

Práctica 2 – Visión Artificial 2024 – Pablo Páramo Telle

Esquema global del método

El método propuesto para la detección de un campo de fútbol en una imagen se divide en los siguientes pasos:

1. **Segmentación del terreno de juego.** En este paso se identifica la región de la imagen que corresponde al terreno de juego, eliminando el fondo y los objetos que no pertenecen al campo.
2. **Localización de las líneas de siega.** En este paso se identifican las líneas de siega del campo, que indican la perspectiva de la imagen.
3. **Segmentación de los objetos presentes en el terreno de juego.** En este paso se identifican los objetos presentes en el terreno de juego, como jugadores y árbitros.

Soluciones aportadas a cada subproblema

Segmentación del terreno de juego:

1. Se aplicó un filtro de umbral sobre la imagen en HVS para identificar los píxeles que pertenecen al rango de color verde del césped.
2. Se aplicaron operadores morfológicos y un filtro de medianas para eliminar el ruido y suavizar los bordes del campo.
3. Se aplicó un algoritmo de detección de contornos para identificar el contorno del terreno de juego.
4. Se aproximó un polígono al contorno del terreno de juego para dejar este recortado con los bordes rectos.

Para la realización de este apartado, se han realizado varias aproximaciones empleando los operadores morfológicos y los filtros con diferentes combinaciones, obteniendo resultados cada vez más satisfactorios, la combinación empleada en el resultado final ha sido la que ha conseguido resolver el problema en la mayoría de las imágenes de prueba.

Evaluación de los resultados:

Para la evaluación de esta solución, se han recortado manualmente con el software ImageJ 10 imágenes escogidas aleatoriamente de las proporcionadas y posteriormente se ha utilizado la función `evaluar_marcar_campo()` del código, la cual nos devuelve la siguiente matriz de confusión realizada sobre las diferentes imágenes.

Verdaderos Positivos (VP): 2827187	Falsos Negativos (FN): 141082
Falsos Positivos (FP): 43282	Verdaderos Negativos (VN): 6204449

Fracción Falsa 97.05%: Esta métrica nos indica el porcentaje de las regiones que han sido correctamente segmentadas por el algoritmo. En otras palabras, solo un pequeño porcentaje de las regiones (2.95%) han sido incorrectamente segmentadas.

Precisión 97.78%: Esta métrica nos indica el porcentaje de píxeles identificados correctamente como pertenecientes a la región de interés del total de píxeles que el algoritmo clasificó en dicha región. En otras palabras, de todos los píxeles que el algoritmo clasificó como pertenecientes a la región de interés, el 97.78% realmente pertenecen a ella.

Sensibilidad 99.31%: Esta métrica nos indica el porcentaje de píxeles que realmente pertenecen a la región de interés y que fueron correctamente identificados por el algoritmo. Es decir, de todos los píxeles que realmente pertenecen a la región de interés, el algoritmo identificó correctamente el 99.31% de ellos.

Coefficiente de Similitud de Dice 0.9854: Esta métrica nos indica la similitud entre la segmentación producida por el algoritmo y la segmentación verdadera. Un valor más alto indica una mayor similitud, por lo que el valor resultante es realmente bueno.

Para evaluar correctamente hay que tener en cuenta que la precisión del recorte manual de las imágenes puede no coincidir exactamente con el recorte hecho por el algoritmo, evitando esta mínima diferencia que puede influir en los resultados, estos son bastante satisfactorios.

Identificación de problemas y posibles mejoras:

En ciertas imágenes donde hay césped detrás de las porterías, este algoritmo no cumple muy bien su función ya que detecta estos segmentos como parte del terreno de juego, una posible solución sería implementar este problema mediante un detector de bordes, para detectar las líneas del campo pegadas al borde ya detectado y recortar a partir de estas el terreno de juego.

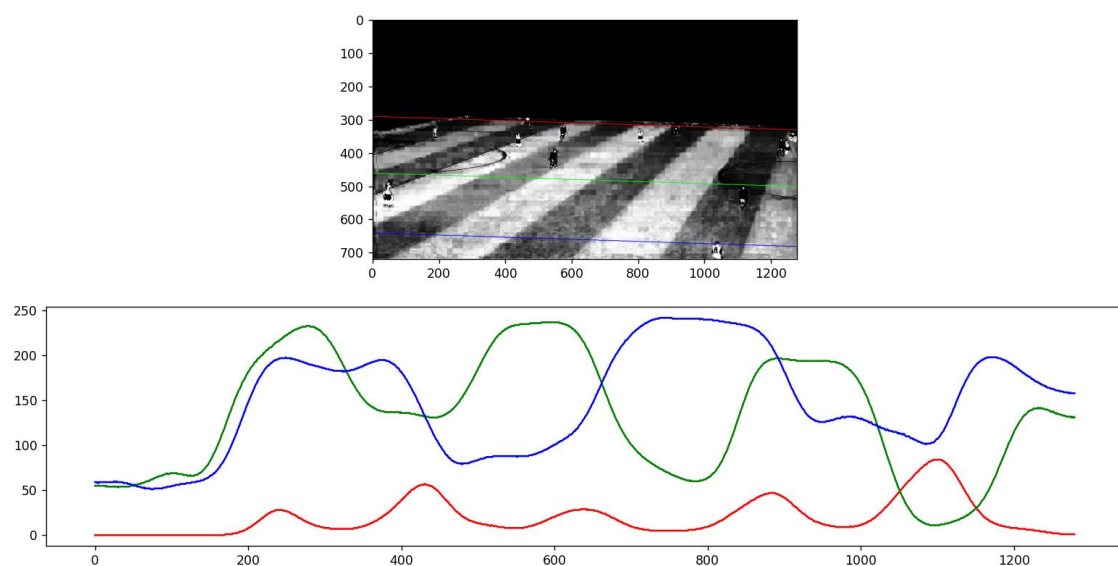
Localización de las líneas de siega:

1. Se aplicó un filtro Gaussiano y otro de medianas sobre la imagen en HSV para eliminar el ruido general del campo y dejar solamente las transformaciones de color de las diferentes segas.
2. Se aplicó el algoritmo de Canny sobre la componente S para detectar los bordes entre las diferentes segas.
3. Se aplicó el algoritmo de la transformada de Hough para detectar las rectas creadas por los bordes entre las segas detectados en el paso anterior.
4. Se calculó la media de la orientación de cada recta, teniendo en cuenta solamente las rectas que superan los 30° por encima o por debajo de la horizontal, de esta manera se evitó una media ensuciada por los valores de las rectas horizontales.
5. Se pintaron las rectas resultantes solamente si cumplían una serie de requisitos, para calcularlos se siguieron una serie de pasos:
 - a. Se estiraron las rectas para conocer sus puntos de corte con los bordes de la imagen.
 - b. Se pintaron solamente las rectas que superaron una diferencia en el eje Y de más de 200px y que la orientación de estas estaba en un rango de 34° por encima o por debajo de la orientación media.
6. Se eliminó la parte del campo recortada en el problema anterior para dejar solamente las rectas marcadas en la zona del terreno de juego

Para la realización de este apartado, se consideró primeramente realizar un algoritmo similar al descrito anteriormente, pero debido a la cantidad de problemas que surgieron a la hora de desarrollarlo, se planteó la idea de desarrollar el detector de las segas de otra manera.

Esta aproximación se basaba en la idea de los detectores de bordes vistos en la asignatura, pero aplicada a este problema y realizada desde cero. Esta implementación, se basó en capturar la recta superior del polígono del primer problema, para así tener la pendiente de la perspectiva del campo. Una vez identificada esta recta en el polígono fijándose en la que tenía

las coordenadas de los vértices de esta en el eje Y más bajas (la situada más arriba en el polígono de recorte), habría que calcular la pendiente de la recta y dividir el espacio que formaba desde el punto de corte con el borde de la imagen en el eje Y más bajo en varios fragmentos. En cada coordenada Y resultante de la división, la idea se basaba en tirar otras rectas con la misma pendiente para así tener rectas paralelas con la perspectiva del campo en diferentes puntos. A partir de aquí, con el componente S de la imagen en HVS y realizada una ecualización del histograma a este para separar más las tonalidades de las segas, la idea era coger el nivel de gris en cada punto x por donde pasara cada recta, así se separarían en diferentes arrays de (1 x ancho_imagen) los niveles de gris de cada recta. Como las gráficas tenían mucho ruido, se decidió aplicarles un filtro gaussiano para reducirlo. Después la idea era seleccionar los cambios en la tonalidad del gris de las diferentes rectas para marcar esas coordenadas y calcular la recta que pasara por esos puntos, para conseguir las diferentes rectas de las segas.



El problema surgió cuando se observó que el primer tramo de tonalidades más claras en cada recta no hacía referencia a la misma sega, sino a diferentes, debido a la perspectiva. En este momento se abandonó esta aproximación y se continuó con la primera idea que había surgido, la cual ha resultado la elegida finalmente para abordar este problema.

Evaluación de los resultados:

Para la evaluación de los resultados de este apartado, se han contado manualmente los Verdaderos Positivos, los Falsos Positivos y los Falsos Negativos ya que no se ha encontrado un método automático para realizarlo.

Verdaderos Positivos (VP): 126	Falsos Negativos (FN): 59
Falsos Positivos (FP): 19	Verdaderos Negativos (VN): N/C

(los resultados del conteo manual pueden variar de la realidad en una pequeña escala)

Sensibilidad 68.11%: Esta métrica nos indica el porcentaje de segas reales que el modelo ha identificado bien.

Precisión: 86.90%: Esta métrica nos indica el porcentaje de las segas que ha identificado el modelo que realmente eran segas reales.

Para evaluar correctamente hay que tener en cuenta que la precisión del conteo manual de las segas puede no coincidir exactamente con la realidad, evitando esta mínima diferencia que puede influir en los resultados, estos no son del todo satisfactorios.

Fijándonos en la sensibilidad, esta debería ser más elevada para obtener unos resultados óptimos ya que el modelo actualmente no pinta un 32.89% de las segas que sí existen realmente.

Identificación de problemas y posibles mejoras:

El principal problema de este modelo está en como se detectan los bordes antes de aplicar la transformada de Hough, estos no son del todo claros y al aplicar la transformada no se consiguen distinguir todas las rectas de las segas del terreno de juego. Para mejorarlo se deberían de poder distinguir mejor los bordes entre las diferentes segas antes de aplicar el algoritmo de Canny para que después fuera mas sencillo detectar las rectas con la transformada de Hough.

Segmentación de los objetos presentes en el terreno de juego

1. Se aplicó la transformación top-hat y black-hat sobre la componente V del terreno de juego ya recortado previamente en el problema 1 y se combinaron en la misma imagen para marcar más en blanco sobre el terreno de juego los jugadores con camisetas claras y los jugadores con camisetas oscuras.
2. Se aplicó un k-means para separar las zonas claras (los jugadores) del resto del fondo.
3. Se aplicaron operadores morfológicos para eliminar el ruido de fondo y las líneas del campo.
4. Se aplicó un algoritmo de detección de contornos en la imagen resultante.
5. Se filtraron los contornos por tres características, el rango del área debería estar comprendido entre $140px^2$ y $2000px^2$, el ancho debería ser menor que $80px$ y el alto menor que $150px$, así se dejaron solamente los contornos que pertenecen a jugadores y no a las líneas que aun sobrevivían a los operadores morfológicos, por ejemplo.
6. Se calcularon los cuadrados alrededor de los contornos y fueron pintados sobre la imagen original.

Para la realización de este apartado se probó con diferentes características a la hora de filtrar la imagen antes de pasar el detector de contornos y también se probó con diferentes parámetros a la hora de filtrar los contornos localizados hasta dar con unos adecuados que detectaban correctamente la mayoría de los jugadores en las imágenes de prueba.

Evaluación de los resultados:

Para la evaluación de los resultados de este apartado, se han contado manualmente los Verdaderos Positivos, los Falsos Positivos y los Falsos Negativos ya que no se ha encontrado un método automático para realizarlo.

Verdaderos Positivos (VP): 245	Falsos Negativos (FN): 31
Falsos Positivos (FP): 24	Verdaderos Negativos (VN): N/C

(Los resultados del conteo manual pueden variar de la realidad en una pequeña escala)

(Los jugadores que están muy pegados y han sido detectados como uno solo, se ha contado como un Verdadero Positivo y un Falso Negativo)

Sensibilidad 91.08%: Esta métrica nos indica el porcentaje de jugadores reales que el modelo ha identificado bien.

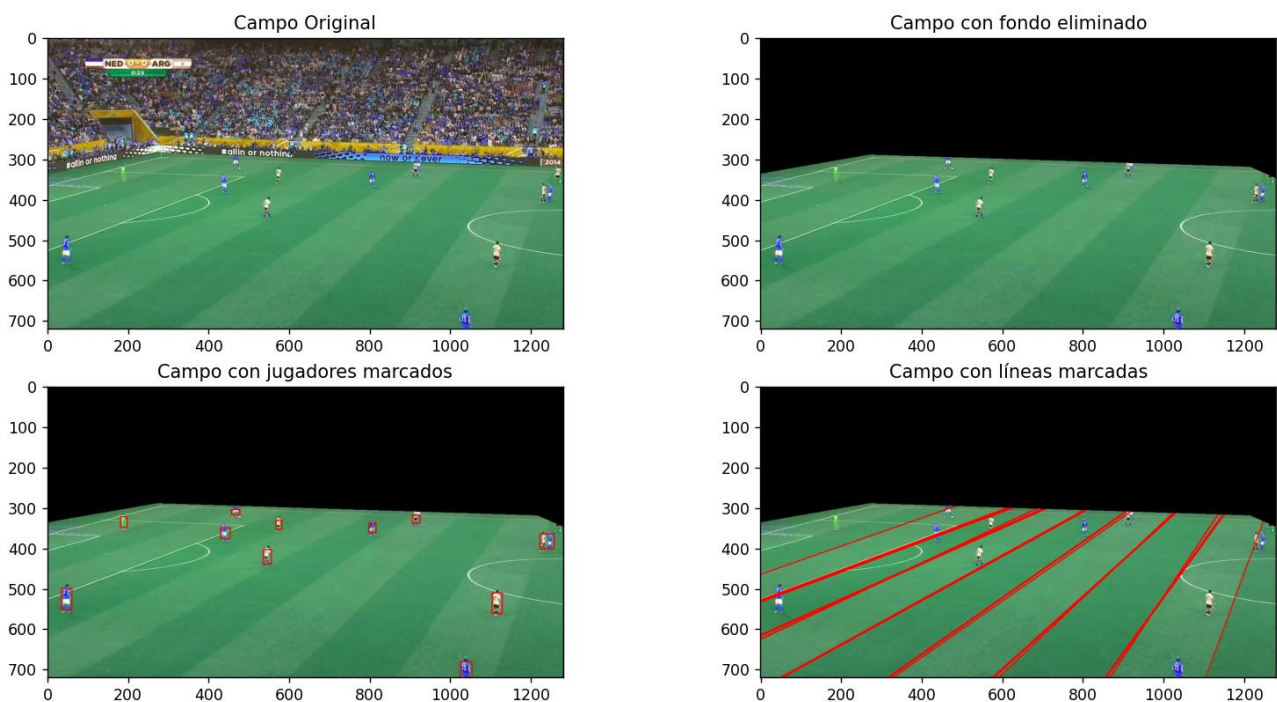
Precisión: 88.77%: Esta métrica nos indica el porcentaje de las jugadores que ha identificado el modelo que realmente eran jugadores reales.

Los resultados a priori son mejorables, sobre todo hablando en términos de precisión ya que falla a la hora de detectar algunos jugadores con las camisetas de los tonos parecidos al del campo. Por otra parte, también habría que mejorar la parte de detección de jugadores cuando están muy juntos, ya que los detecta solamente como un jugador, no como dos.

Identificación de problemas y posibles mejoras:

Para que la detección de los jugadores sea mucho más eficiente, una posible mejora sería aplicar menos filtros a la imagen y detectar más falsos positivos, pero agregando una CNN después que hiciera el trabajo de identificar si en ese cuadro hay un jugador o no. También se podría agregar la funcionalidad de que, en vez de ser binaria, detectara la cantidad de jugadores en el cuadro, para diferenciar los jugadores que están muy pegados en el campo. Por la parte de detectar el balón, no ha sido posible la implementación del algoritmo para ello ya que no se ha dado con una solución mínimamente aceptable.

Resultado final del método:



(Comportamiento de la práctica en la fotografía 0 de las imágenes proporcionadas de prueba)