

UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR



Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

TRABAJO FIN DE GRADO

DETECCIÓN JERÁRQUICA DE VEHÍCULOS

Raúl Palacios Ramos
Tutor: Álvaro García Martín
Ponente: José María Martínez Sánchez

Junio 2018

DETECCIÓN JERÁRQUICA DE VEHÍCULOS

Raúl Palácios Ramos

Tutor: Álvaro García Martín

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

Junio 2018

Trabajo parcialmente financiado por el Gobierno de España bajo el proyecto
TEC2014-53176-R (HAVideo)



Resumen

La implantación de sistemas visuales de captación en espacios tanto públicos como privados, hace que sea muy complicado el análisis y control a nivel de procesamiento. Pero debido a la gran importancia de estos aspectos, será necesario estudiar y desarrollar sistemas que actualicen lo que se ha establecido hasta el momento.

Los algoritmos actuales de detección de vehículos, en general obtienen buenos resultados en espacios controlados. Por esto se ha propuesto un sistema que permita la localización de vehículos en variabilidad de entornos. Particularmente vamos a centrarnos en el caso de oclusiones entre modelos.

El objetivo principal será generar un detector que permita localizar vehículos, mediante la obtención de información a partir de distintas configuraciones mecánicas y aumentando el área de búsqueda de su alrededor. Pudiendo estar definido el vehículo a nivel general o utilizando alguna de sus partes.

Finalmente se evaluará el algoritmo propuesto sobre secuencias de vídeo, demostrando que aumenta el rendimiento en la detección de vehículos gracias a las mejoras realizadas.

Palabras clave

Detección de vehículos, DTDP, oclusiones, detección jerárquica, HDCG, configuraciones, jerarquía.

Abstract

The establishment of visual capture system in both public and private spaces, makes analysis and control at the processing level very complicated. Owing to the great importance of these aspects, it would be necessary to study and expand systems that update what has been settled down so far.

The current vehicle detection algorithms, in general, earn good results in controlled spaces. But this property is undeniable that it is not usually fulfilled. Due to that, a system that allows the location of vehicles in variability of environments has been proposed. Specially, we are going to focus on the case of occlusions between models.

The main objective will be to develop a detector that enables locating vehicles, by obtaining information from different mechanical configurations and increasing the search area around them. Coulding be the vehicle defined at a general level or using some of its parts.

By last the proposed algorithm on video sequences will be reviewed, showing there is an increment in the detection of vehicles thanks to the improvements made.

Keywords

Vehicle detection, DTDP, occlusions, hierarchical detection, HDCG, configurations, hierarchy.

Agradecimientos

Lo primero de todo agradecer el esfuerzo y dedicación que he obtenido por parte de mi tutor Álvaro García Martín.

Dar las gracias a a mis padres, Manuel Palacios Rasco y Pilar Ramos Ortiz, y a mi hermana Cristina Palacios Ramos, ya que sin ellos no podría haber conseguido todos los objetivos que me he propuesto. Además de guiarme para ser el hombre en el que me he convertido.

A todos los compañeros que han estado conmigo durante toda la carrera (sabéis quienes sois...), con vosotros he compartido los mejores momentos que recuerdo en estos últimos cinco años (los peores ya no cuentan).

Finalmente gracias Lucía Gómez Prieto por estar ahí siempre, por tu apoyo incondicional y por haber vivido conmigo estos cinco años de carrera como si fueran los tuyos.

Índice general

Resumen	V
Abstract	VII
Agradecimientos	IX
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Organización de la memoria	3
2. Estado del arte	5
2.1. Detector de vehículos	5
2.1.1. Histogram of Oriented Gradient (HOG)	6
2.1.2. Discriminatively Trained Part Based Models (DTDP)	7
2.2. Non-Maximum Suppression (NMS)	12
2.3. Hierarchical Detection of Persons in Groups (HDPG)	13
2.3.1. Jerarquía	13
2.3.2. Detector	16
2.3.3. Postprocesado	17
3. Diseño y desarrollo	19
3.1. Introducción	19
3.2. Diseño del sistema	19
3.3. Desarrollo del sistema	20
3.3.1. Modelo	20
3.3.2. Jerarquías	21
3.3.3. Detector	24
3.3.4. Postprocesado	25

4. Integración, pruebas y resultados	27
4.1. Introducción	27
4.2. Marco de evaluación	27
4.2.1. Dataset	27
4.2.2. Métricas	28
4.3. Resultados	29
4.3.1. <i>Análisis QRIDR Synchronized</i>	30
4.3.2. <i>Análisis VXUSD Synchronized</i>	36
5. Conclusiones y trabajo futuro	41
5.1. Conclusiones	41
5.2. Trabajo futuro	42
A. Gráficas XUSD Synchronized	47
B. Ejemplos visuales	51

Índice de figuras

2.1. Diagrama de etapas de un detector basado en Modelos Deformables. . . .	6
2.2. Representación HOG.	7
2.3. Modelo DPM parkingcar	8
2.4. Pirámide de Resoluciones.	9
2.5. Proceso detector DTDP	11
2.6. Proceso NMS dado dos Boxes	12
2.7. Modelo INRIA person 2007 rc16	13
2.8. Desplazamiento de Jerarquía de grupos para modelo de personas.	15
2.9. Configuraciones para HDPG.	16
2.10. Distribución de probabilidad dada una configuración.	17
3.1. Diagrama de etapas del modelo HDCG.	20
3.2. Modelo DPM Parking Car.	21
3.3. Configuraciones propuestas para el análisis del HDCG.	23
3.4. Área de búsqueda para el sistema HDCG.	24
4.1. Ejemplos de imágenes del dataset [6]	28
4.2. Máscara de detecciones para QRIDR.	30
4.3. Análisis gráfico del sistema DTDP para QRIDR Synchronized.	31
4.4. DTDP vs HDCG . Configuración 11 para diferentes porcentajes de per- misividad de NMS.	32
4.5. DTDP vs HDCG . Configuración 12 para diferentes porcentajes de per- misividad de NMS.	33
4.6. DTDP vs HDCG . Configuración 13 para diferentes porcentajes de per- misividad de NMS.	34
4.7. Configuración 14 para el sistema HDCG.	37
4.8. DTDP vs HDCG . Configuración 14 para diferentes porcentajes de per- misividad de NMS.	38
A.1. DTDP vs HDCG . Configuración 11 para diferentes porcentajes de per- misividad de NMS.	48

A.2. DTDP vs HDCG . Configuración 12 para diferentes porcentajes de permisividad de NMS.	49
A.3. DTDP vs HDCG . Configuración 13 para diferentes porcentajes de permisividad de NMS.	50
B.1. Ejemplos visuales para las primeras 19 detecciones. HDCP (izquierda) vs DTDP (derecha).	52

Índice de tablas

4.1. AUC sistema DTDP.	30
4.2. AUC sistema HDCP. Configuraciones 11, 12 y 13	35
4.3. DTDP vs HDCG. Porcentajes de eficacia.	36
4.4. AUC y porcentajes de eficacia para el sistema HDCP con respecto DTDP. Configuración 14	37

Capítulo 1

Introducción

1.1. Motivación

En la actualidad el uso de sistemas de vídeo, fundamentalmente para el control de actividades de videovigilancia, ha adquirido un auge de forma exponencial. Esto es debido principalmente al avance de dispositivos tanto a nivel de procesamiento y captación como en la disminución del coste en cuanto a las herramientas manejadas.

La utilización de dichos sistemas es cada día más común y poder detectar de forma veraz y confiable vehículos (coches, furgonetas, autobuses, etc.), ofrece multitud de ventajas en el ámbito de la administración, seguridad o control. Conseguir identificar estos medios de transporte es una tarea complicada y afanosa, aunque jugamos con la labor computacional del análisis automático.

Detectar estos vehículos es una de las tareas fundamentales en los sistemas de monitorización de tráfico o aparcamientos. Pero es indiscutible que las necesidades de la aplicación dependerán de los requerimientos que se encuentren presentes a la hora de implantar el detector.

En los últimos años se han realizado grandes avances en esta tarea, consiguiendo buenos resultados en escenarios simples y/o controlados, modelando los vehículos de forma individual mediante información de apariencia y/o movimiento.

La principal desventaja que puede acarrear este análisis individual viene dada por la influencia del entorno. Esto se debe a que la obtención de resultados satisfactorios se verá en decrecimiento debido a las diferentes opciones en cuanto a la colocación de nuestras fuentes de vídeo (generalmente en un lugar elevado), cambios de iluminación, condiciones

atmosféricas y principalmente las oclusiones entre modelos. Es debido a estos factores que existe una gran variabilidad a la hora de definir el análisis, ya que nos encontramos con diferentes apariencias técnicas y puntos de vista que contribuirán al déficit de estas detecciones.

Por todo lo involucrado anteriormente, en este proyecto nos centraremos en la creación de un algoritmo de detección de vehículos donde se analizarán secuencias de vídeos con grandes aglomeraciones. De tal forma que, teniendo en cuenta las desventajas de los detectores convencionales, trataremos de implementar una forma alternativa que nos permita mejorar su rendimiento mediante la utilización de jerarquía de grupos y con la ayuda de configuraciones mecánicas de los vehículos.

1.2. Objetivos

El objetivo de este trabajo fin de grado es el desarrollo de un algoritmo de detección jerárquica de vehículos en entornos con alta densidad, de tal forma que no se centre únicamente en la detección de forma individual, sino que aproveche la información de detección de múltiples vehículos para mejorar los resultados obtenidos en este tipo de escenarios.

El algoritmo se centrará en la información mecánica de los automóviles utilizándose todas ellas en su conjunto, escogiendo algunas de sus partes (maletero, capó, techo...), o incluso combinaciones de las mismas para obtener mejores detecciones.

Para evaluar los resultados, los experimentos se realizarán sobre secuencias de vídeo con alta densidad. En particular, para nuestro escenario de análisis, una alta ocupación hace que en zonas como ruedas y parte inferior de los vehículos, quedan parcial o totalmente ocluidas por modelos de su alrededor.

Finalmente se realizará una variación del sistema HDPG (Hierarchical Detection of Persons in Groups), que permita la detección de vehículos adaptándose a nuestro problema.

Partiendo de estos propósitos, el desglose de los objetivos es el siguiente:

1. **Investigación del estado del arte actual.** Se realizará un análisis en profundidad de los algoritmos y procedimientos utilizados para detección de objetos teniendo en cuenta:
 - Investigaciones y resultados obtenidos actualmente en el uso de localización

de objetos.

- El estudio en detecciones de personas y grupos de personas determinado por el detector DTDP (Discriminatively Trained Part Based Models).
- Búsqueda y posibles modificaciones de los modelos establecidos para la detección.
- Generación de mapas de confianzas así como su utilización general en las zonas oclusivas de las imágenes.
- Adaptación de las detecciones mediante NMS (Non-Maximum Suppression), que permita la eliminación de detecciones en situaciones de alta similitud.

2. **Aprendizaje de herramientas y bibliotecas.** Utilizar de forma eficaz y óptima las herramientas y bibliotecas necesarias para el correcto ejercicio de este trabajo: Matlab.
3. **Desarrollo en la adaptación del algoritmo.** Estudio y adaptación del algoritmo HDPG (Hierarchical Detection of Persons in Groups), observando las principales diferencias que se pueden aplicar al algoritmo a desarrollar, estableciendo parámetros adecuados y sacando resultados que influyan para el correcto funcionamiento en el tratamiento de vehículos.
4. **Selección del Dataset.** Búsqueda de una base de datos de imágenes que represente las oclusiones entre vehículos y que nos permita la evaluación adecuada del algoritmo propuesto.
5. **Evaluación y análisis de los resultados.** Comparar la eficacia del algoritmo en búsqueda de las ventajas, limitaciones y posibles mejoras, estableciendo una métrica adecuada para conseguir una evaluación correcta de los resultados.

1.3. Organización de la memoria

La memoria del proyecto se divide en los siguientes capítulos:

- Capítulo 1. Motivación y objetivos del proyecto.
- Capítulo 2. Estado del arte: además del estudio genérico en cuanto a detección de objetos, fundamentalmente nos centraremos en las características y peculiaridades de dos algoritmos concretos: DTDP y HDPG.

- Capítulo 3. Diseño y desarrollo: descripción de la adaptación y programación del sistema de detección de vehículos utilizando información procedente de alrededor del objeto. Estará compuesto por cuatro etapas diferenciadas: elección del modelo, utilización de jerarquías, detección y postprocesado. .
- Capítulo 4. Evaluación del sistema: estudio del criterio de evaluación y análisis de los resultados obtenidos para el detector base DTDP y nuestro sistema de detección utilizando información jerárquica.
- Capítulo 5. Conclusiones y trabajo futuro: conclusiones obtenidas tras el análisis de resultados del trabajo presentado. Trabajo pendiente y trabajo futuro.
- Bibliografía.

Capítulo 2

Estado del arte

En la actualidad, la detección es uno de los retos de mayor dificultad en cuanto al análisis computacional. Esto es generalmente debido a la variabilidad de modelos, oclusiones, colores, puntos de vista, partes móviles... que posee un vehículo, empeorando fundamentalmente cuando tenemos en cuenta espacios reales como pueden ser estacionamientos o aglomeraciones en vías públicas con un tránsito elevado de automóviles.

Hoy en día hay multitud de aproximaciones para abarcar el análisis y detección en imágenes independientes o en secuencias de vídeos, que engrandece el trabajo realizado hasta el momento en el ámbito de la localización.

En este capítulo se va a detallar el estado del arte, centrándonos en las partes principales que se deben considerar a la hora de conseguir detectar objetos pero focalizando fundamentalmente en vehículos. De tal forma que se explicarán los diferentes algoritmos desarrollados hasta el momento, los cuales ayudarán a vislumbrar el funcionamiento básico y facilitará la comprensión de nuestro algoritmo a desarrollar.

2.1. Detector de vehículos

Para lograr abordar nuestro algoritmo se ha tenido que considerar un detector al cual poder mejorar su rendimiento. En nuestro caso se ha propuesto como punto de partida el DTDP [4], que está disponible en las librerías de Matlab [7].

Vamos a explicar cada una de las partes diferenciadas que se pueden observar en este detector. Fundamentalmente son cuatro, ver Figura 2.1.

- Implantación del modelo generado mediante HOG (Histogram of Oriented Gra-

dient).

- Pirámide de características, manipulando las posibles configuraciones mecánicas que tiene un vehículo (DTDP).
- Detecciones mediante la obtención y utilización de los mapas de características.
- Postproducción, eliminando las detecciones fuertemente ocluidas mediante NMS.

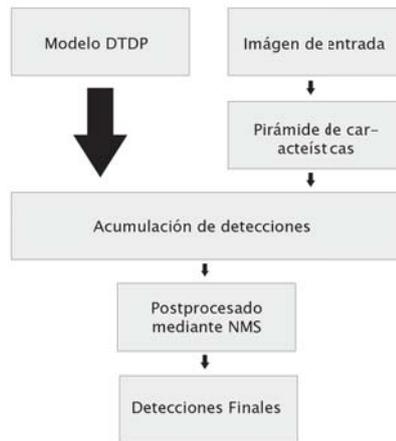


Figura 2.1: Diagrama de etapas de un detector basado en Modelos Deformables.

2.1.1. Histogram of Oriented Gradient (HOG)

Los modelos que se utilizan para evaluar las detecciones se obtienen mediante el algoritmo de HOG [4]. Este algoritmo se basa en la obtención de imágenes características a partir de las cuales se clasifican las regiones de interés (Region of Interest, ROI), indicando si contienen vehículos y descartando aquellas que no los contengan. De esta manera se consigue desprestigiar determinados falsos positivos que puedan aparecer en las imágenes.

El método de HOG consiste en la creación de histogramas locales normalizados de las orientaciones de los gradientes de una imagen. Este método se implementa de acuerdo a una serie de pasos:

- Primero se realiza una división de la imagen en un conjunto de celdas uniforme (cells).

- A continuación se estima la orientación del gradiente de cada uno de los píxeles que conforman un cell.
- Finalmente se almacena la información en un histograma de orientaciones de k-bins. La figura 2.2 representa los descriptores HOG para $k=9$.



Figura 2.2: Representación HOG.

Para obtener una mayor eficacia en espacios con grandes cambios de iluminaciones, ya sea por luces o por sombras, se suele utilizar un proceso de normalización localizado antes de proceder a manipular las imágenes. Esto se consigue teniendo en cuenta áreas de mayor dimensión. El procedimiento de normalización consiste en acumular la energía de los histogramas locales sobre la región de mayor superficie y usar este resultado para aplicárselo a todos los cells que estén contenidos en dicho espacio. Con este proceso conseguimos lo que se denominan descriptores HOG.

En definitiva, HOG se encarga de conseguir las características más importantes de una imagen. Además, es muy eficaz en cuanto a variaciones de traslación y rotaciones, lo que hace que sea uno de los descriptores más utilizados a la hora de detectar coches o cualquier objeto o persona en general.

2.1.2. Discriminatively Trained Part Based Models (DTDP)

El modelo DTDP [8] está muy extendido en la actualidad. Principalmente se debe a que obtiene resultados satisfactorios en situaciones muy dispares. Se caracteriza por la utilización de un sistema multiescala de detección basado en la mezcla de las diferentes estructuras y deformidades del modelo objeto o persona a detectar.

Este sistema consiste en que dado un conjunto de imágenes, se seleccionará el modelo que describa lo mejor posible el patrón a detectar (HOG). El modelo vendrá determinado mediante una serie de estructuras definidas por el root y una serie de filtros correspondiente a cada sección, además de sus modelos deformables. Ver Figura 2.3.

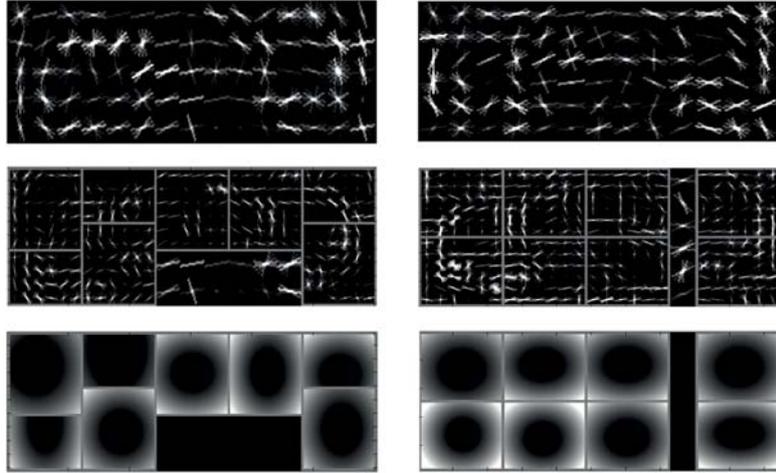


Figura 2.3: Modelo DPM parkingcar

El detector propone N posibles partes ($n=0 \dots N$) que se sitúan alrededor del root principal ($n=0$). Generando las características del root a la mitad de resolución que las partes localizadas del modelo, ya que este especifica las secciones más importantes o destacadas.

Todos los modelos de estudio se caracterizan por la utilización de filtros lineales aplicados a una serie de mapas de características que describen una parte de la imagen (HOG). Generalmente DTDP utiliza una variación del método de HOG, definiendo una plantilla rectangular (matriz de vectores de puntuación). El resultado en una posición determinada (x,y) de la imagen, que será el producto escalar de la matriz de características junto con el filtro de puntuaciones.

Plasmando este método en multiescala, es decir, utilizando una pirámide a partir del filtrado y submuestreo de la imagen original, conseguimos una pirámide de características que permitirán observar las diferencias entre posiciones y escalas de la imagen. De esta manera obtendremos para cada uno de los niveles (Figura 2.4), matrices de puntuaciones a distintas resoluciones.

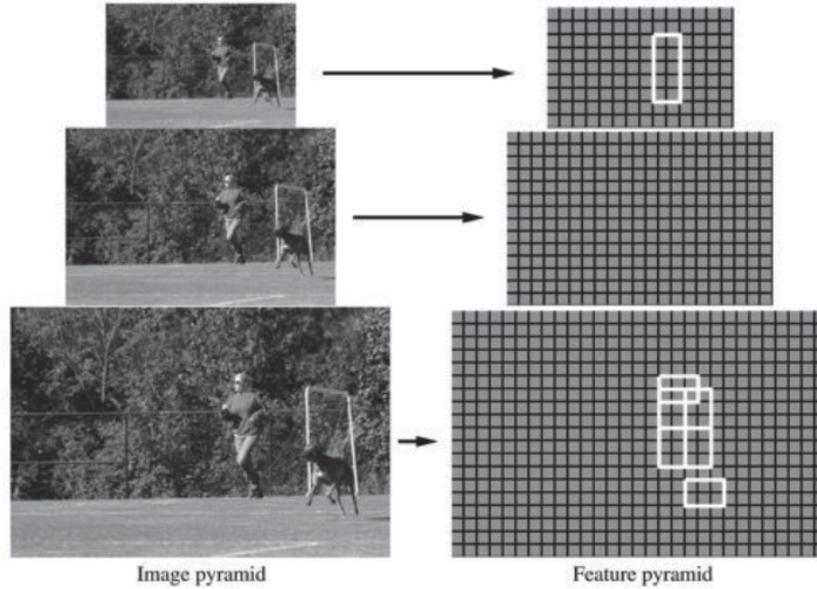


Figura 2.4: Pirámide de Resoluciones.

El root al estar a la mitad de resolución obtienen mejores resultados en imágenes escaladas, que para cada una de las partes [8].

Si tenemos en cuenta cada una de las partes mencionadas anteriormente y nos centramos sólo en cada una de ellas, estas se definirán con una serie de parámetros o variables:

- F_n , que especifica la respuesta al filtro HOG.
- v_n , será un vector de dos dimensiones que indica la posición relativa con respecto al root.
- d_n , función cuadrática que define la deformación de cada parte mediante un vector de cuatro dimensiones.

La puntuación de cada sección se obtiene mediante las ecuaciones 2.1, 2.2 y 2.3.

$$BP_n(x, y, s) = F_n(x, y, s) - (dn, \phi(dx_n, dy_n)) \quad (2.1)$$

$$(dx_n, dy_n) = (x_n, y_n) - (2(x_0, y_0) + v_n, 0) \quad (2.2)$$

$$\phi(dx_n, dy_n) = (dx, dy, dx^2, dy^2) \quad (2.3)$$

Siendo:

- BP_n la puntuación de un pixel a nivel multiescala (s), en una posición determinada (x, y) para cada una de las partes (n).
- F_n la respuesta al filtro de HOG, mencionada anteriormente, pero aplicado a cada uno de los pixeles multiescala.
- (dx_n, dy_n) el desplazamiento en ambos ejes de cada una de las partes.
- $\phi(dx_n, dy_n)$ la distribución de las deformaciones para cada una de las N partes.

El resultado será el sumatorio ($n=0 \dots N$) de las puntuaciones tanto de cada una de las partes como del root, de manera que valoraciones elevadas confirmarán la presencia de vehículos mientras que puntuaciones bajas reafirmarán la usencia de ellos. Cabe destacar la aparición de falsos positivos y detecciones muy similares alrededor de un mismo modelo. El método de supresión se explicará más adelante.

Para aclarar la explicación se presenta la Figura 2.5, donde se puede observar que a partir del modelo junto con una imagen, se obtienen las puntuaciones (responses) a partir de las características de la imagen [3] y de las distintas partes del modelo. De tal forma que, acumulando cada una de las puntuaciones, conseguimos obtener la imagen resultante de detecciones.

Como se ha comentado anteriormente, esté método es de los más manejado por la gran cantidad de posibilidades que proporciona, ya que tan solo variando el modelo nos permite la localización de cualquier vehículo.

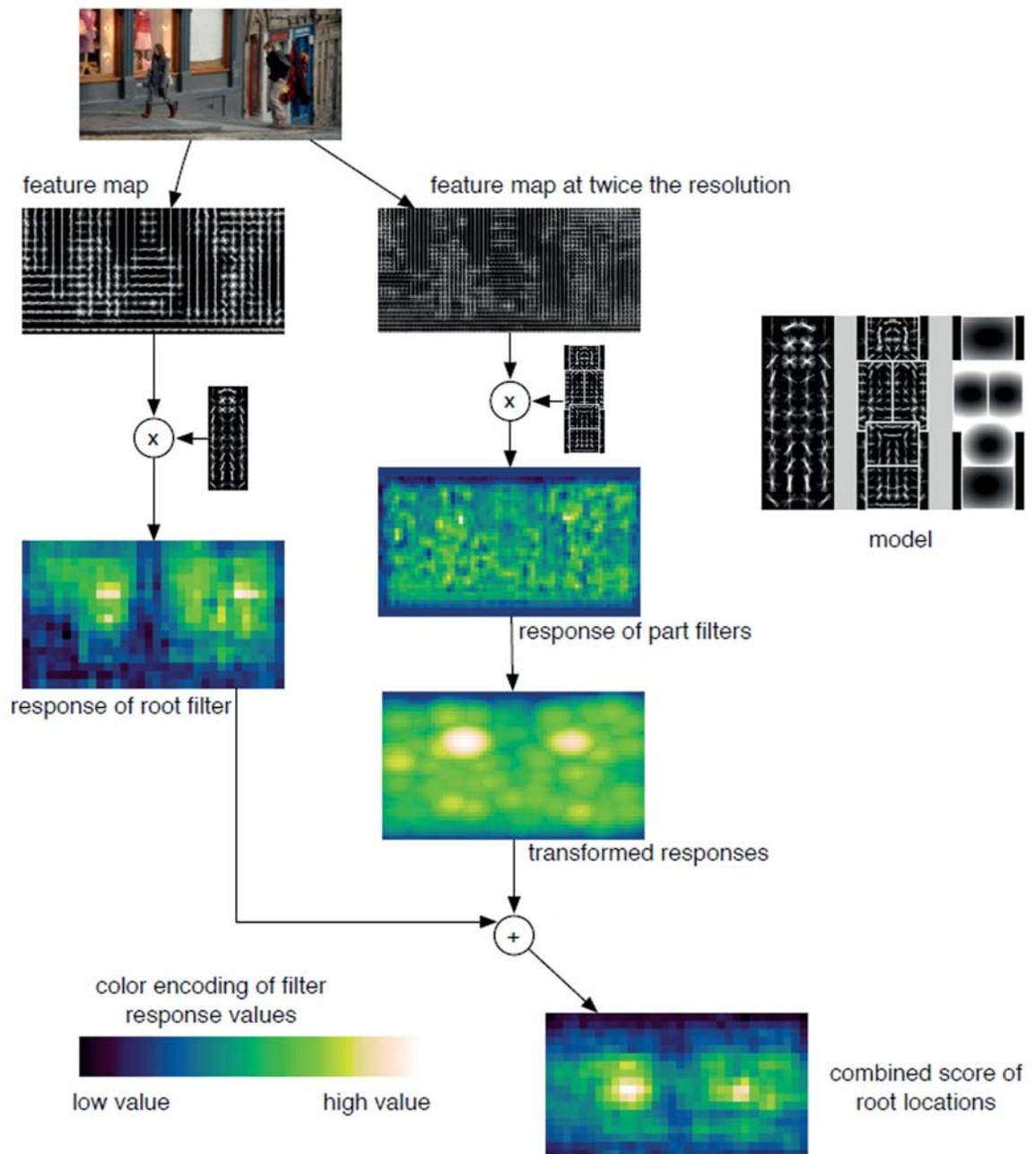


Figura 2.5: Proceso detector DTDP

2.2. Non-Maximum Suppression (NMS)

En el método DTDP, al igual que muchos otros detectores, destacan la aparición de falsos positivos o aglomeraciones de detecciones alrededor de un mismo objeto o persona. Es por esto que las detecciones obtenidas necesitarán un post-procesado. El método más arraigado, que utiliza el algoritmo base DTDP, y el cual hemos considerado punto de partida para nuestro algoritmo a desarrollar, es el conocido como NMS [9].

A nivel general el proceso NMS se encarga de, teniendo en cuenta una determinada detección con una puntuación elevada, eliminar detecciones que se encuentran solapadas con ésta en un determinado porcentaje y que tienen una puntuación menor. Ver figura 2.6.

En cuanto al procesamiento se realiza en dos pasos. Primero una ordenación de las cajas de detecciones del objeto/persona completo (boxes) teniendo en cuenta su puntuación; y en segundo lugar se suprimen las puntuaciones que se solapen más de un 50 %, quedando únicamente el box con mayor puntuación.

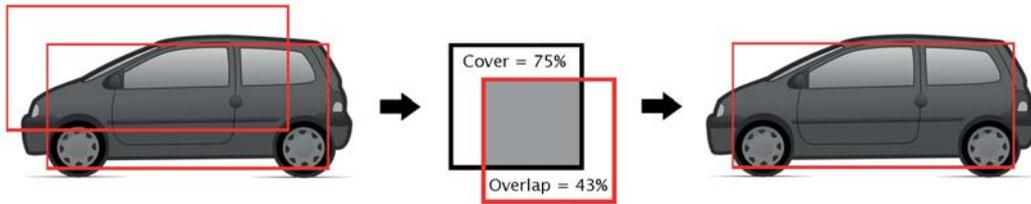


Figura 2.6: Proceso NMS dado dos Boxes

La principal desventaja de este método es el uso de una única configuración, es decir, únicamente va a detectar objetos/personas mediante la utilización de un único modelo. Es por esto que cuando tenemos un escenario en el que tenemos, por ejemplo, personas posicionadas de forma independiente y personas que se encuentran ocluidas unas con otras, el proceso de detección mediante DTDP únicamente obtendrá resultados satisfactorios para una opción, nunca para ambas.

2.3. Hierarchical Detection of Persons in Groups (HDPG)

Para poder desarrollar el algoritmo propuesto en este ensayo, se ha llevado una investigación exhaustiva del funcionamiento del HDPG [2]. Este parte del proceso anteriormente descrito e intenta solucionar las desventajas que pueden acarrear la oclusión de grupos de personas en espacios con grandes aglomeraciones.

Principalmente se desarrollarán en este apartado las mejoras y variaciones fundamentales que propone el algoritmo HDPG con respecto el anteriormente descrito. Ya que son éstas las que tendremos que tener muy en cuenta a la hora de implementar nuestro detector.

2.3.1. Jerarquía

El modelo que se propone en el algoritmo es INRIA person 2007 rc16 [4] (Figura 2.7). Este básicamente consta de dos partes: de frente y de espaldas, cada una de ellas con su respectivas 8 partes del cuerpo y el root. Las partes del cuerpo se relacionan con el root a través de los anchors, que es un vector de tres coordenadas el cual especifica la posición (x, y, r) con respecto al root. Tomando r valores de 0 o 1 según pertenezca o no al root.

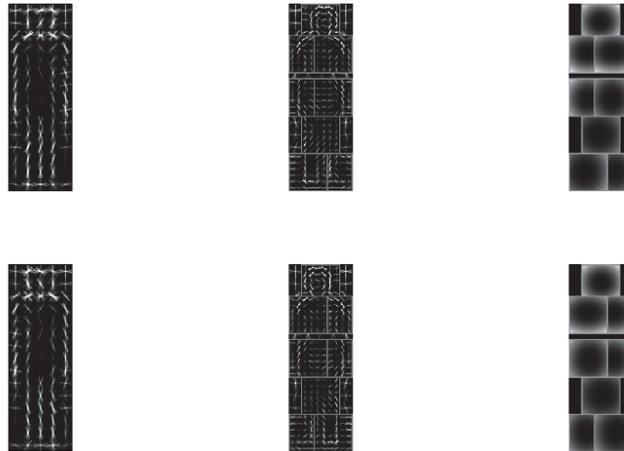


Figura 2.7: Modelo INRIA person 2007 rc16

Se visualiza de izquierda a derecha al root, configuración de partes y sus respectivas deformaciones.

Las diferentes formas de utilizar o modificar el modelo a la hora de aplicarlo en una

imagen para obtener las diferentes detecciones es a lo que nos referimos como jerarquías. HDPG propone dos procedimientos nuevos de jerarquías, a nivel de grupo y a nivel de partes.

2.3.1.1. Jerarquía de grupos

La jerarquía de grupos plantea la modificación del modelo genérico para adaptarlo a las distintas situaciones de análisis.

Es importante resaltar dos aspectos fundamentales a la hora de trabajar con jerarquía de grupos:

- La persona principal (Main Person, MP) será el núcleo principal a partir del cual se detectarán las personas secundarias (Secondary Person, SP).
- Los `anchors_shift`, que indicarán el desplazamiento o posición relativa que tendrán las distintas SP con respecto al MP, serán vectores bidimensionales (A_x , A_y). Representarán las zonas de búsqueda.

El objetivo fundamental será obtener la detección principal del MP acumulando la información obtenida por cada una de las personas que forman el grupo, ver Figura 2.8. De esta manera las personas que estén más ocluidas mejorarán su puntuación gracias al resto de personas.

El sistema desarrollado en el algoritmo HDPG se basa en el desplazamiento vertical y horizontal de un `anchor_shift` entre -6 y +6. Se seleccionan estos valores por los `anchors` de las distintas configuraciones de las partes del cuerpo que presenta el modelo INRIA. De esta forma la oclusión de las SP a nivel horizontal es de medio cuerpo y, a nivel vertical, el tamaño de la cabeza.

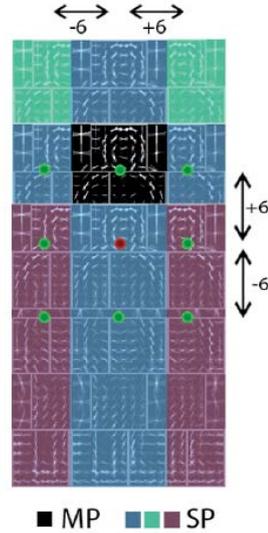


Figura 2.8: Desplazamiento de Jerarquía de grupos para modelo de personas.

En definitiva, los autores proponen un modelo general definido por N modelos individuales. Al igual que en el modelo genérico, cada modelo individual se compondrá del root, los filtros de cada partes y las deformaciones de cada una de ellas.

2.3.1.2. Jerarquía de partes

La jerarquía de partes expone aplicar distintas configuraciones para cada uno de los modelos establecidos por los anchors_shift. Define, como se presenta en la Figura 2.9, cinco posibles configuraciones tanto con o sin root. Siendo cada configuración la selección de un conjunto específico de partes del cuerpo en el modelo.

Dado que se quiere mejorar resultados en situaciones con grandes aglomeraciones de personas, este método se propuso debido a que las SP siempre eran ocluidas por la MP de una forma muy característica, y por lo general obtenían menores puntuaciones en las detecciones. De esta forma, dado un anchor_shift, se estableció que:

- Si $A_x > 0$, se considera un desplazamiento horizontal hacia la derecha, de tal forma que al SP se le aplicaría el modelo del MP pero sólo con las partes del cuerpo que quedan a la derecha del eje longitudinal.
- Si $A_x < 0$, se considera un desplazamiento horizontal hacia la izquierda, de tal

forma que se aplicaría el procedimiento inverso al anteriormente descrito.

- Si $A_x = 0$, se considera un desplazamiento vertical donde solo se utilizan las partes de la cabeza y las correspondientes a ambos hombros.

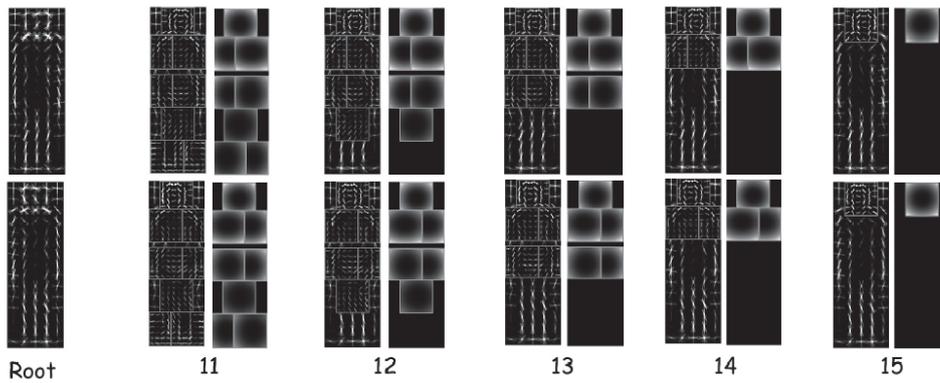


Figura 2.9: Configuraciones para HDPG.

2.3.2. Detector

A la hora de obtener la detección final correspondiente con una persona, se debe de tener en cuenta las puntuaciones obtenidas tanto en el modelo MP, como las obtenidas por los modelos secundarios SP desplazados un determinado `anchor_shift`. Obteniendo entonces los mapas de detecciones o confianza para cada modelo, tanto de las ocho partes del cuerpo como las del root. De esta manera la puntuación total de cada detección será la suma de cada una de las detecciones correspondiente al grupo de análisis (MP + SPs).

Otra particularidad que añade este detector es la forma de sumar las detecciones, ya que realiza una suma ponderada a partir de mapas de probabilidad [1]. Estos mapas se obtienen a partir de una configuración determinada del modelo, es decir, se calcula para cada configuración su distribución de probabilidad. Esta distribución, como muestra la Figura 2.10, indica la probabilidad de que una determinada detección sea una persona.

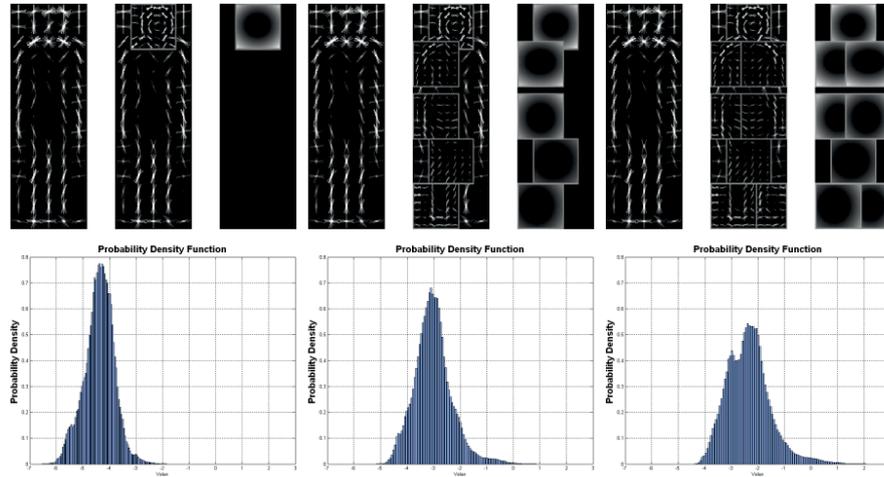


Figura 2.10: Distribución de probabilidad dada una configuración.

Posteriormente, la suma total de probabilidades se vuelve a pasar a puntuación para poder evaluar el algoritmo con respecto al DTDP de forma satisfactoria.

2.3.3. Postprocesado

Para este algoritmo fue necesario variar el proceso de postprocesado que se aplicaba al DTDP original, ya que en este caso se obtenían en un mismo individuo multitud de detecciones producidas por la adición de los distintos modelos con sus configuraciones independientes. Reducir estas detecciones con el proceso NMS habitual contribuiría a reducir las detecciones que podríamos haber conseguido en personas que se encontraban ocultas.

El proceso que se implementó consiste en dos fases:

- En primer lugar se hace un NMS teniendo en cuenta únicamente las detecciones obtenidas a partir de la cabeza. Se realiza este proceso ya que se considera que es poco usual que dos cabezas se oculten. Con lo cual se eliminan las detecciones que estén solapadas mínimamente, adquiriendo como siempre la de mayor puntuación.
- A continuación, a partir de todas las localizaciones restantes se realiza el NMS habitual pero siendo menos estrictos a la hora de aplicar el factor de solapamiento entre detecciones.

De esta forma obtendremos las detecciones finales correspondientes al HDPG, las cuales generalmente son mucho más abundantes con respecto al sistema DTDP.

Capítulo 3

Diseño y desarrollo

3.1. Introducción

Como se ha mencionado anteriormente la base de nuestro algoritmo se asienta fundamentalmente en la detección de vehículos en espacios con grandes aglomeraciones, donde se entorpece notablemente la localización de todos ellos. Esta dificultad se debe a que en estacionamientos donde los vehículos están dispuestos de manera longitudinal, algunas partes mecánicas quedan ausentes a la vista (principalmente ruedas o partes delanteras y traseras de los vehículos), aumentando de manera notable con respecto a la posición de los sistemas de captación de vídeos utilizados. De esta forma, se implantarán diferentes configuraciones junto con áreas de búsqueda que se adecúen lo más posible al escenario de análisis.

Teniendo en cuenta el estado del arte actual y basándonos principalmente en [2], se expondrán las características fundamentales que se han desarrollado a la hora de adaptar dicho algoritmo para la detección de vehículos. La base del sistema a implementar, al igual que HDPG, se asienta en el sistema propuesto por [8].

El algoritmo se ha denominado Detector Jerárquico de Vehículos o Hierarchical Detection of Cars in Groups (HDCG).

3.2. Diseño del sistema

Siguiendo con la dinámica de los detectores mencionados anteriormente, el diseño de nuestra aplicación se constituye por las siguientes etapas, ver Figura 3.1.



Figura 3.1: Diagrama de etapas del modelo HDCG.

- **Elección del modelo.** Esta fase es una de las más importantes. Una correcta descripción del vehículo, la cual se ajuste mejor al entorno de evaluación, ayudará en la mayor parte del proceso para conseguir una mayor proporción de detecciones óptimas. En nuestro caso se ha optado por el DPM Parking Car [8].
- **Utilización de las jerarquías.** La característica fundamental del algoritmo se basa en el uso de jerarquías. Se han tenido que implantar nuevas configuraciones a nivel de partes, además de variar el área de búsqueda en cuanto a MP y SP a nivel de grupo para adaptarlo a nuestro escenario o problema.
- **Detector.** La obtención de los mapas de puntuaciones sigue la misma corriente que el detector HDPG. En general tiene en cuenta mapas de puntuaciones a nivel de multiescalas con diferentes resoluciones.
- **Postprocesado.** El proceso de eliminación de detecciones se ha tenido que variar. Esto se debe a que las configuraciones que se desarrollan en el modelo de personas no se ajusta con el modelo mecánico de vehículos.

3.3. Desarrollo del sistema

3.3.1. Modelo

El modelo seleccionado ha sido el DPM Parking Car (ver Figura 3.2), que se basa en la utilización de descriptores HOG mediante un modelo de partes deformables. Este modelo define dos posibles puntos de vista, los cuales favorecen las detecciones y dependen de la posición de las cámaras de vigilancia.

El primer punto de vista describe vehículos teniendo en cuenta posiciones de videocámaras más inclinadas; en cambio, la segunda pose describe situaciones de cámaras que obtienen perspectivas con menor inclinación. Cada uno de los puntos de vista tendrá asociados sus correspondientes ocho partes mecánicas comprendidas alrededor del root, junto con sus filtros deformables. Además, ambas poses tienen en cuenta posiciones de

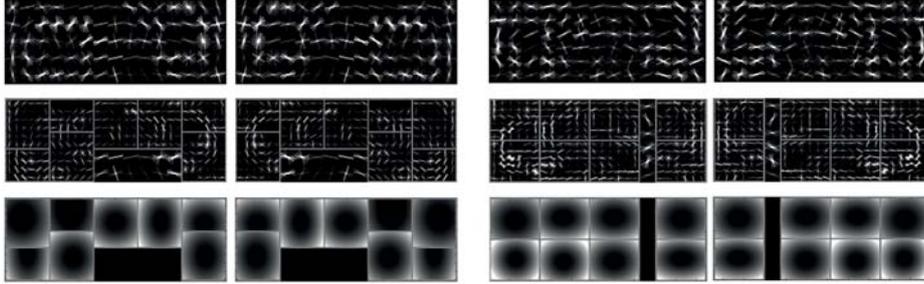


Figura 3.2:

Modelo DPM Parking Car. Primera pose (izquierda) y segunda pose (derecha) del modelo.

vehículos que se encuentran, mirando en su eje longitudinal, hacia la izquierda y hacia la derecha. En definitiva, tendremos un modelo generado a partir de cuatro poses o descriptores diferentes: pose1-derecha, pose1-izquierda, pose2-derecha y pose2-izquierda.

Para cada uno de los puntos de vista, el modelo parte de la definición de tres posibles anchors a tener en cuenta. Asigna un tamaño de $[6 \times 6]$ para todas las partes mecánicas que describen el vehículo. Pero en cambio, para el tamaño del root define una dimensión de $[5 \times 15]$ y $[6 \times 13]$ en cada pose respectivamente. Siendo $[w \times h]$ el ancho y alto establecido por el modelo para el root y cada una de sus partes.

En definitiva, DPM se Parking Car se compone de un total de 36 posibles filtros, 9 (partes + root) teniendo en cuenta los dos poses y los dos puntos de vista (izquierda y derecha) que contempla el modelo.

Como se ha descrito anteriormente, cada pose está constituida de manera autónoma, es decir, la distribución junto con los anchors de cada una de las partes es totalmente distinta. Esto ha dificultado el desarrollo del algoritmo, ya que cada uno de los puntos de vista se ha debido de procesar de forma independiente.

3.3.2. Jerarquías

Al igual que en el sistema HDPG, el algoritmo propuesto se basa en el uso de jerarquías que permitan variar las configuraciones y el área de búsqueda. Su correcta utilización permitirá aproximarse mejor en las detecciones y obtener un mejor rendimiento del detector DTDP.

El algoritmo que estamos implantando se intenta aproximar todo lo posible a unos

resultados óptimos a la hora de su evaluación. Es por esto que se han formulado unas especificaciones a nivel de grupos y de partes que contribuyan a conseguirlo. Pero cabe destacar que HDCG permite multitud de configuraciones que consiguen adaptarse a cualquier situación de análisis.

3.3.2.1. Jerarquía de grupos

El modelo utilizado describe vehículos en posición horizontal, con lo cual la forma para que dos medios de transporte se encuentren ocluidos es que ambos se posicionen en el mismo eje vertical. Esto se produce principalmente en zonas de aparcamiento donde el dispositivo de captación se sitúa en las calles y todos los coches se encuentran aparcados en batería.

Para este escenario de análisis el grupo estará formado únicamente por dos componentes:

- El coche principal (Main Car, MC) que será el núcleo donde se focalizan las detecciones de todo el grupo. Define el vehículo sobre el cual queremos conseguir las detecciones.
- Los coches que se conforman alrededor del MC serán los llamados coches secundarios (Secondary Car, SC), y representarán el área de búsqueda que se pretende abarcar.

Para nuestro escenario de análisis, se ha determinado un SC independiente, pues en un estacionamiento la única oclusión que se tiene en cuenta será el vehículo que se sitúe en la zona más alejada al dispositivo de videovigilancia, es decir, el vehículo posterior al MC.

Otro factor a tener en cuenta es el desplazamiento que se proporciona a los SC con respecto al MC (`anchor_shif`). Este componente se ha inicializado considerando los anchors establecidos por el modelo para cada una de las partes mecánicas del vehículo. En particular se ha definido un factor de +6, teniendo en cuenta únicamente el eje vertical del modelo. Este parámetro consigue que, a la hora de acumular las detecciones, para el SC solo se evalúe la parte superior del vehículo. Cabe destacar que aunque se ha valorado un método que permita evaluar el dataset utilizado, el sistema está capacitado para permitir cualquier área de búsqueda.

3.3.2.2. Jerarquía de partes

Tomando como punto de partida localizaciones referentes a estacionamientos, el análisis del detector se ha centrado en configuraciones que se asemejen, en la medida de lo posible, al escenario a evaluar.

Reconociendo la jerarquía de grupos que se va a desarrollar, se han tenido que implantar nuevas configuraciones a nivel de partes. Considerando que las características técnicas expuestas en el modelo de vehículos son muy limitadas, se han especificado configuraciones que evalúan la zona alta y baja del vehículo. Estas configuraciones serán las posibles opciones que se podrán asignar a cada uno de los `anchors_shift`.

Como se muestra en la Figura 3.3, se establecen un total de tres configuraciones:

- En primer lugar, se define la configuración 11 como el modelo que incluye todo el conjunto de partes. Esta es la disposición que más se va a aproximar a los resultados obtenidos con el método DTDP.
- Basándonos en la hipótesis de que el vehículo que representa al MC, en perspectiva con la cámara de vigilancia, siempre va a ser más largo con respecto al eje horizontal, la zona baja del vehículo que representa el SC nunca se observará. Es por ello que las partes que describen la zona inferior del vehículo SC no supondrán gran ayuda en las detecciones y consecuentemente se representa la configuración 12 como el conjunto de partes que referencian la franja superior.
- Por último, se propone la configuración 13 como igualdad a la configuración 15 del HDPG que estipula la utilización únicamente de la cabeza. Por tanto las detecciones obtenidas de esta distribución solo contendrán la zona que describe el techo del vehículo.

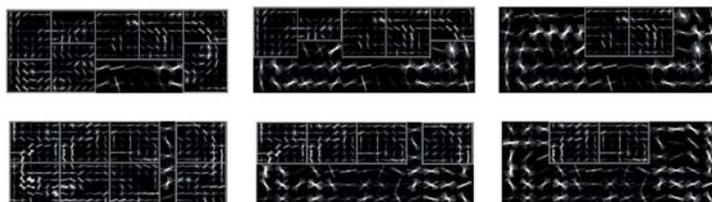


Figura 3.3: Configuraciones propuestas para el análisis del HDCG. Configuraciones 11, 12 y 13 respectivamente.

A la hora de aplicar las diferentes configuraciones antes descritas, y dado que se quieren mejorar situaciones de máxima oclusión entre vehículos, a diferencia del sistema HDPG, el algoritmo implementado propone una única posibilidad de desplazamientos del `anchor_shift`: se considera exclusivamente un desplazamiento vertical ($Ax = 0$). Despreciando en consecuencia los desplazamientos horizontales. De esta forma, si por ejemplo comparamos la configuración 11, tanto para el sistema HDPG como para el HDCG. Ver figura 3.4.

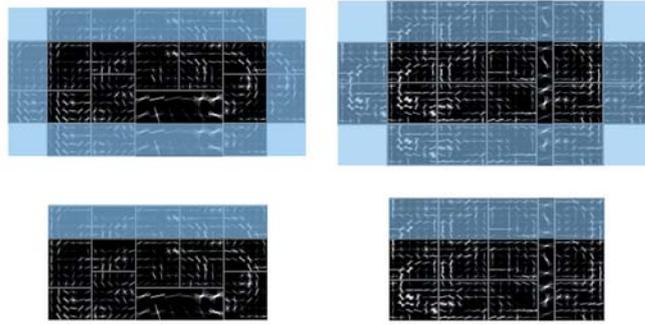


Figura 3.4: Área de búsqueda para el sistema HDCG.

En la parte superior se muestra el área total que puede abarcar el sistema. En la zona inferior el área utilizado para las configuraciones propuestas en el análisis del dataset.

3.3.3. Detector

El proceso de acumulado para cada nivel de la imagen sigue la misma dinámica que el sistema HDPG. Se obtiene la detección final correspondiente con un vehículo teniendo en cuenta las puntuaciones derivadas tanto de nuestro modelo MC, como las obtenidas por el modelo secundario SC, desplazado un determinado `anchor_shift`.

Obtendremos de esta forma los mapas de confianza para cada modelo del root y las ocho partes mecánicas del vehículo. Juntando toda esta información, la localización definitiva será la suma de cada una de las detecciones correspondiente al grupo de análisis (MC + SC).

A diferencia que el sistema HDPG, no se ha llevado a cabo una suma de mapas de detección mediante ponderación de probabilidades. Este proceso no se ha realizado debido a que durante la evaluación empírica se observó que a diferencia del modelo de persona, en nuestro caso para el modelo de coche conserva la misma distribución de puntuaciones para las diferentes configuraciones propuestas, o la diferencia es despreciable.

3.3.4. Postprocesado

Para el algoritmo desarrollado, ha sido necesario variar el proceso de postprocesado que se aplicaba al HDPG. Esto se debe a que las partes del modelo de vehículos no siguen la misma distribución que en el caso de personas. Aun así, se parte de la misma teoría a la hora de suprimir detecciones repetidas.

El proceso que se implementa consiste en dos procesos de NMS sucesivos:

- El primer NMS se realiza teniendo en cuenta las detecciones únicamente del techo. Para ello se ha considerado como techo la parte superior central, que teniendo en cuenta la pose elegida, será representada por dos partes del modelo (configuración 13). Por tanto se eliminarán las detecciones que estén solapadas mínimamente, adquiriendo como siempre la de mayor puntuación.
- A continuación, a partir de todas las localizaciones restantes se realiza el NMS habitual utilizado en el DTDP. Aunque en este caso se es menos estricto a la hora de aplicar el factor de solapamiento entre detecciones, ya que aumentando este factor eliminaríamos detecciones satisfactorias para nuestro sistema.

De esta forma se obtendrán las detecciones finales para cada uno de los frames. En general, el número de localizaciones será mucho más elevado que en el conseguido por el sistema DTDP.

En la siguiente sección se realizarán pruebas a modo de evaluación sobre diferentes secuencias de vídeo, que verificarán el rendimiento que tiene el algoritmo que proponemos con respecto al DTDP definido en el Estado del Arte.

Capítulo 4

Integración, pruebas y resultados

4.1. Introducción

Una vez concluido el análisis y desarrollo del sistema, se ha tenido que analizar el algoritmo para realizar una evaluación objetiva de los resultados. La finalidad de este apartado será poder tener un punto de vista real con respecto al rendimiento que alcanza el detector HDCG en situaciones dispares. Se describirá de este modo el dataset utilizado, el proceso de evaluación que se ha llevado a cabo, la métrica manejada, así como el análisis de los resultados obtenidos.

4.2. Marco de evaluación

4.2.1. Dataset

Para el proceso de evaluación de detección ha sido necesario obtener un dataset que cumpla con la hipótesis que hemos descrito hasta ahora, es decir, secuencias de vídeos donde se refleje la presencia de oclusiones. En nuestro caso se ha escogido la base de datos denominada PLds (Parking Lot dataset) Se trata de un estacionamiento público con imágenes reales (Pittsburgh International Airport), donde muestra el escenario en diversidad de climas e iluminaciones, ver Figura 4.1..

El dataset consta de dos secuencias obtenidas desde distintos puntos de vista del aparcamiento. Siendo un total de 1000 capturas aleatorias, donde cada uno de los cuadros son imágenes con una resolución de 1280x960 píxeles. Además ofrece la posibilidad de trabajar con videos reducidos (100 frames), que son secuencias obtenidas a partir del



Figura 4.1: Ejemplos de imágenes del dataset [6]

Frames obtenidos de la secuencia QRIDR y VXUSD en la parte posterior e inferior respectivamente.

vídeo largo pero con la característica de que ambas cámaras se encuentran sincronizadas.

Por otro lado, para todas las secuencias de imágenes, la base de datos aporta un archivo con las anotaciones de las localizaciones para cada uno de los vehículos (Great Truth, GT). El conjunto de datos junto con las anotaciones GT están públicamente disponibles en [6].

4.2.2. Métricas

El método de evaluación utilizado básicamente consiste en comparar la similitud entre las detecciones obtenidas en nuestro sistema con respecto a las proporcionadas por el dataset, GT.

Se considerarán tres posibles acontecimientos:

- Verdadero positivo, siempre y cuando una determinada detección coincida con alguna de las representadas en GT. Se asume que dos detecciones coinciden cuando el área de solapamiento sea mayor que el 50%.
- Falso positivo, cuando el detector proporciona una localización que no coincide con ninguna de las correspondientes GT.
- Falso negativo, si HDCG no obtiene resultados cuando sí las hay en GT.

Cabe destacar que para el caso de un mismo objeto con muchas detecciones solo consideraremos un verdadero positivo, todas las demás se excluirán como falsos positivos.

La métrica que se aplica es Curve Precision-Recall (CPR) [5], que valora el ajuste de una determinada detección con respecto al número de detecciones obtenidas para un mismo objeto. Las formas de obtener Precision y Recall vienen definidas por las Ecuaciones 4.1 y 4.2.

$$Precision = \frac{\#Verdaderos\ positivos}{\#Verdaderos\ positivos + \#Falsos\ positivos} \quad (4.1)$$

$$Recall = \frac{\#Verdaderos\ positivos}{\#Verdaderos\ positivos + \#Falsos\ negativos} \quad (4.2)$$

Finalmente también se tendrá en cuenta el área que se encuentra bajo la curva (Area Under Curve, AUC), como medida de comparación entre las gráficas obtenidas

4.3. Resultados

En esta sección se ha marcado como objetivo cuantificar la eficacia del uso de jerarquías utilizado en el sistema HDCG, con respecto a DTDP. Para ello se han comparado resultados obtenidos en ambos procesos de localización automática de vehículos.

Para la obtención de resultados se ha realizará un análisis de los dos puntos de vista proporcionados por el dataset (videos QRIDR y VXUSD). Se analizarán las distintas configuraciones teniendo en cuenta tanto el nivel de pose como el nivel general que abarca todo el modelo (Figura 4.1). Proporcionando distintos valores de overlap, ya que se considera en este trabajo el factor clave, en la etapa de postprocesado, que determinará la supresión de detecciones consideradas satisfactorias en escenarios con múltiples oclusiones.

Todo el proceso de evaluación se ha realizado con las secuencias de vídeo Synchronized (100 frames), debido a la alta complejidad y duración en el proceso de detección de vehículos. Cabe destacar que estas secuencias tienen gran cantidad de escenarios aleatorios, en cuanto a número de coches y condiciones climatológicas, lo que hace que el proceso de valoración sea bastante objetivo.



Figura 4.2: Máscara de detecciones para QRIDR. Se aplican máscaras distintas para cada uno de los vídeos del dataset [6].

Además como escenario de análisis solo se considerará el primer carril de vehículos situados de forma longitudinal al frente de la imagen. El proceso es sencillo, se realiza una máscara dónde se eliminan todas aquellas detecciones que se encuentren fuera del área permitida, ver Figura 4.2.

4.3.1. *Análisis QRIDR Synchronized*

Esta secuencia de vídeo se caracteriza por una posición de la fuente de vídeo más baja. Lo que caracteriza muy bien las oclusiones entre vehículos.

En primer lugar se extrajo un análisis de resultados del sistema DTDP a nivel de poses (Figura 4.3 y Tabla 4.1), para poder tener una referencia en comparativa con nuestros resultados.

	Pose1	Pose2	Genérico
DTDP	0.5941	0.5838	0.6104

Tabla 4.1: AUC sistema DTDP.

En segundo lugar se realizó un proceso de verificación entre ambos sistemas (DTDP vs HDCG), para comparar los resultados alcanzados en HDCG en cada una de las configuraciones propuestas en el diseño de este trabajo. Para ello se ha realizado un barrido con valores de 0.5, 0.6 y 0.7 de overlap, lo que nos permite generalizar el funcionamiento del detector propuesto en la obtención de resultados óptimos.

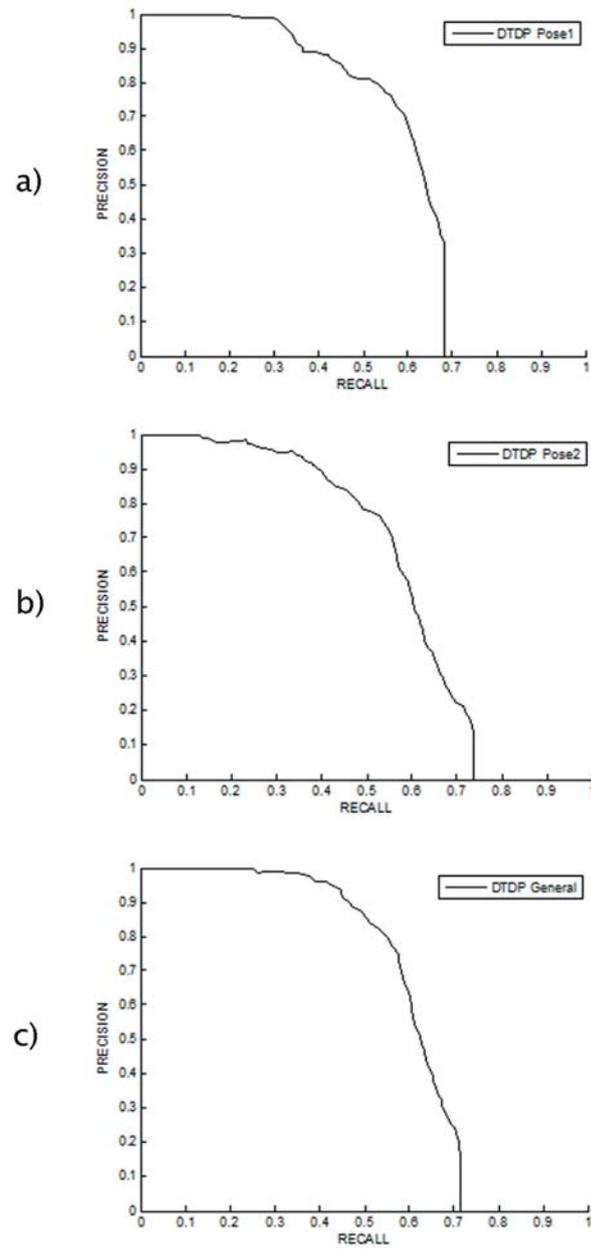


Figura 4.3: Análisis gráfico del sistema DTDP para QRIDR Synchronized. Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

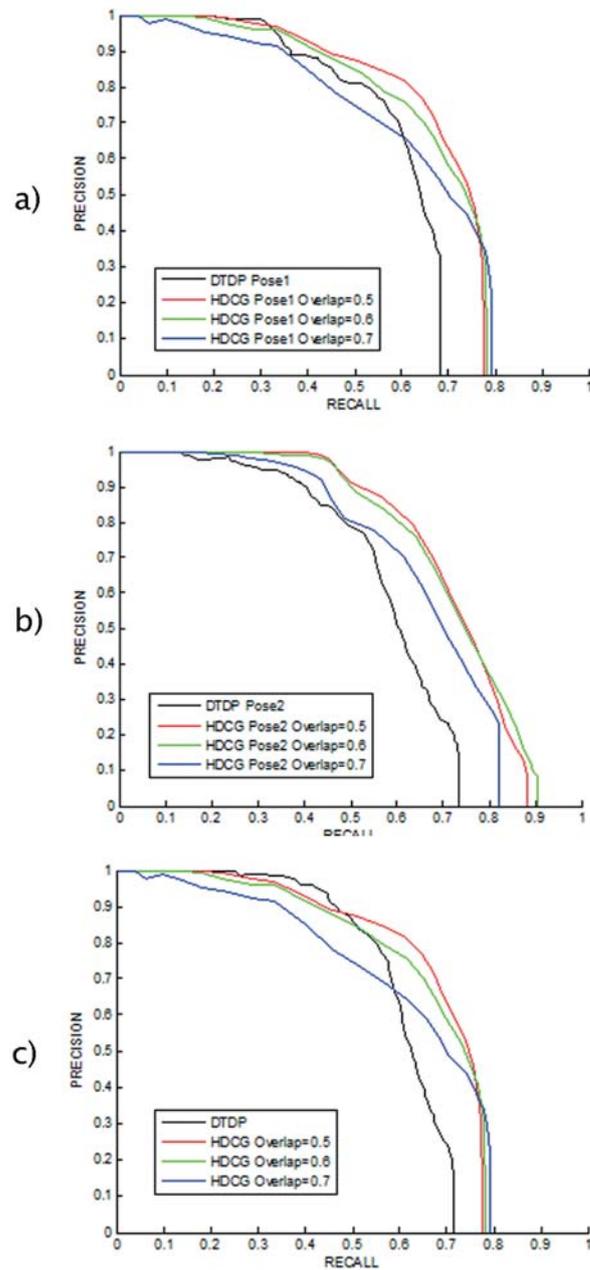


Figura 4.4: DTDP vs HDCG . Configuración 11 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

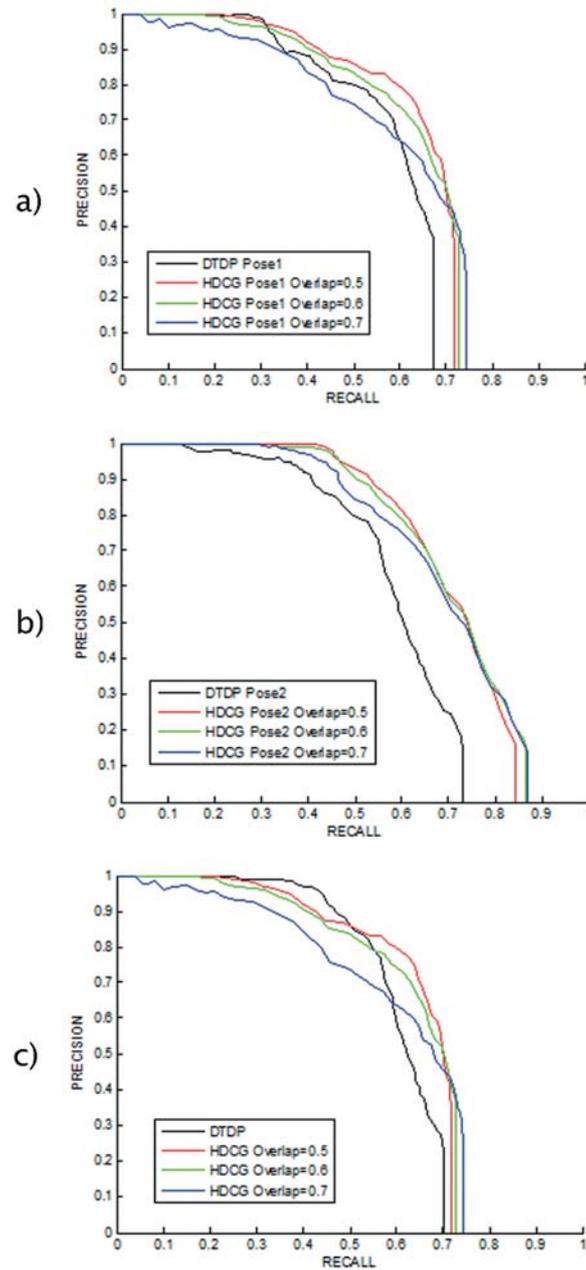


Figura 4.5: DTDP vs HDCG . Configuración 12 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

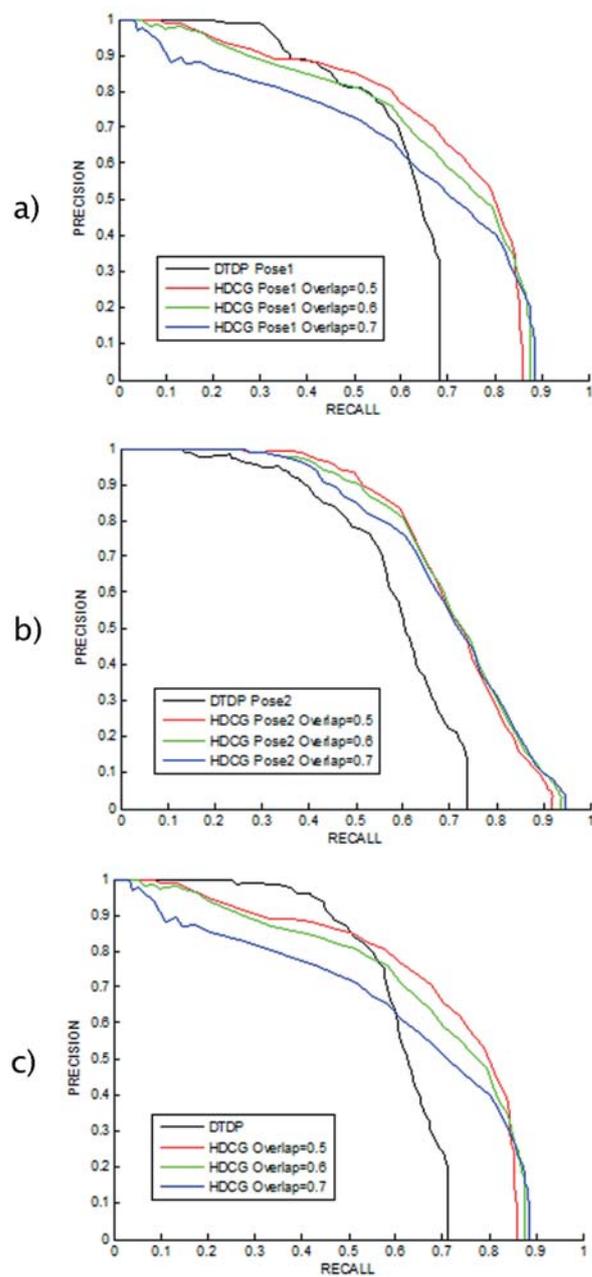


Figura 4.6: DTDP vs HDCG . Configuración 13 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

Los resultados obtenidos en las Figuras 4.4, 4.5 y 4.6, muestran que para las distintas poses y modelo genérico, variando el valor overlap obtenemos resultados similares, es decir, intensificando dicho valor las curvas de las figuras tienden a disminuir en cuanto a Precision y aumentar en Recall. Esto indica que el aumento progresivo del overlap contribuye en el aumento de las localizaciones de vehículos. Pero en consecuencia, favorece la aparición de falsos positivos.

Para tener un punto de vista más objetivo se ha realizado las medidas de AUC de manera que se pueda especificar la configuración que más se aproxime a nuestro escenario de análisis. Ver Tabla 4.2.

NMS	Configuración 11			Configuración 12			Configuración 13		
	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
HDCG Pose1	0.6797	0.6648	0.6236	0.6421	0.6353	0.6015	0.6990	0.6775	0.6197
HDCG Pose2	0.7136	0.7144	0.6641	0.6945	0.6963	0.6824	0.7076	0.7076	0.6961
HDCG Modelo Genérico	0.6807	0.6659	0.6227	0.6475	0.6368	0.6002	0.7009	0.6788	0.6169

Tabla 4.2: AUC sistema HDCP. Configuraciones 11, 12 y 13

Teniendo en cuenta estos resultados se puede determinar que:

- En cuanto a la pose 1 del modelo de vehículos, la configuración 13 para un valor de 0.5 de overlap obtiene los mejores resultado. Coincidiendo de la misma forma, cuando aplicamos el modelo genérico DPM Parking Car completo.
- Para el caso de la pose 2, se obtienen mejores resultados aplicando la configuración 11 con un factor de overlap de un 0.6.

Si tenemos en cuenta todos los análisis obtenidos anteriormente, la pose dos del modelo DPM Parking Car obtiene de las respuestas más favorables, ya que en todo momento se encuentra sobrepasando el umbral establecido por el DTDP y obtiene un valor más elevado del AUC. Lo que conlleva a especificar que dicha pose, junto con el uso de la configuración 11, realiza detecciones jerárquicas más sustanciales, debido principalmente a que en esta configuración se utilizan un mayor número de partes del modelo.

Para concluir con este apartado, se ha elaborado la Tabla 4.3 que evalúa el porcentaje de aumento en cuanto a la eficacia del detector DTDP con respecto al algoritmo

propuesto. Demostrando de forma cuantitativa un incremento del 17.65 % realizando el análisis utilizando la pose 1, de un 22.37 % utilizando la pose 2 y un 14.82 % aplicando el modelo completo DPM Parking Car.

NMS	Configuración 11			Configuración 12			Configuración 13		
	0.5	0.6	0.7	0.5	0.6	0.7	0.5	0.6	0.7
HDCG Pose1	14,40 %	11,9 %	4,96 %	8,07 %	6,93 %	1,24 %	17,65 %	14,03 %	4,30 %
HDCG Pose2	22,23 %	22,37 %	13,75 %	18,96 %	19,27 %	16,88 %	21,20 %	21,20 %	19,23 %
HDCG Modelo Genérico	11,51 %	9,092 %	2,01 %	6,07 %	4,32 %	-1,67 %	14,82 %	11,20 %	1,06 %

Tabla 4.3: DTDP vs HDCG. Porcentajes de eficacia.

4.3.2. *Análisis VXUSD Synchronized*

Para esta secuencia de vídeo juega un factor muy importante el posicionamiento de la fuente de captación a la hora de valorar nuestro sistema, ya que en este caso se encuentra a una altura considerable y no favorece en las oclusiones entre vehículos. Esta característica junto con el modelo DPM Parking Car que describe vehículos visto desde un lateral, ha hecho imposible la obtención de resultados satisfactorios utilizando la hipótesis que se ha seguido hasta el momento. Ver Anexo A.

Esta particularidad, no ha hecho que se deseché nuestro sistema, si no que se han buscado nuevas configuraciones que consigan adaptarse al escenario de análisis. De esta manera queda demostrada la flexibilidad que proporciona el detector HDCG.

Se ha establecido la configuración 14 (Figura 4.7), dónde la parte superior del modelo representa al MC o root y la parte inferior del modelo representa al SC. Para esta situación se tiene en cuenta el factor contrario a la hipótesis que se estableció en un principio, es decir, permite la localización de vehículos con poca oclusión entre ellos.

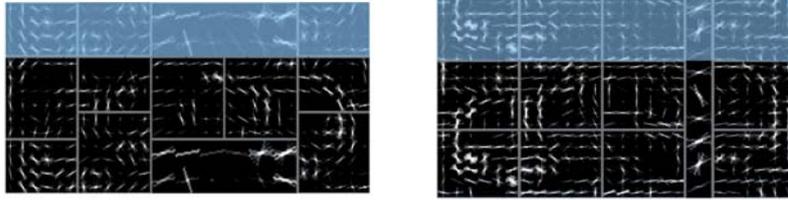


Figura 4.7: Configuración 14 para el sistema HDCG.

Formulando esta nueva hipótesis obtenemos la evaluación mostrada en la Figura 4.8, que al igual que en el apartado anterior, se ha realizado un barrido entre valores de overlap de 0.5, 0.6 y 0.7 como factor principal del NMS en la etapa de postprocesado.

Observando los resultados, se puede apreciar que ocurre la contraposición del análisis obtenido en las configuraciones 11, 12 y 13, ya que el aumento del overlap permite un mayor solapamiento, es decir, en este caso permitimos una oclusión entre coches pequeña y no grande, lo que conlleva la pérdida del factor Precision en el análisis CPR. Con lo cual, conforme disminuimos este valor la curva tiende a sobrepasar el umbral del sistema DTDP, generando de este modo mejores localizaciones y disminuyendo los falsos positivos. Si realizamos un análisis del área que arroja bajo la curva junto con los porcentajes de eficacia, ver Table 4.4, se puede concluir que se obtiene mediante esta configuración un aumento de la eficacia de hasta un 16.84 % con respecto DTDP.

	DTDP	Configuración 14			Evaluación		
NMS	-	0.5	0.6	0.7	0.5	0.6	0.7
Pose1	0.5983	0.6387	0.5930	0.5615	6.75 %	-0.88 %	-0.15 %
Pose2	0.6296	0.6982	0.6979	0.7333	10.89 %	10.84 %	16.47 %
Modelo Genérico	0.6693	0.6679	0.7186	0.6612	-0.2 %	7.36 %	-1.21 %

Table 4.4: AUC y porcentajes de eficacia para el sistema HDCP con respecto DTDP. Configuración 14

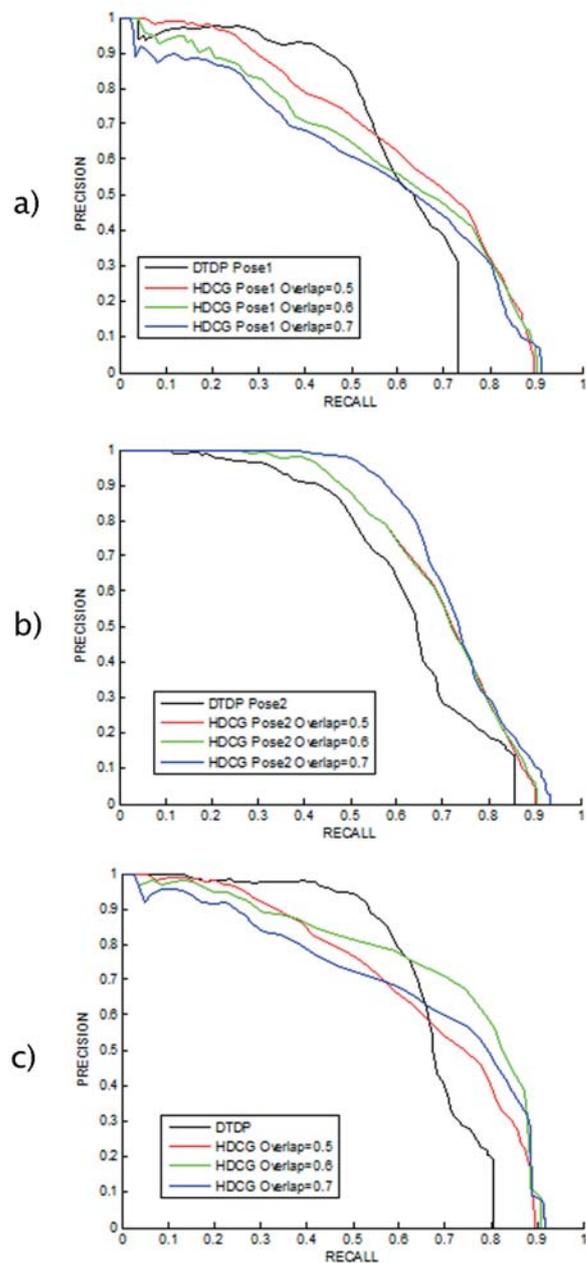


Figura 4.8: DTDP vs HDCG . Configuración 14 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

Cabe destacar que la pose 2 utilizada durante todo el proceso de detección, obiene los mejores resultados tanto para la hipótesis de oclusión como para la de no oclusión. Esto es debido a que describe vehículos desde un punto de vista caracterizado por una cierta verticalidad, lo que indica que se adapta a la perfección en nuestro escenario de análisis.

Capítulo 5

Conclusiones y trabajo futuro

5.1. Conclusiones

Desde que se ha comenzado el trabajo, se ha referenciado el auge y expansión que se ha producido con las tecnologías de vídeo. El análisis a nivel de detección se ha convertido en un factor clave, pero la gran cantidad de información de la que se dispone y las distintas maneras de abarcarlo hace muy difícil encontrar un sistema que se aproxime en todos los aspectos. Por esto fijamos el objetivo en la creación de un sistema basado en la búsqueda jerárquica de vehículos, donde se presentaran grandes aglomeraciones y oclusiones entre ellos.

Después de haber implementado el diseño y desarrollo del algoritmo, previamente realizado el estudio exhaustivo del estado del arte, si hacemos una evaluación a la vista de los resultados, se pueden concluir los siguientes puntos:

- El modelo DPM utilizado a la hora de describir el objeto a localizar es un factor clave, ya que su correcta descripción determinará el aumento o disminución de las posibles detecciones.
- La funcionalidad jerárquica que ofrece HDCG, es decir, el uso de configuraciones que se adecuen al proceso de evaluación, así como el área de búsqueda o anchors-shift utilizados, determinarán la correcta obtención de resultados.
- Por último a nivel de postprocesado, la elección de un overlap idóneo, que permita la disminución en cuanto a presencia de falsos positivos, pero que su permita la

detección de vehículos ocluidos, será la característica principal a la hora de obtener una respuesta eficaz satisfactoria.

Como conclusión final, mencionar que el sistema propuesto durante todo el trabajo ha cumplido el objetivo principal de mejorar el rendimiento con respecto al detector DTDP. La utilización de jerarquías como proceso de localización será la clave en cuanto al procesado automático de señales de vídeo.

5.2. Trabajo futuro

Una vez realizado todo el proceso de investigación, desarrollo y evaluación de resultados, son notables las múltiples aplicaciones de cara al futuro tanto para conseguir mejores resultados o aproximaciones como para encontrar nuevas utilidades a estos sistemas.

A continuación se detallan algunas de las posibilidades que se podrían tener en cuenta a la hora de afrontar nuevos retos en este área de investigación.

- En primer lugar, se podría analizar el algoritmo teniendo en cuenta nuevos métodos de descripción del modelo, es decir, aplicando modelos diferentes al utilizado en este trabajo (DPM Parking Car).
- Se podría ajustar el algoritmo teniendo en cuenta coches de alrededor, de tal forma que se estudiara la jerarquía de grupos utilizando la información de más de un coche, es decir, a nivel de parejas, tríos, etc.
- Estudiar la relación que existe en el aumento o disminución de los `anchor_shift`, permitiendo así variar el área de búsqueda.
- Estudiar la relación que existe entre el uso o no del `root` en el modelo DPM, ya que puede ocurrir que su utilización influya en las detecciones para vehículos fuertemente ocluidos.
- Otra vía de estudio sería comparar los resultados obtenidos con otros algoritmos desarrollados, que tengan en cuenta secuencias de vídeos con solapamiento de vehículos.
- La realización de un proceso de seguimiento o tracking que permitiera despreocuparse más falsos positivos con respecto al proceso de postprocesado utilizado en nuestro sistema.

- Sería interesante estudiar el algoritmo con otras secuencias de videos, es decir, analizar el sistema para distintos dataset. De tal forma que permitiese ver el comportamiento del detector y analizar las ventajas y desventajas con respecto al entorno.
- Por último lugar, sería conveniente un examen en profundidad a nivel de código, que permitiese disminuir en gran medida el tiempo de ejecución ya que ha sido un gran factor que ha jugado en nuestra contra a la hora de analizar el algoritmo.

Bibliografía

- [1] R. H. Evangelio A. Garcia-Martin and T. Sikora. Multi-configurations for part-based person detectors. 2014.
- [2] Jose M. Martinez Alvaro Garcia-Martin, Ricardo Sanchez-Matilla. Hierarchical detection of persons in groups. 2017.
- [3] T. Malisiewicz C. Vondrick, A. Khosla and A. Torralba. Hoggles: Visualizing object detection features. pages 1–8, 2013.
- [4] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. pages 886–893, 2005.
- [5] J. Davis and M. Goadrich. The relationship between precision-recall and roc curves. pages 232–241, 2006.
- [6] <http://www.vpu.eps.uam.es/DS/PLds/>.
- [7] <http://es.mathworks.com/> Matlab.
- [8] D. McAllester P. F. Felzenszwalb, R. B. Girshick and D. Ramanan. Object detection with discriminatively trained part-based models. 32:1626–1646, 2005.
- [9] Alexander G. Hauptmann Jose M. Martinez Rafael Martin-Nieto, Alvaro Garcia-Martin. Automatic vacant parking places managements system using multicameta vehicle detection. pages 3–4, 2018.

Apéndice A

Gráficas XUSD Synchronized

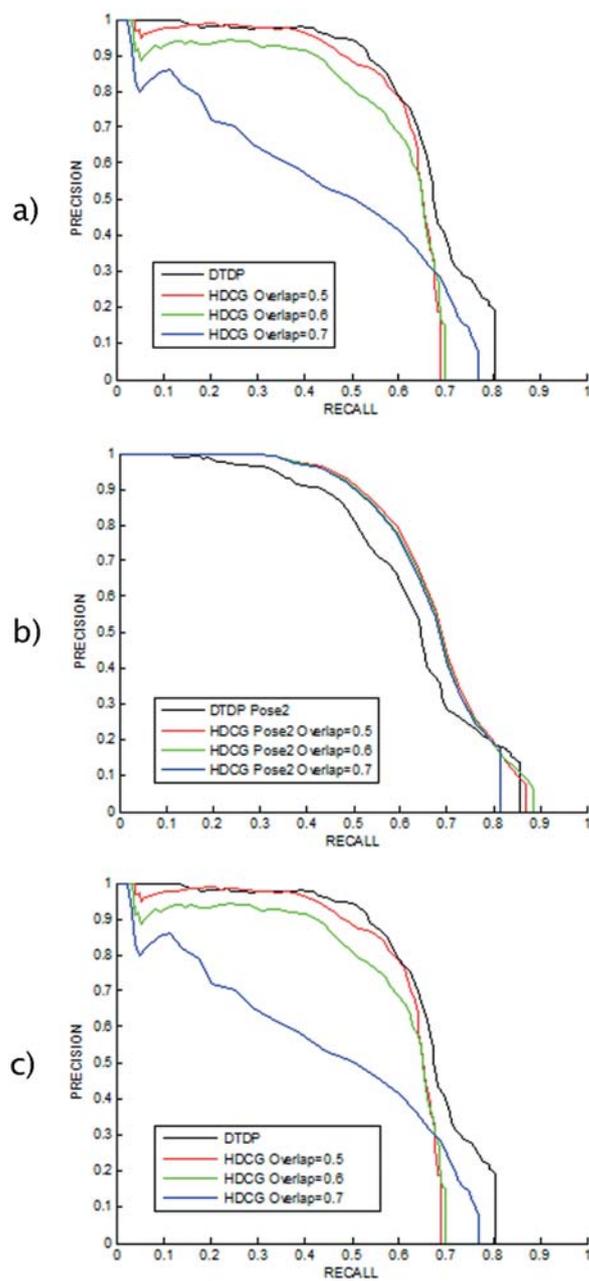


Figura A.1: DTDP vs HDCG . Configuración 11 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

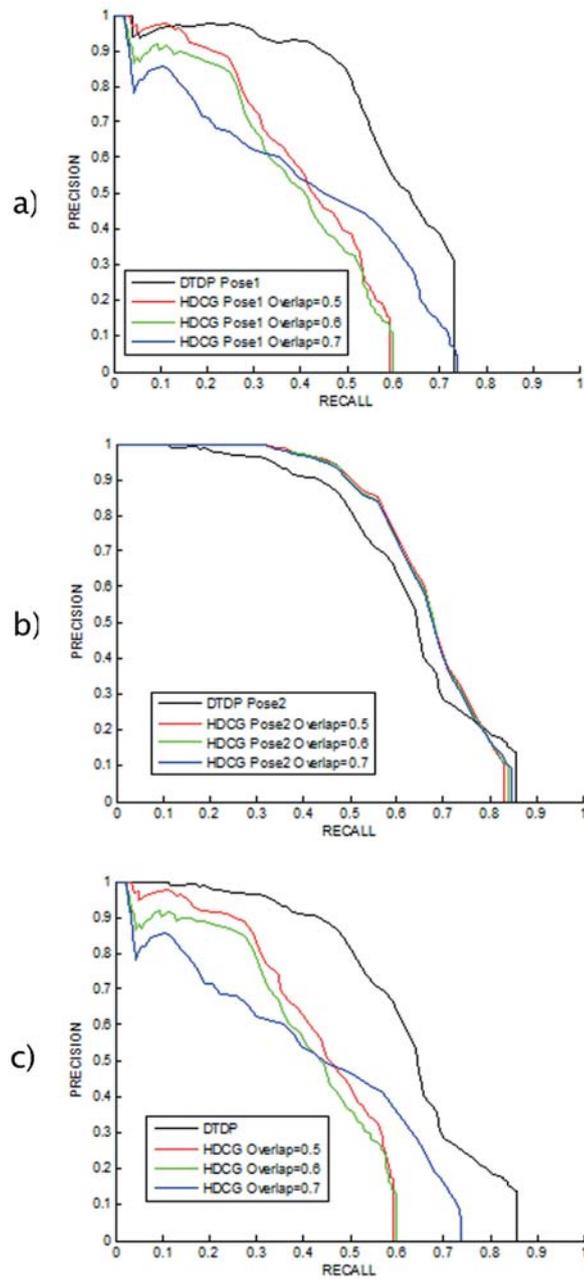


Figura A.2: DTDP vs HDCG . Configuración 12 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

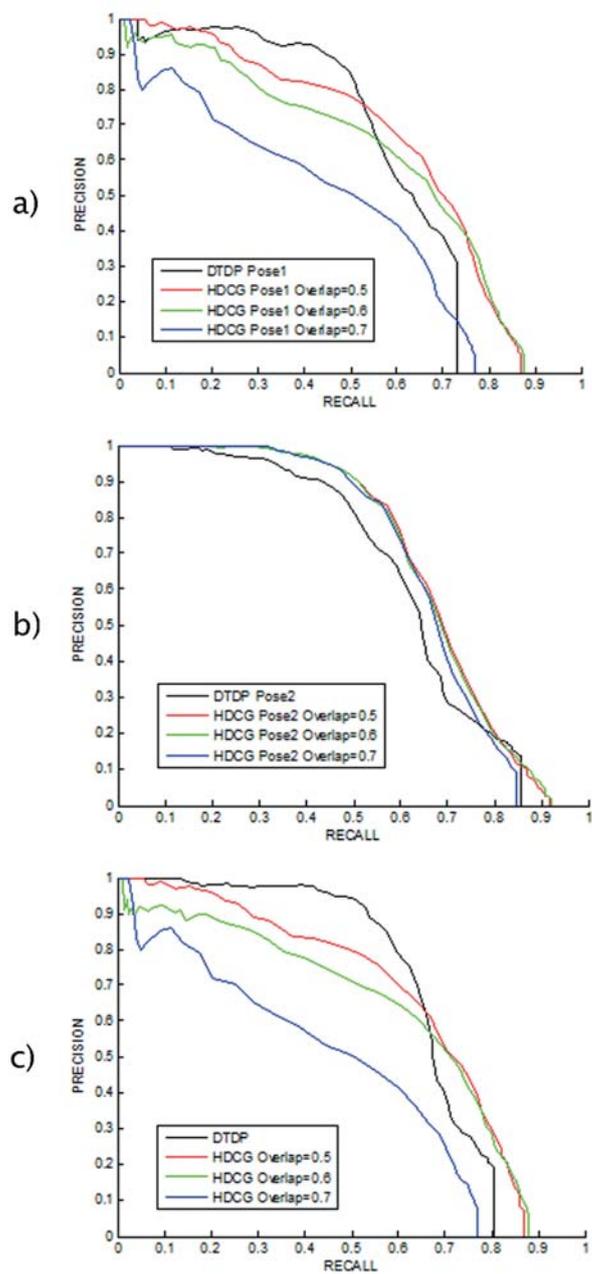


Figura A.3: DTDP vs HDCG . Configuración 13 para diferentes porcentajes de permisividad de NMS.

Se muestra a) Pose 1 del modelo; b) Pose 2 del modelo; c) Modelo genérico DPM Parking Car. Ambas poses.

Apéndice B

Ejemplos visuales

