

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Histograma del gradiente con múltiples orientaciones (hog-mo) detección de personas

Histogram of the Gradient with Multiple Orientations (Hog-Mo) People Detection

Marco Javier Flores Calero¹ Darío Javier Robayo Jácome² Darío Alexander Saa Escobar³

Para citar este artículo: M. J. Flores; D. J. Robayo, y D. A. Saa (2015). Histograma del gradiente con múltiples orientaciones (hog-mo) detección de personas. *Revista Vínculos*, 12(2), 138-147.

Recibido: 15-01-2015 / **Modificado:** 20-03-2015 / **Aprobado:** 07-05-2015

Resumen

En el campo de la visión por computador el problema de la clasificación de personas aún permanece como un desafío abierto de investigación. Por lo tanto, en este trabajo se realizan las siguientes aportaciones. Primero se introduce un nuevo método de extracción de características basado en el descriptor HOG (Histogram of Oriented Gradient) con múltiples orientaciones del gradiente, calculado sobre partes del cuerpo humano, denominado HOG-MO. Luego se construye un clasificador utilizando HOG-MO y SVM, se verifica su desempeño al compararlo con otras propuestas del estado del arte mediante las curvas ROC, logrando un adecuado equilibrio entre tiempo de cómputo y tasa de clasificación. En seguida se construye un sistema mono-cámara de detección de personas que trabaja en múltiples resoluciones, en el espectro visible, bajo condiciones variables de iluminación y de escala. Este sistema ha sido probado sobre una base de datos de personas en ambientes urbanos, en el espectro visible (BD-AU), creada para el desarrollo de aplicaciones en sistemas inteligentes de transporte (SIT) para la detección de peatones.

Palabras clave: clasificación de personas, vector de características, HOG-MO, múltiples orientaciones, SVM, SIT.

Abstract

In the field of computer vision, the people classification problem has not yet been fully resolved and there are still many unknowns to solve. This paper presents a new version of the HOG descriptor, which is based on multi-orientation and human body parts, using the SVM classifier. The ROC curves show the discrimination power of each case analyzed, compared with the state of the art algorithms. Finally, a system for people detecting in urban environments is presented. This detector has been tested on a database created for applications in Intelligent Transport Systems (ITS).

Keywords: People classification, characteristic vector, HOG-MO, multi-orientation, SVM, ITS.

- 1 Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas-ESPE, Av. Gral. Rumiñahui, s/n, PBX 171-5-231B, Sangolquí (Pichincha), Ecuador. E-mail: mjflores@espe.edu.ec
- 2 Escuela de Ingeniería en Sistemas, Pontificia Universidad Católica del Ecuador, Av. Manuela Sáenz s/n y Ernesto López, Ambato (Tungurahua), Ecuador. E-mail: drobayo@pucesa.edu.ec
- 3 Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas-ESPE, Av. Gral. Rumiñahui, s/n, PBX 171-5-231B, Sangolquí (Pichincha), Ecuador. E-mail: dasaa@espe.edu.ec

1. INTRODUCCIÓN

Al amparo de la tecnología de visión por computador, las técnicas de detección y clasificación de personas son importantes para la construcción de un sinnúmero de aparatos tecnológicos en los campos de la robótica, de la interacción humano-computador, por ejemplo: los sistemas de protección de peatones (SPP) [1-8], los sistemas de video-vigilancia [7,9], entre otros. Estos aparatos deben vigilar continuamente la zona de interés para detectar a las personas que puedan aparecer en la escena, sin imponer ningún tipo de restricción, como pueden ser: de forma (vestimenta, tamaño en función de la distancia), de condiciones ambientales (exceso o baja iluminación, lluvia, niebla o neblina, etc.), u otras [10]. Por la complejidad del problema sus soluciones trabajan parcialmente y bajo ciertas condiciones, cometen altas tasas de falsas detecciones [11] y su fiabilidad aún es dudosa.

Es así que se han presentado propuestas bajo visión monocular (2D) [2, 12-14] y visión estéreo (3D) [1,11,15]. En ambos casos esta tecnología es relativamente económica [1], pero tiene el gran inconveniente de requerir de un algoritmo altamente complejo para extraer la información útil de la escena [3], lo que involucra importantes tiempos de cómputo.

Por todos estos argumentos en este artículo se presenta un nuevo método de extracción de características basado en el descriptor HOG, calculado sobre partes del cuerpo humano, con multi-orientación, usando el algoritmo SVM. Además, se implementan algunas variantes del descriptor HOG con el objetivo de contrastar los resultados y verificar la calidad de esta nueva propuesta, basado en el análisis de las curvas ROC. Luego, como extensión de este trabajo, se construye un prototipo para aplicaciones en seguridad vehicular con el objetivo de establecer un sistema de alerta temprana para el conductor ante una posible situación de atropellamiento.

El documento está formado de cinco partes: la sección uno presenta la introducción y los lineamientos generales del trabajo investigativo; en la sección dos se exhibe el estado del arte, donde se describen

las principales técnicas utilizadas para resolver la problemática de clasificación de personas; en la sección tres se expone el trabajo desarrollado durante la investigación y sus principales resultados experimentales; en la sección cuatro se describe un sistema para detectar personas en ambientes urbanos en el espectro visible; finalmente, en la última parte se presentan las conclusiones y los trabajos a futuro.

2. ESTADO DEL ARTE EN LA CLASIFICACIÓN DE PERSONAS

En el campo del reconocimiento de patrones, usando visión por computador, numerosos investigadores han hecho esfuerzos importantes para buscar los detalles que mejor puedan representar al objeto de interés, en este caso la forma humana, y han utilizado instrumentos como los wavelets Haar [16], el filtro de Gabor, el descriptor HOG [17-19], la imagen de bordes [15] o una combinación de ellos con importantes resultados experimentales.

Para la caracterización de personas, el descriptor HOG es ampliamente utilizado desde su introducción por Dalal y Triggs en 2005 [18-19], cuya propuesta inicia con el cálculo de los gradientes, vertical y horizontal, a partir de los cuales se construyen las matrices de orientación y magnitud. Inmediatamente se divide la imagen de tamaño 64x128 píxeles en bloques y celdas rectangulares, un bloque se compone de 2x2 celdas y una celda de 8x8 píxeles. Luego se construye el histograma acumulando la magnitud bajo la discretización de la orientación sobre cada celda, seguido de un proceso de normalización sobre cada bloque con la norma L_2 , para finalizar en un único vector de características concatenando los histogramas de los bloques. Por otra parte, Gerónimo *et al.* [1], [16] han propuesto una versión simplificada para acelerar el cómputo denominada SHOG, donde se introduce el uso de la imagen integral para almacenar la magnitud en cada categoría de la orientación y posteriormente proseguir con el cálculo de los histogramas. Otra propuesta es presentada por Li *et al.* [20] donde se toma en cuenta la información espacial, esta se extiende al caso de

múltiples escalas para la construcción de los denominados bloques variables que van desde 16x16 píxeles hasta 64x128 píxeles, utilizando la norma L_1 y la imagen integral para reducir el tiempo de cómputo en la discretización de los gradientes. Cao *et al.* [21] han introducido el concepto de HOG de segundo orden, donde se construyen relaciones entre los histogramas de las celdas usando las funciones: media armónica, mínimo y producto. He *et al.* [22] han presentado el método denominado SS-HOG, donde utilizando la teoría de espacio de escalas, calculan el descriptor HOG sobre cada escala (una escala es una nueva imagen generada a partir de la imagen original convolucionada con un núcleo gaussiano de varianza σ , donde σ se denominada la escala). Por otra parte, Fernández *et al.* [23] han dividido la imagen del cuerpo humano en distintas piezas para caracterizar por separado cada una de ellas (cabeza, brazos, piernas y tronco), para extraer las características de estas partes han utilizado una combinación de: filtro de Canny, Haar wavelets, HOG y la textura. Finalmente, con la misma metodología Shashua *et al.* [2] han introducido un método que divide el cuerpo humano en 13 regiones; cada una de ellas se ha fraccionado cuatro (2x2) celdas espaciales sobre las cuales se ha calculado el histograma del gradiente utilizando ocho orientaciones, luego, usando la regresión Ridge, reducen la dimensión del vector de características. Otro punto importante es la elección del algoritmo de clasificación. Las investigaciones previas manifiestan que las técnicas de redes neuronales artificiales, máquinas de vectores soporte (SVM, Support Vector Machine), Adaboost y el análisis de componentes

principales (PCA, Principal Component Analysis) están entre las más citadas [2,12,15,16,24-26], generando resultados aceptables para la etapa de discriminación. Entre los principales inconvenientes son que aún presentan están los altos costes computacionales e importantes tasas de falsas detecciones.

3. DESARROLLO DEL ESTUDIO EXPERIMENTAL EN LA CLASIFICACIÓN DE PERSONAS EN AMBIENTES URBANOS

La estructura del estudio de la clasificación de personas que se presenta en este documento se compone de las siguientes etapas: I) base de datos en ambientes urbanos, II) base de datos de entrenamiento y validación, III) construcción del vector de características, IV) entrenamiento del algoritmo de clasificación.

3.1 Base de datos en ambientes urbanos

Para esta investigación se ha construido una base de datos en ambientes urbanos, en el espectro visible, denominada BD-AU (base de datos en ambientes urbanos), generada con un sistema mono-cámara. BD-AU contiene un conjunto de imágenes que fueron tomadas en las calles de la ciudad de Latacunga, Ecuador, bajo distintos escenarios durante el día donde interactúan personas con su ambiente en sus actividades cotidianas.

Actualmente, BD-AU está compuesta por cuatro secuencias de imágenes, cada una formada por 1600 fotogramas. La figura 1 muestra algunos ejemplos de diferentes escenarios durante el día.



Figura 1. Ejemplos de imágenes de BD-AU.

Fuente: elaboración propia.

4. BASE DE DATOS DE ENTRENAMIENTO Y VALIDACIÓN

Para desarrollar los experimentales se ha construido una base de datos con 24 920 imágenes, de las cuales 8246 corresponden a personas y el resto a no-personas, manteniendo así la relación 1:2 entre ejemplos positivos y negativos [5], [10]. Esta base se ha dividido aleatoriamente en dos partes, una base de entrenamiento (70%) y una base de validación (30%). En la figura 2 se presentan varios ejemplos positivos y negativos.



Figura 2. Ejemplos de (A) personas y (B) no personas.

Fuente: elaboración propia.

BD-AU ha sido creada con la finalidad de aportar al conocimiento universal, por tanto, pronto será liberada para el uso de la comunidad académica.

5. CONSTRUCCIÓN DEL VECTOR DE CARACTERÍSTICAS BASADO EN PARTES DEL CUERPO HUMANO

La extracción de características es el primer paso para la clasificación de objetos y la construcción de los algoritmos de reconocimiento de patrones. En el caso particular del problema de clasificación de personas, el descriptor HOG y algunas de sus variantes han demostrado ser competitivos frente a otros métodos [18].

En esta investigación, el procedimiento de extracción de características consiste en: a) extracción de las

características que mejor representan a una persona sobre una imagen basado en el descriptor HOG con múltiples orientaciones, b) entrenamiento del algoritmo de clasificación y c) análisis para elegir la mejor opción en función de las curvas ROC.

6. DESCRIPTOR HOG

Este descriptor es una transformación del espacio 2D (imagen) al espacio 1D (histograma), al dividir la imagen en varios bloques para calcular el histograma de su intensidad basados en las orientaciones del gradiente [3, 18, 19]. Los pasos a seguir son los siguientes:

I. Calcular las imágenes gradiente, I_x y I_y utilizando (1) y (2) sobre la imagen I :

$$G_x(x, y) = I(x + 1, y) - I(x - 1, y) \quad (1)$$

$$G_y(x, y) = I(x, y + 1) - I(x, y - 1) \quad (2)$$

Donde (x, y) es la luminosidad en el pixel (x, y) .

II. Obtener la norma (M_G) y la orientación del gradiente (θ_G) utilizando (3) y (4)

$$M_G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \quad (3)$$

$$\theta_G(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (4)$$

Calcular el histograma mediante la discretización de θ_G en N_b intervalos sobre cada celda de tamaño $m \times m$ píxeles. La discretización de θ_G se realiza en el intervalo $[-\pi, \pi]$. Los parámetros $m=8$ y $m=8$ varían entre autores. Por ejemplo, Dalal y Triggs [18, 19] utilizan $m=8$ y $m=8$, Xu y Gao [6] eligen $m=8$ y $m=8$, Shashua et al. [2] recurren a $m=2$ y $m=2$.

III. Utilizando la norma en L_2 ($\|x\|_2 = \sqrt{|x_1|^2 + \dots + |x_n|^2}$), normalizar el vector de características para obtener:

$$v = \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon}} \quad (5)$$

Donde $\varepsilon = 0,001$ sirve para evitar divisiones por cero. Este proceso reduce los impactos de la iluminación.

7. ENTRENAMIENTO DEL ALGORITMO DE CLASIFICACIÓN

En este trabajo se han utilizado los clasificadores SVM lineal y no lineal.

7.1 SVM

SVM es parte de los sistemas basados en la teoría de aprendizaje estadístico [27] donde se desea encontrar un hiperplano, H , dado por:

$$H: \langle \omega, x \rangle + b = 0 \quad (6)$$

Donde b es la normal al hiperplano y b es la distancia del hiperplano al origen.

Para resolver esta ecuación SVM se plantea a través de un problema de optimización convexa, garantizando así una única solución. En su forma dual se tiene:

$$\begin{aligned} \min_{\alpha} J(\alpha) &= \frac{1}{2} \alpha Q \alpha^T - e^T \alpha \\ \text{s.t.} & \\ 0 \leq \alpha_i \leq C, \quad & i = 1, \dots, m \\ y^T \alpha &= 0 \end{aligned} \quad (7)$$

Donde $\alpha \in R^m$ son los multiplicadores de Lagrange, $\alpha \in R^m$ es el vector de unos, C es una cota superior que penaliza el error, Q es una matriz de tamaño $m \times m$ con $Q_{ij} = y_i y_j K(x_i, x_j)$, $x_j \in R^m$ es la función núcleo, $x_j \in R^m$ con $j = 1, \dots, n$, es el vector de características, $y_i = 1$ es el vector que representa las clases, en este caso, $y_i = 1$ para persona y $y_i = -1$ para no-persona, $i = 1, \dots, m$.

La función núcleo determina la complejidad de la frontera de decisión pudiendo ser lineal o no lineal.

En la opción lineal $K(x_i, x_j) = x_i^T x_j$. En el caso no-lineal $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$, que se denomina RBF (Radial Basic Function), y γ especifica la tasa de decaimiento de $K(x_i, x_j)$ hacia cero.

Para optimizar (7) es necesario conocer los mejores valores de C y γ , para ello se utiliza el método de búsqueda de cuadrícula usando validación cruzada [27-28]. En el caso lineal, únicamente, se busca el valor del parámetro C .

7.2 Descriptor HOG con múltiples orientaciones sobre regiones estratégicas (HOG-MO) del cuerpo humano

El cuerpo humano presenta predominantemente una forma vertical y se encuentra compuesto de varias partes [2, 23]. Cada una de ellas tiene una forma particular, por ejemplo, la cabeza y las piernas tienen distinta forma. Tomando en consideración esta información, en esta propuesta se ha decidido dividir el cuerpo humano utilizando la estructura de Shashua *et al.* [2], ver figura 3, más un conjunto de orientaciones divididas en 4, 6, 8, 10, 12 y 14 para categorizar el ángulo del gradiente θ_G , sobre imágenes de tamaño 18x36 píxeles; a partir de donde se extraerá la mejor opción para calcular el histograma.

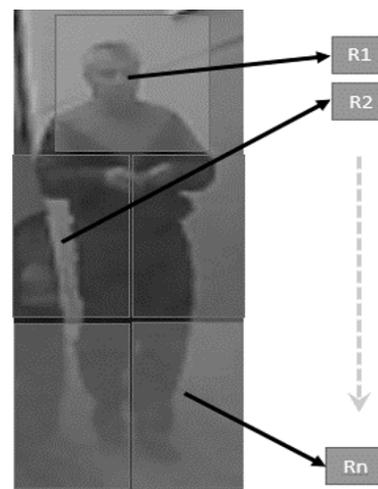


Figura 3. División en partes del cuerpo humano en varias regiones con distintas orientaciones.

Fuente: elaboración propia.

Para elegir el mejor tamaño del histograma, es decir, tomando como variable la orientación, sobre cada una de las regiones se ha desarrollado un trabajo experimental con todas las combinaciones posibles de los ángulos utilizando SVM lineal. Los resultados se presentan en la tabla 1, en ella se puede observar que las opciones predominantes son las de 12 y 14 orientaciones.

Tabla 1. Parámetros del clasificador SVM por regiones, con multi-orientación.

Región	1	2	3	4	5	6	7
Número de orientaciones del histograma	12	12	14	12	12	12	14
Región	8	9	10	11	12	13	
Número de orientaciones del histograma	14	10	14	14	12	12	

Fuente: elaboración propia.

Luego de la elección de las orientaciones para cada región se ha generado un vector de características unificado sobre el cual se ha entrenado el clasificador SVM, en las opciones lineal y no lineal. Los resultados se presentan en la tabla 2. Adicionalmente, en este trabajo se han implementado varias variantes del descriptor HOG presentes

en el estado del arte. La tabla 2 muestra los resultados y sus parámetros para los casos de Shashua *et al.* [2], Ludwig *et al.* [29] y Xu y Gao [6], y la figura 4 presenta las respectivas curvas ROC [30] para comparar el poder de clasificación en cada caso, donde se observa que la propuesta HOG-MO es superior. Todos los algoritmos de clasificación presentados en este trabajo han tomado varias semanas de trabajo computacional en calcular los parámetros, en especial la propuesta presentada de HOG-MO por regiones.

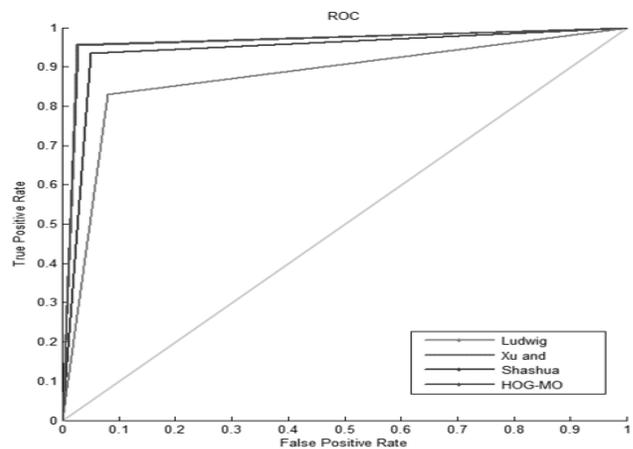


Figura 4. Curvas ROC de los métodos implementados en esta investigación, con SVM Lineal.

Fuente: elaboración propia.

Tabla 2. Métodos para el estudio experimental para la clasificación de personas.

Método	Clasificador	Parámetros	Precisión (%)
HOG-MO	SVM Lineal	C=3946,47	97,12
	SVM RBF	C=32 y $\gamma = 0,707101$	98,34
Shashua <i>et al.</i>	Adaboost con regresión Ridge		96,62
Ludwig <i>et al.</i>	SVM Lineal	C=26	92,54
	SVM RBF	C=4 y $\gamma = 1,414214$	98,23
Xu and Gao	SVM Lineal	C=65536	98,328
	SVM RBF	C=32 y $\gamma = 0,007813$	98,547

Fuente: elaboración propia.

7.3 Tiempos de cómputo

En las aplicaciones de visión por computador es importante trabajar con métodos eficientes tanto en el poder de clasificación como en el tiempo de cómputo. Así, el tiempo promedio de ejecución de HOG-MO es de 8,5 milisegundos en el caso lineal y de 13 milisegundos en el caso no lineal. Estos tiempos son inferiores a los descritos en el estado del arte, por ejemplo, Xu y Gao [6] requiere de 255 milisegundos. Estos experimentos se realizaron en un computador portátil Core-I7 a 1, 8GHz con 8Gb de RAM. El programa se implementó en Visual Estudio 2008 usando C/C++ y OpenCv.

7.4 Detección de personas en ambientes urbanos en el espectro visible

Como parte de esta investigación se ha desarrollado un algoritmo para la detección de personas en ambientes urbanos, en el espectro visible, utilizando información mono-cámara. Se destaca que en este tipo de propuestas no existe ningún tipo de restricción, por ejemplo, las personas pueden aparecer de manera inesperada en la escena.

En esta construcción se ha utilizado una estrategia denominada SVM en dos etapas (2SVM) [8], usando la nueva propuesta HOG-MO. La primera fase consiste en utilizar SVM lineal y el método de desplazamiento por ventana (SW, sliding window) [1,18], para obtener un conjunto de candidatos a personas. En la segunda etapa se usa SVM no-lineal con RBF para mejorar la precisión en la clasificación. SW es un método exhaustivo de búsqueda que trabaja sobre todas las localizaciones y escalas y ha demostrado ser efectivo pero costoso computacionalmente [1,18]; sin embargo, al desarrollar una búsqueda focalizada en función de algunos parámetros, por ejemplo, la distancia, es posible reducir significativamente el tiempo de cómputo.

En este trabajo se ha implementado una modificación de SW bajo la propuesta de mundo plano [1] en conjunto con el tamaño promedio de las personas en función de la distancia. Para calcular estos

tamaños se ha usado la base de datos BD-AU cuyos resultados se presentan en la tabla 3.

En la salida del método anterior se obtienen muchas hipótesis alrededor del verdadero objeto de interés, estas múltiples hipótesis han sido agrupadas por escala utilizando el algoritmo de agrupamiento dado por [1], utilizando (8) para medir la distancia entre regiones B_1 y B_2 .

$$d(B_1, B_2) = \frac{\text{area}(B_1 \cap B_2)}{\text{area}(B_1 \cup B_2)} \quad (8)$$

Si $d(B_1, B_2)$ es menor a un umbral dado, las regiones se agrupan y forman un conglomerado y así sucesivamente hasta obtener todas las personas presentes en la escena. De esta manera se completa el procedimiento para construir el sistema para detectar personas en distintas resoluciones.

Tabla 3. Tamaño promedio de las personas, en función de la distancia, sobre imágenes de tamaño 320x240 píxeles, calculados en BD-AU.

Distancia (metros)	Tamaño promedio (píxeles)
0-10	65x135
10-20	43x90
20-50	13x26

Fuente: elaboración propia.

Actualmente, el tiempo de procesamiento es de ocho imágenes por segundo. La figura 5 muestra varios ejemplos del funcionamiento del método propuesto.

8. CONCLUSIONES

En esta investigación se han realizado los siguientes aportes novedosos: en primer lugar, se ha presentado un estudio detallado usando los principales algoritmos del estado del arte basados en el descriptor HOG. A continuación, se ha construido una nueva base de datos denominada BD-AU, la misma que ha sido creada con la finalidad de aportar al conocimiento universal y el uso de la comunidad académica en el desarrollo de aplicaciones en SIT.



Figura 5. Ejemplos de detección de personas en escenas urbanas, en el espectro visible.

Fuente: elaboración propia.

Posteriormente, se ha introducido una nueva propuesta denominada HOG-MO basada en el descriptor HOG y distintas orientaciones del gradiente, evaluado sobre regiones del cuerpo humano; los experimentos desarrollados demuestran, a través de las curvas ROC, la capacidad de discriminación entre personas y otro tipo de objetos en ambientes urbanos durante el día. Además, se ha logrado un equilibrio entre el tiempo de cómputo y la tasa de clasificación, debido a que HOG-MO presenta tiempos menores y resultados similares a Xu y Gao [6].

Luego, se ha construido un detector de personas mono-cámara especializado en localizar individuos en el espectro visible, donde se ha utilizado una doble estrategia. En la primera etapa se usa SW con la hipótesis de mundo plano, con SVM lineal; y en la segunda fase SVM no-lineal. Así se ha logrado mejorar significativamente el desempeño del sistema en su conjunto. Esta propuesta trabaja bajo condiciones desafiantes como variación en la iluminación y la escala.

Finalmente, en el futuro se trabajará en métodos para reducir simultáneamente el número de regiones de interés y el tiempo de cómputo, para llegar a construir un sistema en tiempo real.

9. FINANCIAMIENTO

Esta investigación ha sido financiada por la Pontificia Universidad Católica del Ecuador, sede Ambato, a través de la Primera Convocatoria Proyectos de Investigación, Innovación y Desarrollo 2013.

REFERENCIAS

- [1] D. Gerónimo, "A global approach to vision-based pedestrian detection for advanced driver assistance systems", Tesis PhD, Universidad de Barcelona, 2009
- [2] A. Sashua, A. Gdalyahu y Y. Hayun, "Pedestrian detection for driving assistance systems, single frame classification and system level performance". In Proc IEEE Intelligent Vehicle Symposium. [En línea]. Disponible en <http://www.cs.huji.ac.il/~shashua/papers/iv04-ped.pdf>
- [3] T. Gandhi y M. Trivedi, "Pedestrian protection Systems: Issues, survey and challenges", *IEEE Transactions on Intelligent Transport Systems*, vol. 8, no. 9, sep. 2007.
- [4] APROSYS Project, "Strategies for enhanced pedestrian and cyclist friendly design", septiembre, 2009.
- [5] S. Munder y D. Gavrilu, "An experimental study on pedestrian classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 11, pp. 1863-1868, 2006.

- [6] F. Xu y M. Gao, "Human detection and tracking based on HOG and Particle Filter", *IEEE 3rd International Congress on Image and Signal Processing (CISP2010)*, pp. 1503-1507, 2010.
- [7] X. Wang, M. Wang y W. Li, "Scene-Specific Pedestrian Detection for Static Video Surveillance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 361-374, 2014.
- [8] K. Min, H. Son, Y. Choe y Y. G. Kim, "Real time Pedestrian Detection Based on a Hierarchical Two-Stage Support Vector Machine", *IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 114-119, junio 2013.
- [9] D. Brazey y C. Gout, "An algorithm for automatic people detection from depth map sequences". *European Workshop on Visual Information Processing (EUVIP)*, pp. 1-6, diciembre 2014.
- [10] D. Fernández, "Stereo vision based pedestrian detection system for assisted driving", tesis PhD, Universidad de Alcalá, 2008.
- [11] P. Kelly, N.E. O'Connor y A.F. Smeaton, "A Framework for Evaluating Stereo-Based Pedestrian Detection Techniques", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, pp. 1163-1167, agosto 2008.
- [12] M. Enzweiler y D. Gavrila, "Monocular pedestrian detection: survey and experiments", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no.12, diciembre 2009.
- [13] O. Y. Tetik y B. Bolat, "Detection of pedestrians from still images", *IEEE Conference on Signal Processing and Communication Applications (SIU2011)*, pp. 670-673, junio 2011.
- [14] C. Keller, M. Enzweiler y D. Gavrila, "A New Benchmark for Stereo-Based Pedestrian Detection", *IEEE Intelligent Vehicles Symposium (IV2011)*, julio 2011.
- [15] C. Hilario, "Detección de peatones en el Espectro Visible e Infrarrojo para un Sistema Avanzado de Asistencia a la Conducción", tesis Doctoral, Universidad Carlos III de Madrid, 2008.
- [16] D. Gerónimo, A. Sappa, D. Ponsa y A. Lopez, "2D-3D based on board pedestrian detection system. Computer Vision and Image Understanding", *Computer Vision and Image Understanding*, vol. 114, pp. 583-595, 2010.
- [17] G. Xu, X. Gu, L. Liu y Z. Wu, "Real time pedestrian detection based on Edge factor and histogram of oriented gradient", *IEEE International Conference on Information and Automation (ICIA2011)*, julio 2011.
- [18] N. Dalal y B. Triggs, "Histograms of oriented gradients for human detection", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2005.
- [19] N. Dalal, "Finding People in Images and Videos", tesis PhD, Institut National Polytechnique de Grenoble, 2006.
- [20] Z. Li, Y. Sun, F. Liu y W. Shi, "An Effective and Robust Pedestrian Detecting Algorithm & Symposia", *IEEE International Conference on Intelligent Transportation Systems*. Beijing, pp.12-15, octubre 2008.
- [21] H. Cao, K. Yamaguchi, T. Naito y Y. Ninomiya, "Pedestrian recognition using second order HOG feature", *Proceeding of the 9th Asian Conference on Computer Vision (ACCV'09)*, vol. part II, pp. 628-634, Xian-China, 2010.
- [22] N. He, J. Cao y L. Song, "Scale Space Histogram of Oriented Gradients for Human Detection", *IEEE International Symposium on Information Science and Engineering (ISISE)*, Shanghai, China, pp. 167-170, 2008.
- [23] D. Fernández, I. Parra, A. Sotelo, P. Revenga, S. Alvarez y M. Gavilán, "3D candidate selection method for pedestrian detection on non-planar roads", *IEEE Proceedings of the Intelligent Vehicle Symposium*, pp. 13-15, junio 2007.
- [24] M. Enzweiler, A. Eigenstetter, B. Schiele y D. Gavrila, "Multi-Cue Pedestrian Classification with Partial Occlusion Handling", *IEEE International Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010.
- [25] G. Lian, J. Lai y Y. Yuan, "Fast Pedestrian Detection using a Modified WLD Detector in Salient Region", *International Conference on System Science and Engineering*, junio 2011.

- [26]N. Pal, "Object-Background Segmentation Using New Definitions of Entropy[C]", *IEEE Proc.*, pp. 284-295, 1989.
- [27]N. Cristianini y J. Shawe-Taylor, "An Introduction to Support Vector Machines and other kernel methods", *Journal of Machine Learning Research*, vol. 5, pp. 137-182, 2004.
- [28]C. Chung y C. Lin, "LIBSVM: A library for support vector machines", *ACM Transactions on Intelligent Systems and Technology*, vol. 2, issue 3, pp. 27:1-27:27, 2011.
- [29]O. Ludwig, D. Delgado, V. Goncalves y U. Nunes, "Trainable Classifier-Fusion Schemes: An Application To Pedestrian Detection", *12th International IEEE Conference on Intelligent Transportation Systems*, vol. 1. St. Louis, pp. 432-437, 2009.
- [30]K. Woods y K. Bowyer, "Generating ROC curves for artificial neural networks", *IEEE eTransactions on Medical Imaging*, vol. 16, no. 3, junio 1997.

