



FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA ELECTRÓNICA

**DISEÑO Y SIMULACIÓN DE UN SISTEMA DE DETECCIÓN DE
MOVIMIENTO BASADO EN SEGMENTACIÓN DE IMÁGENES
UTILIZANDO LOS MÉTODOS DE UMBRALIZACIÓN Y
SUSTRACCIÓN DE FONDO APLICADO A UN SISTEMA DE
VIDEOVIGILANCIA**

PRESENTADA POR

NATALIA MAGALY GUTIERREZ MEDINA

MANUEL JESÚS VELAZCO FLORES

ASESORES

ELVIS RIVERA RIVERA

FERNANDO JIMÉNEZ MOTTE

TESIS

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO ELECTRÓNICO

LIMA – PERÚ

2019



CC BY-NC

Reconocimiento – No comercial

Los autores permiten transformar (traducir, adaptar o compilar) a partir de esta obra con fines no comerciales, y aunque en las nuevas creaciones deban reconocerse la autoría y no puedan ser utilizadas de manera comercial, no tienen que estar bajo una licencia con los mismos términos.

<http://creativecommons.org/licenses/by-nc/4.0/>



USMP
UNIVERSIDAD DE
SAN MARTIN DE PORRES

**FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA ELECTRÓNICA**

**DISEÑO Y SIMULACIÓN DE UN SISTEMA DE DETECCIÓN
DE MOVIMIENTO BASADO EN SEGMENTACIÓN DE
IMÁGENES UTILIZANDO LOS MÉTODOS DE
UMBRALIZACIÓN Y SUSTRACCIÓN DE FONDO APLICADO
A UN SISTEMA DE VIDEOVIGILANCIA**

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERÍA
ELECTRÓNICA**

PRESENTADA POR

**GUTIERREZ MEDINA, NATALIA MAGALY
VELAZCO FLORES, MANUEL JESÚS**

ASESORES

**Ing. ELVIS RIVERA RIVERA
Ing. FERNANDO JIMÉNEZ MOTTE**

LIMA, PERÚ

2019

ÍNDICE

	Página
RESUMEN	vii
ABSTRACT	viii
CAPÍTULO I: MARCO TEÓRICO	1
1.1 Antecedentes	1
1.1.1 Sustracción de fondo basada en diferencia de imágenes	3
1.1.2 Basada en la media de los “N” píxeles anteriores	4
1.1.3 Basada en la mediana de los “N” píxeles anteriores	4
1.1.4 Basada en la media móvil	4
1.1.5 Basada en la Gaussiana Simple	5
1.2 Bases teóricas	6
1.2.1 Imagen digital	6
1.2.2 Modos de color	7
1.2.3 Visión computacional	7
1.2.4 Procesamiento de imágenes digitales	8
1.2.5 Segmentación de imagen	9
1.2.6 Sustracción de fondo	10
1.2.7 Función de distribución de probabilidad	11
1.2.8 Función de distribución de probabilidad normal o campana de Gauss	11
1.3 Definición de términos básicos	12
1.3.1 Fondo o fondo unimodal	12
1.3.2 Primer plano	13

1.3.3 Método	13
1.3.4 Técnica	13
1.3.5 Aplicación	13
1.3.6 Sensibilidad	14
1.3.7 Predictividad positiva	14
CAPÍTULO II: METODOLOGÍA	15
2.1 Modelamiento de la obtención de la imagen unidimensional en el vector de Luminancia	15
2.2 Modelamiento del fondo aproximando la Función de Densidad de Probabilidad Gaussiana por cada píxel basado en el principio de Gaussiana Simple	16
2.3 Modelamiento de la detección de primer plano mediante el cálculo de la distancia de Mahalanobis y la estimación del valor umbral	17
2.4 Modelamiento de la segmentación de imágenes para la detección de primer plano	18
2.5 Modelamiento del fondo adaptativo Gaussiano	19
CAPÍTULO III: DESARROLLO DEL PROYECTO	21
3.1 Obtención de la imagen vectorial en el modelo de Luminancia	22
3.2 Función de Densidad de Probabilidad Gaussiana para representar los cambios de luminancia de un píxel en el tiempo	23
3.3 Cálculo de la distancia de Mahalanobis del píxel respecto a los parámetros del modelo de fondo	32
3.4 Implementación de la emisión de alarmas y almacenaje de imágenes por evento	34
CAPÍTULO IV: PRUEBAS Y RESULTADOS	37
4.1 Análisis de un píxel perteneciente al fondo	37
4.2 Análisis de un píxel perteneciente al primer plano	41
CONCLUSIONES	48
RECOMENDACIONES	50
FUENTES DE INFORMACIÓN	51

ÍNDICE DE FIGURAS

Página

Figura 1. Representación de la estructura de una imagen digital	6
Figura 2. Descomposición de una imagen RGB en sus tres componentes	7
Figura 3. Ejemplo de aplicación de la visión computacional	8
Figura 4. Ejemplo de aplicación del procesamiento de imágenes	9
Figura 5. Imagen binarizada, compuesta solo por blanco y negro	9
Figura 6. Ejemplo de una imagen resultante de la sustracción de fondo	10
Figura 7. Ejemplo de una distribución normal o Gaussiana	12
Figura 8. Diagrama de bloques del sistema	21
Figura 9. Píxeles con alto nivel de luminancia	22
Figura 10. Píxeles con bajo nivel de luminancia	23
Figura 11. Píxeles representantes del fondo.	24
Figura 12. Gráfico de línea y FDP aproximada a la Gaussiana píxel P1	25
Figura 13. Gráfico de línea y FDP aproximada a la Gaussiana píxel P2	26
Figura 14. Gráfico de línea y FDP aproximada a la Gaussiana píxel P3	27
Figura 15. Gráfico de línea y FDP aproximada a la Gaussiana píxel P4	28
Figura 16. Gráfico de línea y FDP aproximada a la Gaussiana píxel P5	29
Figura 17. Gráfico de línea y FDP aproximada a la Gaussiana píxel P6	30
Figura 18. Gráfico de línea y FDP aproximada a la Gaussiana píxel P7	31
Figura 19. Píxeles representantes del primer plano y del fondo	32
Figura 20. Resultado de la alarma visual	34
Figura 21. Radio de comunicación de doble vía y el botón PTT	35
Figura 22. Circuito activador de PTT y emisor de audio	36
Figura 23. Valores de la luminancia del píxel 330,190 en 24 horas de muestreo	38
Figura 24. Gráfica de la media y la desviación estándar de la luminancia del píxel 330, 190 en 24 horas de muestreo	39

Figura 25. Valores de la distancia de Mahalanobis del píxel 330, 190 en 24 horas de muestro	40
Figura 26. Gráficas superpuestas calculadas para el píxel 330, 190 en 24 horas de muestreo	41
Figura 27. Valores de la luminancia del píxel 300,588 en 24 horas de muestreo	42
Figura 28. Gráfica de la media y la desviación estándar de la luminancia del píxel 333, 588 en 24 horas de muestreo	43
Figura 29. Valores de la distancia de Mahalanobis del píxel 300, 588 en 24 horas de muestro	43
Figura 30. Gráficas superpuestas calculadas para el píxel 300, 588 en 24 horas de muestreo	44
Figura 31. Nivel de acertividad de la detección de movimiento en horario diurno	46
Figura 32. Nivel de acertividad de la detección de movimiento en horario nocturno	47

ÍNDICE DE TABLAS

Tabla 1. Valores de media y varianza de los píxeles en el tiempo	24
Tabla 2. Resultados del cálculo de la distancia de Mahalanobis para cada píxel.	33
Tabla 3. Cálculo de la Sensibilidad y Predictividad positiva para determinar el Umbral de día.	45
Tabla 4. Cálculo de la Sensibilidad y Predictividad positiva para determinar el Umbral de noche.	45

RESUMEN

Esta investigación cuyo título es Diseño y simulación de un sistema de detección de movimiento basado en segmentación de imágenes utilizando los métodos de umbralización y sustracción de fondo aplicado a un sistema de video vigilancia, es la concreción de un proyecto de Pre-Grado titulado Sistema de video vigilancia multicámara con detección de movimiento y emisión de alarma vía RF. Como objetivo principal, se investigó el estado del arte de los sistemas computacionales de detección de movimiento que ha permitido rediseñar el sistema de video vigilancia. Para la sustracción de fondo, que es la base de la detección de movimiento, se aplican los principios de Gaussiana simple y se calcula la distancia de Mahalanobis que permite determinar la similitud entre imágenes. Para la umbralización, se analizó una gran cantidad de muestras, a fin de obtener una base de datos que permita definir el nivel de umbral que mida la similitud entre imágenes. Los resultados que se obtuvieron ofrecieron altos índices de sensibilidad y predictividad positiva hasta en un 95%.

Palabras clave: Gaussiana simple, distancia de Mahalanobis, umbralización, sensibilidad, predictividad positiva.

ABSTRACT

This research entitled Design and simulation of a motion detection system based on segmentation of images using the thresholding and background subtraction methods applied to a video surveillance system, is the conclusion of a Pre-Grade project entitled Multi-camera video surveillance system with motion detection and alarm output via RF. As a main objective, the state of the art of the motion detection computer systems was investigated, which has allowed us to redesign the video surveillance system. For background subtraction, which is the basis of motion detection, Simple Gaussian is applied and Mahalanobis is calculated to determine the similarity between images. For the thresholding, a large number of samples were analyzed, in order to obtain a database that allows to define the threshold level that measures the similarity between images. The results obtained had high sensitivity and predictivity indexes up to 95%.

Key words: Simple Gaussian, Mahalanobis distance, thresholding, sensitivity, positive predictive

INTRODUCCIÓN

La sensación de inseguridad e incremento de los crímenes que se vive en el país sumado a la dificultad de las autoridades para prevenir el delito, es una constante amenaza a la calidad de vida de los peruanos. Esta sensación, sin embargo, no es exclusiva de las personas sino también de las empresas y organizaciones que tienen la necesidad de implementar sistemas de seguridad que les permita proteger su patrimonio.

Según la publicación Perú: La victimización de las empresas del INEI (2016), el 21.3% de las empresas han sufrido algún hecho delictivo y “el robo o hurto es la forma más común de hecho (delictivo) que afecta a las empresas con 79,8%”. Además, de este universo de 226,116 empresas, el 22.5 % son microempresas y el 23% son pequeñas empresas.

También señala este artículo, que “el 51,5% manifiesta que destinó menos del 3% en gastos de seguridad y/o vigilancia”. Adicionalmente indica que “para el 2016, de las empresas que utilizaron medidas de seguridad, el 70,1% de empresas utilizó sistema de video y captura de imágenes”. Otros incluyen radios de comunicación y personal de seguridad.

Pese a que los índices de seguridad son elevados y que las principales víctimas son las empresas pequeñas, la mayoría destinan pocos recursos en su seguridad. Además, la mayoría de empresas cuenta con sistemas existentes, por lo que incrementar los niveles de seguridad

sin incurrir en gastos mayores es prioritario. Las cámaras y las radios que conforman sus sistemas de seguridad, son herramientas independientes que no se comunican entre sí y pueden ser mal utilizadas por el personal de seguridad.

El sistema de video vigilancia multicámara con detección de movimiento y emisión de alarma vía RF, resuelve el problema de falta de optimización de recursos integrando dos sistemas: cámaras y radios de comunicaciones en banda libre. El sistema detecta movimiento analizando las imágenes recibidas de las cámaras, mientras que emite las alarmas de voz a través de las radios. Agrega además las funcionalidades de registro de alarmas e imágenes con la finalidad de que el usuario tenga acceso a la información para los fines que ellos estimen convenientes.

El sistema funciona y cumple con su objetivo, sin embargo, los métodos utilizados para detectar movimiento no permitían la medición de los resultados más allá de lo que la experiencia pudiera permitir. Para resolver este problema, se investigaron las técnicas de sustracción de fondo aplicadas a la detección de movimiento teniendo en cuenta que la aplicación final es la video vigilancia.

Las técnicas más sencillas y conocidas calculan la diferencia entre imágenes consecutivas o entre una imagen de fondo estática y una imagen actual, donde el modelamiento de la imagen de fondo se obtiene con una imagen inicial fija. En un nivel intermedio, están las técnicas que utilizan variables estadísticas como la media o la mediana de las imágenes anteriores para modelar el fondo de la escena. Por último, los métodos más robustos asocian el valor de intensidad de cada píxel en el tiempo con una (o varias) distribución unimodal Gaussiana realizando el modelado del fondo por cada píxel.

El objetivo general consiste en diseñar un sistema de visión computacional que mediante el análisis de imágenes digitales recibidas

desde más de dos cámaras IP, permita la detección de movimiento aplicando la sustracción de fondo y la estimación de un valor de umbral para aplicarse a un sistema de video vigilancia de interiores en plataforma .Net C Sharp.

Los objetivos específicos son:

- Investigar y analizar el estado del arte de los métodos aplicados para la detección de movimiento en imágenes digitales tomadas en ambientes de interiores.
- Aplicar el método de umbralización de imágenes, fijando el umbral a partir de las pruebas y resultados obtenido en diferentes casos.
- Aplicar el método de sustracción de fondo a partir de la técnica que según el estado del arte ofrezca resultados aceptables para esta aplicación y que facilite también la segmentación de imágenes.
- Comprobar que los métodos pueden ser aplicados en la video vigilancia con respuestas en tiempo real.

La presente tesis comprende cuatro (4) capítulos. El primero trata sobre el estado del arte y las bases teóricas comprendidas dentro del marco teórico. En el segundo, se describen la metodología utilizada para la implementación del sistema, además de detallar los materiales empleados. El tercero describe el desarrollo del proyecto incluyendo el modelamiento de cada uno de los métodos, así como el análisis de los métodos seleccionados para la detección de movimiento. En el cuarto se presentan las pruebas realizadas sobre una muestra de imágenes y los resultados obtenidos. Para finalizar, se presentan las conclusiones y recomendaciones de la presente investigación.

CAPÍTULO I

MARCO TEÓRICO

1.1 Antecedentes

Para el desarrollo de esta tesis, se investigó el estado del arte de las técnicas que se utilizan para la sustracción de fondo y umbralización que son los métodos mediante los cuales se realiza la detección de movimiento.

Existen varios algoritmos para la detección de movimiento que por ejemplo, se basan en el promedio temporal de una secuencia de imágenes como Friedman y Russell (2013) o Ridder, Munkelt y Kirchner (1995). Están también los que utilizan la estimación Gaussiana Adaptiva como Köhle, Merkl y Kastner (1997) o la estimación de parámetros basados en procesos de píxeles como Sun, Feng y Tan (2000) o Grimson, Stauffer, Romano y Lee (1998).

Algunos sistemas proponen actualizar las imágenes de fondo basados en la mezcla de distribuciones Gaussianas aplicada para cada valor del píxel con el fin de adaptarse a la variación de la iluminación y alteración de los fondos como Stauffer y Grimson (1999). Otros autores proponen la obtención y actualización del fondo a nivel de píxel, a nivel región y a nivel de trama como Toyama, Krumm, Brumitt y Meyers (1999). En otros casos, se propone un modelo estadístico para representar cada píxel con los valores de intensidad máxima, mínima y la diferencia de intensidad máxima entre tramas consecutivas analizadas durante el

período de inicialización y actualizadas periódicamente como Haritaoglu, Harwood y Davis (2000).

Existen muchas clasificaciones a las técnicas de sustracción de fondo para objetos en movimiento como Picardi (2004), Cheung y Kamath (2004) o Cucchiara (2006). Cada uno de estos estudios analiza y clasifica las técnicas con diferentes criterios y las categorías en que se realiza la clasificación varían dependiendo de la aplicación. De manera general, se sabe que en interiores se diseñan algoritmos sencillos y de rápida ejecución, como el caso de Diferencia de imágenes de Leng y Dai (2007) o Media Móvil de Tang, Miao y Wan (2007) que son sensibles al ruido. En el caso de ambientes de exteriores, se requiere de la aplicación de técnicas más robustas como la Mezcla de Gaussianas como indican Stauffer y Grimson (1999) o Estimación de Densidad Kernel según Elgammal, Harwood y Davis (2000).

La sustracción de fondo es el método más utilizado para la detección de movimiento y las técnicas más utilizadas en el procesamiento de imágenes para la sustracción de fondo se listan a continuación:

- Sustracción de fondo basado en diferencia de imágenes.
- Sustracción de fondo basado en la media de los “N” píxeles anteriores.
- Sustracción de fondo basado en la mediana de los “N” píxeles anteriores.
- Sustracción de fondo basado en la media móvil.
- Sustracción de fondo basado en la Gaussiana simple.
- Sustracción de fondo basado en mezcla de Gaussianas.
- Sustracción de fondo basado en estimación de densidad del Kernel.
- Sustracción de fondo basado en estimación mediante la técnica mean-shift.

- Sustracción de fondo basado en coocurrencias de variaciones en la imagen.
- Sustracción de fondo basado en Autobackgrounds ("Eigenbackgrounds").

Para el caso de uso en interiores de los sistemas de video vigilancia, los primeros cinco métodos listados son lo más utilizados según el estado del arte, ya que todos ellos tienen un bajo costo computacional.

Piccardi (2004) realiza una evaluación de las técnicas más robustas para la sustracción de fondo, parte de las que menciona y que son aplicables a interiores se listan a continuación.

1.1.1 Sustracción de fondo basada en diferencia de imágenes

El fondo se actualiza con la imagen obtenida en el instante t anterior:

$$F_{(t)} = I_{(t-1)} \quad \text{Ecuación 1}$$

Donde $F_{(t)}$ es la imagen de fondo e $I_{(t-1)}$ representa la imagen anterior. Las imágenes se evalúan en el tiempo y son clasificados como fondo si se cumple la siguiente desigualdad:

$$I_{(t)} - F_{(t)} < U \quad \text{Ecuación 2}$$

Donde $I_{(t)}$ representa la imagen actual, $F_{(t)}$ es la imagen de fondo y U es un valor de umbral. Esta técnica puede aplicarse de dos formas: La primera con imágenes consecutivas $I_{(t-1)}$ e $I_{(t)}$, donde la imagen anterior $I_{(t-1)}$ se considera la imagen de fondo asumiendo que la imagen de fondo es una imagen estática que se va actualizando cada cierto periodo T .

1.1.2 Basada en la media de los “N” píxeles anteriores

Según esta técnica, el fondo se modela mediante el cálculo de la media de los “N” píxeles anteriores, aplicando la ecuación 3.

$$F_{(t)} = \frac{\sum_{i=1}^{i=t} I_{(t)}}{t} \quad \text{Ecuación 3}$$

Donde $F_{(t)}$ es la imagen de fondo e $I_{(t)}$ representa a la imagen actual. Los píxeles se van evaluando en el tiempo y son clasificados como fondo si se cumple la desigualdad mostrada en la ecuación 2.

1.1.3 Basada en la mediana de los “N” píxeles anteriores

Esta técnica es similar a la mencionada previamente, con la diferencia que el fondo se modela mediante el cálculo de la mediana de los “N” píxeles anteriores, aplicando la ecuación 4.

$$F_{(t)} = \text{mediana} (I_{(t-1)}, I_{(t-2)}, \dots, I_{(t-n)}) \quad \text{Ecuación 4}$$

Donde $F_{(t)}$ es la imagen de fondo e $I_{(t-1)}$ e $I_{(t-2)}$ representa las imágenes anteriores en tiempos anteriores diferentes. Las imágenes se evalúan en el tiempo y son clasificados como fondo si se cumple la desigualdad mostrada en la ecuación 2.

1.1.4 Basada en la media móvil

Con esta técnica el fondo se modela y actualiza según la Ecuación 5.

$$F_{(t)} = \alpha I_{(t)} + (1 - \alpha) F_{(t-1)} \quad \text{Ecuación 5}$$

Donde $F_{(t)}$ es la imagen de fondo actual, $I_{(t)}$ representa la imagen actual, α es el factor de aprendizaje cuyo valor debe estar entre 0 y 1 y $F_{(t-1)}$ es la imagen de fondo anterior.

1.1.5 Basada en la Gaussiana Simple

Cuando se analizan los valores que ha tomado un píxel en una secuencia de imágenes por un período, su Función de Densidad de Probabilidad (FDP) se puede aproximar una Función de Densidad de Probabilidad Gaussiana que tiene una media (μ) y una varianza (σ^2) como se muestra en la Ecuación 6.

$$FDP(x) = \eta(x - \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(\mu-x)^2}{2\sigma^2}} \quad \text{Ecuación 6}$$

Según Wren, Azarbayejani, Darrell y Pentland (1997), se puede estimar la media y la varianza en función de los valores que el píxel ha tomado en el tiempo y del valor actual. En cada instante t se evalúa el valor del píxel y se determina si está dentro de una región determinada. De ser así es considerado como fondo, de lo contrario es primer plano.

Varios autores utilizan el principio de la Gaussiana Simple como en el caso de Gordon, Darrell, Harville y Woodfill (1999) en cuyo caso lo utilizan para segmentar el movimiento de la imagen estimando para cada componente de color R, G y B de un píxel una distribución Gaussiana. Otro uso del mismo principio es el filtro de Kalman que en palabras de Cervera (2010) consiste en “una técnica recursiva de seguimiento lineal en la que los valores de los píxeles pueden modelarse por una distribución Gaussiana. La media y la varianza del fondo se actualizan utilizando filtros adaptativos para acomodar los cambios de iluminación” (p.16).

1.2 Bases teóricas

Existen algunos conceptos básicos que es necesario definir antes de iniciar con la explicación detallada del desarrollo de esta investigación.

1.2.1 Imagen digital

Una imagen digital es un arreglo que puede ser representada de forma matricial o vectorial y que contiene información digitalizada en cada uno de los píxeles que la conforman.

En forma discreta, los índices de las filas y las columnas indican la posición del píxel en la imagen y cada píxel en la matriz tiene un determinado valor de intensidad o brillo independientemente del modelo de color elegido. El tamaño de la imagen digital se sabe es de $M \times N$ píxeles, M denota el número de columnas y N el de filas. Para el caso de la imagen de la Figura N°1 su dimensión es de 440×443 píxeles.



Figura 1. Representación de la estructura de una imagen digital.

Fuente: Bahamón (2011)

1.2.2 Modos de color

Los modelos o modos de color son modelos matemáticos que permiten representar los colores en forma numérica. Utilizan normalmente tres o cuatros componentes de color. Por ejemplo, para el modelo RGB, cada píxel posee información de los tres componentes de color rojo, verde y azul. Cada componente posee 8 bits de información, es decir, el píxel posee 24 bits en total. Para cada píxel existen más 16 millones de posibles combinaciones de color.



Figura 2. Descomposición de una imagen RGB en sus tres componentes.

Fuente: Bahamón (2011)

1.2.3 Visión computacional

Los sistemas de Visión Computacional intentan simular la capacidad del ser humano de interpretar, analizar y brindar resultados precisos a partir de las imágenes captadas por cámaras permitiendo resolver problemas o situaciones con igual o mayor exactitud que el ojo humano.

La visión computacional está altamente relacionada con el procesamiento de imágenes. Este último es quien se encarga de hacer las mejoras de las imágenes que se van a procesar y que serán utilizadas en el proceso de obtención e interpretación de datos que se encarga de realizar la visión computacional.

Una aplicación de la visión computacional se muestra en la figura 3 en la que se determina la distancia entre un objeto y la cámara que lo registra.

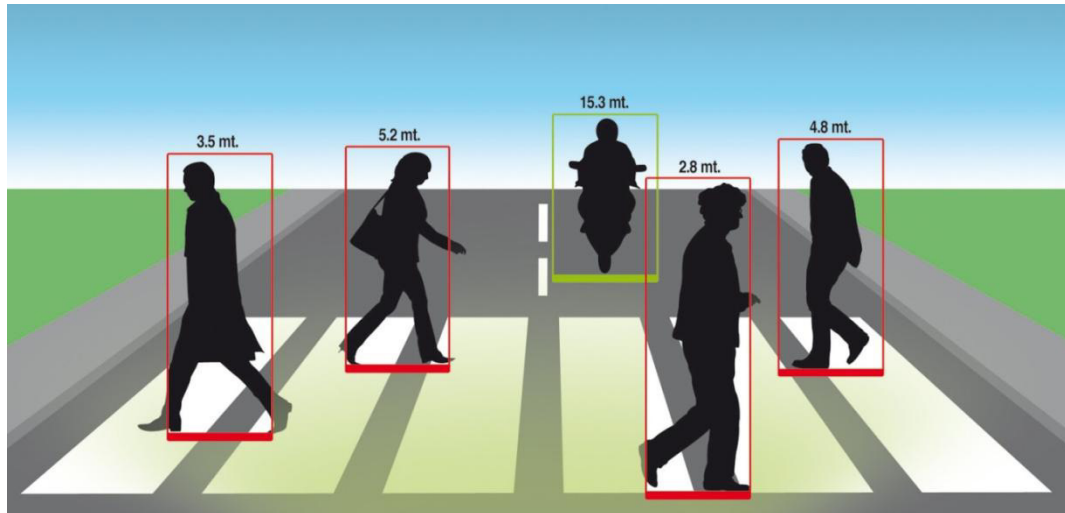


Figura 3. Ejemplo de aplicación de la visión computacional.

Fuente: Guevara (2013)

1.2.4 Procesamiento de imágenes digitales

El objetivo del procesamiento digital de imágenes es obtener una imagen resultante que permita obtener información relevante para una aplicación específica. Quiere decir que los métodos no son siempre los mismos en todos los procesamientos, sino que dependen de los resultados que se esperan. En general, el resultado esperado debe mejorar la calidad de la imagen o ayudar la búsqueda de la información.

Dentro de las técnicas más utilizadas en el procesamiento de imágenes encontramos la aplicación de filtros para eliminación de ruido (mediana, media, gaussiano), realce de bordes, ecualización de imágenes, etc.

Por ejemplo, en la figura 4 se muestra la aplicación del filtro promedio. En la imagen a) se observa la imagen

original, en b) está la imagen con un filtro de promedio de 3 x 3 y en c) se muestra imagen con un filtro de promedio de 5 x 5

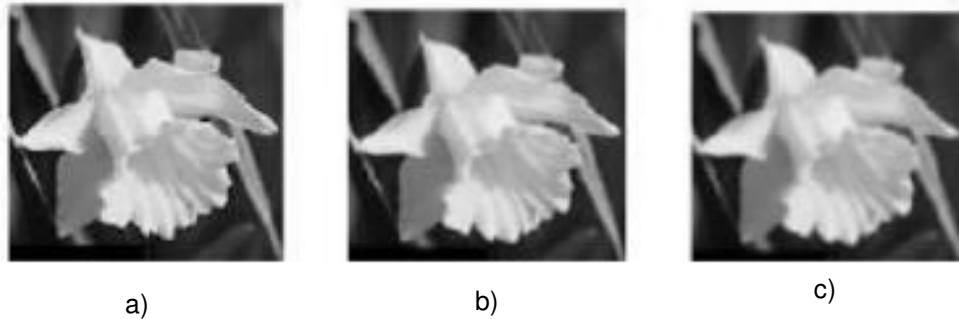


Figura 4. Ejemplo de aplicación del procesamiento de imágenes.

Fuente: Álvarez, Barbará Rodríguez (2010)

1.2.5 Segmentación de imagen

La segmentación de imágenes es el método más conocido del procesamiento de imágenes. Consiste en seleccionar partes de la imagen y elegir aquello que es relevante para un determinado objetivo. Dentro de las técnicas más conocidas se encuentra por ejemplo la binarización como se muestra en la figura N° 5.



Figura 5. Imagen binarizada, compuesta solo por blanco y negro.

Fuente: Valverde-Rebaza (2009)

1.2.6 Sustracción de fondo

La sustracción de fondo es un método muy utilizado para detectar objetos en movimiento sobre una secuencia de imágenes digitales ya que permite diferenciar en una imagen el fondo y el primer plano. Según este método, los elementos de una imagen que son estáticos son considerados como fondo, mientras que todos los elementos móviles son considerados primer plano, es decir, se tratan de los nuevos objetos en la imagen.

Para definir el fondo se debe pasar por un primer proceso de inicialización de parámetros que se obtiene luego de un tiempo de observación inicial, también conocido como entrenamiento. De esta manera, se logra modelar un fondo que permitirá determinar si los píxeles observados pertenecen al fondo o al primer plano. La figura N° 6 muestra un método muy sencillo de sustracción de fondo donde se realiza la resta aritmética píxel a píxel de los valores de luminancia.

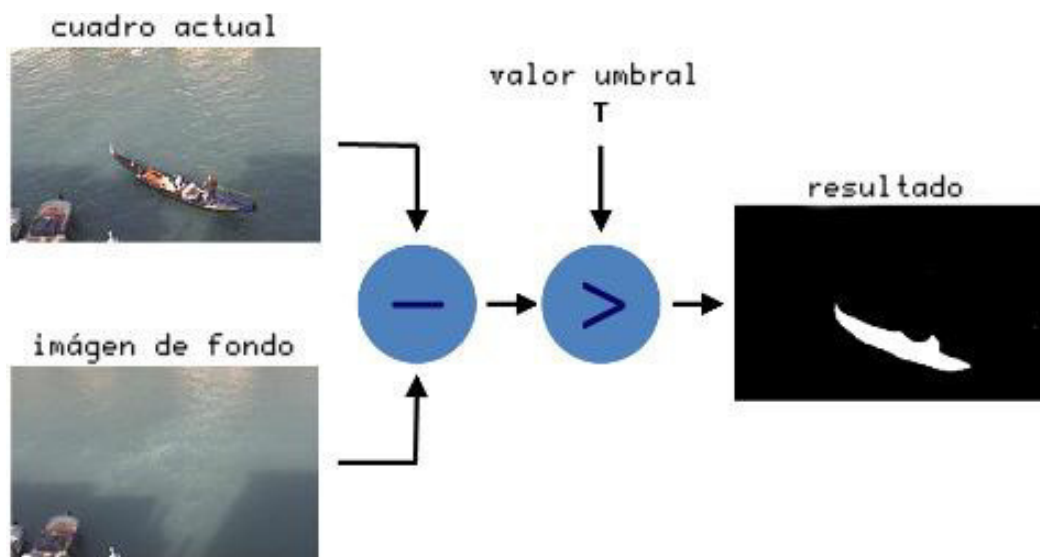


Figura 6. Ejemplo de una imagen resultante de la sustracción de fondo.

Fuente: Yabo A. (2015)

El procesamiento para el modelado de fondo se realiza a partir de los valores de los píxeles, expresados en su valor de luminancia. De esta forma, la decisión de pertenencia o no al fondo se realizará por cada píxel independientemente del valor que puedan tener sus píxeles vecinos.

1.2.7 Función de distribución de probabilidad

Según Rondón (2018), la distribución de probabilidad de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable aleatoria la probabilidad de que dicho suceso ocurra. Está definida sobre el conjunto de todos los sucesos, cada uno de los sucesos es el rango de valores de la variable aleatoria.

1.2.8 Función de distribución de probabilidad normal o campana de Gauss

Según García (2010) la distribución normal es una distribución de probabilidad de variable continua que describe los datos que se agrupan en torno a un valor central. Todo proceso en el que solo existan causas aleatorias de variación sigue una ley de distribución normal. Esta condición que aparece con frecuencia en fenómenos naturales (de ahí que se la denomine “normal”), puede obtenerse en los procesos industriales si los procesos se llevan a un estado en el que solo existen causas comunes de variación.

Una distribución normal se caracteriza porque los valores de las mediciones tienden a agruparse alrededor de un punto central correspondiente a la media, la representación de los datos es simétrica a ambos lados de la media y las desviaciones estándares quedan situadas a igual distancia unas de otras.

Como se muestra en la figura N° 7, la proporción de mediciones situada entre la media y las desviaciones es una constante en la que la media \pm una vez la desviación estándar cubre el 68,3% de los casos, la media \pm dos veces la desviación estándar cubre el 95,5% de los casos y la media \pm tres veces la desviación estándar cubre el 99,7% de los casos.

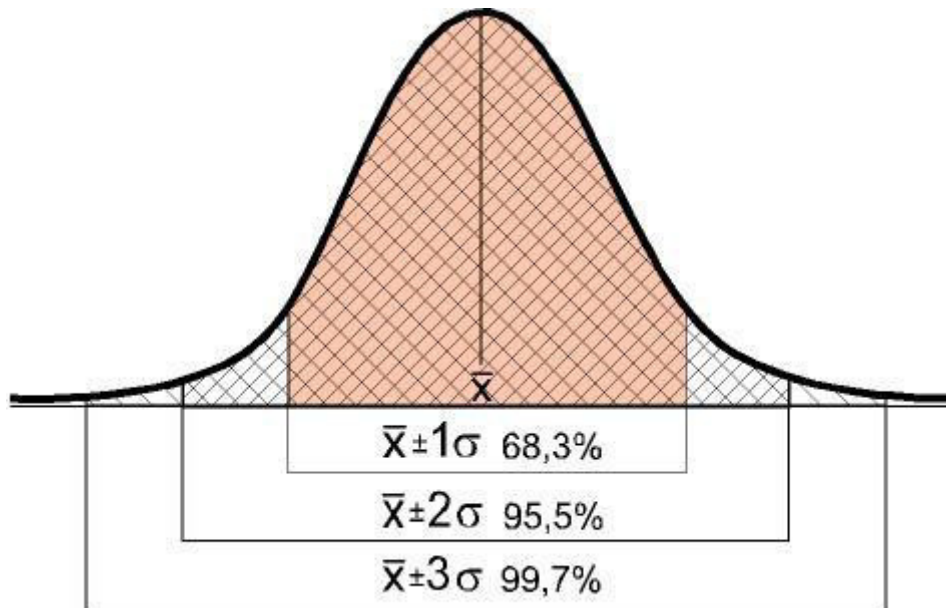


Figura 7. Ejemplo de una distribución normal o Gaussiana.

Fuente: García (2010).

1.3 Definición de términos básicos

Durante la lectura de la presente investigación, se encontrarán términos que es necesario aclarar para un mejor entendimiento de la lectura.

1.3.1 Fondo o fondo unimodal

Se define como fondo o fondo unimodal a todos aquellos píxeles que pueden describirse de una misma manera y que prácticamente se mantienen constantes durante un período de tiempo.

1.3.2 Primer plano

Es el conjunto de píxeles que está formado por los elementos que son nuevos en la imagen y que representan movimiento en la escena.

1.3.3 Método

El método es una forma particular de ejecutar una acción de forma organizada, sistemática, y/o estructurada. Se refiere a una o varias técnicas que se ejecutan en orden para realizar una función. Para la presente investigación, llamaremos método a la sustracción de fondo y a la umbralización, que son los métodos seleccionados para la detección de movimiento.

1.3.4 Técnica

Una técnica es conjunto de procedimientos que tiene como finalidad obtener un resultado proyectado. Para la presente investigación, se verán que hay diferentes técnicas para la sustracción de fondo que, como se explicó, es uno de los métodos para la detección de movimiento.

1.3.5 Aplicación

Para este caso se llamará aplicación a la detección de movimiento. Se refiere al uso final que tendrá los métodos de sustracción de fondo y umbralización en la presente tesis.

1.3.6 Sensibilidad

Según Pita y Péregas (2010), la sensibilidad es la probabilidad de clasificar correctamente a un individuo enfermo, es decir, la probabilidad de que para un sujeto enfermo se obtenga en la prueba un resultado positivo. La sensibilidad es, por lo tanto, la capacidad del test para detectar la enfermedad. Para la aplicación, es la capacidad de detectar movimiento.

1.3.7 Predictividad positiva

Según Pita y Péregas (2010), la predictividad positiva es la probabilidad de padecer la enfermedad si se obtiene un resultado positivo en el test. El valor predictivo positivo puede estimarse, por tanto, a partir de la proporción de pacientes con un resultado positivo

CAPÍTULO II METODOLOGÍA

Para la implementación de esta investigación se utilizaron los siguientes materiales:

- 02 cámaras IP TrendNet modelo TV-IP501P (día)
- 01 cámara IP TrendNet modelo TV-IP762IC (día/noche)
- 01 cámara IP TrendNet modelo TV-IP100 (día)
- 03 radios Motorola Two-Way modelo MD200R
- 01 PC procesador Intel Core i5, cuarta generación de 2.6 GHz, 8 Gb de memoria RAM con Sistema Operativo Windows 7 Professional Server Pack 1 de 64 bits.

Como parte de la metodología empleada para el desarrollo de la presente investigación, se modelaron cada una de las etapas aplicadas a la detección de movimiento y emisión de alarmas, una vez definiendo el método de sustracción de fondo basado en la Gaussiana Simple. Cada una de las fórmulas indicadas en los siguientes puntos, se aplicaron al sistema para obtener el resultado deseado.

2.1 Modelamiento de la obtención de la imagen unidimensional en el vector de Luminancia

Como se había indicado en el Capítulo I, las imágenes digitales en modelo RGB están constituidas por un arreglo de píxeles y cada píxel a su vez, está formado por tres (3) componentes el Rojo, Verde y Azul.

Cada uno de los componentes del píxeles cuenta con 8 bits de información, quiere decir que cada píxel tiene 24 bits de información para su procesamiento.

En síntesis, una imagen de resolución 640 x 352 tiene en total 225,280 píxeles y 5'406,720 bits de información. De considerarse el procesamiento de uno de los componentes, se deberían de procesar 1'802,240 bits, la tercera parte de la información de la imagen original.

Por lo tanto, es preferible realizar el procesamiento de un arreglo unidimensional cuyos valores de píxel sean representativos de las tres (3) componentes de color. Se calculó la luminancia de cada píxel para realizar sobre estos datos los posteriores procesamientos.

El modelo de luminancia se obtiene aplicando la Ecuación 7 sobre la imagen original.

$$L = 0.299 I_R + 0.587 I_G + 0.114 I_B \quad \text{Ecuación 7}$$

Donde I_R es el valor del componente rojo de la imagen, I_G es el valor del componente verde de la imagen, I_B es el valor del componente azul de la imagen y L es la imagen unidimensional de luminancia resultante.

2.2 Modelamiento del fondo aproximando la Función de Densidad de Probabilidad Gaussiana por cada píxel basado en el principio de Gaussiana Simple

Cuando se analizan los arreglos que se pueden crear a partir de los valores de luminancia que toma un píxel en una secuencia de imágenes unimodales, cada uno de estos arreglos se puede aproximar a una Función de Densidad de Probabilidad (FDP) Gaussiana, como lo postula Wren, Azarbajani, Darrell y Pentland (1997).

La Función de Densidad de Probabilidad Gaussiana, como se expresó en la Ecuación 6, es una función dependiente de la media (μ) y la

varianza (σ^2) de los datos, que se obtienen aplicando las Ecuaciones 8 y 9:

$$\mu_t(x) = \sum_{i=1}^{i=t} \frac{L_i(x)}{t} \quad \text{Ecuación 8}$$

$$\sigma_t^2(x) = \sum_{i=1}^{i=t} \frac{L_i^2(x)}{t} - \mu_t^2(x) \quad \text{Ecuación 9}$$

Donde μ_t es la media de la Función de Densidad de Probabilidad del píxel en el instante t , L_i es el valor del píxel luminancia en cada instante t , y σ_t^2 es la varianza en el momento t .

2.3 Modelamiento de la detección de primer plano mediante el cálculo de la distancia de Mahalanobis y la estimación del valor umbral

La distancia de Mahalanobis es una medida de probabilidad estadística basada en la correlación entre las variables, que permite determinar una similitud estadística entre una muestra desconocida y conocida.

$$D_{Mh}(x, y) = \sqrt{(x - y)^T \Sigma^{-1} (x - y)} \quad \text{Ecuación 10}$$

Donde Σ es la matriz de covarianzas, x e y son las variables multidimensionales a las que se desea aplicar la distancia de Mahalanobis y D_{Mh} es la distancia de Mahalanobis.

La matriz de covarianzas es equivalente a la varianza de la variable unidimensional. Para el caso presentado y tratándose de variables aleatorias unidimensionales (valor del píxel), la Distancia de Mahalanobis se define como:

$$D_{Mh}(x, y) = \sqrt{(x - y)^T \sigma^2^{-1} (x - y)} \quad \text{Ecuación 11}$$

En notación escalar, la Ecuación 11 se reduce a:

$$D_{Mh}(x, y) = \sqrt{(x - y)^2 \sigma^{-2}} \quad \text{Ecuación 12}$$

Finalmente, la distancia de Mahalanobis queda reducida a:

$$D_{Mh}(x, y) = \frac{(x - y)}{\sigma} \quad \text{Ecuación 13}$$

La aplicación de esta fórmula en la presente investigación se define como:

$$D_{Mh}(x, \mu) = \frac{|(L(x) - \mu)|}{\sigma} \quad \text{Ecuación 14}$$

Donde $L(x)$ es el valor del píxel actual, μ es la media obtenida de modelamiento del fondo y σ es la desviación estándar del modelamiento del fondo.

Se espera que los valores de luminancia de píxel que estén dentro de una región delimitada por un umbral estimado a partir del cálculo de la distancia de Mahalanobis, sean los píxeles que modelan el fondo, mientras que los píxeles del primer plano se deben encontrar fuera de dicho margen.

2.4 Modelamiento de la segmentación de imágenes para la detección de primer plano

Para obtener el resultado deseado, se requiere de realizar la segmentación de las imágenes que permita la umbralización de la imagen a partir de la fijación de un nivel de umbral que permita determinar si un píxel corresponde al primer plano o al fondo y luego reemplazar los píxeles que han sido detectados como primer plano por píxeles saturados de rojo sin presencia de los componentes verde y azul. El resultado es obtener los segmentos de la escena que representan movimiento en color rojo permitiendo al usuario una rápida detección del evento. Esto

representa en el sistema la alarma visual. La segmentación se puede representar matemáticamente bajo la aplicación de las Ecuaciones 15, 16 y 17:

$$\begin{array}{l}
 IPP_R(x) \quad \left\{ \begin{array}{ll} 255, & L(x) \geq U \\ I_R, & L(x) < U \end{array} \right. \quad \text{Ecuación 15} \\
 \\
 IPP_G(x) \quad \left\{ \begin{array}{ll} 0, & L(x) \geq U \\ I_G, & L(x) < U \end{array} \right. \quad \text{Ecuación 16} \\
 \\
 IPP_B(x) \quad \left\{ \begin{array}{ll} 0, & L(x) \geq U \\ I_B, & L(x) < U \end{array} \right. \quad \text{Ecuación 17}
 \end{array}$$

Donde IPP_R , IPP_G e IPP_B representan a los componentes rojo, verde y azul respectivamente de la imagen de primer plano I_R , I_G e I_B representan las componentes rojo, verde y azul de la imagen original en modelo de color RGB, $L(x)$ es el valor de luminancia del píxel actual en evaluación y U es el valor de umbral.

2.5 Modelamiento del fondo adaptativo Gaussiano

El modelo del fondo de la imagen debe tener la capacidad de adaptación en el tiempo para que ante la aparición de nuevos objetos que se mantengan estáticos en la escena puedan incluirse en el modelo de fondo.

Por lo tanto, la media (μ) y la varianza (σ^2) deben ir adaptándose en el tiempo aplicando las Ecuaciones 18 y 19 formuladas por Wren, Azarbayejani, Darrell y Pentland (1997).

$$\mu_{(t)} = (1 - \alpha)\mu_{(t-1)} + L_{(t)} \alpha \quad \text{Ecuación 18}$$

$$\sigma^2_{(t)} = \alpha(L_{(t)} - \mu_{(t-1)})^2 + (1 - \alpha)\sigma_{(t-1)}^2 \quad \text{Ecuación 19}$$

Donde μ es la media del modelo de fondo, L es el valor de luminancia del píxel actual, σ^2 es la varianza del modelo de fondo y α es la velocidad de adaptación del sistema.

CAPÍTULO III

DESARROLLO DEL PROYECTO

Se aplicaron los modelamientos presentados a una muestra limitada de imágenes seleccionadas que cumplieran la función de representar un fondo y un primer plano para cada caso. En el mismo orden en que se presentan se aplican en el sistema. En la figura N° 8 se muestra el diagrama de bloques general del sistema que se resume a dos bucles o ciclos de programación.

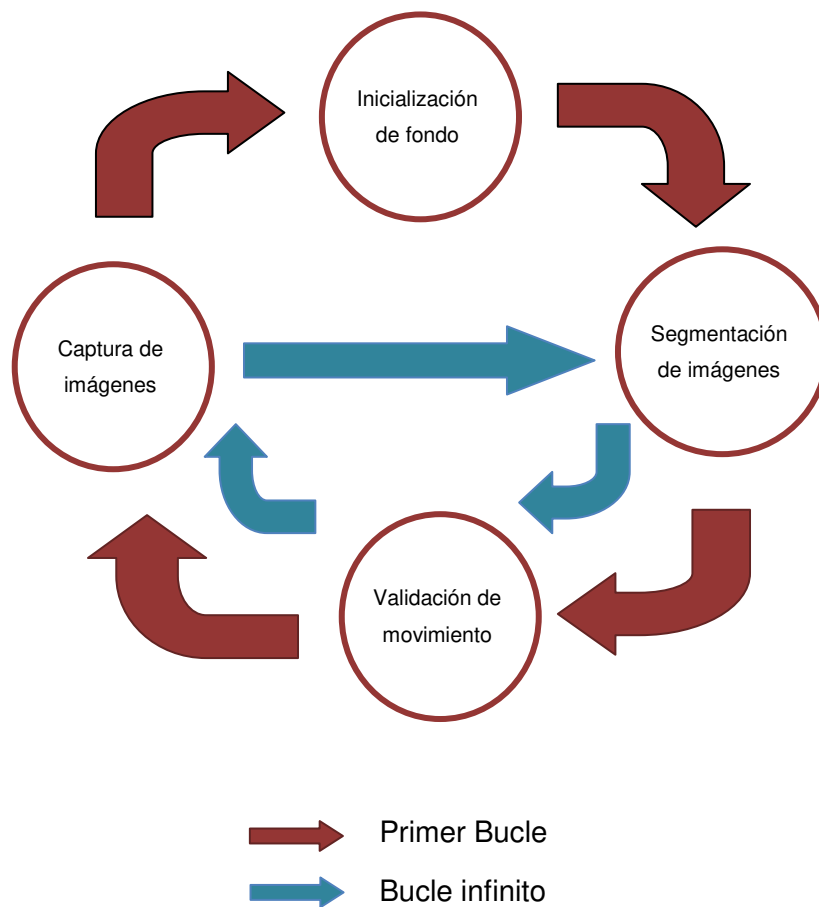


Figura 8. Diagrama de bloques del sistema
Elaboración: Los autores

3.1 Obtención de la imagen vectorial en el modelo de Luminancia

Cada una de las imágenes obtenidas que forman parte de una secuencia de video, están formadas por píxeles que a su vez tienen tres (03) componentes de color, como se explicó en el punto Modelamiento de la obtención de la imagen unidimensional en el modelo de Luminancia.

La ecuación 7 es aplicada a cada una de las imágenes para obtener los vectores Luminancia. Por ejemplo, se muestra el resultado de transformar píxeles RGB a luminancia con alto y bajo nivel de la misma en las figuras 9 y 10. En la figura 9 se muestra a la izquierda, la imagen en RGB y los valores de cada componente de color del píxel evaluado P. A la derecha se muestra la imagen de luminancia resultante y el valor de luminancia del mismo píxel P.

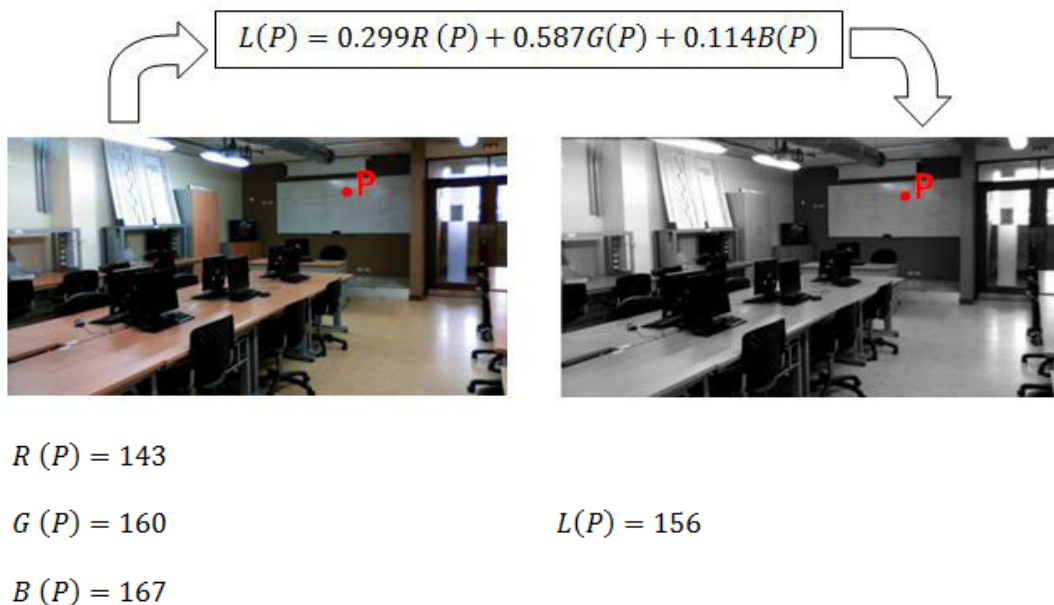


Figura 9. Píxeles con alto nivel de luminancia
Elaboración: Los autores

En la figura 10, se muestra a lado izquierdo la imagen en RGB y los valores de cada componente de color del píxel evaluado P. A la derecha se muestra la imagen de luminancia resultante y el valor de luminancia del mismo píxel P.



$$R(P) = 4$$

$$G(P) = 6$$

$$B(P) = 4$$

$$L(P) = 5$$

Figura 10. Píxeles con bajo nivel de luminancia
Elaboración: Los autores.

3.2 Función de Densidad de Probabilidad Gaussiana para representar los cambios de luminancia de un píxel en el tiempo

Los valores de luminancia que toma un píxel del fondo en una secuencia de imágenes, puede aproximarse a una Función de Densidad de Probabilidad Gaussiana. Para comprobarlo, se realizaron pruebas sobre cuarenta (40) imágenes de resolución 640 x 352 píxeles tomadas en un periodo T. La figura 11 es una representación de estas imágenes y de los píxeles que se eligieron como representantes del fondo de la imagen. Los siete (07) píxeles seleccionados presentan cambios de luminancia en el tiempo y tienen las siguientes ubicaciones en las coordenadas de las abscisas y ordenadas: P1: 100, 195; P2: 262,184, P3: 330, 70, P4: 80, 440, P5: 180, 440, P6: 270, 440, P7: 300, 440.

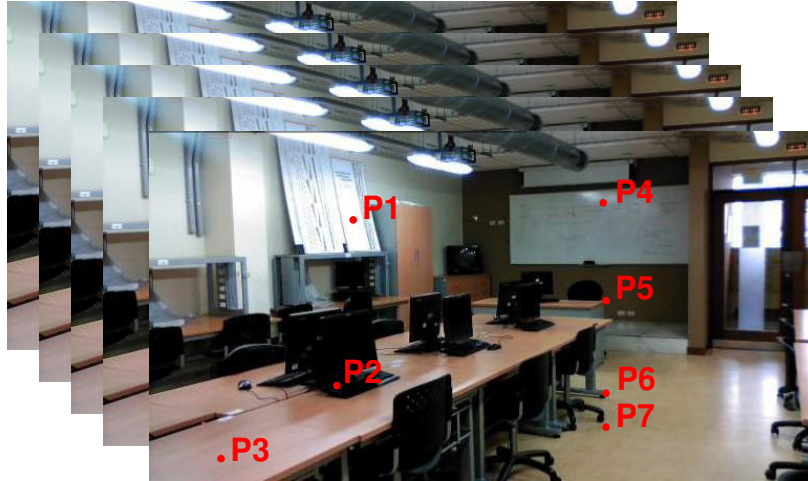


Figura 11. Píxeles representantes del fondo.

Elaboración: Los autores

Pasado el periodo T, se obtienen siete (07) arreglos o vectores de cuarenta (40) valores cada uno correspondiente a los valores de luminancia de cada píxel. Cada uno de estos arreglos, que corresponden a un conjunto de variables aleatorias, puede aproximarse a una Función de Densidad de Probabilidad Gaussiana mediante el cálculo de la media (μ) y la varianza (σ^2).

Aplicando las Ecuaciones 8 y 9 en cada uno de los arreglos, se obtienen los siguientes valores de media (μ) y varianza (σ^2) mostrándose en la tabla N°1.

Tabla 1. Valores de media y varianza de los píxeles en el tiempo

Arreglo de Píxel	Media (μ)	Varianza (σ^2)
P1	254.49	0.25
P2	6.00	11.08
P3	160.56	0.96
P4	153.21	4.33
P5	40.72	5.90
P6	155.74	15.22
P7	135.33	1.25

Elaboración: Los autores

Para comprobar que el valor del píxel en el tiempo se puede representar mediante una función de densidad de probabilidad (FDP) Gaussiana, se utilizó la herramienta Excel de Microsoft Office y se generaron dos gráficos para cada caso. El primero, corresponde al gráfico de línea de los valores de luminancia del píxel y el segundo es el resultado de aplicar la fórmula en Excel de FDP en el mismo arreglo dependientes de la (μ) y la varianza (σ^2) indicadas en la Tabla N°1.

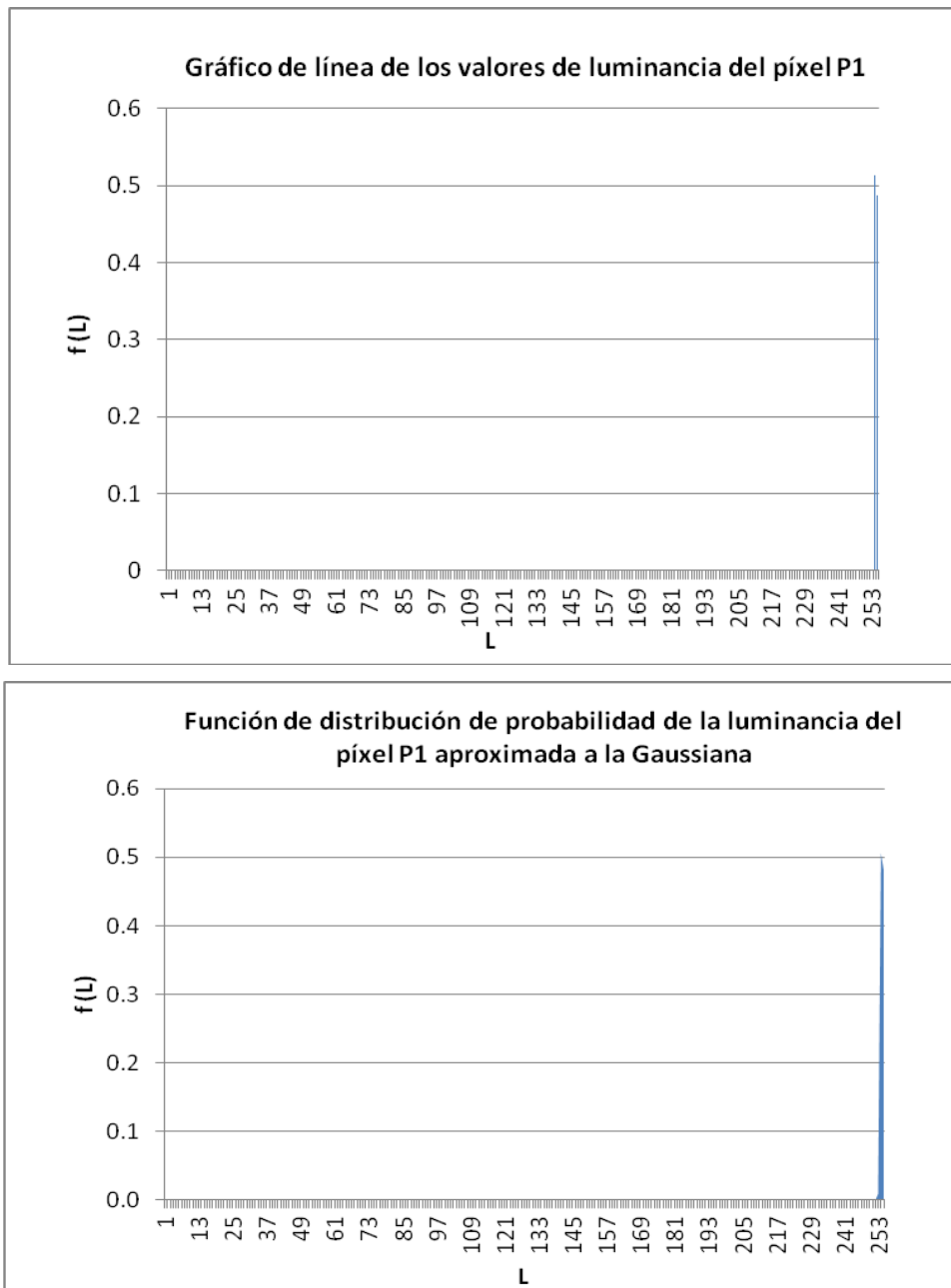


Figura 12. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P1

Elaboración: Los autores

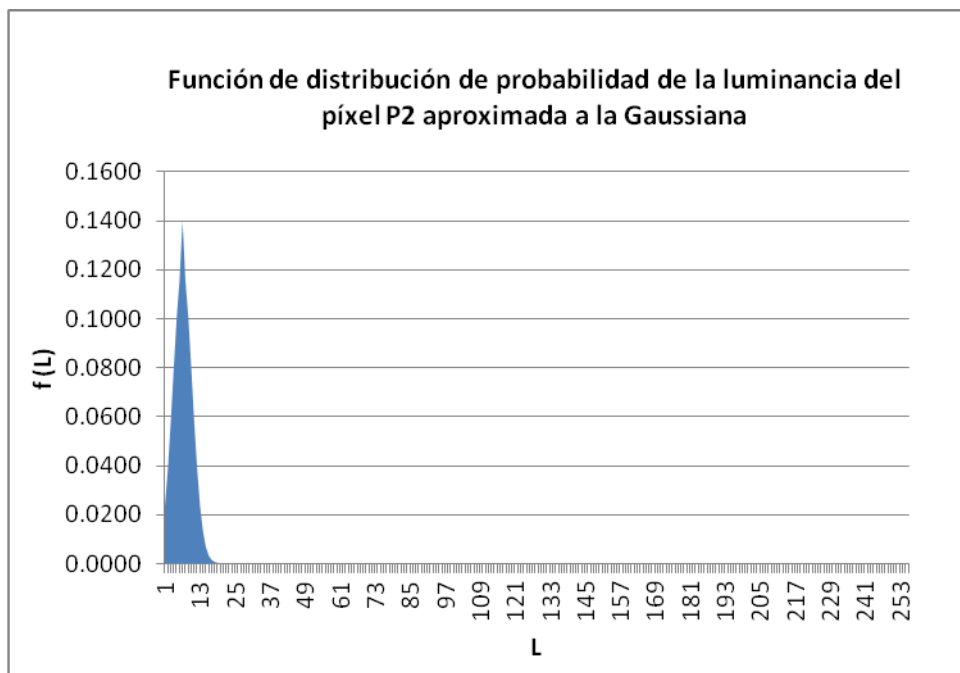
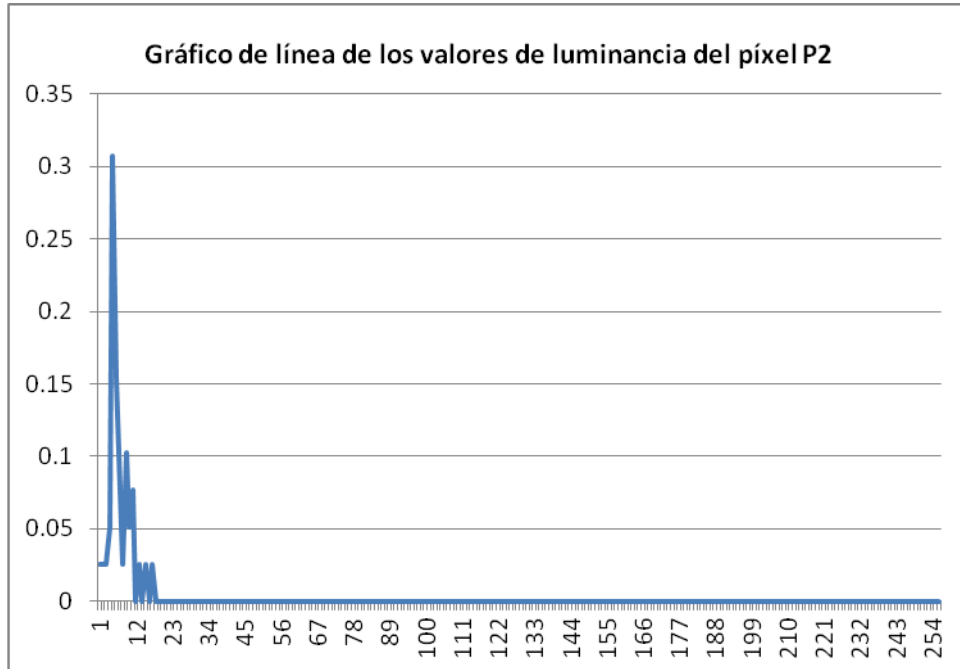


Figura 13. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P2

Elaboración: Los autores

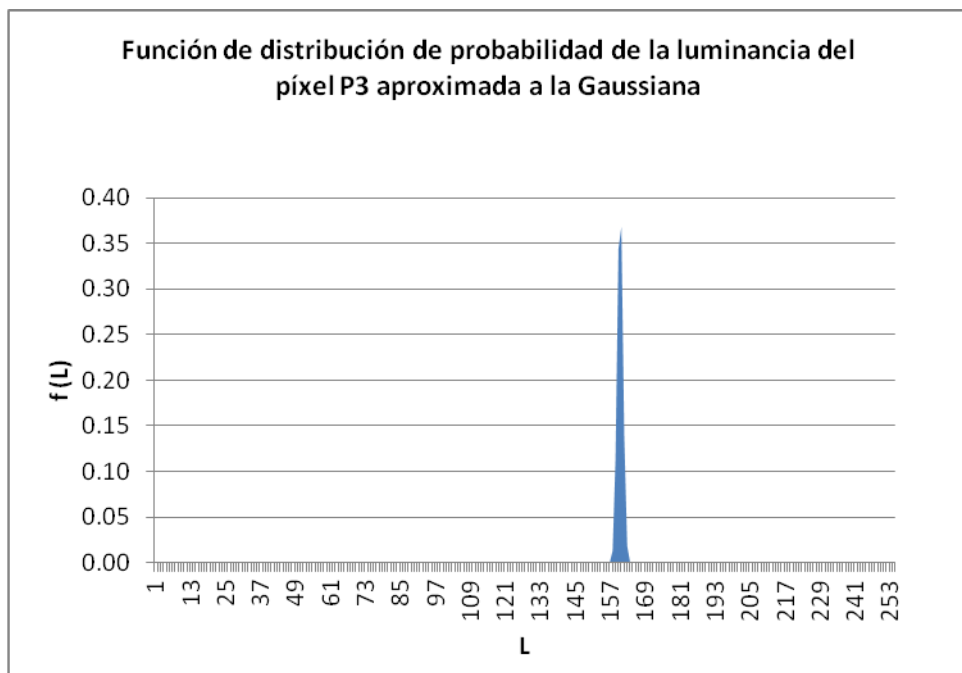
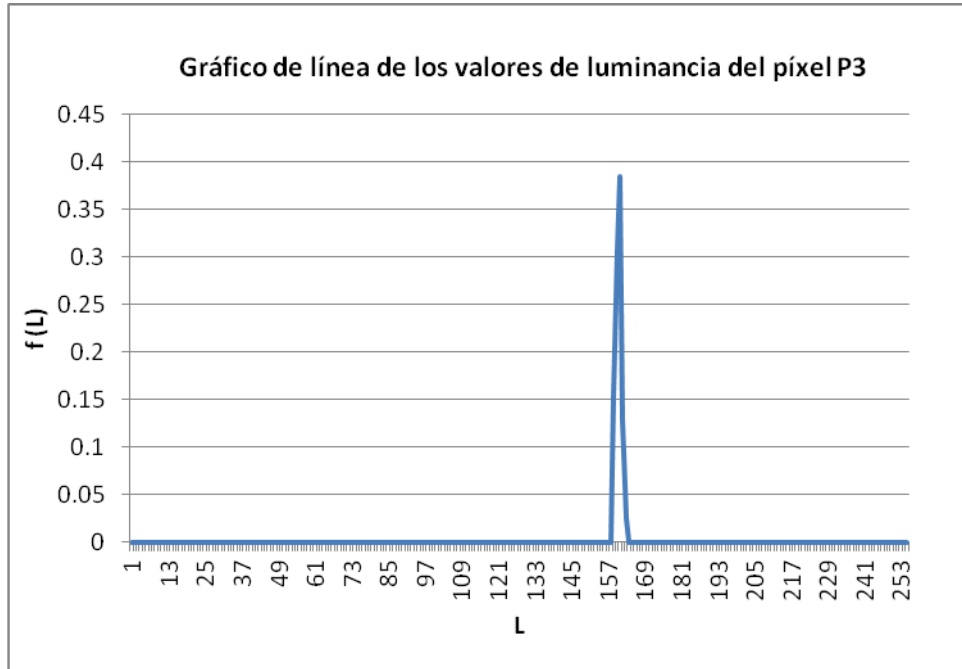


Figura 14. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P3

Elaboración: Los autores

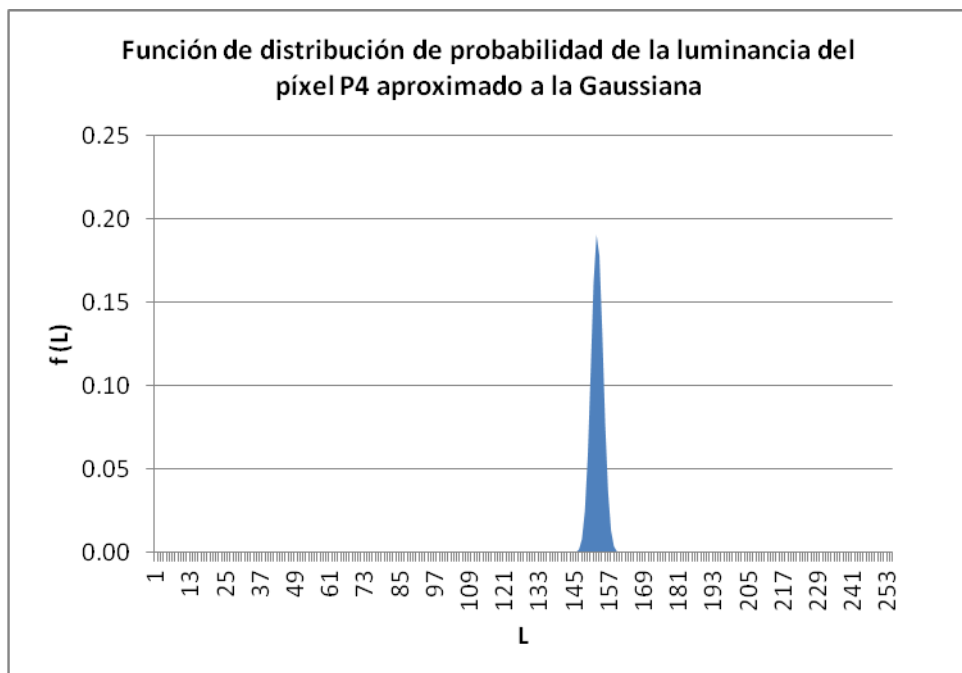


Figura 15. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P4

Elaboración: Los autores

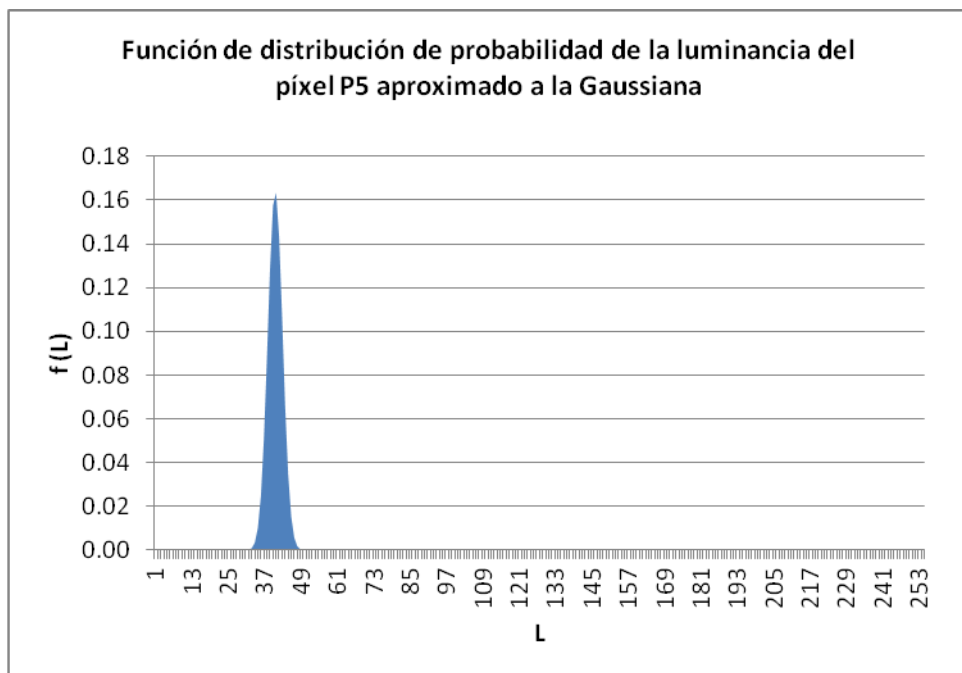
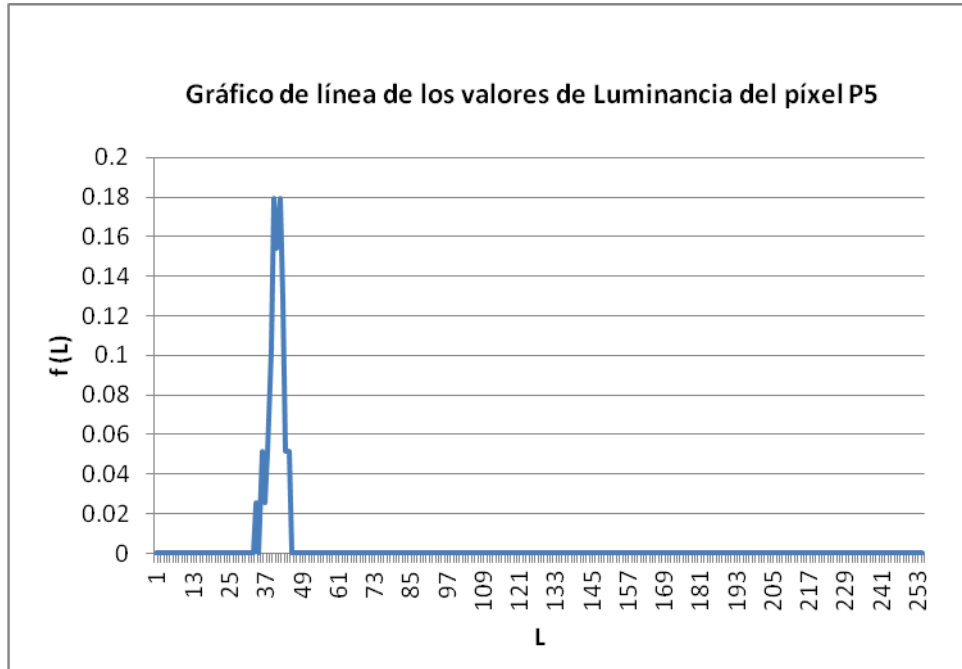


Figura 16. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P5

Elaboración: Los autores

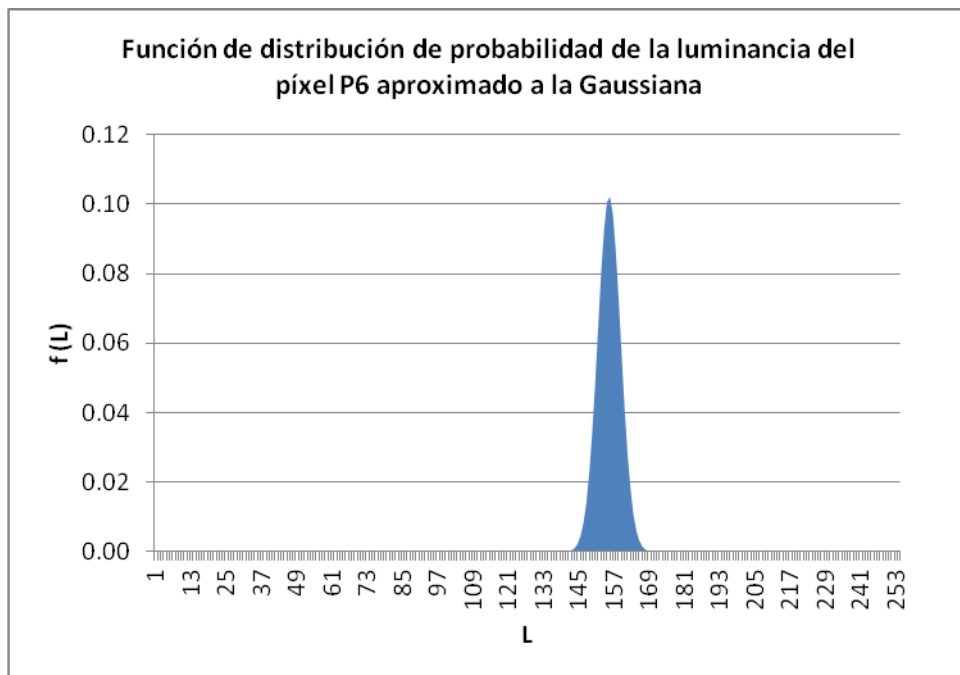
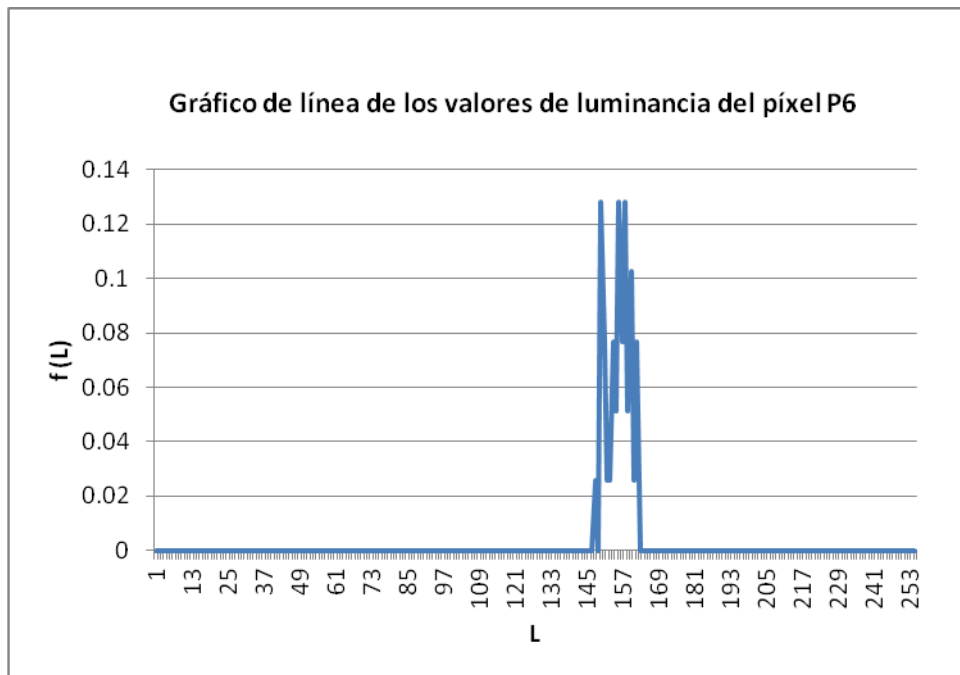


Figura 17. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P6

Elaboración: Los autores

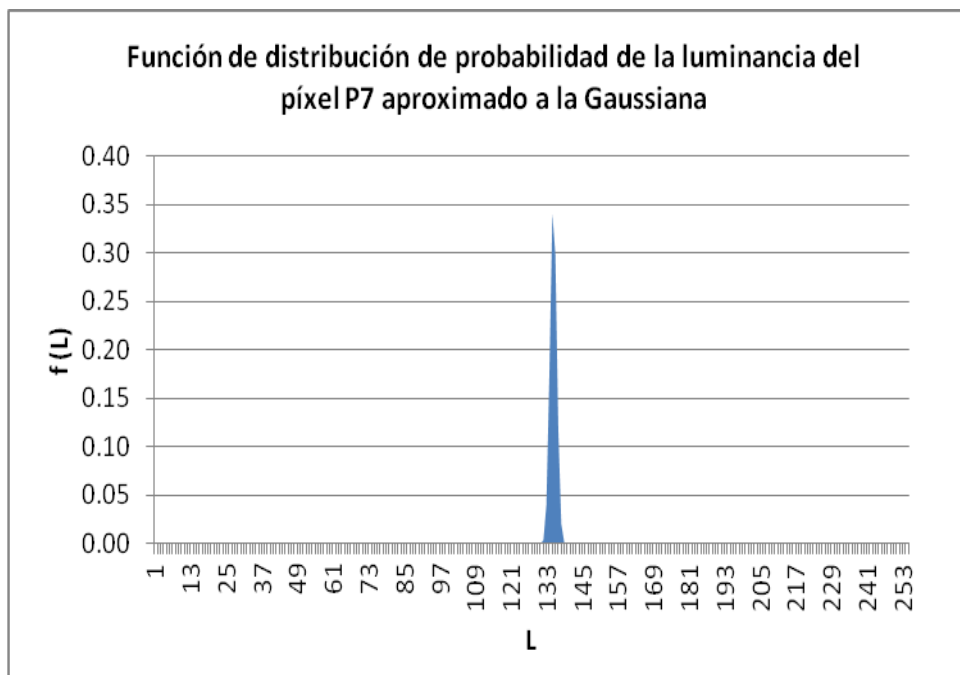
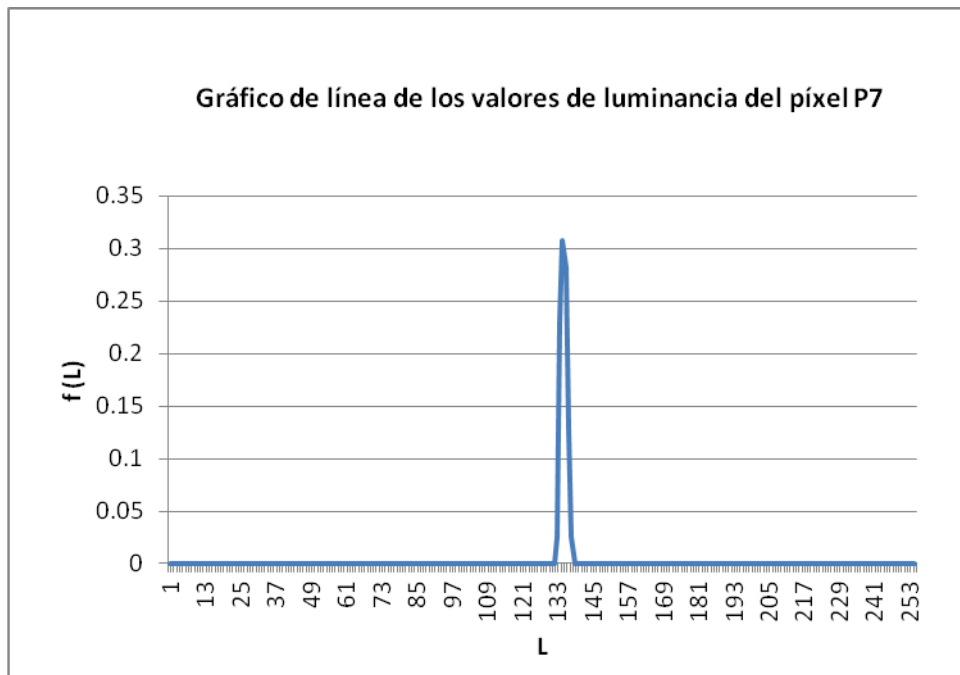


Figura 18. Gráfico de línea y FDP aproximada a la Gaussiana del píxel P7

Elaboración: Los autores

3.3 Cálculo de la distancia de Mahalanobis del píxel respecto a los parámetros del modelo de fondo

Se indicó en la sección de modelamiento que mediante el cálculo de la distancia de Mahalanobis es posible medir la similitud entre dos variables estadísticas. Para el presente caso se trata de medir la similitud del píxel actual y el fondo.

Se realizó para ello la selección de siete (7) píxeles de la Figura N° 19 representantes de la imagen actual de los cuales M1, M2 y M3 pertenecen a la imagen de fondo y M4, M5, M6 y M7 pertenecen al objeto de primer plano. Con los valores de luminancia de estos píxeles, se calcula la distancia de Mahalanobis respecto del fondo modelado en el punto anterior aplicando la Ecuación 14.

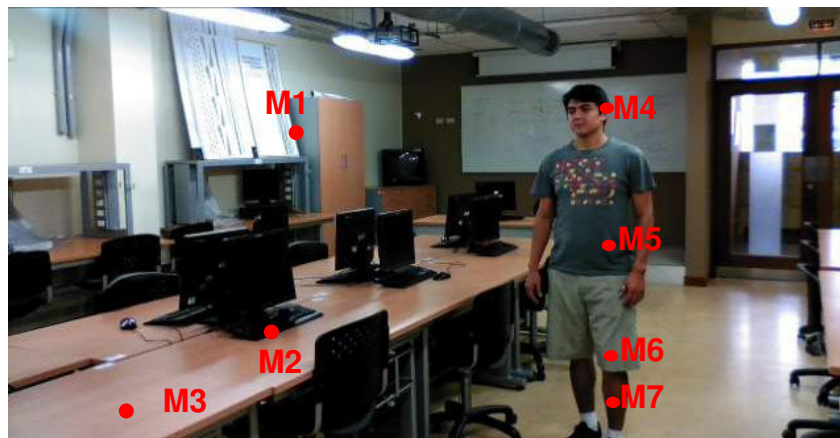


Figura 19. Píxeles representantes del primer plano y del fondo

Elaboración: Los autores

Con los valores de la media y la varianza del modelo de fondo de cada píxel, se realizó el cálculo de la distancia de Mahalanobis de los píxeles M1, M2, M3, M4, M5, M6 y M7. Los resultados se muestran en la Tabla N°2.

Tabla 2. Resultados del cálculo de la distancia de Mahalanobis para cada píxel.

	Píxel (x,y)	Valor actual del píxel (Vpíxel)	Media (μ) del fondo	Varianza (σ^2) del fondo	Distancia de Mahalanobis (D_{Mh}) del píxel
PÍXELES DE FONDO	M1 (100,195)	254	254.49	0.25	0.97
	M2 (262,184)	13	6	11.08	2.10
	M3 (330,70)	160	160.56	0.96	0.57
PÍXELES DE PRIMER PLANO	M4 (80,440)	5	153.21	4.21	72.10
	M5 (180,440)	40	40.72	5.89	0.30
	M6 (270,440)	30	155.74	15.22	32.24
	M7 (300,440)	20	135.33	1.25	103.25

Elaboración: Los autores

Se observa que para los píxeles M1, M2 y M3 que pertenecen al fondo, la distancia de Mahalanobis, es un valor menor que el calculado para los píxeles M4, M6 y M7 que pertenecen al primer plano. Por lo tanto, se deduce que una distancia de Mahalanobis baja, significa una mínima diferencia entre el píxel evaluado y el fondo, mientras que una distancia de Mahalanobis alta, implica gran diferencia entre el píxel evaluado y el fondo.

Se evidencia también que existen píxeles de primer plano cuya distancia de Mahalanobis es baja. Es el caso del píxel M5 cuya distancia de Mahalanobis es de 0.30. Se entiende que el resultado está dentro del margen de error propio del sistema.

La distancia de Mahalanobis es el parámetro utilizado en el programa desarrollado para definir el valor de umbral y que, por tanto, ha permitido definir la pertenencia de un píxel al fondo o al primer plano.

3.4 Implementación de la emisión de alarmas y almacenaje de imágenes por evento

Para la emisión de alarmas se tuvo en consideración que se requiere de un sistema que envíe dos diferentes tipos de alarmas: visuales y audibles. La finalidad es que el sistema no requiera de un permanente monitoreo o que el monitoreo sea el mínimo necesario.

La alarma visual consiste en la aplicación del modelamiento de la segmentación de imágenes para la detección de primer plano presentado en el capítulo anterior, en el que los píxeles con una distancia de Mahalanobis superior al nivel de umbral sean reemplazados por un píxel saturado de rojo. El efecto del procesamiento de la imagen de esta forma es presentar zonas rojas en la imagen que sean de fácil detección por el ojo humano, como se muestra en la figura N° 20.



Figura 20. Resultado de la alarma visual

Elaboración: Los autores

La alarma audible consiste en el uso de audios con mensajes predefinidos donde se indica el nombre del área vulnerada. La emisión se realiza a través de radios portátiles de dos vías que operan en banda

libre. Estos radios incluyen un botón llamado Push To Talk o PTT ubicado normalmente a un lado del equipo que debe ser presionado para aperturar la comunicación, como se muestra en la figura N° 21.

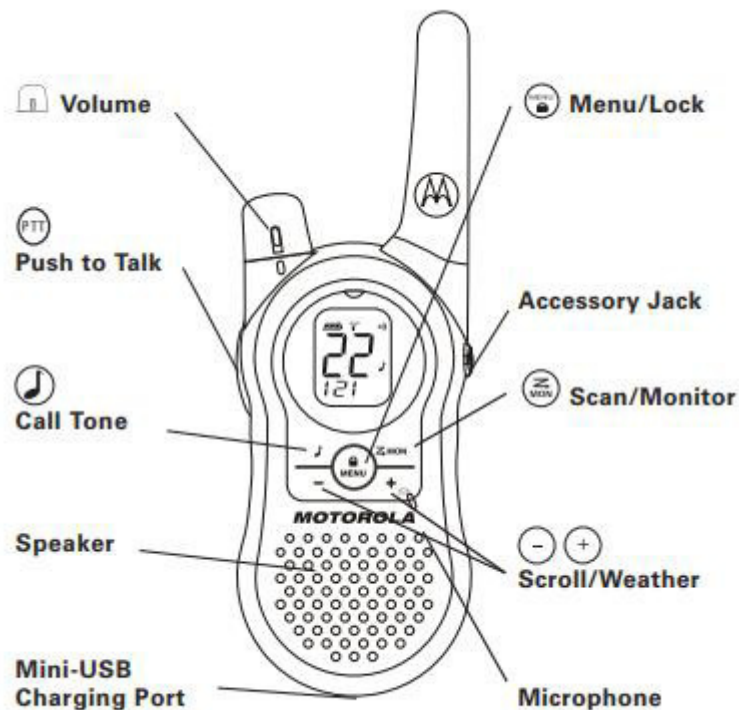


Figura 21. Radio de comunicación de doble vía y el botón PTT

Fuente: Motorola

Para este sistema son utilizadas dos radios: una emite la señal de audio y debe conectarse al computador para este fin y el segundo es el que porta el usuario. Por esta razón se desarrolló una interfaz física que habilita automáticamente la comunicación con el resto de radios. El circuito electrónico elaborado para esta interfaz se presenta en la figura 22.

La comunicación entre la radio emisora y el computador se logra mediante el uso de dos componentes: el puerto paralelo y la salida de audio del computador. El primero, permite abrir la comunicación de la radio emisora a través del envío desde el computador de una señal digital que activa el botón PTT de la radio emisora de estado normalmente

abierto. La interfaz asegura la apertura de la comunicación para la emisión de la alarma de voz y el cierre automático al finalizar el mensaje evitando que se mantenga la comunicación abierta cuando no se emite el audio.

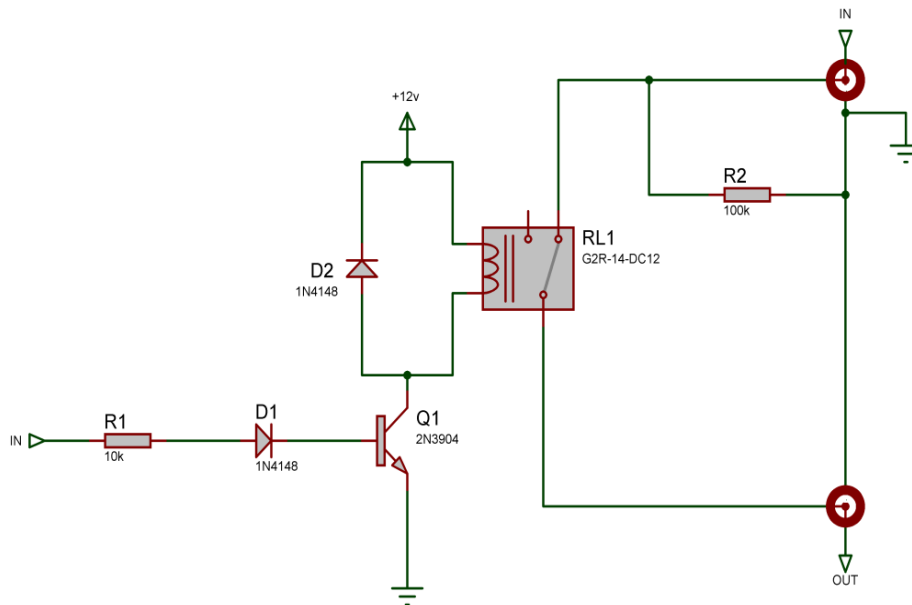


Figura 22. Circuito activador de PTT y emisor de audio

Elaboración: Los autores

Finalmente, el almacenaje de los eventos consiste en la capacidad del sistema de guardar en una parte de la memoria definida por el usuario dos tipos de registros. Por un lado están las imágenes capturadas por las cámaras y por otro, el registro de eventos almacenado en un procesador de texto. En ambos casos, el almacenamiento incluye el registro de las fechas y horas de los eventos.

CAPÍTULO IV

PRUEBAS Y RESULTADOS

Para el desarrollo de este capítulo, se utilizaron 1426 imágenes de resolución 640 x 352 tomadas durante veinticuatro (24) horas en un mismo escenario donde se generó movimiento controlado con la finalidad de evaluar el comportamiento de un píxel de fondo y de primer plano en el tiempo. Las imágenes fueron capturadas con la cámara TV-IP7621C a razón de 1 imagen por minuto aproximadamente.

4.1 Análisis de un píxel perteneciente al fondo

Se consideró analizar los valores de luminancia de un píxel que no presenta variaciones en el periodo T de duración 24 horas. El píxel evaluado y ubicado en la posición con coordenadas 330,190 y guarda información de la imagen de un reloj de pared en interiores. Durante la prueba, la imagen estuvo afectada de los cambios de la luz natural por el paso de las horas y en ausencia absoluta de la luz artificial. La Figura N° 23 muestra el registro de los cambios de luminancia del píxel durante este período. Este muestreo se inició a las 00:00 horas y culminó a las 23:59 horas de un mismo día.

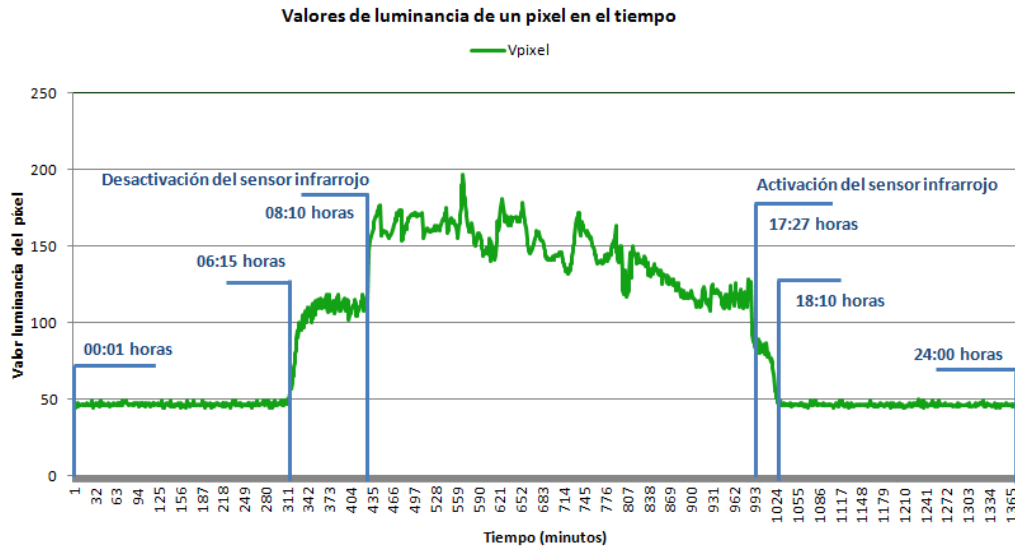


Figura 23. Valores de la luminancia del píxel 330,190 en 24 horas de muestreo.

Elaboración: Los autores

En la Figura N° 23, se observa que los valores de luminancia del píxel evaluado tienen un comportamiento estocástico en el tiempo. La luminancia del píxel en el período comprendido desde las 00:00 hasta las 06:14 horas se mantiene prácticamente constante con algunas mínimas variaciones y desde las 06:15 hasta las 14:17 horas presenta fluctuaciones de hasta 20 veces el valor inicial. Las variaciones más significativas ocurren en cuatros diferentes momentos: cuando la luz del sol empieza a incidir en la imagen analizada a las 06:15 horas, cuando se desactiva el sensor de infrarrojos a las 08:10 horas, cuando se vuelve a activar el sensor de infrarrojos a las 17:27 horas y cuando empieza a oscurecerse la escena hasta las 18:10 horas, momento en el que el valor de luminancia del píxel nuevamente se estabiliza.

La Figura N° 24 muestra los valores de la media y la desviación estándar calculados a partir de los valores de luminancia que con el que cuenta el píxel en el período T. Se observa que la varianza de los datos aumenta en los períodos en los que la luminancia del píxel presenta cambios significativos.

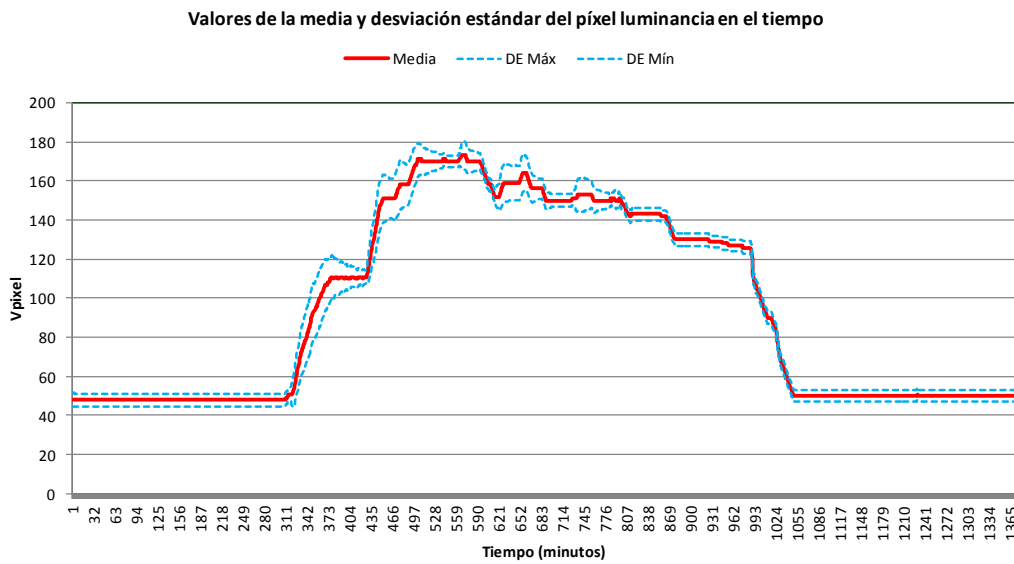


Figura 24. Gráfica de la media y la desviación estándar de la luminancia del píxel 330, 190 en 24 horas de muestreo
Elaboración: Los autores

En la Figura N° 25 se muestra el gráfico de la distancia de Mahalanobis en el período T calculada aplicando la ecuación 14 en cada instante t. De la gráfica se observa que los valores más altos de la distancia de Mahalanobis ocurren cuando el valor de la luminancia del píxel varía en más del 20%.

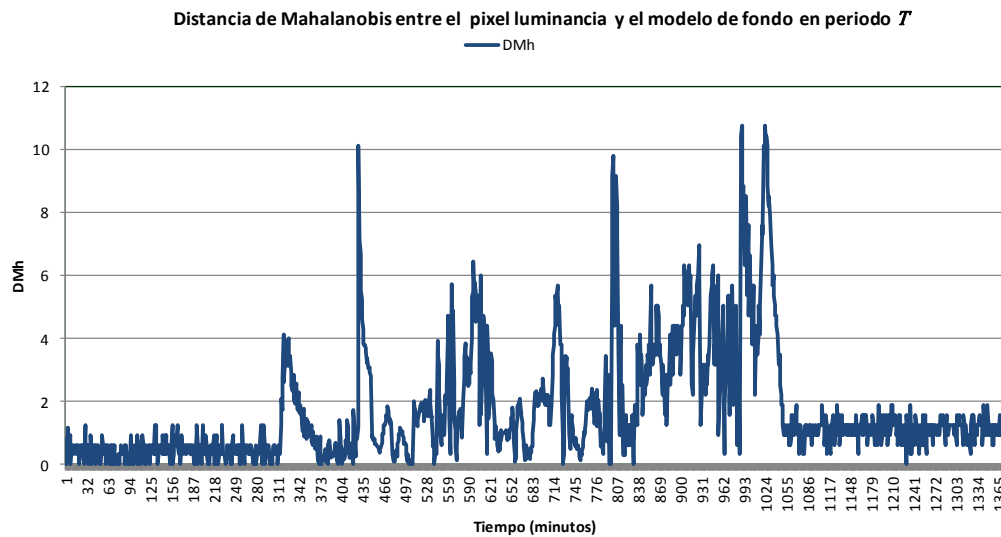


Figura 25. Valores de la distancia de Mahalanobis del píxel 330, 190 en 24 horas de muestro.

Elaboración: Los autores

Finalmente, en la Figura N° 26 se muestran las gráficas superpuestas de los cálculos mencionados. Se observa que la Distancia de Mahalanobis aumenta en relación con el incremento de la luminancia del píxel.

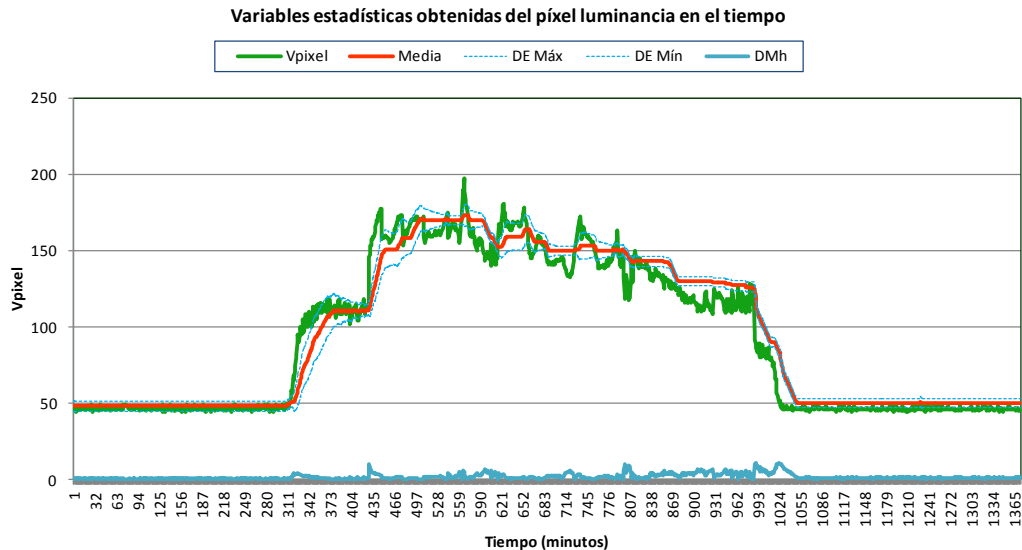


Figura 26. Gráficas superpuestas calculadas para el píxel 330, 190 en 24 horas de muestreo

Elaboración: Los autores

La distancia de Mahalanobis debe tener valores muy cercanos al cero cuando se trata de un píxel de fondo y valores muy altos para píxeles de primer plano. En esta evaluación, se sabe que el píxel es parte del fondo, pero la variación del nivel de luminancia puede ejercer efecto sobre este cálculo dando resultados diferentes al esperado. Un mal cálculo de la estimación del nivel de umbral en base a la distancia de Mahalanobis puede generar un sistema que emita falsas alarmas.

4.2 Análisis de un píxel perteneciente al primer plano

Se analizaron los valores de luminancia de un píxel que presenta movimiento programado en cada hora, basados en la presencia de la manecilla del reloj conocida como minuterero. La Figura N° 27 registra los cambios de luminancia del píxel ubicado en las coordenadas 300, 588, por el que pasa el minuterero del reloj cada 60 minutos.

Este muestreo inició desde las 00:00 hasta las 23:59 horas de ese mismo día.

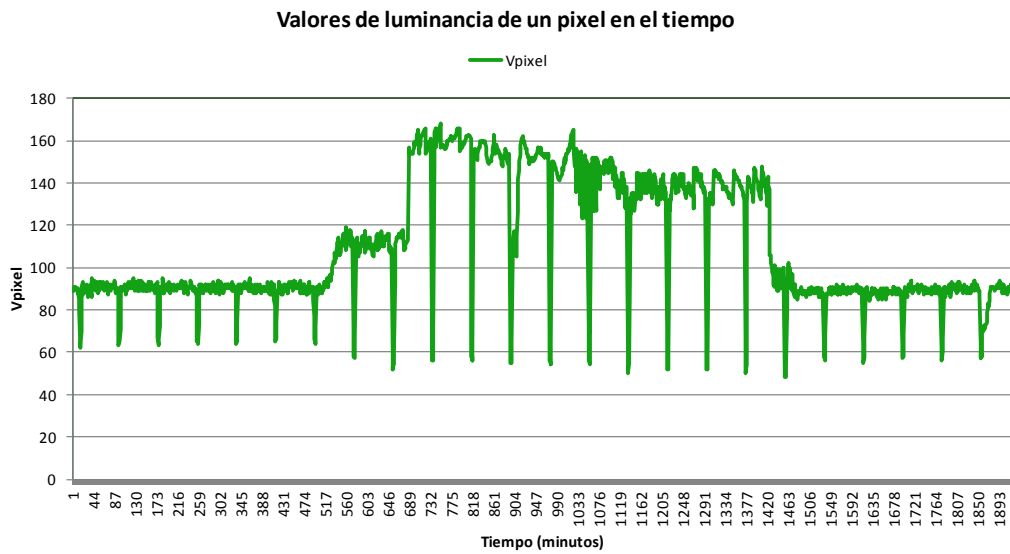


Figura 27. Valores de la luminancia del píxel 300,588 en 24 horas de muestreo.

Elaboración: Los autores

Como se observa en la Figura N° 26, los valores de luminancia del píxel presentan cambios de luminancia hasta en 65.2% de su valor previo como se demuestra en el instante $t= 735$ ó $t= 948$.

En la Figura N° 28, se muestran los valores de la media y la desviación estándar del mismo píxel. Aplicando las ecuaciones 18 y 19, la media y la varianza se van actualizando en función de la luminancia, media y varianza calculados hasta el instante anterior. Por ello, la gráfica obtenida de la media se mantiene constante en ciertos períodos hasta que el cambio de luminancia del píxel la modifica llegando a ser del 13.9% de su valor anterior. Esto ocurre por ejemplo en el período 976 como se muestra. Como no se trata de un cambio de luminancia permanente, la media vuelve a su estado estacionario.

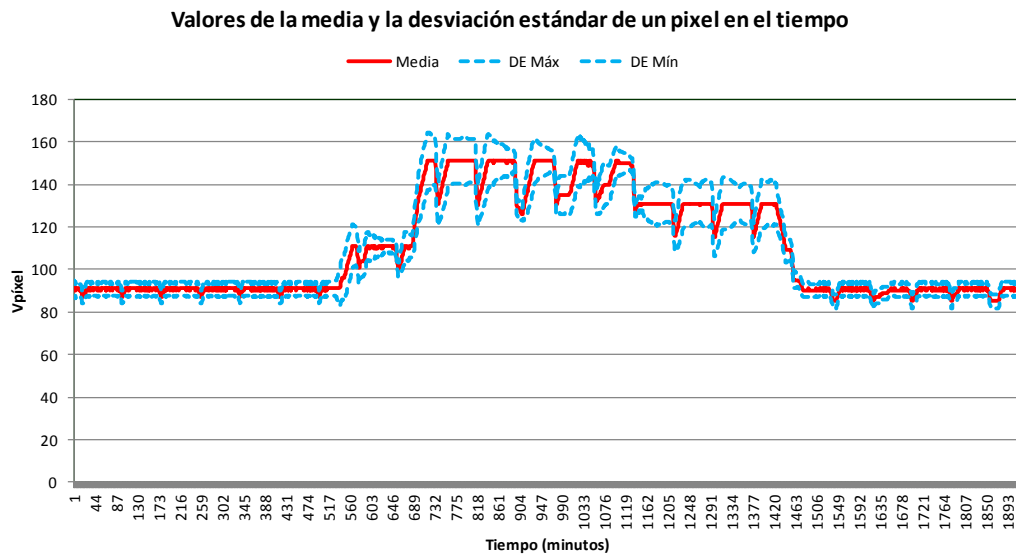


Figura 28. Gráfica de la media y la desviación estándar de la luminancia del píxel 333, 588 en 24 horas de muestreo

Elaboración: Los autores

En la figura N° 29 se observa como la distancia de Mahalanobis aumenta en función del incremento de la luminancia del píxel.

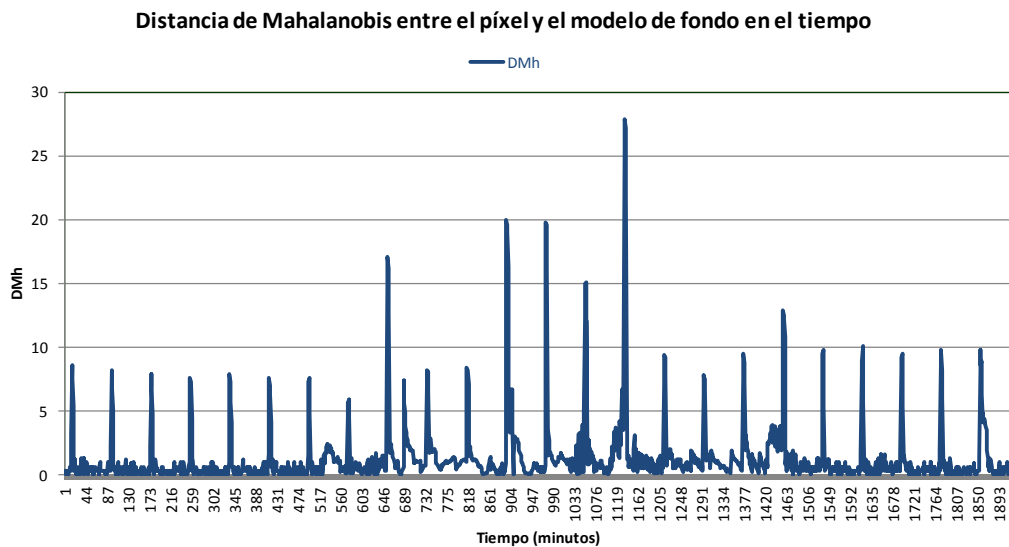


Figura 29. Valores de la distancia de Mahalanobis del píxel 300, 588 en 24 horas de muestro.

Elaboración: Los autores

Finalmente, en la Figura N° 30 se muestran las gráficas superpuestas de los cálculos mencionados anteriormente. Se observa que la distancia de Mahalanobis aumenta en relación con el incremento de la luminancia del píxel.

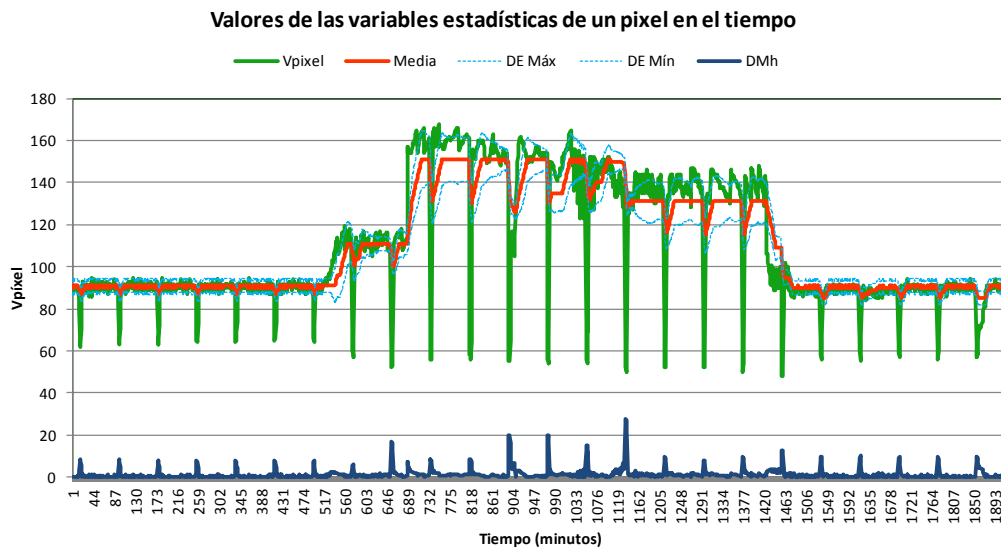


Figura 30. Gráficas superpuestas calculadas para el píxel 300, 588 en 24 horas de muestreo

Elaboración: Los autores

En base a los resultados obtenidos en estas pruebas, se determinó el valor de umbral que brinde altos valores de sensibilidad y predictividad positiva. Las ecuaciones 20 y 21 muestran las fórmulas aplicadas.

$$Sensibilidad = \frac{QVP}{(QVP + QFN)} \quad \text{Ecuación 20}$$

$$Predictividad\ positiva = \frac{QVP}{(QVP + QFP)} \quad \text{Ecuación 21}$$

QVP representa la cantidad de verdaderos positivos y se refiere a la cantidad de píxeles de primer plano detectados como primer plano. QVN o cantidad de verdaderos negativos, es la cantidad de píxeles que no son de primer plano no detectados como primer plano. QFP o

cantidad de falsos positivos es la cantidad de píxeles que no son de primer plano y que fueron detectados como de primer plano. QFN o cantidad de falsos negativos, es la cantidad de píxeles que son de primer plano y que no fueron detectados como primer plano.

El cálculo de estos parámetros se indican en las Tablas N°3 y N°4.

Tabla 3. Cálculo de la Sensibilidad y Predictividad positiva para determinar el Umbral de día.

D _{Mh} /Umbral	2	3	4	4.5	5	6
QVP	46	46	46	45	44	44
QVN	616	661	676	681	683	686
QFP	76	31	16	11	9	6
QFN	0	0	0	1	2	2
Sensibilidad	100.00 %	100.00 %	100.00 %	97.83%	95.65%	95.65%
Predictividad positiva	37.70%	59.74%	74.19%	80.36%	83.02%	88.00%

Elaboración: Los autores

Tabla 4. Cálculo de la Sensibilidad y Predictividad positiva para determinar el Umbral de noche.

D _{Mh} /Umbral	2	3	4	4.5	5	6
QVP	75	70	69	63	61	52
QVN	1061	1088	1105	1109	1110	1110
QFP	48	22	5	1	0	0
QFN	1	5	6	12	14	23
Sensibilidad	98.68%	93.33%	92.00%	84.00%	81.33%	69.33%
Predictividad positiva	60.98%	76.09%	93.24%	98.44%	100.00%	100.00%

Elaboración: Los autores

Se requiere de un valor umbral que permita tener un nivel de sensibilidad y predictividad positiva que minimice las falsas detecciones, es decir, los falsos positivos y los verdaderos negativos.

De acuerdo a la Tabla N°3, el valor umbral adecuado para la detección de movimiento en imágenes de día es de seis (06) y de acuerdo a la Tabla N°4, para el caso de las imágenes de noche el valor umbral es de cuatro (04).

Aplicando los valores de umbral a los dos casos presentados en este capítulo, se calcula el porcentaje de acierto, encontrándose que hasta en el 99% de los casos los píxeles se detectan correctamente, es decir, se detectan los píxeles etiquetados como verdadero negativo y verdadero positivo en su mayoría. Esto se demuestra en las figuras N° 31 y N° 32 correspondientes al caso de día y noche respectivamente.

RESULTADOS DE SENSIBILIDAD Y PREDICTIVIDAD DE LA DETECCIÓN DE MOVIMIENTO - DÍA

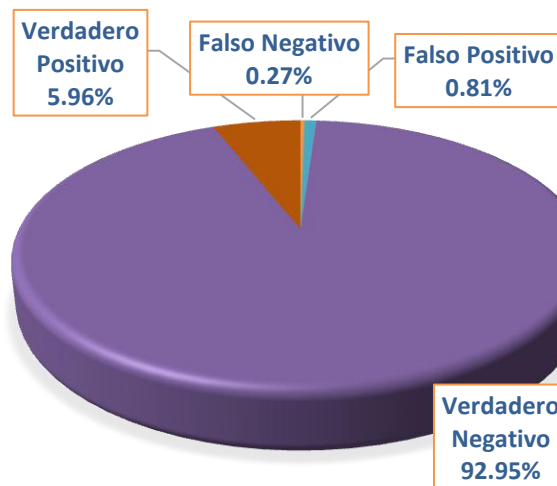


Figura 31. Nivel de acertividad de la detección de movimiento en horario diurno.

Elaboración: Los autores

**RESULTADOS DE SENSIBILIDAD Y PREDICTIVIDAD DE LA
DETECCIÓN DE MOVIMIENTO - NOCHE**

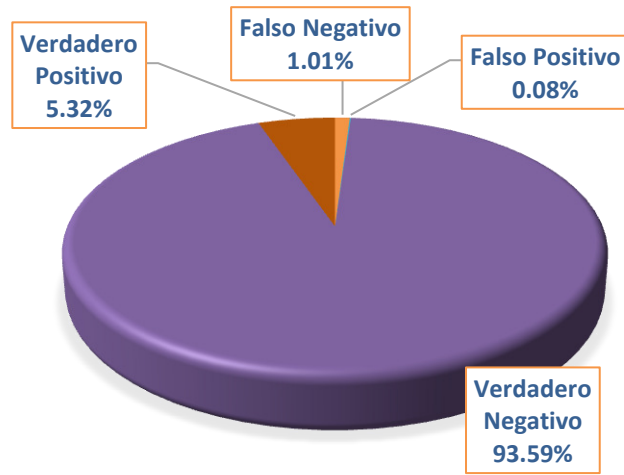


Figura 32. Nivel de acertividad de la detección de movimiento en horario nocturno.

Elaboración: Los autores

CONCLUSIONES

1. Se logró concluir con el diseño del sistema de detección de movimiento por visión computacional preparada para procesar las imágenes de hasta 5 cámaras IP.
2. La detección de movimiento se realizó aplicando dos métodos; el primero llamado sustracción de fondo basado en el principio de la Gaussiana Simple, y el segundo llamado umbralización basado en el cálculo de la distancia de Mahalanobis para la segmentación de las imágenes.
3. Se demuestra que los valores de luminancia del píxel en el tiempo puede aproximarse a una distribución de probabilidad Gaussiana.
4. El principio de la Gaussiana Simple permite mantener el fondo actualizado de acuerdo a los cambios de luminancia, mientras que la distancia de Mahalanobis es un parámetro que permite medir la similitud entre los píxeles y, por lo tanto, entre las imágenes.
5. Un bajo valor de la distancia de Mahalanobis, significa una mínima diferencia entre el píxel evaluado y el fondo, mientras que una distancia de Mahalanobis alta implica gran diferencia entre el píxel evaluado y el fondo.

6. El uso de la umbralización como método para la segmentación de las imágenes requiere de un nivel de umbral cuyo valor se estimó en base a las pruebas realizadas con 1426 muestras.
7. Se determinan dos niveles de umbral para el sistema; para el horario diurno el umbral es de seis (6) mientras que para la noche, el umbral es de cuatro (4).
8. Los niveles de sensibilidad y predictividad positiva calculados según los niveles de umbral seleccionados alcanzan valores de hasta el 97%, lo cual demuestra que los resultados son muy aceptables.
9. Aplicando el nivel de umbral especificado, se logra alcanzar hasta un 99% de efectividad del sistema.
10. Los sistemas de alarma generan registros de eventos en imágenes y textos que detallan la fecha y la hora de cada suceso.
11. La interfaz física diseñada para la emisión de alarmas de voz por radio permite el uso de otras radios en la misma banda seleccionada pues incluye el control del botón PTT, encargado de dar apertura a la comunicación en este tipo de comunicación.
12. El sistema ha sido probado en tiempo real y es capaz de emitir alarmas visuales y audibles, por lo que puede ser utilizado en sistemas de seguridad por video vigilancia.

RECOMENDACIONES

1. Es recomendable incluir un análisis de especificidad además de las realizadas en esta investigación de sensibilidad y predictividad positiva.
2. Hay varias mejoras que podrían realizarse al sistema, como por ejemplo, la capacidad de capturar imágenes de otras marcas de cámaras o la capacidad de envíos de alarma por correo o mensaje de texto. Cualquiera de ellos debe asegurar baja carga computacional que permita que el sistema siga operando en tiempo real.
3. En caso de implementarse el sistema, es necesario brindarle las condiciones de respaldo correspondientes a un sistema crítico. Esto es, por ejemplo, el uso de UPS y grupo electrógeno para asegurar continuidad de la operación a pesar de los cortes de energía.

FUENTES DE INFORMACIÓN

Bibliográficas

Cervera P. (2010) *Integración de información de movimiento en la segmentación de secuencias de vídeo basada en el modelado de fondo*, Madrid.

Cheung S. C., Kamath C. (2004) *Robust techniques for background subtraction in urban traffic video*. Video Communications and Image Processing. SPIE Electronic Imaging, San Jose, UCRL-JC-153846, UCRL-CONE-200706

Cucchiara R. (2006), *People Surveillance*, VISMAL Palermo

Leng B., Dai Q. (2007), *Video object segmentation based on accumulative frame difference*, PCE Picture Coding Symposium, Beijing.

Elgammal A., Harwood D., and Davis L. (2000), *Non-parametric model for background subtraction*. In Proceedings of the 6th European Conference on Computer Vision-Part II, pages 751–767.

Friedman N, Russell S. (2013), *Image segmentation in video sequences: a probabilistic approach*, California (UAI 97)

Gordon G., Darrell T., Harville M., Woodfill J. (1999), *Background Estimation and Removal Based on Range and Color*. CVPR, IEEE Comp Society Conf. on Comp Vision and Patt Recog (CVPR'99), California.

Grimson W. E. L., Stauffer C., Romano R., and Lee L. (1998), *Using adaptive tracking to classify and monitor activities in a site*, in Proc. IEEE Conf. Computer Vision and Pattern Recognition, California.

Haritaoglu I., Harwood D., Davis and L. S. (2000), *W: Real-time surveillance of people and their activities*, IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 809–830.

Instituto Nacional de Estadística e Informática INEI (2017), *Perú: La victimización en las empresas, 2015 – 2016*, Lima.

Köhle M., Merkl D., and Kastner J. (1997), *Clinical gait analysis by neural networks: Issues and experiences*, in Proc. IEEE Symp. Computer-Based Medical Systems, pp. 138–143.

Picardi M. (2004), *Background subtraction techniques: a review*, IEEE International Conference on Systems, Man and Cybernetics, Maryland

Ridder C., Munkelt O., and Kirchner H (1995), *Adaptive background estimation and foreground detection using Kalman-filtering*, in Proc. Int. Conf. Recent Advances in Mechatronics, pp. 193–199.

Stauffer C. and Grimson W. (1999) *Adaptive background mixture models for real-time tracking*, in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Massachusetts vol. 2, pp. 246–252.

Sun H. Z., Feng T., and Tan T. N. (2000), *Robust extraction of moving objects from image sequences*, in Proc. Asian Conf. Computer Vision, Taiwan, R.O.C., pp. 961–964.

Tang Z., Miao Z., Wan Y. (2007), *Background Subtraction Using Running Gaussian Average and Frame Difference*. ICEC 2007, Berlín.

Toyama K., Krumm J., Brumitt B., and Meyers B. (1999), *Wallflower: principles and practice of background maintenance*, in Proc. Int. Conf. Computer Vision, pp. 255–261.

Wren C.R., Azarbayejani A., Darrell T., Pentland A.P. (1997), *Pfinder: Real-time tracking of the human body*.

Electrónicas

Álvarez Y., Barbará E., Rodríguez O. (2010), *Filtrado digital en el procesamiento de imágenes empleando Matlab*, descargado de https://www.researchgate.net/publication/308787583_FILTRADO_DIGITAL_EN_EL_PROCESAMIENTO_DE_IMAGENES_EMPLEANDO_MATLAB

Bahamón N., (2011), *Restauración de imágenes mediante un modelo matemático basado en las técnicas de detección de bordes y propagación de texturas*, recuperado de http://www.bdigital.unal.edu.co/4988/1/8300062011_Parte1.pdf

García J. (2010). *La curva de distribución normal o campana de Gauss*, recuperado de <https://jesusgarciaj.com/2010/01/22/la-curva-de-distribucion-normal/>

Guevara O. (2013). *Visión computacional para asistencia en conducción*, descargado de <http://nightcrowling.blogspot.com/2013/01/computer-vision-based-driving-assistance.html>

Motorola (2012) *Two-Way Radio User's Guide* descargado de <https://www.motorolasolutions.com/content/dam/msi/docs/products/two-way-radios/consumer/mh230-uy-pa-us-user-guide.pdf>

Rondón M. (2018). *Distribuciones de Probabilidad Discretas y Continuas*. 2019, de RPubS Sitio web: <https://rpubs.com/mrondon/422093>

Pita Fernández S., Pértegas Díaz S. (2010). *Pruebas diagnósticas: Sensibilidad y Especificidad*. Sitio web: https://www.fisterra.com/mbe/investiga/pruebas_diagnosticas/pruebas_diagnosticas.asp

Valverde-Rebaza J. (2009), *Binarización de imágenes*, descargado de <http://jc-info.blogspot.com/2009/03/binarizacion-de-imagenes-codigo.html>

Yabo A., (2015). *Video-detección vehicular para sistemas inteligentes de transporte (SIT)*, Sitio web: https://www.researchgate.net/figure/Figura-41-Esquema-de-sustraccion-de-fondo-12_fig3_305775549