

## Prototipo de un dispositivo de conteo automático de personas

GARCÍA-TRINIDAD, Enrique\*†, MATA-GONZÁLEZ, José, VILLALOBOS-ACOSTA, José y RAMÍREZ-RAMOS, Gabriela.

*Universidad Tecnológica Fidel Velázquez, Instituto Politécnico Nacional – ESIME Azcapotzalco*

Recibido Abril 13, 2016; Aceptado Junio 29, 2016

### Resumen

El presente trabajo describe el desarrollo de un prototipo de dispositivo de conteo automático de personas, el cuál es un sistema portátil que posee una cámara de vídeo que captura imágenes y una computadora que analiza éstas ejecutando el algoritmo de Detección rápida de objetos usando Cascadas Impulsadas de Características Simples que detecta los rostros humanos presentes en el campo de visión de la cámara. Se ejecuta después un algoritmo de aprendizaje automático llamado Regresión lineal multivariable que permite reducir el porcentaje de error de detección de rostros a cero en 350 iteraciones con las imágenes probadas. Los algoritmos son implementados en una tarjeta *Raspberry Pi 3* y el número de personas que la cámara detecta es mostrado en una pantalla de cristal líquido. El trabajo contribuye al desarrollo de sistemas de conteo de personas portátiles basados en visión por computadora.

**Visión por computadora, aprendizaje automático, sistemas embebidos, Adaboost, Regresión lineal multivariable**

### Abstract

This paper describes the development of a prototype of an automatic people counting device, which is a portable system that has a video camera that captures images and a computer that analyses them running the Rapid object detection using a boosted cascade of simple features algorithm that detects human faces present in the field of view of the camera. Then, the computer runs a machine learning algorithm called multivariable linear regression which reduces the error rate face detection to zero in 350 iterations using tested images. The algorithms are implemented in a *Raspberry Pi 3* system and the number of people detected by the camera is displayed on a liquid crystal display. This work contributes with the development of automatic people counting portable devices based on computer vision.

**Computer Vision, Automatic Learning, Embedded Systems, Adaboost, Multivariate Linear Regression**

**Citación:** GARCÍA-TRINIDAD, Enrique, MATA-GONZÁLEZ, José, VILLALOBOS-ACOSTA, José y RAMÍREZ-RAMOS, Gabriela. Prototipo de un dispositivo de conteo automático de personas. *Revista de Prototipos Tecnológicos* 2016, 2-4: 66-75

\* Correspondencia al Autor (Correo Electrónico: [phd.enrique.garcia@ieee.org](mailto:phd.enrique.garcia@ieee.org))

† Investigador contribuyendo como primer autor.

## Introducción

El presente trabajo describe un sistema de conteo automático de personas el cuál es un dispositivo portátil que posee una cámara de video que captura imágenes y una computadora que analiza las imágenes provenientes de dicha cámara y ejecuta el algoritmo de Detección rápida de objetos usando Cascadas Impulsadas de Características Simples (Viola, Jones, 2001) que detecta los rostros humanos presentes en el campo de visión de la cámara. Se ejecuta después un algoritmo de aprendizaje automático que permite reducir el porcentaje de error de detección de rostros. El sistema cuenta de nuevo los rostros e imprime el número de personas en una pantalla de cristal líquido. Además, el sistema envía el dato del número de personas contadas a través del puerto de comunicación serial del dispositivo con la finalidad de conectar una computadora, un Controlador Lógico Programable o cualquier dispositivo que posea un puerto de comunicación serial.

De acuerdo al estado del arte y los antecedentes, no existen aparatos o dispositivos iguales al presentado en este trabajo, la patente que tiene más relación es la llamada *People counting systems and methods* (Ku, Steenburgh, 2010). Ku desarrolló un método que permite contar personas en una imagen. Para cada imagen individual en un conjunto de imágenes, uno o más *blobs* en la imagen individual están representados por uno o más símbolos en un histograma, estos símbolos son analizados en el histograma para contar a las personas en la imagen. Ku da a conocer algunas técnicas basadas en imágenes, que tratan de distinguir una imagen de la otra. Estas técnicas obtienen los *blobs* de diferentes imágenes y aplican reglas para conocer si en el *blob* está presente un rostro humano. Sin embargo, el método de Ku presenta una desventaja.

Ku propone una metodología donde la precisión en la detección de los rostros es muy baja debido a la aplicación de reglas de tratamiento de imágenes sencillas, además presenta resultados aceptables siempre y cuando la iluminación del lugar donde se desea contar a las personas sea óptima.

Se ha desarrollado un dispositivo de conteo automático de personas que usa un algoritmo de aprendizaje automático por medio de regresión multivariable con la finalidad de complementar los métodos de detección de rostros humanos presentados por Viola y Ku. La precisión en el conteo de personas es optimizada y además se ha propuesto su implementación en un sistema portátil basado en una tarjeta Raspberry 3 con la finalidad de facilitar su traslado a cualquier ubicación donde se desee el conteo de personas.

Los sistemas de conteo automático de personas que funcionan en base a componentes electrónicos, y que hacen uso de las técnicas de visión por computadora y usan como principal referencia la detección de rostros como manera de contar a las personas requieren una instalación masiva en el lugar o área donde se requiera un conteo de personas. Estos sistemas no pueden ser trasladados de un lugar a otro sin requerir un desmontaje y montaje posterior. También la eficacia del conteo depende en gran medida de las condiciones de iluminación del lugar. Se establece la hipótesis en que si se desarrolla un sistema embebido portátil y la precisión del conteo es optimizada usando un algoritmo de aprendizaje automático se puede solucionar los problemas de portabilidad del sistema y reducción en el error de conteo de personas.

En la sección 2 del presente artículo se describe el estado del arte del trabajo, describiendo el algoritmo de detección rápida de objetos usando cascadas impulsadas de características simples.

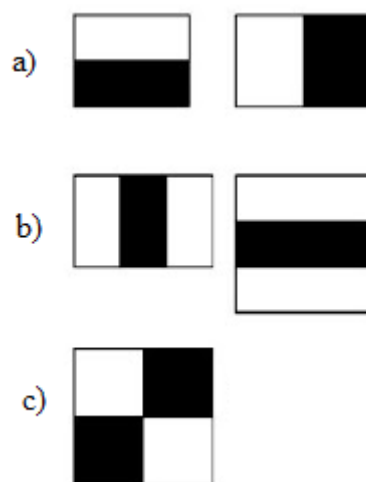
En la sección 3 se describe el algoritmo propuesto en este trabajo basándose en el algoritmo de regresión lineal multivariable. En la sección 4 se propone una implementación de éstos algoritmos y se desarrolla la experimentación. En la sección 5 se exponen los resultados y posteriormente se dan las conclusiones del presente trabajo.

### Algoritmo de detección rápida de objetos usando cascadas impulsadas de características simples

Este trabajo trata sobre el desarrollo de un sistema de conteo automático de personas el cual es un dispositivo electrónico portátil que incluye una cámara de video que captura imágenes, las cuáles son analizadas por una computadora que se encarga de detectar los rostros de las personas presentes en el campo de visión de la cámara. El dispositivo cuenta a las personas en base a la detección de sus rostros. La computadora analiza las imágenes provenientes de la cámara usando una técnica de visión por computadora llamada algoritmo de Detección rápida de objetos usando Cascadas Impulsadas de Características Simples y cuenta los rostros detectados de acuerdo a la detección de características. El procedimiento de detección de objetos clasificando imágenes se basa en el valor de características simples. Hay varias razones para usar características en vez de utilizar el análisis de píxeles directamente. La razón más común es que las características pueden actuar para codificar conocimiento en un dominio hecho a la medida, lo que puede ser difícil de aprender para un sistema que usa una cantidad finita de datos de entrenamiento. Otra razón es que los sistemas de detección basados en características operan más rápido que los sistemas basados en píxeles (Papargeorgiou et al, 1998). En este trabajo se utilizan características de rectángulos, descritos por Viola, ya que pueden ser calculados rápidamente usando una representación para la imagen llamada imagen integral. Más específicamente se ocupan tres tipos de características.

El valor de una característica de dos-rectángulos es la diferencia entre la sumatoria de los píxeles dentro de las regiones de los dos rectángulos. Las regiones tienen el mismo tamaño y forma así como son horizontal o verticalmente adyacentes como se muestra en la Figura 1. Una característica de tres-rectángulos calcula la sumatoria de los dos rectángulos en los extremos de un rectángulo central y finalmente una característica cuatro-rectángulos calcula la diferencia entre pares diagonales de rectángulos. La imagen integral en un punto de la imagen  $x, y$  contiene la sumatoria de píxeles arriba y a la izquierda de  $x, y$  por lo que:

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



**Figura 1** a) Característica dos-rectángulos o borde, b) Característica tres-rectángulos o línea y c) Característica cuatro-rectángulos. La suma de los píxeles que se encuentran dentro de los rectángulos blancos es sustraída de la suma de los píxeles de los rectángulos negros.

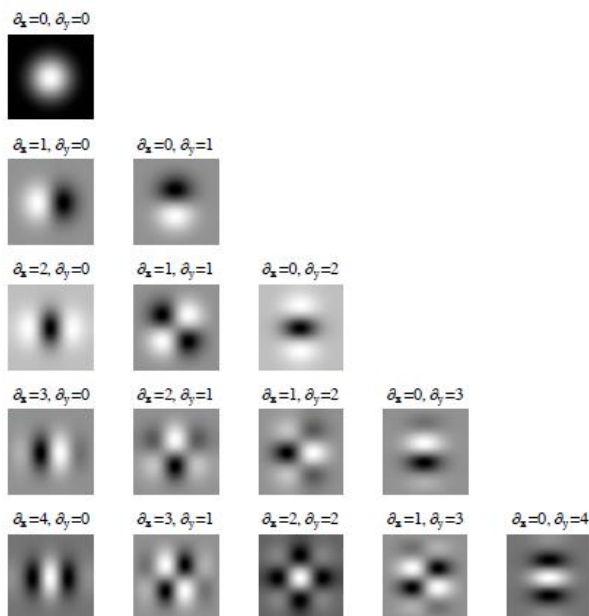
Donde  $ii(x, y)$  es la integral de la imagen y  $i(x, y)$  es la imagen original. Entonces sí:

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (2)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (3)$$

Donde a partir de las ecuaciones 1, 2 y 3  $s(x, y)$  es la suma de acumulativa de las filas,  $s(x, -1) = 0$  y  $ii(-1, y) = 0$  la integral de la imagen puede ser calculada en sólo una iteración a partir de imagen original.

Usando la imagen original  $\delta_x = 0, \delta_y = 0$  se puede realizar el cálculo en un arreglo de cuatro posiciones  $\delta_x = 0, \delta_y = 4$  como se muestra en la Figura 2.



**Figura 2** La razón de que Viola utiliza filtros Haar, es que pueden ser evaluados de manera muy eficiente. Se resta la suma de los píxeles cubiertos por la región negro del filtro de la suma de los píxeles cubiertos por la región blanca, y puesto que las regiones son rectangulares, la suma de los píxeles de cada región se puede calcular de manera eficiente de la imagen integral correspondiente.

Dado un conjunto de entrenamiento con respuestas positivas y negativas de rostros humanos, cualquier algoritmo de aprendizaje automático se aproxima al aprendizaje de una función de clasificación. Viola utiliza una variante del algoritmo *AdaBoost* (Freund, Shapire, 1999) el cual es usado como un algoritmo de rápida eficiencia que permite clasificar las imágenes. Freund prueba que el error se aproxima a cero de manera exponencial en un cierto número de iteraciones.

En el caso de que la resolución base para la detección de rostros humanos sea de 24 x 24 píxeles y que el número de características llegue a las 180000-rectángulos es posible encontrar clasificadores a los 576-rectángulos como lo indica Viola.

Un clasificador  $h_j(x)$  que consiste de una característica  $f_j$  y un límite  $\theta_j$  con una paridad  $p_j$  indica si hay un rostro encontrado o no, de acuerdo a la ecuación 4:

$$h_j(x) = 1 \{ \text{si } p_j f_j(x) < p_j \theta_j \} \quad (4)$$

$$h_j(x) = \{0 \text{ en cualquier otro caso} \}$$

Donde  $x$  es una parte de la imagen de 24 x 24 píxeles. El algoritmo de Viola de detección rápida de objetos usando cascadas impulsadas de características simples es:

Dado unos conjuntos de entrenamiento de imágenes  $(x_1, y_1), \dots, (x_n, y_n)$  donde  $y_i = 0, 1$  son ejemplos incorrectos y correctos respectivamente, entonces, se inicializan los pesos  $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$  para  $y_i = 0, 1$  respectivamente, y  $m$  y  $l$  son los números de respuestas incorrectas y correctas respectivamente.

– Para  $t = 1, \dots, T$ :

1. Se normalizan los pesos,

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$$

para que  $w_t$  sea una distribución de probabilidad.

2. Para cada característica  $j$  entrenar al clasificador  $h_j$  el cual es restringido a usar una sola característica. El error es evaluado con respecto a

$$w_t, \epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$

3. Se escoge un clasificador  $h_t$  con el menor error posible  $\epsilon_t$

4. Se actualizan los pesos:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$

Donde  $e_i = 0$  si el ejemplo  $x_i$  es clasificado correctamente y  $e_i = 1$  en cualquier otro caso y

$$\beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$$

- Al final el clasificador es:

$$h_x = \begin{cases} 1 & \text{si } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Donde

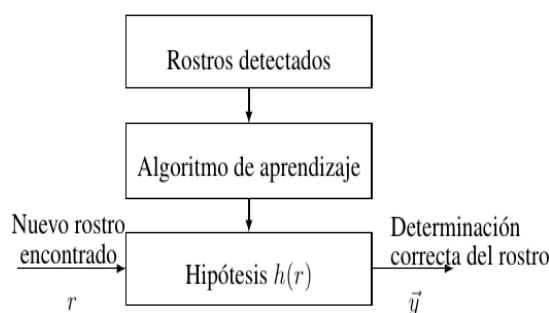
$$\alpha_t = \log \frac{1}{\beta_t}$$

### Algoritmo propuesto usando regresión lineal multivariable

Arthur Samuel definió al aprendizaje automático como el campo de estudio que da a las computadoras la habilidad de aprender sin ser explícitamente programadas (Samuel, 1959). Se puede afirmar entonces, que el problema de detección de rostros es un problema que puede ser resuelto usando aprendizaje automático, ya que el concepto de programación explícita es la de un programa que ejecuta una a una las instrucciones sin posibilidad de cambios en la estructura de éste. Como el problema del aprendizaje de saber cuáles son rostros verdaderos y cuáles no, no puede ser completado por un programa que no se modifique a sí mismo, ya que no sería factible escribir miles de programas, uno para cada rostro similar, es necesario usar otro método como el de dar a la computadora la habilidad de aprender. Tom Mitchel dice que el aprendizaje automático se puede definir mediante el estudio del siguiente problema (Mitchel, 1997): *Un programa de computadora se dice que aprende de la experiencia E con respecto a alguna tarea T y alguna medida del desempeño P, si hay desempeño en T medido por P, entonces se mejora con la experiencia E.*

Definiendo el problema en cuestión, T es la tarea de discriminar cuáles son rostros verdaderos y cuáles no lo son, E es la experiencia de los resultados del algoritmo desarrollado por Viola de predecir cuáles son rostros humanos y cuáles no y P es el número o fracción de los rostros correctamente calculados.

El aprendizaje automático se divide en dos áreas: aprendizaje supervisado y aprendizaje no supervisado. En el aprendizaje supervisado las respuestas correctas son dadas. En el caso de este sistema, las respuestas correctas son el conjunto de rostros correctamente detectados, y que son utilizadas para generar más respuestas correctas (rostros correctamente detectados). En el caso del aprendizaje no supervisado los datos son datos sin saber a qué corresponden y se pretende encontrar alguna estructura. Se llega a la conclusión entonces, que el modelo de aprendizaje utilizado en este trabajo es un modelo de aprendizaje automático supervisado. El algoritmo que se utiliza es el algoritmo de regresión lineal con múltiples variables o multivariable. La regresión lineal con múltiples variables es una función que aproxima un comportamiento de un sistema a una función lineal, de acuerdo a la Figura 3.



**Figura 3** Estimador de rostros correctamente detectados usando regresión línea multivariable.

Si los estados  $x_1, x_2, x_3, x_4, x_5$  y  $x_6$  corresponden a seis características de un rostro correctamente detectado por el algoritmo de Viola: número de ojos, distancia entre ojos, diámetro de ojos, posición de nariz, posición de boca y tamaño de cabeza respectivamente en píxeles y  $m$  es el número de los ejemplos de entrenamiento obtenidos por el algoritmo de Viola entonces el conjunto de ecuaciones 5 representa a los estados del sistema:

$$\begin{aligned}\vec{x}_1 &= [x_1^{(1)} \ x_1^{(2)} \ \dots \ x_1^{(m)}]^T \\ \vec{x}_2 &= [x_2^{(1)} \ x_2^{(2)} \ \dots \ x_2^{(m)}]^T \\ \vec{x}_3 &= [x_3^{(1)} \ x_3^{(2)} \ \dots \ x_3^{(m)}]^T \\ \vec{x}_4 &= [x_4^{(1)} \ x_4^{(2)} \ \dots \ x_4^{(m)}]^T \\ \vec{x}_5 &= [x_5^{(1)} \ x_5^{(2)} \ \dots \ x_5^{(m)}]^T \\ \vec{x}_6 &= [x_6^{(1)} \ x_6^{(2)} \ \dots \ x_6^{(m)}]^T\end{aligned}\quad (5)$$

Lo que se desea es estimar mediante una hipótesis el estado  $\vec{y}$  con la finalidad de encontrar si el rostro encontrado es un rostro verdadero. Si la ecuación 6 propone el vector de estados del sistema:

$$\begin{aligned}\vec{y} &= [y^{(1)} \ y^{(2)} \ \dots \ y^{(m)}]^T, \ x_0 = 1 \\ X &= [\vec{x}_0 \ \vec{x}_1 \ \vec{x}_2 \ \vec{x}_3 \ \vec{x}_4 \ \vec{x}_5 \ \vec{x}_6]\end{aligned}\quad (6)$$

La hipótesis es una función que pasa adecuadamente un plano a través de los datos, la cual puede ser plana y es expresada en la ecuación 7:

$$h_\theta(X) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_5 x_5 + \theta_6 x_6 \quad (7)$$

La finalidad es escoger a  $\theta = [\theta_0 \ \theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5 \ \theta_6]^T$  para que  $h_\theta(X)$  sea lo más cercana a  $y$ . Para esto se necesita que  $h_\theta(X) - y$  sea mínima, además con la finalidad de tener valores positivos ésta se puede reescribirse como  $(h_\theta(X) - y)^2$ .

Para el  $i$ -ésimo ejemplo de entrenamiento se tiene:

$$\sum_{i=1}^m (h_\theta(X^{(i)}) - y^{(i)})^2 \quad (8)$$

Como se quiere minimizar el error medio se puede reescribir la ecuación 8 como:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2 \quad (9)$$

Como  $h_\theta(X)$  puede vectorizarse como  $h_\theta(X) = \theta^T X$

$J(\theta)$  en la ecuación 9 puede ser expresada como se muestra en la ecuación 10.

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (\theta^T X^{(i)} - y^{(i)})^2 = \frac{1}{2m} \sum_{i=1}^m \left( \left( \sum_{j=0}^n \theta_j X_j^{(i)} \right) - y^{(i)} \right)^2 \quad (10)$$

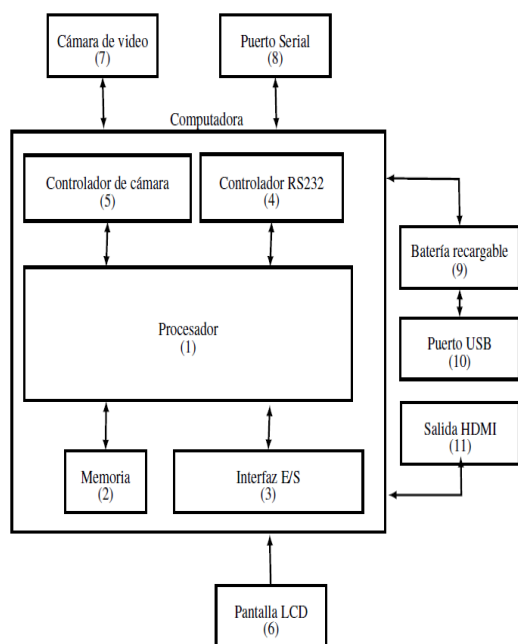
Donde  $J(\theta) = \text{mín}(\theta)$  y se conoce como Función costo. En el caso del algoritmo de aprendizaje se aplica el Algoritmo del Gradiente Descendiente, el cual se expresa en la ecuación 11.

$$\begin{aligned}&\text{Repetir hasta converger}\{ \\ &\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)} \\ &\}\text{Simultáneamente actualizar } \theta_j \text{ para } j = 0, \dots, 6\end{aligned}\quad (11)$$

Esto para el  $j$ -ésimo valor de  $\theta_j$ , y donde  $\alpha$  expresa la velocidad de convergencia para que  $J(\theta)$  se aproxime a cero.

## Experimentación

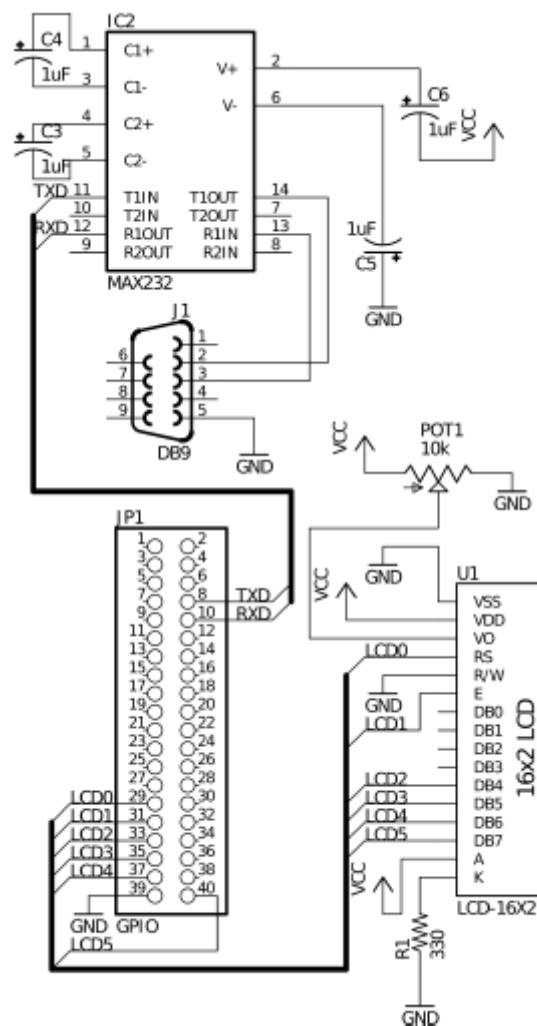
La Figura 4 es un diagrama a bloques simplificado que ilustra los componentes del dispositivo de conteo automático de personas.



**Figura 4** Diagrama a bloques del dispositivo electrónico.

El sistema incluye una cámara de video (7) la cual por medio del controlador de cámara (5) se comunica con la computadora. La computadora incluye un procesador (1), una memoria (2) y una interfaz de entrada/salida (E/S) (3). La computadora es opcionalmente acoplada con otro dispositivo externo vía puerto de comunicación serial (8) usando el controlador RS232 de la computadora. La interfaz E/S es acoplada a una pantalla alfanumérica LCD (6) y a un led ultrabrillante infrarrojo. La alimentación de energía eléctrica se hace a través de una conexión USB (10), la cual también permite descargar actualizaciones del controlador principal del dispositivo así como cargar la batería recargable (9). En caso de que se requiera servicio técnico o se requiera reprogramar el dispositivo se hace uso de una conexión HDMI, que permite conectar el dispositivo a una pantalla para que se pueda tener acceso al Sistema Operativo del dispositivo. Cuando el dispositivo se encuentra trabajando, es necesario que el objetivo (los rostros de las personas) sea ampliamente visible por el lente de la cámara.

La perspectiva y la distorsión angular creada por la posición de la cámara pueden ser compensadas utilizando un mapeo geométrico durante la etapa de procesamiento de imagen. Además de esto, la cámara puede aumentar la apertura y el tiempo de exposición para incrementar la recepción de luz. Adicionalmente, en el caso de que el área donde se coloque el dispositivo de conteo automático de personas carezca de iluminación, el led infrarrojo puede ser empleado para iluminar dicha área, con el objetivo de que se pueda contar a las personas sin necesidad de que éstas noten este conteo, por ejemplo, durante la reproducción de una película en el cine.



**Figura 5** Circuito de comunicación e impresión de datos en pantalla.



La computadora que se encuentra dentro del dispositivo de conteo automático de personas es una *Raspberry Pi 3*, la cual posee un *procesador QuadCore Broadcom ARMv8* el cual es un procesador de 64-bit y que ejecuta el sistema operativo *Raspbian*, una variante de *Ubuntu Linux*. El software que se ejecuta en el sistema operativo como programa principal está escrito en *Python* y utiliza módulos de librerías de código abierto *openCV*, cuyo funcionamiento es explicado más adelante. El software cuenta a las personas de manera automática usando como información las imágenes procedentes de la cámara de video, después dicha información sale a través del puerto de la computadora GPIO hacia la pantalla LCD alfanumérica y al circuito integrado MAX232 que cambia los valores lógicos de la computadora *Raspberry Pi 3* a valores lógicos de una computadora personal, tal y como se muestra en la Figura 5. Antes de utilizar el dispositivo de conteo automático de personas, el cual es mostrado en la Figura 6, la computadora y la cámara deberán ser inicializadas, el tiempo de espera para que el dispositivo funcione adecuadamente es de 60 segundos, esto para que el sistema operativo (*Raspbian*) sea cargado en la memoria de la computadora (*Raspberry Pi 3*). La computadora realiza un análisis de las imágenes capturadas por la cámara de video utilizando las librerías de procesamiento de imagen *openCV*, usando el lenguaje de programación *Python*, esto en base al algoritmo de *Viola*. El número de personas en la escena es determinado por la cantidad de rostros reconocidos por el proceso de procesamiento de imagen. La computadora graba los seis estados del algoritmo propuesto que son descritos anteriormente y que son patrones geométricos que relacionan los ojos, narices, labios y tamaño de la cabeza. Para determinar la presencia de un rostro humana, un análisis de color en los rostros encontrados es realizado.

Aunque los seres humanos tenemos una amplia variedad de colores de piel, en el espacio de colores éstos caen cerca del rojo, con variaciones de blanco y negro. El color de piel de los seres humanos no es verde, ni azul, en un espacio de colores RGB (Red/Rojo, Green/Verde, Blue/Azul). Este espacio de colores es una combinación de los colores previamente mencionados. El sistema computarizado cambia éste espacio de colores al espacio de colores HSI (Hue-Saturación-Intensidad). El componente *Hue* representa el color de la imagen, la saturación representa la blancura de la imagen y la intensidad representa a la imagen en escala de grises. En el espacio del componente *Hue*, la piel humana está cerca del valor 255 (en una escala de 0-255), la cual corresponde al color rojo. Una vez que los rostros son detectados, el espacio *Hue* e Intensidad son analizados para confirmar que realmente se trata de un rostro humano. Además es ejecutado un algoritmo para determinar el histograma de cada una de los rostros encontrados con la finalidad de producir diferentes características faciales. Las características son guardadas como una serie de vectores en la memoria de la computadora para que posteriormente se ejecute el algoritmo propuesto, con la finalidad de reducir el error en la detección de rostros, y por consecuencia del conteo de personas. Este proceso es mostrado en la Figura 7.



**Figura 6** Partes del prototipo



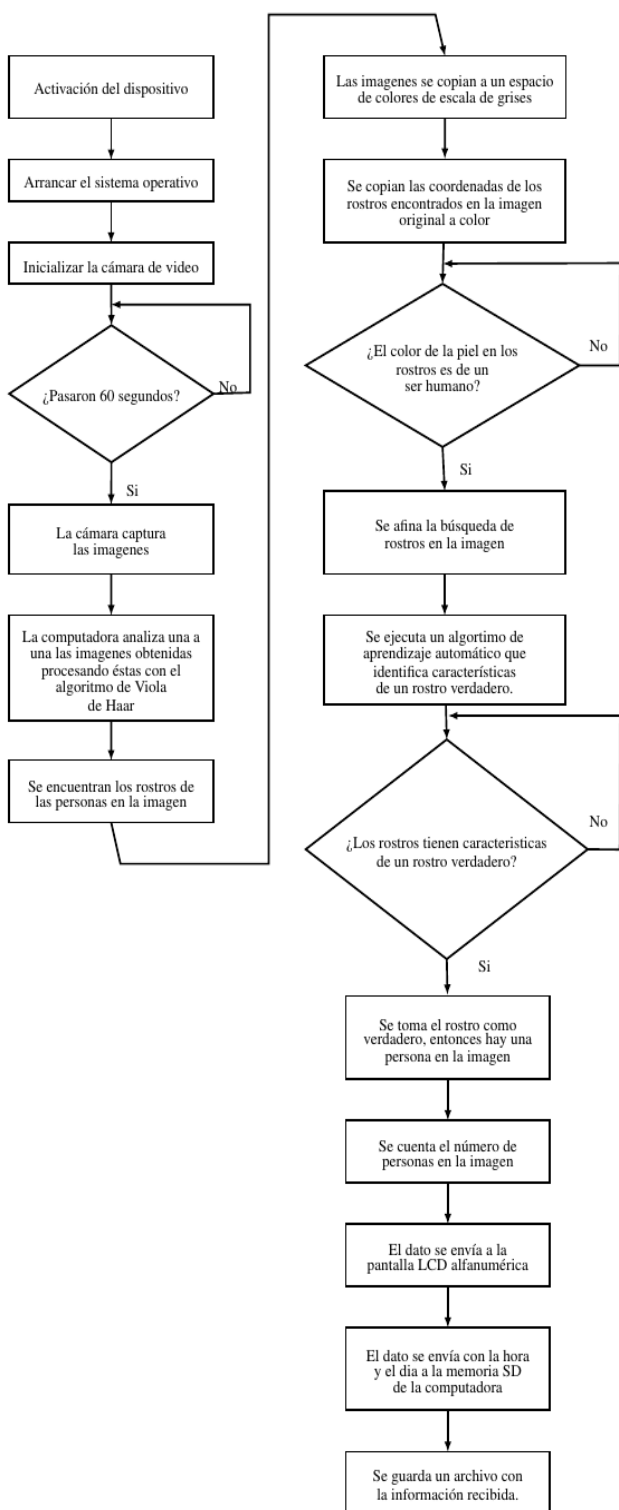


Figura 7 Diagrama de flujo del código.

Resultados

De acuerdo a la imagen presentada en la Figura 8 se ejecuta el algoritmo desarrollado y se obtiene el conteo automático de personas como se muestra en la Figura 9.



Figura 8 Imagen a tratar por el dispositivo.

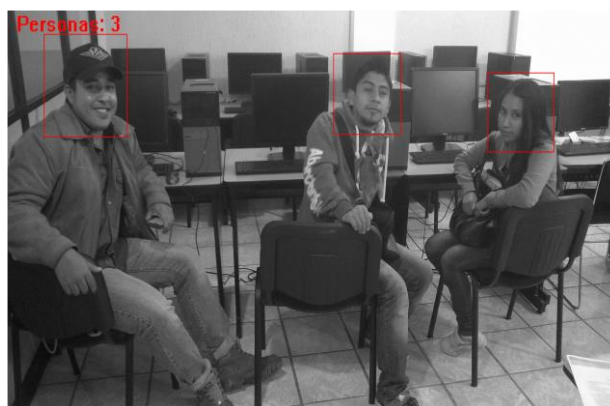
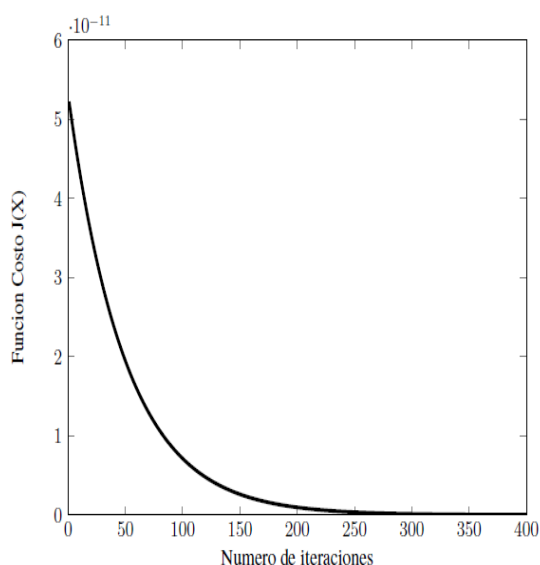


Figura 9 Conteo automático de personas.

El error usando el algoritmo de aprendizaje automático se reduce a cero a las 350 iteraciones del algoritmo propuesto, como se muestra en la Figura 10.



**Figura 10** Error en el conteo de personas, para la imagen propuesta.

### Conclusiones

Se ha desarrollado el prototipo propuesto obteniendo resultados de rendimiento favorables. Los trabajos de Viola y Freund son un referente preciso que permite el desarrollo de sistemas de detección de rostros humanos. En este trabajo se ha contribuido con la implementación de dichas técnicas y se propuesto un algoritmo para disminuir el error. El prototipo se encuentra en un estado de registro ante el Instituto Mexicano de la Propiedad Industrial como modelo de utilidad industrial. En el futuro se pretende modificar el algoritmo propuesto para introducir más características y estados que permitan, en imágenes más complejas, tender el error del conteo de personas a cero.

### Agradecimientos

Se agradece al Consejo Mexiquense de Ciencia y Tecnología COMECyT por los apoyos dados para la realización de esta investigación.

### Referencias

- Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition 2001*, (1), 511-518. doi: 10.1109/CVPR.2001.990517
- Ku, S. and Steenburgh, M. (2006). People counting systems and methods. US Patent. 10/949, 295.
- Papageorgiou, C.P. and Oren, M. and Poggio, T. (1998). A general framework for object detection. *Sixth International Conference on Computer Vision*, (1), 555-562. doi: 10.1109/ICCV.1998.710772.
- Freund, Y. and Shapire, R. (1999). A Short Introduction to Boosting. *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, (1), 1401-1406. doi: 10.1.1.107.3285.
- Samuel, A. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210.
- Mitchel, T. (1997). *Machine Learning*. McGraw-Hill, (1), 2.