

Visión por Computador (2018-2019)
Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

Descriptores HOG en la detección de
peatones



Ignacio Aguilera Martos y Diego Asterio de Zaballa
14 de enero de 2019

Índice

1. Descripción del problema y enfoque de la resolución	3
1.1. Problema a resolver	3
1.2. Descriptores HOG	3
1.3. Fases de la resolución	4
2. Valoración de los resultados	4
3. Trabajo futuro: propuestas de mejora	4

1. Descripción del problema y enfoque de la resolución

En la resolución de este trabajo nos hemos planteado el siguiente problema a resolver, dada una imagen de una persona andando por la calle (un peatón), ¿cómo podemos reconocer que es un peatón?

Esta misma cuestión fue planteada por Navneet Dalal y Bill Triggs en su paper “Histogram of Oriented Gradients for Human Detection” en el que explican el desarrollo de unos descriptores que aplicados a su dataset de personas obtienen unos resultados muy buenos en la detección de las mismas.

Estos descriptores son los descriptores HOG o descriptores de histogramas basados en gradientes. Como veremos a lo largo del trabajo se han empleado distintas variaciones a la hora de hallar los descriptores por Dalal y Triggs, quedándonos nosotros con las elecciones que han resultado más fructíferas para ellos en su análisis.

Para comenzar hay que saber que la detección de personas es un problema difícil de abordar y que actualmente resulta muy interesante en aplicaciones por ejemplo en coches, de forma que si detecta un peatón andando por delante del vehículo este entienda que tiene que detenerse.

Las herramientas usadas en este proyecto han sido:

- Python para la implementación.
- OpenCV y NumPy para las operaciones de los algoritmos.
- El módulo de SVM incluido en OpenCV para poder predecir la existencia o no de un humano en una imagen en base a los descriptores calculados.
- El dataset dado por los investigadores empleado en la elaboración de su artículo.

A continuación hacemos una descripción un poco más elaborada del problema propuesto.

1.1. Problema a resolver

El problema consiste en, dada una imagen que contiene o no un peatón, debemos determinar si es que lo contiene y además una región aproximada en la que se encuentra el mismo.

Para la elaboración de la solución del problema hemos utilizado varios ingredientes entre los que tenemos la fase de creación de descriptores y entrenamiento de la SVM y la fase de test.

En la fase de entrenamiento de los descriptores obtenemos un dataset ya formado por parte del grupo de investigación con parches de personas y con parches que no son personas todos con el mismo tamaño. De esta forma podemos obtener descriptores comparables entre sí para las imágenes positivas y negativas con lo que entrenaremos una SVM.

La segunda fase es la de test que comparte dificultad e importancia con la primera. Como explicaremos más en detalle posteriormente resolvemos el problema de dada

una imagen, ¿cómo la tenemos que analizar para poder identificar si hay o no personas y dónde están? En secciones siguientes veremos que la solución adoptada ha sido una ventana deslizante con algunas modificaciones importantes cuyos parches extraídos se pasan a la SVM y se predice sobre ellos para poder detectar el área de la imagen en la que obtenemos respuestas positivas a la pregunta de si hay o no una persona.

1.2. Descriptores HOG

El problema mencionado anteriormente se ha enfrentado desde distintos puntos de vista lo que ha dado lugar a diversos descriptores. Nuestra implementación se basa en el paper de Navneet Dalal y Bill Triggs en el que desarrollan un descriptor basado en el gradiente de una imagen llamado descriptor HOG. El método se basa en calcular los gradientes de una imagen y a continuación construir histogramas locales basado en la orientación de dichos gradientes.

La idea básica sobre la que se desarrolla la técnica es que las características de forma y apariencia de un objeto de forma local se pueden resumir gracias a la distribución de la orientación de los gradientes de la imagen.

A continuación se describe la implementación de la técnica a grandes rasgos. En primer lugar la ventana de la imagen sobre la que se va a construir el vector de características se divide en celdas. Estas celdas son pequeñas regiones de la imagen. En cada una de estas celdas se calcula el histograma de la dirección de los gradientes que en ella se encuentran. Esto define la re A continuación estas celdas se agrupan en bloques para que sea posible normalizarlas.

El descriptor HOG es un vector de características La construcción del descriptor HOG se apoya fundamentalmente en el gradiente de la imagen.

1.3. Fases de la resolución

2. Valoración de los resultados

3. Trabajo futuro: propuestas de mejora