

# Desarrollo de un Sistema de Vigilancia Portable

Manuel Alejandro Jiménez Vivas, Código 032008, Fundación Universitaria San Martín, Facultad de Ingeniería.

**Resumen**— En el presente artículo se estudian las principales características en la captura de video de los sistemas de vigilancia, de los cuales se selecciono e implemento un algoritmo de generación de video donde se modifica la frecuencia de muestreo, de tal manera que utilice la menor cantidad de recursos, posteriormente se evaluó el códec de video y se calcularon los requerimientos mínimos para su implementación en dispositivos portables.

**Palabras clave**—Algoritmo, binarización, cámara, centro de masa, códec, compresión, distancia euclidiana, escala de grises, fotograma, frecuencia de muestreo, grabación, imagen, memoria, pixel, pixel umbral, reconocimiento de patrones, resolución, RGB, sistema de vigilancia, velocidad, video.

## I. INTRODUCCIÓN

LA seguridad ha ganado gran importancia durante la última década dentro de la sociedad, debido a la constante amenaza de grupos terroristas y grupos de delincuencia común, por tal motivo los sistemas de vigilancia se hacen cada vez más indispensables dentro de los entes gubernamentales, la industria y las personas del común. Los sistemas de vigilancia modernos no solo deben cumplir con la función de registrar sucesos dentro de una secuencia entendible de video, sino también estar en la capacidad de contar con ciertas exigencias adicionales, como analizar las imágenes capturas, el uso de técnicas de compresión y transmisión eficiente de las imágenes, esto con el fin de incrementar la rapidez de alertar a las autoridades si llegase a existir algún peligro que atentase contra la seguridad y/o la integridad.

En los sistemas de vigilancia visual modernos es indispensable que los sucesos registrados sean grabados en imágenes de alta calidad, para que estas imágenes adquieran gran veracidad y credibilidad ante la sociedad y las cortes de justicia, en el momento de hacer una identificación ágil y certera de los objetos, rostros y movimientos de las personas incidentes en la secuencia de video grabada. Esto hará que la sociedad tenga más confianza en los entes gubernamentales y privados que prestan el servicio de brindar protección y con ello se aumentará el sentido de la seguridad para las personas que integran esta sociedad.

Este proyecto propone un sistema de vigilancia visual el cual comprime la secuencia de video capturada disminuyendo la

rata de muestreo de los fotogramas y no haciendo compresión sobre estos últimos, conservando así la calidad visual de las imágenes y de esta manera pueda ser implementado en dispositivos portables, principalmente en teléfonos celulares.

## II. FUNCIONAMIENTO DE LOS SISTEMAS DE VIGILANCIA USADOS EN LA ACTUALIDAD

Los sistemas de vigilancia que son utilizados en la actualidad se basan en arreglos de cámaras que deben estar en condiciones de contar con ciertos parámetros en su arquitectura para poder cumplir con la función de prestar vigilancia visual en entornos predeterminados para cada sistema. Los parámetros que deben contener las arquitecturas de los sistemas de vigilancia se evidenciaron en el desarrollo del marco referencial. En este aparte se hará referencia a los parámetros de mayor relevancia para la implementación de un sistema de vigilancia visual.

Un sistema de vigilancia moderno típico está basado en arreglos de cámaras móviles con las que se permite hacer control y énfasis en los eventos u objetivos de interés que ocurran en el escenario, estos sistemas deben también estar en las condiciones de poder transmitir el registro de sucesos del campo visual durante las 24 horas del día los 7 días a la semana, proveer controles para adaptar el sistema a las condiciones climáticas y de iluminación del entorno, con el fin de no perder el registro de eventos de importancia, para poder cumplir con estas exigencias, los sistemas de vigilancia modernos deben realizar ciertas tareas procesamiento fundamentales, como lo son [9] [4]:

Calibración de la cámara con respecto a un plano cartesiano de referencia local.

Adquisición de la imagen de la escena.

Adaptar un modelo del fondo de la escena.

Detectar eventos en las regiones del primer plano.

Registrar los sucesos.

Teniendo implementadas estas tareas fundamentales en un sistema de vigilancia, se puede considerar este sistema de este tipo para hacerlo más robusto, al cual se le puedan adicionar tareas de procesamiento adicionales que puedan prestar servicios adicionales, como:

Detección de movimiento.

Mejoramiento de la resolución de la imagen.

Compresión de los datos.

### III. ALGORITMOS PARA EL RECONOCIMIENTO

El reconocimiento de objetos es una tarea fundamental para cualquier sistema de visión y en escenarios controlados se hace fácil su implementación y la construcción de modelos en tres dimensiones de los objetos, pero en escenarios no estructurados se hace inviable aplicar técnicas de reconocimiento debido a las dificultades de segmentar adecuadamente el objeto incidente en el escenario.

El problema del reconocimiento se puede abordar de dos formas principalmente, la primera es la forma geométrica basándose en las características y la segunda forma es fotométrica basándose en su apariencia visual. Y las técnicas de mayor uso han sido el análisis de los componentes principales (PCA) y la técnica de análisis lineal discriminante (LDA).

#### A. Análisis de Componentes Principales (Principal Component Analysis, PCA)

PCA, comúnmente referida al uso de Eigenfaces, es la técnica impulsada por Kirby & Sirovich en 1988. Con PCA, el sondeo y la galería de imágenes deben ser del mismo tamaño y deben ser normalizadas previamente para alinear los ojos y bocas de los sujetos en las imágenes. La aproximación de PCA es luego utilizado para reducir la dimensión de los datos por medio de fundamentos de compresión de datos y revela la más efectiva estructura de baja dimensión de los patrones faciales.

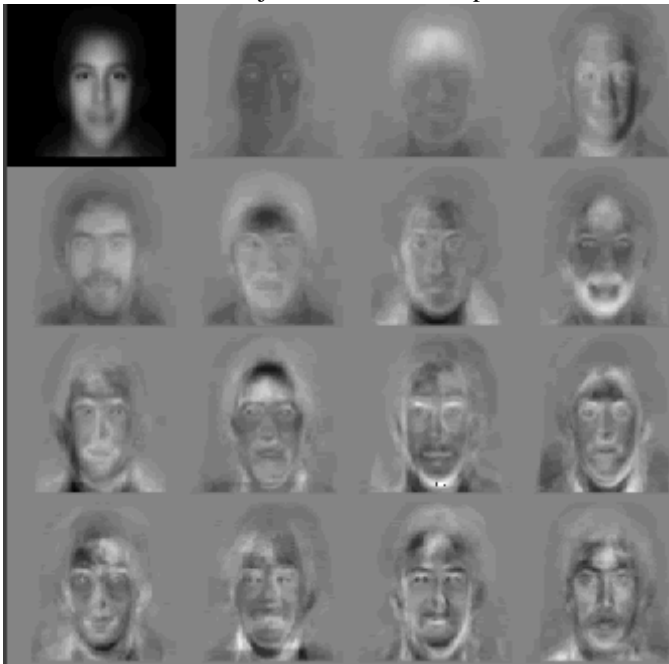


Fig. 1. Eigenfaces estándar.

Esta reducción en las dimensiones quita información que no es útil y descompone de manera precisa la estructura facial en componentes ortogonales (no correlativos) conocidos como Eigenfaces. Cada imagen facial puede ser representada como una suma ponderada (vector de rasgo) de los eigenfaces, las cuales son almacenadas en un conjunto de una dimensión.

Una imagen de sondeo es comparada con una galería de imágenes midiendo la distancia entre sus respectivos vectores de rasgos. La aproximación PCA típicamente requiere la cara completa de frente para ser presentada cada vez; de otra forma la imagen dará un resultado de bajo rendimiento. La ventaja primaria de esta técnica es que puede reducir los datos necesarios para identificar el individuo a 1/1000 de los datos presentados [10].

#### B. F Análisis Lineal Discriminante (Linear Discriminant Analysis, LDA)

Format and save your graphic images using a suitable graphics processing program that will allow you to create the images as PostScript (PS), Encapsulated PostScript (EPS), or Tagged Image File Format (TIFF), sizes them, and adjusts the resolution settings. If you created your source files in one of the following you will be able to submit the graphics without converting to a PS, EPS, or TIFF file: Microsoft Word, Microsoft PowerPoint, Microsoft Excel, or Portable Document Format (PDF). LDA es una aproximación estadística para clasificar muestras de clases desconocidas basadas en ejemplos de entrenamiento con clases conocidas. Esta técnica tiene la intención de maximizar la varianza entre clases (ej. Entre usuarios) y minimizar la varianza de cada clase (Ej. De cada usuario). En la Figura 15 cada bloque representa una clase, hay grandes variaciones entre clases, pero pequeñas en cada clase. Cuando se trata con datos faciales de alta dimensión, esta técnica enfrenta el problema de muestras de tamaño pequeño que surge donde hay un número pequeño de ejemplos de entrenamiento comparados a la dimensionalidad del espacio de muestra.



Fig. 2. Ejemplo de seis clases usando LDA.

#### C. Estándares para el Reconocimiento

La estandarización es una porción vital del avance del mercado y el estado del arte. Mucho trabajo se ha realizado en los estándares para facilitar la interoperabilidad y los formatos de intercambio de datos, lo que ayudará a facilitar el avance de la tecnología en una plataforma estandarizada. Las normas ANSI/INCITS (M1) 385-2004 e ISO 19794-5 de formato de intercambio de datos de reconocimiento facial son los mayores estándares en esta área y están dirigidas al examen humano detallado de imágenes de rostros, verificación de identificación humana, e identificación y verificación facial automatizada.

Los estándares han establecido una imagen frontal definida

y han irrumpido en sub-secciones dando tratamiento a imágenes frontales y no frontales (definiendo una imagen frontal a cinco grados del centro y una imagen no frontal está definida por la ubicación de los ojos). Estos estándares dejan otras imágenes tales como semiperfil indefinidas, pero aseguran que las imágenes enrolladas alcanzarán el estándar de calidad necesario tanto para reconocimiento automatizado de rostros como para inspección humana de imágenes de rostros. Estos estándares también facilitan el uso de información de rostros en aplicaciones que tienen capacidad de almacenamiento limitada (Ej. Pasaportes, visas, licencias de conducir). Otros estándares como INCITS 398-2005 (Common Biometric Exchange Formats Framework, CBEFF), o Marco de trabajo de Formatos para Intercambios Comunes de Biometría, tratan específicamente con los elementos utilizados para describir los datos de biometría de forma [10].

La iluminación de interiores razonable controlada dada, la tecnología de punta de reconocimiento facial es de verificación del 90% a una tasa de falsa aceptación de 1%. El uso de modelos moldeables, los cuales mapean una imagen 2D sobre una grilla 3D en un intento de superar radiaciones posturales y de iluminación, puede mejorar significativamente el reconocimiento facial no frontal. El rendimiento de la lista de vigilancia decrece como función del tamaño de una galería el rendimiento utilizando listas de vigilancia más pequeñas es mejor que utilizando las más grandes.

#### IV. DETERMINAR LA VELOCIDAD DE LOS OBJETOS DENTRO DEL VIDEO

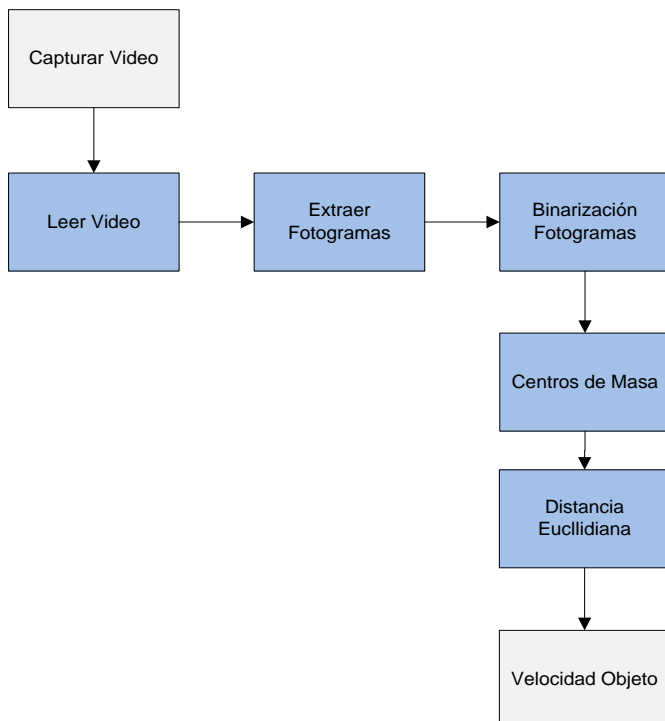


Fig. 3. Diagrama del proceso para encontrar la velocidad de los objetos.

Para calcular este parámetro fue necesario primero conocer

la posición del objeto de interés, para después calcular el centro de masa del objeto y finalmente trazar un vector velocidad entre dos fotogramas binarizados.

##### A. Proceso de Binarización de la Imagen

En el proceso de binarización de la imagen se comparan dos fotogramas, el primero es una imagen fija en un instante de tiempo  $t_0$  antes que el objeto entre en escena, y el segundo fotograma es tomado en un instante de tiempo  $t$  considerado el estado actual del objeto [6].

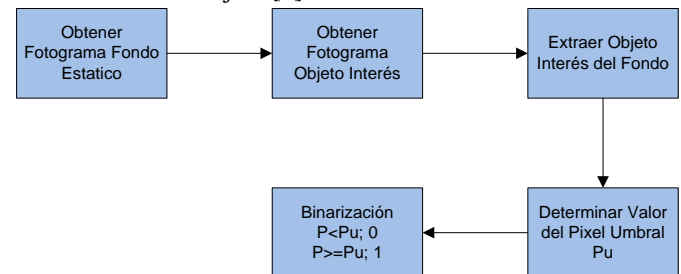


Fig. 4. Proceso para binarizar fotogramas.

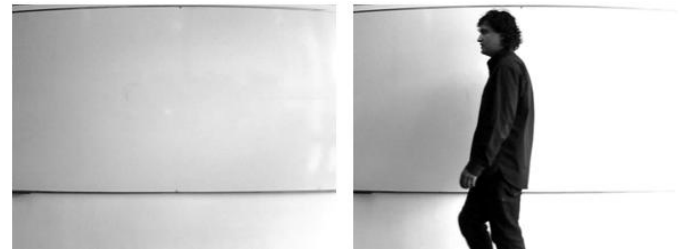


Fig. 5. (a) Fotografía capturada en  $t_0$ , (b) Fotografía capturada en  $t$ .

Al sustraer la imagen en  $t$  de la imagen  $t_0$  se pueden identificar los pixeles que se modifican cuando el objeto entra en la escena.



Fig. 6. Imagen resultado de la resta de los fotogramas  $t_0-t$ .

Para el proceso de binarización se le asigna un valor de luminancia de '1' a los pixeles pertenecientes al objeto incidente en la escena, y a los demás pixeles de la imagen un valor de '0'. Para determinar los pixeles que componen el objeto y darles el valor de '1' se halla un valor de pixel umbral

dentro de ese contorno [8].

Luminancia del pixel (0=negro  $\rightarrow$  1=blanco).

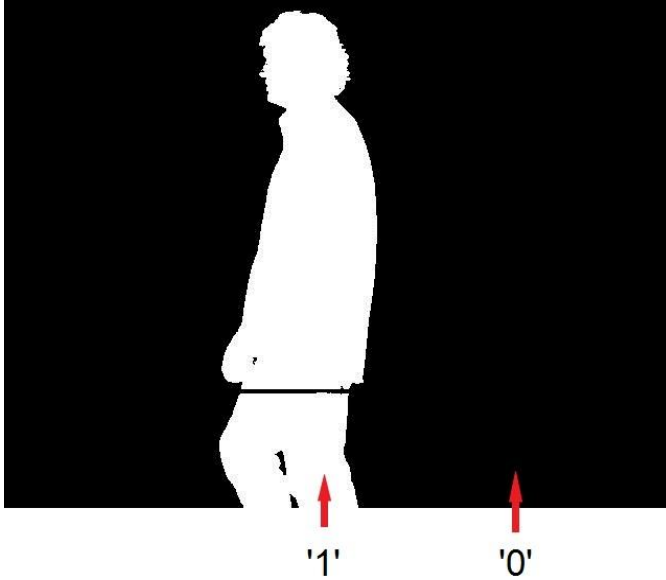


Fig. 7. Imagen binarizada.

### B. Cálculo centro de Masa.

El centro de masa es hacia donde tiende la acumulación de la energía del objeto de interés, el hallar las coordenadas del centro de masa de un fotograma capturado en un instante de tiempo  $t$  y comparar esta información con la de otro fotograma capturado en un instante de tiempo  $t+n$ , esta comparación nos proporciona información, como la dirección de movimiento del objeto y la velocidad a la que se desplaza.

Para obtener las coordenadas del centro de masa se utiliza la siguiente ecuación.

$$x_c = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} x \times I_B(x, y)}{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I_B(x, y)} ; y_c = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} y \times I_B(x, y)}{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I_B(x, y)}$$

(1)

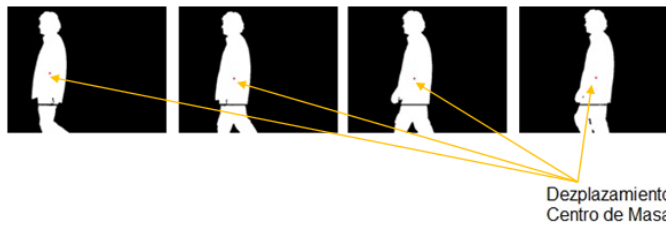


Fig. 8. Desplazamiento centro de masa de un objeto.

La secuencia de fotogramas mostrada en Fig. 8, las imágenes fueron extraídas de una secuencia de video que se encuentra grabada a una tasa de 15fps (fotogramas por segundo), de derecha a izquierda se aprecian, el fotograma 60, el fotograma 65, el fotograma 70 y el fotograma 75, de la secuencia de video mencionada.

### C. Cálculo Velocidad del Objeto.

Para el calcular la velocidad instantánea de los objetos

dentro del video se hace necesario conocer la distancia euclidiana (2) entre dos centros de masa del objeto de interés, lo cuales son calculados en dos fotogramas capturados consecutivamente y el otro parámetro que se debe conocer es el periodo de captura de los fotogramas. Con esta información se obtiene la velocidad del objeto en pixeles/segundo [5].

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2)$$

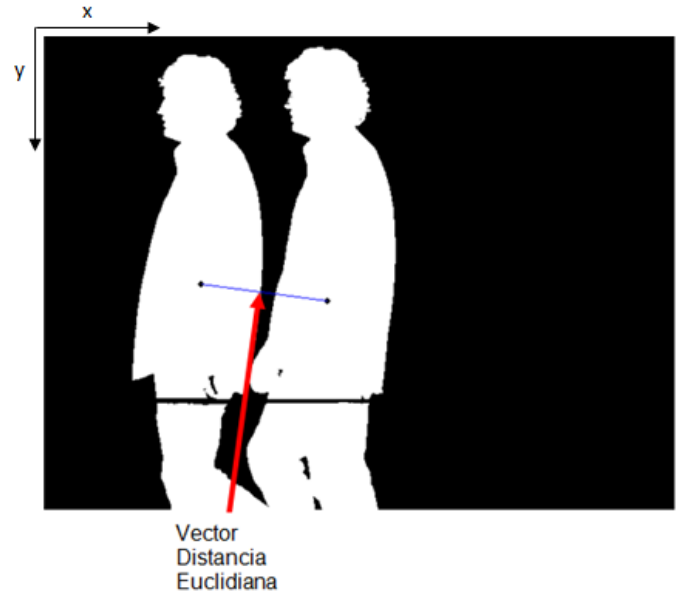


Fig. 9. Gráfica vector de la distancia euclidiana.

En la Fig. 9 se observa el gráfico del vector de la distancia euclidiana entre fotograma 60 y el fotograma 75 referenciados en la Fig. 8, este vector proporciona la información de la cantidad de desplazamiento del objeto de interés entre un fotograma y otro, en un tiempo determinado.

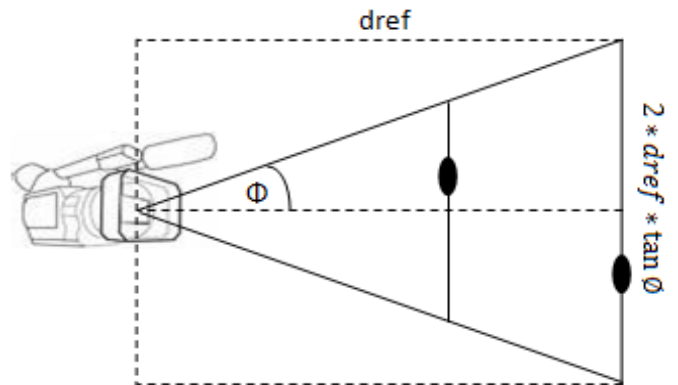


Fig. 10. Espacio que capta la cámara.

Como la distancia que se obtiene se encuentra en una magnitud de pixeles por segundo, se debe hallar su equivalente a metros por segundo, lo cual se obtiene a partir de la distancia horizontal que cubre la cámara la cual depende de la distancia referencial que existe entre la cámara y el objeto de interés.

De la Fig. 10 se tiene que la distancia recorrida se puede

obtener en función de la distancia referencial entre la cámara el objeto de y se puede expresar como:

$$n - \text{pixeles} = 2 * d_{ref} * \tan \theta \quad (3)$$

En donde  $n$  es la resolución horizontal del fotograma capturado.

Finalmente la velocidad de la distancia recorrida por el objeto de interés en metros es expresada como:

$$V = \frac{d}{T} = \frac{d}{T} \times \frac{2 * d_{ref} * \tan \theta}{\text{ancho fotograma}} \text{ metros/segundo} \quad (4)$$

En donde  $d$  es la distancia euclidiana,  $T$  es periodo de captura de los fotogramas y  $d_{ref}$  es la distancia referencial entre el objeto de interés y la cámara.

#### V. DETERMINAR FRECUENCIA DE MUESTREO DE LOS FOTOGRAMAS

Con los resultados obtenidos de las pruebas realizadas, se dio paso a la determinación de de la rata de muestreo para la captura de fotogramas. Teniendo como parámetros para la selección los siguientes:

Velocidad de los objetos incidentes en el escenario de grabación.

Distancia referencial entre la cámara y el objeto de interés.

La limitante de capturar los fotogramas en recintos cerrados, debido a las condiciones de iluminación inestables en espacios al aire libre.

La resolución adecuada de los fotogramas para que se pueda implementar reconocimiento de objetos.

Tamaño del fotograma capturado en el disco (ocupación de memoria).

En las pruebas de realizadas para calcular la velocidad de objetos de interés dentro de video, las cuales son las de mayor importancia, arrojaron los siguientes datos; la velocidad a la cual se desplazan las personas caminado se encuentra alrededor de los 0.60 metros por segundo, el área aproximada de las locaciones que serán los futuros escenarios de grabación son en estimación de 10 metros<sup>2</sup> y el tamaño de ocupación de memoria de un fotograma de 640x480 pixeles en escala de grises sin ningún tipo de compresión es de aproximadamente 300Kbits, variando este valor con respecto a la incidencia de luz dentro del escenario.

Tomando como promedio el tiempo que le toma a una persona recorrer la mitad distancia máxima (5 metros), el cual es de 3 segundos y el tamaño promedio de cada fotograma. Se tomo la decisión de establecer la rata de muestreo de los fotogramas en la secuencia de video, en 2 fotogramas por segundo, logrando de esta manera capturar al menos 6

imágenes para hacer análisis al objeto de interés.

El decidir una rata de muestro como esta permite hacer la compresión planteada por el proyecto, ya que los estándares existentes plantean ratas de muestreo de mínimo 15 fotogramas por segundo y resoluciones de fotogramas más pequeñas, como se muestra en la siguiente tabla.

TABLA I  
ESTÁNDARES DE COMPRESIÓN [7]

Aplicación	Estándar	Tamaño	Rata de Fotogramas
Videoconferencia Tasa baja de bit	H.263	128x96 176x144	15-30
Video Básico Telefonía	H.261	176x144 352x288	10-20
Videoconferencia	H.261	352x288	15-30
Multimedia	MPEG-1	352x240	15-30
TV Digital – NTSC	MPEG-2	720x480	30
TV Alta Definición	MPEG-2	1200x800	30-60

#### VI. SELECCIÓN ALGORITMO PARA CAPTURA LA CAPTURA DE FOTOGRAMAS

En la Fig. 11 se describe el proceso propuesto para la captura de los fotogramas, en donde como primer paso es el acceder al hardware del dispositivo, es decir establecer una conexión con la cámara, cuando ya se establece esta conexión se adquiere una vista del escenario a grabar, mientras se encuentre activa esta vista preliminar solo se almacenará en la memoria volátil un solo fotograma el cual se borrara si no se activa la grabación, para darle espacio al siguiente fotograma de vista preliminar, de esta forma se aprovechan los recursos de memoria del dispositivo, al momento de activar la grabación de la secuencia, cada fotograma capturado se almacena en la memoria a una frecuencia de dos por segundos y sin ningún tipo de compresión.

Los fotogramas son guardados sin formato, ya que las imágenes son para un propósito de análisis de la mismas posterior a la grabación, entonces no se hace necesario reproducir el video dentro del dispositivo.

#### VII. SELECCIÓN DEL CÓDEC DE REPRODUCCIÓN DE LOS FOTOGRAMAS

Después de haber adquirido los fotogramas se determino la forma en reproducirlos en una secuencia de video, para este fin se estudiaron varias arquitecturas de códec para reproducir video, finalmente se selecciono el estándar de AVI (Audio Video Interleave), por razones expuestas a continuación.

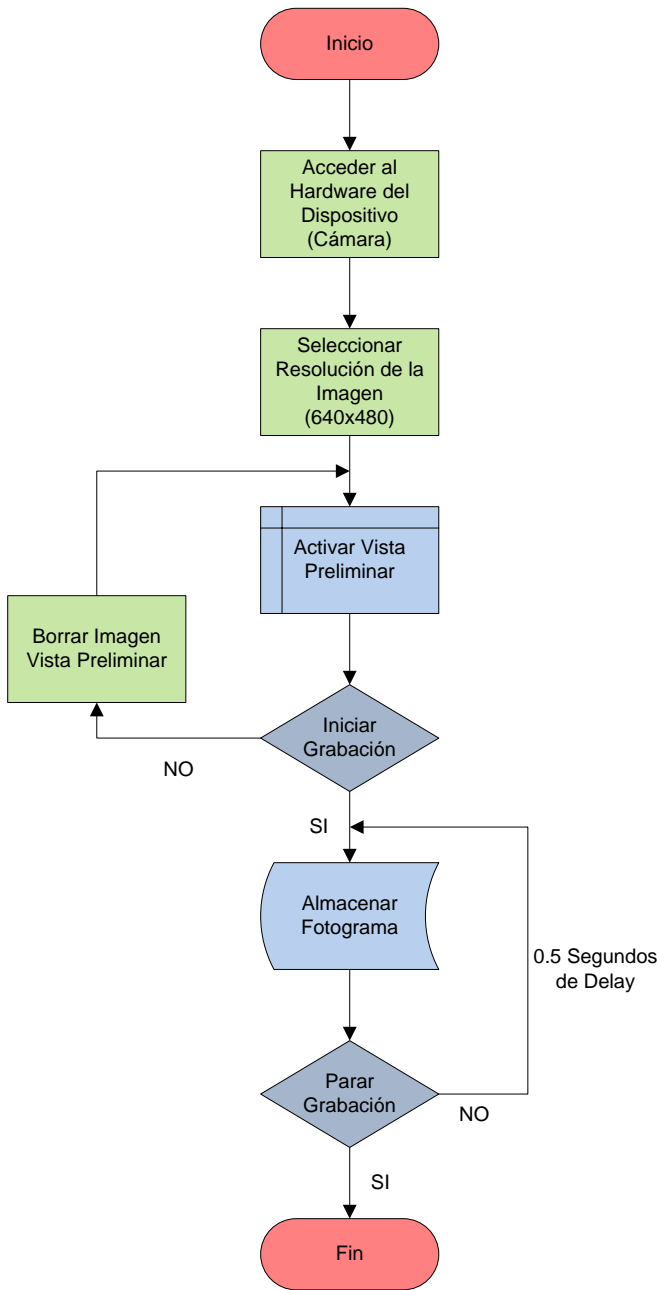


Fig. 11. Algoritmo para la captura de fotogramas.

AVI no es más que un formato de archivo desarrollado por la compañía Microsoft que puede guardar datos en su interior codificados de diversas formas y con ayuda de diversos codecs los cuales aplican diversos factores de compresión, pero también existe la posibilidad de almacenar los ficheros en un formato AVI "raw" o crudo, es decir, sin compresión.

Como uno de los propósitos específicos de este proyecto es el de no hacer ningún tipo de compresión a los fotogramas capturados no conviene crear una secuencia de video con un códec que haga compresión al reproducirlos. Con el reproductor AVI raw podemos montar los fotogramas en secuencia sin comprimirlos y es de uso gratuito, por estas razones se selecciono el reproductor AVI.

### VIII. REQUERIMIENTOS TÉCNICOS MÍNIMOS NECESARIOS

Los requerimientos técnicos mínimos que deben contener los dispositivos portables en los cuales pueda ser implementado el sistema se determinaron a partir de la resolución de los fotogramas.

El tamaño de ocupación de memoria de un solo fotograma esta dado por:

$$\text{tamaño} = \text{mapa de color en bytes} \times \text{tamaño fotograma} \quad (5)$$

Este resultado se da en bytes haciéndose necesario convertir el mapa de color de bits a bytes.

Entonces un fotograma sin ningún tipo de compresión con características como las anteriormente mencionadas tiene un tamaño en disco de:

$$\text{tamaño} = 1 \text{ byte} \times 640 \times 480 = 307.2 \text{ Kbytes} \quad (6)$$

El resultado anterior fue para un solo fotograma, así mismo si se graba una secuencia de video de un minuto con una frecuencia 30 fotogramas por segundo sin compresión el video el tamaño del video es de:

$$\text{tamaño} = 1 \text{ byte} \times 640 \times 480 \times 30 \text{ fotogramas} \times 60 \text{ segundos} = 552 \text{ Mbytes} \quad (7)$$

Como se observa en el resultado el tamaño de un video sin compresión grabado con una rata de muestreo de fotogramas de las más usadas es muy grande por lo cual se hace necesario la compresión del video.

Como se determinó en este proyecto una compresión mediante la reducción de la rata de muestreo y esta se estableció en 2 fotogramas por segundo, se calcularon las requerimientos mínimos de memoria para una grabación continua de 30 minutos, de la siguiente forma:

En la memoria volátil del dispositivo siempre debe estar almacenado un fotograma correspondiente tiempo presente, por ende como mínimo esta memoria debe ser de 307.2 Kbytes pero se recomienda un tamaño mayor a este.

Para una grabación de 30 minutos se requiere una memoria de almacenamiento de:

$$\text{tamaño memoria} = 1 \text{ byte} \times 640 \times 480 \times 2 \text{ fotogramas} \times 60 \text{ segundos} \times 30 \text{ minutos} \\ \text{tamaño de memoria} = 1.10592 \text{ Giga byte} \quad (8)$$

Requiriéndose como mínimo una capacidad de almacenamiento en el dispositivo portable de 1 Gigabyte.

## IX. PRUEBAS Y RESULTADOS

### A. Prueba 1

Para esta prueba se utilizó el laboratorio de electrónica básica de la Fundación Universitaria San Martín (FUSM), en donde se grabó un video de 6 segundos con una tasa de muestreo de 15 fotogramas por segundo y resolución de 640x480 píxeles.

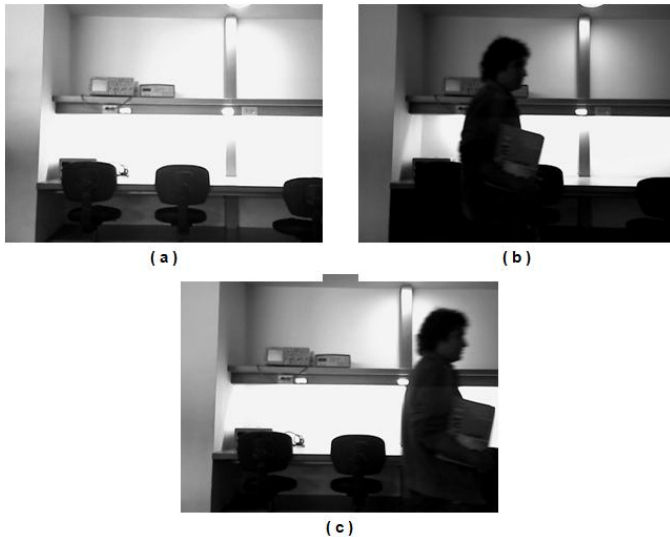


Fig. 12. Fotogramas extraídos para la Prueba 1.

En la Fig. 12, se hace alusión a las siguientes imágenes (a) Fondo estático. (b) Fotograma 60 de la secuencia de video capturada. (c) Fotograma 75 de la secuencia de video capturada, en donde se puede apreciar el inconveniente de los cambios en la iluminación en los tres fotogramas de la figura anterior esto se debe a que las cámaras tienen por defecto auto ajustar la apertura del diafragma cuando hay objetos incidentes en el escenario, la apertura del diafragma permite controlar la incidencia de luz dentro de la grabación, por tal motivo en el fotograma 60 se nota que existe menos luz a comparación de la luz incidente en el fotograma 75, lo cual indica que existió un retardo en la apertura del diafragma cuando el objeto entro en el escenario.

Estas deficiencias de iluminación no permitieron hacer los cálculos con certeza en cuanto a la posición y distancia recorrida por el objeto, tal como se evidencia en la Fig 13.

La imagen (a) de la Fig. 13 es el resultado de la resta entre el fotograma fondo estatico y el fotograma 60 y en la imagen (c) su equivalente binarizado, en la parte derecha de la figura la imagen (b) corresponde a la imagen resultante de la resta entre el fotograma 75 y el fondo estatico y en la parte inferior su equivalente binarizado. Como se puede apreciar en esta figura no es posible deducir con certeza cual es la ubicación del objeto en las imágenes debido a los cambios de iluminación que se dieron durante la grabación del video, por lo cual surge el problema de valores diferentes en los píxeles que forman el fondo del escenario entre un fotograma y otro, por esta razón no se puede hacer una resta lineal entre las imágenes. Como se

aclara en el apartado de desarrollo se necesita binarizar el fotograma para hallar el centro de masa de los objetos y en esta prueba no es posible encontrar el valor de pixel umbral, ya que se presentan valores similares de píxeles dentro del objeto y en el escenario, por eso al hacer la binarización de los fotogramas brinda información errónea acerca de la ubicación del objeto de interés.

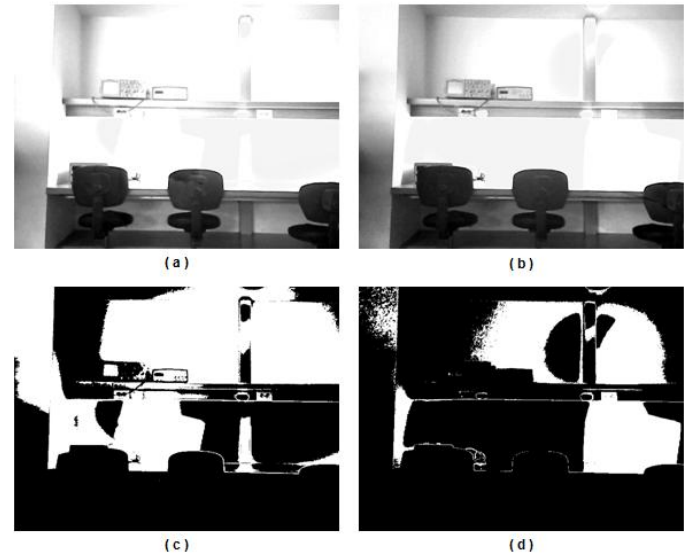


Fig. 13. Registro de errores de la Prueba 1.

### B. Prueba 2

Para esta prueba se mejoraron las condiciones de luz incidente en el escenario y se hizo una mejora en código de matlab para la manipulación de los fotogramas.

La locación para la grabación de esta prueba fue el laboratorio de comunicaciones de la FUSM. En donde se consiguió mejorar las condiciones de iluminación.

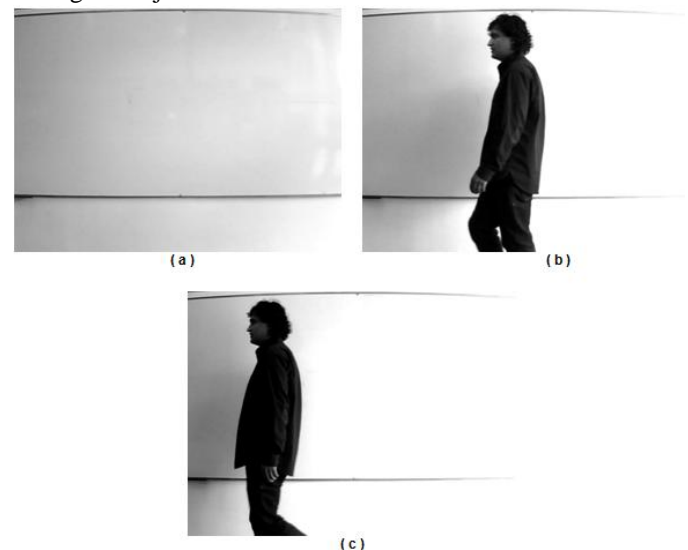


Fig. 14. Fotogramas extridos para la Prueba 2.

En la Fig. 14 se muestran las siguientes imágenes. (a) Fondo estático. (b) Fotograma 60 de la secuencia de video capturada. (c) Fotograma 75 de la secuencia de video capturada, en donde

se hacen más evidente las diferencias existentes entre los fotogramas 60 y 75 con respecto al fondo del escenario, estas diferencias permiten facilidad al determinar la existencia de un objeto incidente en el escenario. El video fue grabado con una frecuencia de muestreo de 15 fotogramas por segundo y resolución de 640x480 pixeles y un tiempo de duración de 10 segundos.

### 1) Binarización de los Fotogramas

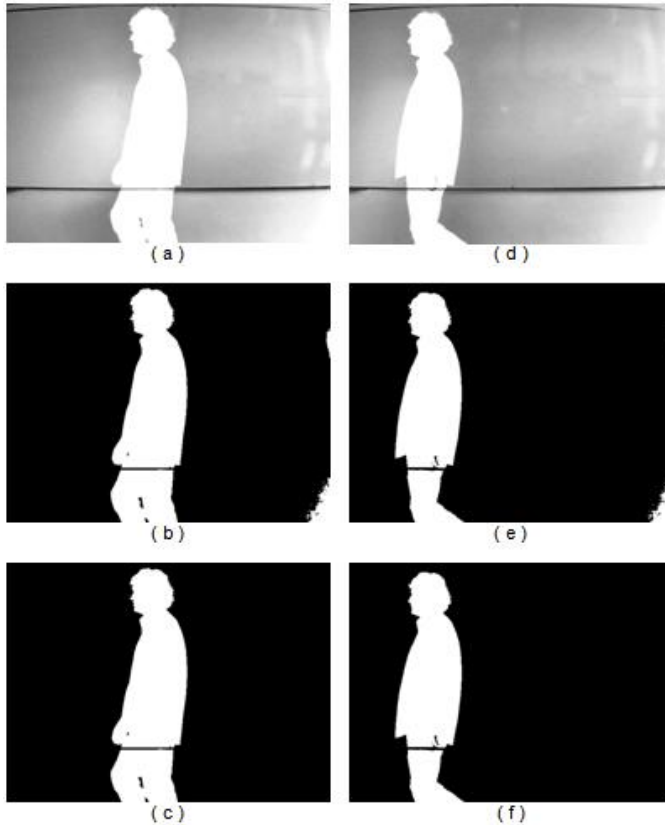


Fig. 15. Resultado de la binarización.

La parte derecha de la Fig. 15 hace alusión al proceso de binarización del fotograma 60 extraído de la secuencia video capturada para esta prueba, en la imagen (a) se observa como al restar el fotograma 60 del fondo estático se resalta la silueta del objeto incidente en el escenario en el segundo 4 de grabación, de igual manera se hace para el fotograma 75 de la secuencia y el resultado se muestra en la imagen (d). En las imágenes (b) y (e) se muestran las fotogramas binarizados correspondientes al número 60 y al número 75 respectivamente de secuencia de video, en estas imágenes se puede observar que aun se presentan problemas en la binarización ya que en la esquina inferior derecha de cada fotograma algunos pixeles quedaron con valor de '1' esto se registra como si existiera un objeto incidente en estas partes, lo cual no sucedió, esto se debe a los cambios en la incidencia de luz sobre la lente de la cámara, y si se hiciera el cálculo del centro de masa con estas imágenes se obtendría información inexacta.

El problema expuesto anteriormente se resolvió para las

imágenes (c) y (f) de los fotogramas 60 y 75 respectivamente, escogiendo un valor del pixel umbral más cercano a los pixeles que forman la silueta del objeto de interés, de esa manera se eliminaron las formas que no pertenecían propiamente al objeto.

### 2) Cálculo de la distancia euclidiana.

Para esta prueba el tamaño de las imágenes sobre las cuales se halló el centro de masa es de 640x480 pixeles. Para el fotograma 60 los cálculos arrojaron las coordenadas (288.033, 268.95) del centro de masa del objeto, puesto que los valores de las posiciones de los pixeles están dadas en valores enteros entonces se escogieron los valores de mayor proximidad a los valores del resultado, finalmente las coordenadas del centro de masa del objeto de interés en el fotograma 60 son (288, 269) y en el fotograma 75 son (159.99, 252.46) siguiendo el mismo criterio quedan (160, 252). Estas coordenadas se graficaron en la Fig. 16.

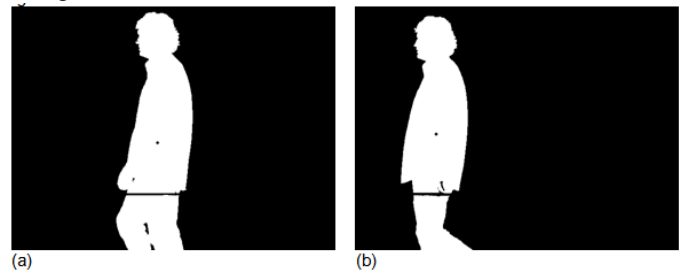


Fig. 16. Gráficas de los centros de masa.

Teniendo las coordenadas de los centros de masa del objeto de interés en dos fotogramas de una secuencia de video se emplea la ecuación de la distancia euclidiana para calcular la distancia recorrida por el objeto.

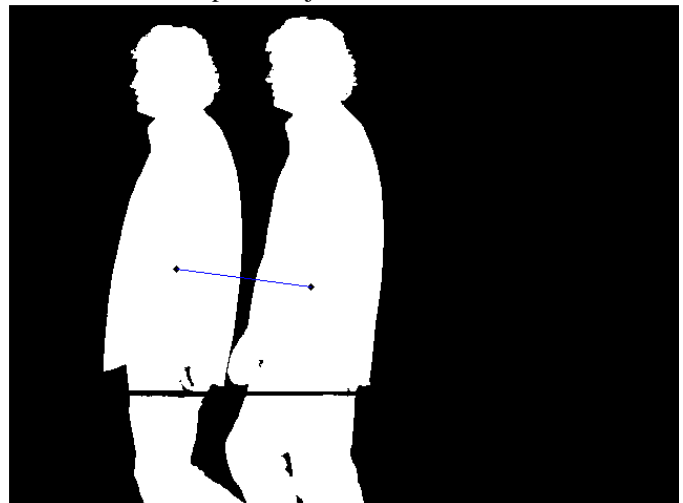


Fig. 17. Gráfica de la distancia recorrida por el objeto de interés.

La distancia recorrida por el objeto de interés la describe el comportamiento del vector graficado de color azul en la Fig. 17 y tiene una magnitud de 129.0998 pixeles.

La distancia referencial desde la cámara hasta el objetivo es de 2.20 mts. y el ángulo de apertura del lente de la cámara con

respecto a la distancia referencial es de 70.332°. Con estos datos se puede ya calcular el equivalente de la distancia recorrida en metros. La distancia recorrida por el objeto de interés fue de 0.625 mts. Como la frecuencia de muestreo de los fotogramas es de 15 fotogramas por segundo, entonces el tiempo transcurrido entre el fotograma 60 y el fotograma 75 fue de un segundo, finalmente se tiene que la velocidad del objeto de interés en esta prueba fue de 0.62 metros por segundo.

## X. CONCLUSIONES

A través del desarrollo del proyecto se documentó como los sistemas de vigilancia usados en la actualidad basan el manejo de memoria bajando la resolución de los fotogramas utilizado codecs complejos como M-PEG, minimizando así la posibilidad de reconocer objetos dentro de la secuencia de video.

El algoritmo propuesto para la grabación de la secuencia de video saca provecho los recursos de la memoria RAM de los dispositivos ya que solamente tiene que almacenar un fotograma, debido a que el periodo para la toma fotogramas al ser de 0.5 segundos durante el cual ese fotograma se guarda directamente en la memoria ROM y se elimina de RAM para dale paso al fotograma siguiente.

El proceso de binarización es muy confiable para determinar el área de objetos interés incidentes en imágenes y para este proceso el valor del pixel umbral es fundamental para hacer cálculos erróneos acerca de la posición del objeto.

La incidencia de luz en el escenario de grabación es un factor crítico en la grabación de de una secuencia de video y se hace incontrolable en espacios al aire libre debido a los constantes cambios, lo cual reduce la calidad de las imágenes adquiridas.

El calcular la velocidad de objetos dentro de una secuencia de video fue fundamental para determinar la rata de muestreo de fotogramas.

En las pruebas realizadas del movimiento de para una distancia referencial de 3.2 metros entre el la cámara y el objeto de interés, con un ángulo 70° de apertura del lente de cámara, arrojaron una velocidad 0.6 metros por segundo promedio del movimiento de las personas.

Si es posible hacer compresión de una secuencia de video disminuyendo la rata de muestreo, como se evidencio en desarrollo del proyecto ya que un video de 1 minuto sin compresión en los fotogramas a una tasa normal de muestreo de 30 fotogramas por segundo, es de un tamaño de 500 Megabytes, y con la tasa de muestreo propuesta en este proyecto se puede 30 minutos de grabación continua con una capacidad de almacenamiento de 1Gigabyte, obteniendo una mejor calidad en los fotogramas adquiridos.

Para la reproducción de la secuencia de fotogramas adquirida se selecciono el reproductor AVI raw, ya que en este brinda la posibilidad de almacenar ficheros sin compresión y

es de uso gratuito.

No es necesario el color en una imagen para el reconocimiento de patrones ya que los algoritmos utilizados para este propósito utilizan imágenes en 2 dimensiones. El trabajar con una imagen en escala de grises si facilita su análisis y velocidad de procesamiento.

Se selecciono una resolución de fotogramas de 640x480 pixeles para un campo visual de captura de 10 metros, ya que las técnicas de reconocimiento estudiados reducen los datos necesarios para identificar un individuo a 1/1000 de los datos presentados.

El algoritmo propuesto para la grabación de video permite almacenar cantidades finitas de fotogramas, no obstante debe estar supervisado de forma manual por el usuario para evitar, ya que se podrían saturar los recursos de memoria del dispositivo.

## REFERENCES

- [1] Chaudhuri, S., & Taur, D. (2005). High-Resolution Slow-Motion Sequencing: How generate a slow-motion sequence from a bit stream. *IEEE Signal Processing Magazine: Surveillance Networks*, 22(2), 16-24.
- [2] Collins, R., Lipton, A., Fujiyoshi, H., & Kanade, T. (2001). A system for video surveillance and monitoring. *Proc. IEEE*, 89(10), 1456-1477.
- [3] E.U.I.T.Telecomunicación, D. d. I. A. y. C. (2008). Retrieved marzo de 2009, 2009, from [www.diac.upm.es/acceso\\_profesores/asignaturas/tdi/tdi/transformadas/pdf/dct.pdf](http://www.diac.upm.es/acceso_profesores/asignaturas/tdi/tdi/transformadas/pdf/dct.pdf)
- [4] Foreti, G., Micheloni, C., Snidaro, L., Remagnino, P., & Ellis, T. (2005). Active video-Based Surveillance System: The low-level image and video processing techniques needed for implementation. *IEEE Signal Processing Magazine: Surveillance Networks*, 22(2), 26-36.
- [5] Gonzalez, R., Woods, R., & Eddins, S. (2003). *Digital Image Processing Using Matlab* (Primera ed., pp. 624): Prentice Hall.
- [6] Hampampur, A., Brown, L., Connel, J., Hass, N., Merlk, H., Pankanti, S., et al. (2005). Smart Video Surveillance: Exploring the concept of multiscale spatiotemporal tracking. *IEEE Signal Processing Magazine: Surveillance Networks*, 22(2), 38-45.
- [7] Herman, K. (2007). *CCTV Surveillance:Video Practices and technology*. Burlington: Elsevier.
- [8] Kapur, J. N., Sahoo, P. K., & Wong, A. K. C. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, Image Processing*. 29(3), 38-45.
- [9] Kruegle, H. (2007). *CCTV Surveillance:Video Practices and technology*. Burlington: Elsevier.
- [10] (NSTC), N. S. a. T. C., Technology, C. o., & Biometrics, S. o. (2006). *Face Recognition*: Executive Office of the President of the United States.
- [11] Parker, J. R. (2001). Gray-level thresholding in badly illuminated images. *IEEE Trans. PattenAnal. Machine Inteli.*, 13(8), 813-819.
- [12] Westwater, R., & Furht, B. (1997). *Real-Time Video Compression: Techniques and Algorithms*. Massachusetts: Academic Publishers.
- [13] Winkler, S. (2005). In W. Sussex (Ed.), *Digital Video Quality: Vision Models and Metrics* (Primera ed.). England: Willey.