

Detección vehicular mediante técnicas de visión de máquina

Robinson Jiménez Moreno*

Oscar Fernando Avilés S.**/ Dario Amaya Hurtado*** Camilo Gordillo C**** / Fabio Espinosa V.*****
robinson.jimenez@unimilitar.edu.co, oscar.aviles@unimilitar.edu.co, dario.amaya@unimilitar.edu.co,
U1801074@unimilitar.edu.co,

Resumen:

El presente artículo esboza los resultados del diseño de un detector de objetos mediante clasificadores Haar, los cuales operan en función a descriptores rectangulares relacionados con la intensidad de una región en una imagen. Este clasificador se entrena para la detección de automóviles, con el objetivo de establecer la cantidad de flujo vehicular en una vía, soportados en la información proveniente de cámaras de videovigilancia. Se realiza el entrenamiento del clasificador obteniendo un porcentaje de aciertos en la detección del 92.9%, y se comparan los resultados frente a técnicas de visión de máquina como lo es el flujo óptico, presentando un desempeño superior en más del 30%. Los tiempos de procesamiento obtenidos son en promedio de 40 milisegundos.

Palabras clave: flujo vehicular, clasificador haar, visión de máquina.

Abstract

This paper outlines the results of design of a Haar classifier, which operate according to rectangular descriptors related to the intensity of an image region, for the detection of cars in order to establish the amount of vehicular traffic on a road, supported on the information from video surveillance cameras. The training of the classifier takes place obtaining a percentage of correct detection of 92.9%, and compared the results against machine vision techniques such as optical flow, showing superior performance in more than 30%. Processing times obtained are average of 40 milliseconds.

Keywords: traffic flow, haar classifier, machine vision.



¹Recibido: 29 de febrero de 2012

Aprobado: 17 de abril de 2012

* Ingeniero Electrónico de la Universidad Distrital Frco. José de Caldas; Magister en automatización Industrial de la Universidad Nacional de Colombia. Dirección de Correo: robinson.jimenez@unimilitar.edu.co. Teléfonos: 6942301-3003240012

** **Co-autor:** Ingeniero Electrónico de Universidad Antonio Nariño; Magister en Sistemas Automáticos de Producción y PhD en Ingeniería Mecánica. Dirección de Correo: oscar.aviles@unimilitar.edu.co;

*** **Co-autor:** , Ingeniero Electrónico de Universidad Antonio Nariño; Magister en Teleinformática y PhD en Ingeniería Mecánica. Dirección de Correo: dario.amaya@unimilitar.edu.co;

**** **Co-autor:** , Ingeniero Mecatrónico de Universidad Militar. Dirección de Correo: U1801074@unimilitar.edu.co;

***** **Co-autor:** Ingeniero Mecatrónico de Universidad Militar. Dirección de Correo: U1801067@unimilitar.edu.co;

1. Introducción

Los problemas de movilidad que afrontan ciudades capitales como Bogotá, generan incrementos en los tiempos de desplazamiento, que afectan tanto la productividad como la competitividad de estas [1]. Una solución a esta problemática se encuentra en la implementación de sistemas inteligentes y autónomos que puedan gestionar el flujo vehicular [2]. A fin de lograr el control del tráfico en una ciudad se debe identificar el nivel de flujo vehicular en cada una de las vías, para ello de forma comercial se encuentran sensores del tipo capacitivo e inductivo [3], que requieren remoción de la capa asfáltica para su instalación y los cuales se ven sometidos a altas cargas mecánicas y de temperatura al encontrarse en operación, lo que reduce su vida útil.

Por otro lado, las técnicas de visión de máquina cobran cada vez más fuerza en la automatización de sistemas. Por medio de una cámara, con una protección adecuada, se puede tener un dispositivo de detección vehicular no intrusivo, de fácil instalación y mantenimiento. Algunos trabajos previos en el área han aplicado técnicas de procesamiento de imagen para la detección de flujo vehicular, en [4] se presenta una técnica mediante operaciones morfológicas que arroja una precisión de 85% en la detección del flujo vehicular.

Los clasificadores tipo Haar se han venido utilizando cada vez más para la detección de objetos de interés, en [5] se presenta la técnica principal para implementar un algoritmo de detección de objetos, la cual es aplicada en este caso para la detección de rostros, en [6] es utilizada esta técnica a fin de segmentar los ojos y la boca, de un conductor frente al volante para analizar características de fatiga. De forma tal que es posible entrenar un clasificador de este tipo para diferentes fines y en el caso de interés, identificar vehículos en el escenario de una vía arterial.

Este artículo presenta en la Sección 2 las características fundamentales del clasificador, en la Sección 3 el diseño del mismo, en la Sección 4 los resultados obtenidos y finalmente en la sección 5 las conclusiones obtenidas.

Características del clasificador.

Mediante técnicas de procesamiento de imagen, la representación de un objeto se puede clasificar en: representación global, representación basada en partes y representación por plantillas. Si se usa una representación global, los objetos estarán representados no por datos crudos de una imagen sino por características (features), que se obtienen de aplicar preprocesamiento a está o a secuencias de imágenes. Donde una característica denota una pieza de información que es relevante para resolver tareas. En [5] Viola-Jones presentan el primer detector basado en características que proporcionó tasas de detección competitivas en tiempo real.

Esté sistema utiliza grupos de características Haar para llevar a cabo la detección, las cuales están basadas en el cambio de intensidad en la imagen. El valor de una de estas características se obtiene de la diferencia entre la suma de las intensidades de los pixeles dentro de dos o más regiones rectangulares, tal y como se muestra en la Figura 1.

De forma tal que para la obtención de los valores de estas características sobre una imagen, se toma cada una de estas (verticales, horizontales, oblicuas, de 2 o más rectángulos) y se recorre toda la imagen, aumentando y disminuyendo el tamaño de los rectángulos, como se aprecia en la Figura 2.

Como se puede inferir, el número de las sumas de intensidades de los pixeles dentro de cada rectángulo dependerá del tamaño de estos, lo cual



Figura 1. Características Haar básicas.



Figura 2. Desplazamiento de las características Haar en la imagen.

generara una tasa de procesamiento inconstante y posiblemente baja [7, 8]. Para solucionar este inconveniente, en [8] se propone usar una forma de representación llamada imagen integral denotada como ii , la cual se relaciona con la imagen original i , en cada punto (x', y') estará mediante la ecuación 1.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad 1$$

De esta forma se puede logra que el tamaño de los rectángulos (Haar) no influya en el número de sumas, por lo tanto se puede mantener una tasa de procesamiento constante [8, 9].

Ahora, sí se sabe que se tiene una imagen de 24x24 pixeles, el número de posibles características que se pueden ubicar dentro de esta imagen fácilmente supera las 180.000. Aunque el cálculo de cada característica es realmente eficiente, construir el clasificador con todas estas características, sería muy costoso computacionalmente. En [8], se ha encontrado que no es necesario el conjunto completo de características para poder representar una clase de objetos eficazmente.

Para encontrar el nuevo subconjunto de características altamente descriptivas de la clase de objeto que se quiere identificar, se hace uso de un algoritmo de aprendizaje de maquina iterativo llamado Boosting (Adaboost), que pretende construir un clasificador fuerte a partir de varios clasificadores débiles, como el observado en la Figura 3.

La ecuación 2 describe un clasificador débil, donde x es una imagen de resolución base, f es el valor de la característica j aplicada a la imagen, p es la polaridad de la característica (orientación) y θ es el umbral que decide si la característica eva-

luada es representativa sobre la imagen en uso. Donde al hablar de clasificador débil se habla también de característica, por lo que se usaran ambos términos indistintamente [9].

$$h_j(x) = \begin{cases} 1 & p \cdot f_j(x) > p \cdot \theta_j \\ -1 & \text{en otro caso} \end{cases} \quad 2$$

Una vez obtenido el clasificador fuerte se podría iniciar el proceso de detección a voluntad del usuario. Este proceso a grandes rasgos consta de los siguientes pasos:

- Se toma la imagen donde se quiere realizar la detección y se crea una nueva imagen convirtiendo está a escala de grises.
- Se calcula la representación en imagen integral de esta nueva imagen.
- Se crea una ventana de detección con un tamaño inicial igual al de los tamaños de los ejemplos de entrenamiento positivos.
- Se recorre la imagen integral con la ventana de detección, calculando las características que se puedan encontrar sobre esta ventana.
- Se compararan las características obtenidas, con las que están presentes dentro del clasificador fuerte y se realiza la clasificación sobre si la ventana contiene o no el objeto de interés.
- Se continúa moviendo la ventana de detección hasta que haya recorrido la imagen completamente.
- Se escala la ventana de detección en un factor determinado y se inicia nuevamente el recorrido de esta por sobre la imagen. Esta es otra razón por la que este sistema es más rápido que los demás sistemas de detección, ya que en estos se escala la imagen lo cual es menos eficiente y requiere de más procesos a realizar.

Figura 3. Clasificador fuerte compuesto de clasificadores débiles.



Figura 4. Estructura en cascada: cada imagen es un clasificador fuerte, aumento en complejidad de izquierda a derecha

De forma tal que la estructura final del clasificador genera una cascada de mayor complejidad en las últimas etapas como lo muestra la Figura 4.

Diseño del clasificador

Para el diseño del clasificador se deben considerar ciertas características fundamentales que darán por resultado los parámetros óptimos con que ha de operar, éstas son: conjunto de características del clasificador débil, tipo de clasificador a utilizar, tamaño del patrón de entrada, relación entre conjunto de entrenamiento y si se utilizara simetría vertical. El calculo de cada uno de estos parámetros se realiza en las siguiente subsecciones.

Características del clasificador.

En el diseño del clasificador se pueden considerar el conjunto de características Haar básicas o el conjunto extendido propuesto en [10], para la realización de las pruebas de caracterización se utiliza un conjunto de 6364 imágenes las cuales contienen 10050 ejemplos positivos, es decir, con vehículos previamente identificados manualmente. Al probar el clasificador débil con ambos conjuntos de características se obtienen los resultados de la tabla 1, en la cual se aprecia una tasa de aciertos similar, pero la presencia de falsos positivos representa casi 71% de reducción para el conjunto extendido, mostrando así mejor desempeño.

Tabla 1. Comparación del desempeño de características Haar

Conjunto de características	Tasa de aciertos	Falsos positivos	Falsos negativos
Básico	0.843	0.350	0.157
Extendido	0.834	0.244	0.165

Tipo de Clasificador

Existen tres tipos de clasificadores débiles característicos para el entrenamiento, los cuales en esencia varían el número de características: con una característica (Stump), con dos característi-

cas (CART Classification And Regression Tree) y el último con cuatro características. En la tabla 2, se observa que el clasificador CART2 presenta una reducción en la presencia de falsos positivos de un 45.7% respecto al clasificador Stump y tan solo una reducción en la tasa de aciertos del 11.64% respecto al mismo. Dando como resultado un mejor desempeño del clasificador con dos características.

Tabla 2: Comparación de desempeño entre clasificadores débiles.

Clasificador	Tasa de aciertos	Falsos positivos	Falsos negativos
Stump	0.834	0.244	0.165
CART2	0.737	0.132	0.262
CART4	0.696	0.122	0.303

Patrón de Entrada

El tamaño de la ventana a utilizar es fundamental en los tiempos de procesamiento para establecer una diferencia máxima entre los tamaños de los vehículos que se encuentran y que se han especificado en la imagen, con los encontrados por el detector. En la tabla 3 se evidencia que el mejor desempeño se presenta en tamaños de 18x18 y 20x20 pixeles, ya que presentan menor porcentaje de incidencia de falsos positivos y mayor porcentaje de vehículos detectados.

Tabla 3. Efectos de la variación del tamaño en el patrón de entrada.

Tamaño	Tasa de aciertos	Falsos positivos	Falsos negativos
18X18	0.834	0.244	0.165
20X20	0.744	0.119	0.255
24X24	0.713	0.432	0.286
20X18	0.745	0.141	0.254

Relación Entre Conjunto de Entrenamiento

En el entrenamiento hay que involucrar imágenes que no solo corresponden a casos positivos, es decir, vehículos en la imagen que son el objeto de interés, sino también de objetos que se pueden

presentar en la imagen y no se deben clasificar (casos negativos). En la tabla 4 se puede observar cómo influye el tamaño de los conjuntos de ejemplos positivos y negativos sobre el desempeño de la detección, además de los efectos que conlleva tener más ejemplos de uno o de otro conjunto. Se observa que a mayor número de ejemplos positivos el clasificador tiene mayor probabilidad de detectar un número mayor de automóviles, pero también una mayor incidencia de falsos positivos en las detecciones; sucede lo contrario si se aumenta el número de ejemplos negativos, ya que el número de automóviles detectados es menor, pero el número de falsas detecciones también se reduce.

Tabla 4. Desempeño frente a la relación imágenes positivas/negativas.

Tamaño (+/-)	Tasa de aciertos	Falsos positivos	Falsos negativos
1:1	0.834	0.244	0.165
1:2	0.776	0.175	0.223
2:1	0.810	0.221	0.189

Simetría vertical

De igual forma, la tabla 5 presenta la manera en que influye o no la presencia de simetría vertical, es decir, que el objeto poseerá las mismas características pero reflejadas con respecto a la división central vertical sobre este. En los ejemplos positivos de vehículos usados en el entrenamiento sobre el proceso de detección se esperaba que un entrenamiento sin la suposición de simetría, otorgara mayor robustez al detector frente a variaciones de rotación del vehículo, esta hipótesis no fue valida dado que la tasa de aciertos supera en el caso de sin simetría en un 6.3% al caso con simetría.

Tabla 5. Desempeño con y sin simetría vertical.

Simetría vertical	Tasa de aciertos	Falsos positivos	Falsos negativos
con	0.771	0.220	0.228
sin	0.834	0.244	0.165



Figura 5. Rotaciones del objeto de interés.

La Figura 5 muestra las condiciones de rotación típicas en el conjunto de imágenes positivas, que dieron lugar a considerar la posibilidad de simetría vertical.

Resultados obtenidos

La implementación del algoritmo fue realizada en un equipo de computo de 2GB de memoria RAM, con procesador Intel Core Duo de 1.6 GHz, programación en C bajo Visual Studio 2010 Edición Express y utilizando la librería de visión de máquina OpenCV.

El clasificador final quedó parametrizado como se observa en la tabla 6. Al evaluarlo sobre una base de datos de 152 imágenes en la que se encuentran 634 vehículos, presento una tasa de acierto del 92.9%. Finalmente, la figura 6 permite observar el resultado de detección del clasificador final, una característica importante de este es que se puede evaluar con un solo sensor varios carriles, determinando todo el flujo vehicular en una vía con solo una cámara. De igual se puede forma resalta el tiempo de detección final por imagen, el cual se encuentra en promedio a los 40 mseg, permitiendo trabajar secuencias de video de hasta 27 cuadros por segundo.

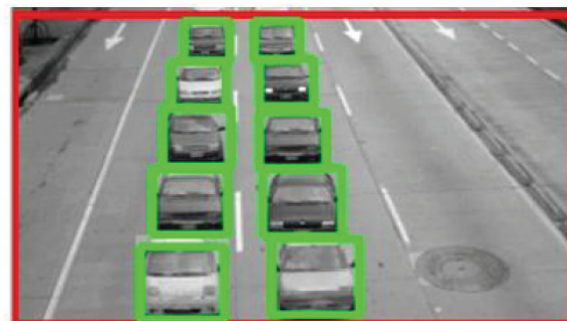


Figura 6. Resultado de detección del clasificador.

Tabla 6. Características del clasificador final.

Parámetro de entrenamiento	Valor óptimo
Conjunto de Características	Extendido
Tipo de clasificador débil	CART2
Tamaño patrón entrada	18X18 & 20X20
Relación conjunto de entrenamiento	1:2
Simetría vertical	Si

Para validar el clasificador obtenido frente a otras técnicas de visión de máquina, se implementó un algoritmo de flujo óptico para detección de vehículos. Al probar cada algoritmo, la base de datos del clasificador ofreció un mejor desempeño que el algoritmo de flujo óptico, como se aprecia en la tabla 7. Esto se debe a que el algoritmo de flujo óptico no se comportan bien bajo ciertas condiciones como lluvia debido a su incapacidad de ser selectivos, es decir, reconoce los reflejos como un elemento más, como se muestra en la figura 7a. Por otro lado, como se muestra en la figura 7b, este tipo de algoritmos no es capaz de reconocer movimientos independientes cuando estos se encuentran muy cercanos entre sí. A diferencia del clasificador que bajo estas mismas

Tabla 7. Desempeño algoritmos de detección.

	Aciertos	Falsos negativos	Cantidad de vehículos
Clasificador (Viola - Jones)	589	45	634
	92.9%	7%	
Flujo óptico	392	242	
	61.8%	38.1%	



Figura 7. Desempeño Flujo óptico



Figura 8. Desempeño del clasificador

condiciones se muestra robusto, según se aprecia en la figura 8.

Conclusiones

Se encontraron los parámetros óptimos de diseño del clasificador, los cuales permiten asegurar un desempeño de aciertos del 92.9%, cuando las imágenes presentan una luminosidad suficiente como las características de la luz de día.

Los tiempos de procesamientos obtenidos mediante esta técnica de detección de objetos, los cuales se establecen cerca a los 40 mseg, permiten su implementación en sistemas de tiempo real, lo cual hace viable su desarrollo en hardware autónomo.

La comparación del clasificador frente a técnicas como flujo óptico arroja una mejora de más del 30%, reduciendo errores por efectos de ruido como el provocado por la lluvia, destacando así la robustez lograda bajo condiciones de operación normal durante el día.

Referencias

[1] D. Robles, P. Nãñez, N. Quijano, Control y Simulación de Tráfico Urbano en Colombia: Estado del Arte. Revista de Ingeniería No. 29, Universidad

de los Andes. Bogotá, Colombia. ISSN: 0121-4993. 2009.

[2] M. Wiering, J. Van Veenen, J. Vreeken, A. Koopman, Intelligent Traffic Light Control. Institute of Information and Computing Sciences. Utrecht University. 2004.

[3] Consulta enero de 2012. Disponible en: <http://www.quadrex.es/fichaC.php?c=18&area=A>.

[4] A. Ajmal, I. Hussain, Vehicle Detection Using Morphological Image Processing Technique.

[5] P. Viola and M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features, IEEE. 2001.

[6] R. Jiménez, F. Prieto y V. Grisales, Detection of the Tiredness Level of Drivers Using Machine Vision Techniques. CERMA IEEE. pp 97-102. 2011.

[7] R. Krishna Radha, Speeding up Adaboost Object Detection with Motion Segmentation and Haar Feature Acceleration.

[8] P. Viola, M. Jones, Robust Real-time Object Detection.

[9] O. H. Jensen, Implementing the Viola-Jones Face Detection Algorithm.

[10] L. Rainer, A. Kuranov, P. Vadim, Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection.