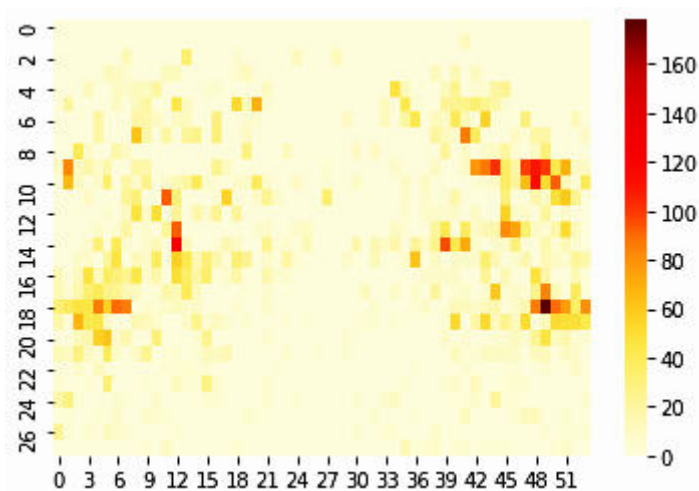


Figura 7. Mapas de calor para uno de los equipos del partido analizado.

El siguiente resultado es la representación del mapa de calor correspondiente a uno de los equipos, solapados con una cancha de básquetbol en vista superior (ver Figura 8). Este primer prototipo del sistema fue diseñado para funcionar con un video capturado con una cámara fija con rotaciones horizontales, como las utilizadas para la televisión en una vista central. La segmentación de equipos es parametrizable, dado que funciona en base a los colores de sus camisetas, y puede ser adaptado a cualquier equipo. De igual manera, la cancha puede ser parametrizada y permite la aplicabilidad en otras canchas.

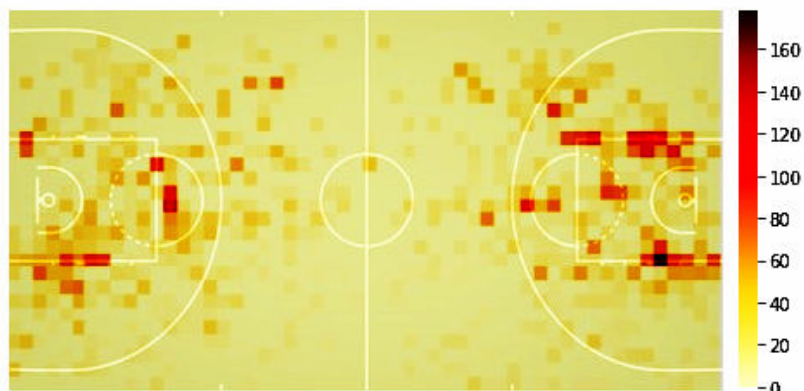


Figura 8. Mapas de calor resultante para uno de los equipos analizados, visto sobre la cancha.

Como trabajo futuro se prevé evaluar este método con videos de mayor resolución, ya que es esperable que mejoren tanto el algoritmo de registración como el algoritmo de detección de personas. Además, se pretenden implementar filtros para mejorar los resultados visuales de la cancha.

Otra de las tareas a implementar es la mejora de la generación de la vista superior de la cancha, realizando para esto la corrección de errores que se discutieron previamente: la detección de los bordes del campo de juego y la reducción del ruido presente debido a la presencia de sombras de los jugadores. A mediano plazo, se propone incorporar algoritmos de tracking que permitan el seguimiento de cada jugador para realizar estadísticas personalizadas.

Agradecimientos

Los autores desean agradecer al instituto sinc(i) UNL-CONICET, a UNL (con CAI+D 50620190100145LI)

Referencias

1. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. In the 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16) (pp. 265-283).
2. Chen, K., & Wang, M. (2014, July). Image stitching algorithm research based on OpenCV. In Proceedings of the 33rd Chinese Control Conference (pp. 7292-7297). IEEE.
3. Cheshire, E., Halasz, C., & Perin, J. K. (2013). Player tracking and analysis of basketball plays. In European Conference of Computer Vision.

4. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381-395.
5. Howse, J., & Minichino, J. (2020). *Learning OpenCV 4 Computer Vision with Python 3: Get to grips with tools, techniques, and algorithms for computer vision and machine learning*. Packt Publishing Ltd.
6. Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in science & engineering*, 9(03), 90-95.
7. Kim, J., Imani, S., de Araujo, W. R., Warchall, J., Valdés-Ramírez, G., Paixão, T. R., ... & Wang, J. (2015). Wearable salivary uric acid mouthguard biosensor with integrated wireless electronics. *Biosensors and Bioelectronics*, 74, 1061-1068.
8. Laganière, R. (2017). *OpenCV 3 Computer Vision Application Programming Cookbook*. Packt Publishing Ltd.
9. Lu, W. L., Ting, J. A., Little, J. J., & Murphy, K. P. (2013). Learning to track and identify players from broadcast sports videos. *IEEE transactions on pattern analysis and machine intelligence*, 35(7), 1704-1716.
10. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
11. Seshadri, D. R., Li, R. T., Voos, J. E., Rowbottom, J. R., Alfes, C. M., Zorman, C. A., & Drummond, C. K. (2019). Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *NPJ digital medicine*, 2(1), 1-16.
12. Sekine, Y. et al. A fluorometric skin-interfaced microfluidic device and smartphone imaging module for in situ quantitative analysis of sweat chemistry. *Lab. Chip* 18, 2178–2186 (2018).
13. Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2019). Estimation of knee joint forces in sport movements using wearable sensors and machine learning. *Sensors*, 19(17), 3690.
14. Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021.
15. Wen, P. C., Cheng, W. C., Wang, Y. S., Chu, H. K., Tang, N. C., & Liao, H. Y. M. (2015). Court reconstruction for camera calibration in broadcast basketball videos. *IEEE transactions on visualization and computer graphics*, 22(5), 1517-1526.