

Tipo de artículo: Artículo original  
Temática: Procesamiento de imágenes.  
Recibido: 29/08/2016 | Aceptado: 16/02/2017

## Seguimiento de múltiples personas en aplicación de videovigilancia con cámaras de baja resolución

### *Multiple people tracking in video surveillance applications with low-resolution cameras*

Aldo Garcés Matilla <sup>1\*</sup>, Fidel Guerrero Peña <sup>2</sup>, Alejandro Garcés Calvelo <sup>3</sup>, Elisa Hernández Silva<sup>4</sup> y Magali Matilla Belett<sup>1</sup>

<sup>1</sup> Universidad de Oriente. Ave. Patricio Lumumba s/n Altos de Quintero, Santiago de Cuba, Cuba.

<sup>2</sup> Universidad Federal de Pernambuco. Av. Prof. Moraes Rego, 1235, Ciudad Universitaria 50.670-901, Recife, Brasil.

<sup>3</sup> Universidad Jaume I. Av. de Vicent Sos Baynat s/n 12071 Castellón de la Plana, España

<sup>4</sup> Centro de Investigaciones de Energía Solar. Micro 3, Reparto "Abel Santamaría", Santiago de Cuba, Cuba.

\* Autor para correspondencia: [agarcesm91@gmail.com](mailto:agarcesm91@gmail.com)

---

#### Resumen

La detección y seguimiento de personas es de vital interés por su amplio potencial de aplicaciones, tales como: videovigilancia, sistemas asistidos de conducción, herramientas de evaluación en el ámbito médico, interacción humano-computador. Precisamente en este trabajo es presentada una aplicación para la realización de tareas de videovigilancia con cámaras de baja resolución. Esta aplicación está basada en secuencias de video capturadas desde un sistema distribuido de visión con múltiples cámaras, permitiendo la detección y seguimiento de personas. Se utilizaron métodos y técnicas bien conocidas, integradas de manera creativa para cubrir las diferentes etapas de procesamiento, donde se introdujeron mejoras para una mayor eficiencia computacional del sistema. Especial atención reciben la sustracción de fondo con el modelo de Mezcla Gaussiana en la detección de objetos y el filtro de Kalman para el seguimiento, además los descriptores Histogramas de Gradientes Orientados y Máquina de Soporte Vectorial en la detección de persona. Para demostrar la validez de los métodos implementados se realizan una gran variedad de experimentos con cámaras de baja resolución, en los cuales se obtuvieron resultados satisfactorios. En la etapa de seguimiento de múltiples personas, solo se cometió error en el 11% de las personas detectadas, motivado fundamentalmente por la ocurrencia de oclusiones.

**Palabras Clave:** detección de objetos, seguimiento, detección de personas y sistema de videovigilancia.

## Abstract

*People detection and tracking is of vital interest for its many applications, such as video surveillance, driver assisted systems, evaluation tools in the medicine, human-computer interaction. In this work is presented an application to carry out tasks of video surveillance with low resolution cameras. This application is based on video sequences captured from a distributed multi-camera vision system, allowing the people detection and tracking. The methods and techniques used in each stage are well known, these were built creatively in the different stages of processing, and these were improved for greater computational efficiency of the system. Within these receive special attention the background subtraction with Gaussian Mixture Model to objects detection and the Kalman filter to tracking, also the descriptors Histograms of Oriented Gradients and Support Vector Machine in people detection. To demonstrate the validity of the methods implemented is realized a variety of experiments with low-resolution cameras, getting satisfactory results. In the tracking stage of multiple persons only occurred error in 11% of detected people, mainly motivated by the occurrence of occlusions.*

**Keywords:** *object detection, tracking, people detection and video surveillance systems.*

---

## Introducción

La necesidad de garantizar la seguridad de personas y bienes en entornos cada vez más masificados, explica la enorme expansión y desarrollo que están experimentando los sistemas de vigilancia en todo el mundo. El análisis de cantidades enormes de información adquiridas por estos sistemas, obliga al desarrollo de técnicas de análisis automático, con una intervención mínima por parte de operadores humanos.

La visión artificial es una de las áreas de la computación que tiene infinidad de aplicaciones en el campo de la seguridad y la vigilancia (Sun & Meng, 2015; Kinjal & Thakore, 2012). El poder automatizar ciertas funciones como el reconocimiento y seguimiento de personas, vehículos y ciertos objetos abre un gran campo de posibilidades para detectar situaciones anómalas en secuencias de video tomadas en un entorno concreto. Un sistema informático con estas capacidades, puede convertirse en una herramienta eficaz para discriminar los sucesos que merecen ser informados a los agentes encargados de los sistemas de vigilancia, reduciendo el número de acciones a realizar por ellos y simplificando el posterior proceso de indagación. Así, por ejemplo, se puede tener decenas de cámaras realizando tareas de vigilancia y muy pocos agentes encargados de monitorizar operaciones como: el control de acceso a áreas especiales, la identificación de personas a distancia, las estadísticas de movimiento de grupos de personas, el análisis de tráfico

(tales como las retenciones y atascos) y la detección de comportamientos anómalos (por ejemplo, los accidentes y actos de violencia. Esto llevaría a un sistema quizás más barato y sin duda más fiable.

En las últimas décadas se ha logrado un avance impresionante en los sistemas automáticos de vigilancia (Aggarwal & Ryoo, 2011; Parekh, Thakore, & Jaliya, 2014; Smeulders, 2014). Sin embargo, la detección automática de personas y seguimiento en secuencias de video aún constituye un problema extremadamente complejo, que requiere de nuevos métodos y herramientas cada vez más fiables, eficientes y eficaces. Además, otro inconveniente es que presenten un funcionamiento estable con secuencias de imágenes de baja resolución, ya que en estas pueden dificultarse la extracción de características y detección de personas u objetos al estar representados por un número de píxeles pequeños.

Precisamente, el objetivo de este trabajo es proponer un método para la realización de tareas de videovigilancia en ambientes no controlados con cámaras de baja resolución. Para lograr este propósito, se planteó la realización de una etapa de experimentación, para seleccionar y formular adecuaciones a importantes técnicas de visión por computadora para el desarrollo de las diferentes tareas de videovigilancia. Como resultado final de la investigación, se proyectó la implementación de una aplicación que incluirá el método propuesto.

Justamente, en este trabajo, se presenta una aplicación eficaz, eficiente y flexible, para la realización de tareas de videovigilancia en ambientes no controlados con cámaras de baja resolución. Esta aplicación está basada en secuencias de video capturadas desde un sistema distribuido de visión con múltiples cámaras, permitiendo la detección y seguimiento de personas. Para ello, se utilizan métodos y técnicas bien conocidas, las cuales fueron integradas de manera creativa para cubrir las diferentes etapas de procesamiento, tales como la sustracción de fondo con el modelo de Mezclas Gaussianas (Stauffer & Grimson, 2000) para la detección de objetos, el filtro de Kalman (Patel & Thakore, April 2013) para realizar el seguimiento. En la etapa de detección de personas se utilizó el método de Dalal y Triggs (Dalal & Triggs, 2005), este fue aplicado solo a las regiones de interés devuelto por la sustracción de fondo y así proporcionar mayor eficiencia en la implementación. Para demostrar la validez de la aplicación propuesta se realizan experimentos, en los cuales se obtuvieron satisfactorios resultados.

## **Trabajos Relacionados**

En este epígrafe se introducen los principales métodos y técnicas relacionadas con el seguimiento de objetos y detección de personas utilizados en este trabajo. Ellos constituyen la base de la herramienta de videovigilancia desarrollada.

## **Seguimiento de objetos**

El seguimiento de objetos mediante videocámaras ha sido un tema muy explorado en el campo de la visión por computador (Parekh, Thakore, & Jaliya, 2014). En el contexto del presente trabajo, el seguimiento puede ser dividido en dos principales etapas (Athanesious & Suresh, October 2012): detección de objetos interesantes en la escena y seguimiento de estos objetos.

La detección de los objetos permite realizar la localización de los mismos en sucesivos fotogramas (Parekh, Thakore, & Jaliya, 2014). Entre otros, especial atención se les ha brindado a los métodos basados en detectores de puntos (Gauglitz, Höllner, & Turk, 2011); en segmentación (Shi & Malik, 2000; Ukinkar & Samvatsar, 2012); en sustracción de fondo (Stauffer & Grimson, 2000; Zhang & Ding, 2012); y en el entrenamiento supervisado (Santhanam, Sumathi, & Gomathi, 2012). Muchos de ellos se comportan dependientes del ambiente y tienen alto costo computacional.

Los métodos de seguimiento de un objetivo tienen como finalidad la implementación de un seguidor encargado de generar la trayectoria del objetivo a lo largo de un periodo de tiempo mediante la localización de su posición en cada fotograma del video (Parekh, Thakore, & Jaliya, 2014). Estos métodos son divididos en tres grandes grupos (Kinjal & Thakore, 2012) seguimiento por puntos (Patel & Thakore, April 2013; Yin, Na, Choi, & Oh, 2011), de núcleo (Mishra, Mahesh, & Nitna, 2012; Athanesious & Suresh, March 2013) y silueta (Cole, Austin, & Cole, 2004; Athanesious & Suresh, March 2013).

En el presente trabajo se le presta especial interés al método de sustracción de fondo con modelo de mezcla gaussiana (Stauffer & Grimson, 2000) para la representación del fondo, debido a que constituye una técnica robusta y adaptable a los cambios de ambiente (Zhang & Ding, 2012). Para la etapa de seguimiento se considera oportuno un método de seguimiento por punto, específicamente el estadístico por correspondencia filtro de Kalman (Welch & Bishop, 2006), dado su eficacia y rapidez (Patel & Thakore, April 2013).

## **Detección de persona**

A la hora de reconocer los objetos como personas se puede recurrir a diferentes enfoques (Dollar, Wojek, Schiele, & Perona, 2012; Paul, Haque, & Chakroborty, 2013). Uno de ellos, son los sistemas basados en detectores de formas, seleccionando un conjunto de rasgos para discriminar los objetos que no son de interés con respecto a los que sí lo son. Otra opción es emplear sistemas basados en descriptores, tales como los Descriptores Basados en el Análisis de Componentes Principales (PCA) (Shlens, 2005) muy utilizados en la detección de peatones, los descriptores Wavelet de Haar (Viola & Jones, 2004) que permiten definir de manera robusta clases de objetos complejos, siendo invariantes

a cambios de color y de textura, y los descriptores SIFT (Lowe, 2004) orientado a la extracción de características distintivas de una imagen en escala de grises, de tal manera que pueda reconocer una misma característica entre diferentes vistas de un mismo objeto o escenas.

Sin embargo, por la naturaleza del problema abordado en este trabajo, se requiere de algún método más robusto a cambios bruscos de iluminación. Una alternativa a los métodos anteriormente mencionados, lo constituyen los Descriptores Basados en Histogramas de Gradientes Orientados (HOG), presentado por Navneet Dalal y Bill Triggs en el Instituto Nacional de Investigación en Informática y Automática (INRIA), en 2005 (Dalal & Triggs, 2005). Además de su buen comportamiento frente a diferentes condiciones de iluminación, los HOG exhiben robustez ante pequeños cambios en el contorno de la imagen, diferentes fondos y escalas, que garantizan buena precisión.

### Método de Detección y Seguimiento de Personas

En este epígrafe se propone un método de detección y seguimiento de personas para el desarrollo de aplicaciones de videovigilancia. El mismo integra algunos métodos y técnicas descritas en el epígrafe 1, e incorpora algunas mejoras en el procesamiento relacionado.

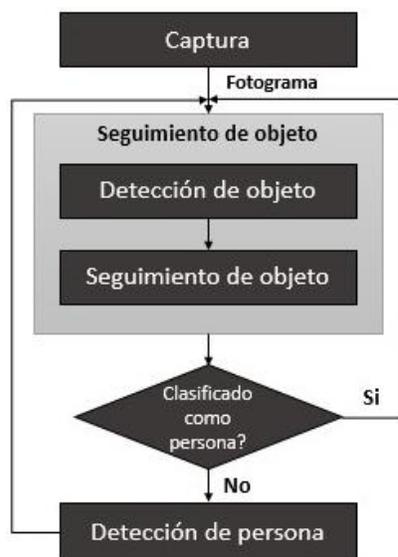


Figura 1. Diagrama general del procesamiento

En la Figura 1 se muestra el esquema general del funcionamiento del método. Como se observa para cada imagen capturada por cámara, se da seguimiento a los objetos de interés encapsulados en este diagrama como un único subproceso. Cada objeto de interés (ROI) que se le da seguimiento pasa por una etapa de clasificación con memoria

para conocer si es “persona” o “no persona”, el cual evita clasificaciones redundantes en imágenes sucesivas y lograr una mejor eficiencia computacional. Esto es debido a que sí el objeto fue identificado como persona en fotogramas anteriores no pasa nuevamente a la etapa de detección de persona y así lograr una mayor eficiencia, como evidencia el esquema.

### **Seguimiento de objeto**

El seguimiento de objetos se divide en dos subprocesos, detección de objeto y seguimiento de objeto. Este método comienza con la extracción de la información de la escena captada mediante una cámara. La detección de los objetos en movimiento es abordada como un problema de sustracción de fondo, debido a la adquisición del video a través de cámaras fijas. Aquí se propone un método de sustracción de fondo basado en el modelo mezcla gaussiana (MOG), que permite el aprendizaje del fondo de la escena mediante la combinación de distribuciones de Gauss. A partir de este aprendizaje y los fotogramas obtenidos en etapa anterior, pueden ser extraídos del fondo aquellos objetos en movimiento, eliminando sombras para facilitar la detección de las personas. En base a la máscara de movimiento devuelto por la sustracción de fondo se van a descartar aquellos contornos con área  $< \delta$  o área  $> \beta$ , ya que no son candidatos a ser personas, donde  $\delta$  y  $\beta$  son valores tomados en dependencia de la distancia de la cámara con respecto a la escena. Estos descartes permitirán deshacerse de objetos según el tamaño de los objetos al ser muy pequeños o demasiados grandes. El resto de los objetos detectados van a ser englobados por rectángulos, siendo estas las Regiones de Interés (ROIs), con las que se va a trabajar durante el resto del procesamiento.

Para realizar el seguimiento de la persona se propone el uso del filtro de Kalman el cual es un potente algoritmo recursivo que se utiliza para estimar la posición de un punto o característica en movimiento y la incertidumbre de la medida en una imagen posterior en una fase de predicción y realiza una fase de corrección en base a la predicción y la medición visual. Este método se le aplica a cada ROI detectado en la fase anterior, los cuales se le dan seguimiento mediante el centroide del objeto de cada uno en la secuencia de fotogramas, para un posterior reconocimiento (clasificar si es persona). En la Figura 2 se muestra el seguimiento en un instante de la secuencia de video, donde cada objeto se le asigna un identificador (ID) que será el mismo en toda la secuencia de video.

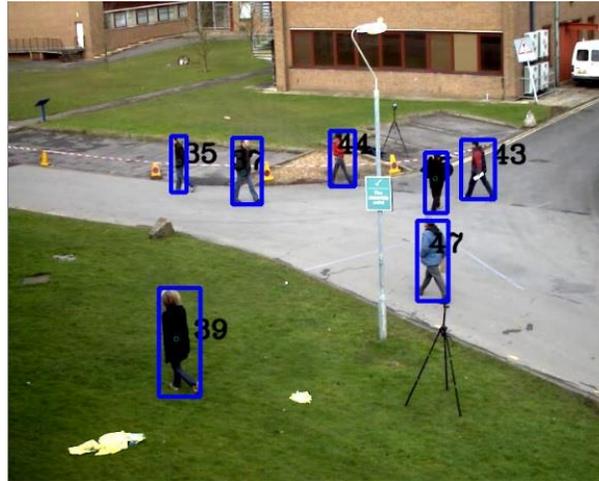


Figura 2. Ejemplo de seguimiento, cada objeto es identificado con un ID

### Detección de persona

Para la detección de personas se propone una variante del método de Dalal y Triggs (HOG + Máquina de Soporte Vectorial, SVM) (Dalal & Triggs, 2005), circunscribiéndolo únicamente a aquellas regiones de interés (ROIs) que contengan a uno o varias personas y así lograr una mayor eficiencia en el sistema.

El detector utiliza una ventana de localización de deslizamiento que se mueve alrededor de la imagen. En cada posición de la ventana del detector, se calcula un descriptor de HOG. Este descriptor es procesado por la SVM (Abe, 2010) entrenada previamente, para su clasificación (como “persona“ o “no persona“).

Antes de calcular las características puede realizarse una fase previa de adaptación de la imagen de entrada para aumentar la probabilidad de encontrar el contorno correctamente y también lograr una mayor eficiencia. Dicha fase puede contemplar la transformación de la imagen de RGB a escala de grises para mejorar el rendimiento y simplificar los cálculos sucesivos, modificar el tamaño de la imagen (por ejemplo, a 64x128) y la ecualización del histograma de la imagen, con lo cual se maximiza el contraste.

El beneficio de realizar la conversión a escala de grises consiste en reducir el número de píxeles que se tienen que computar, ya que se pasa de una imagen de tres canales a una de un canal. Esta simplificación se puede realizar por el hecho que para realizar el cálculo de bordes no son relevantes los valores concretos de intensidad de cada canal, sino más bien la variación de intensidad global que se produce.

## Implementación de un sistema de videovigilancia

El sistema de videovigilancia se basa en el análisis de secuencias de fotogramas capturadas por un sistema distribuido con múltiples cámaras. Estas secuencias pueden ser recibidas mediante cámaras instaladas en un dispositivo de captura (cliente) o en la estación de procesamiento (servidor) como muestra la Figura 3. El cliente se encarga de capturar y enviar las secuencias de fotogramas al servidor por red. En tanto, el servidor presenta un comportamiento dual, pues puede controlar directamente cámaras (comportándose como un cliente) o recibe las secuencias de video enviadas desde los clientes asociados, concentrando el procesamiento y la visualización de los resultados. Con el objetivo de mejorar el análisis de las secuencias de video en el servidor, se utiliza un procesamiento multi-hilo, uno por secuencia de fotogramas capturada por cada cámara, tanto local o remota.

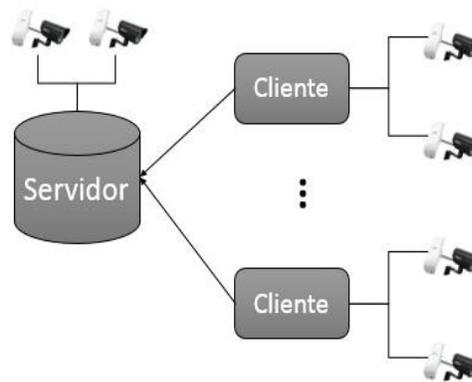


Figura 3. Capturas de las imágenes

El control del sistema, por parte de los agentes de seguridad, es realizada mediante la interfaz que se muestra en la Figura 4. En la parte izquierda se visualizan las cámaras conectadas, tanto local o remota y en la derecha se observa de manera ampliada cualquier captura señalada por el usuario. En la parte inferior de la interface se encuentra un panel de notificaciones del procesamiento. En las capturas las personas son delimitadas por un rectángulo y se le asigna un identificador a cada persona. Se dan diferentes opciones al usuario, entre ellas, que el sistema grabe de manera automática en caso de que se esté dando seguimiento a personas, estas secuencias de videos son guardadas para un posterior análisis. Además, existe la posibilidad de ver el historial de trayectoria que ha llevado la persona en toda la secuencia de las imágenes, como vemos en la captura ampliada que visualizada en la Figura 4.

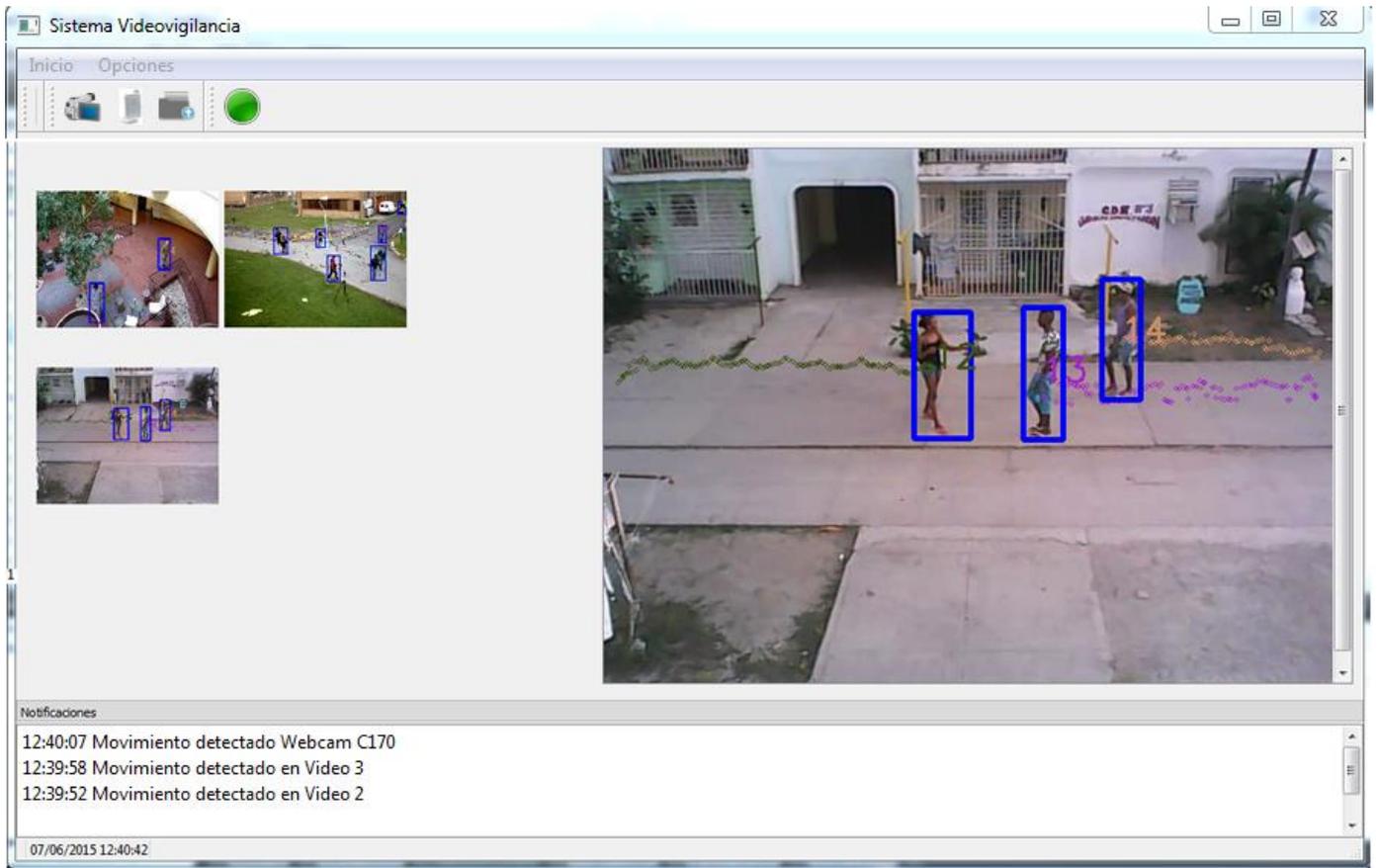


Figura 4. Interfaz del sistema de videovigilancia

## Validación del Sistema de Videovigilancia

Este epígrafe está dedicado a la evaluación de los resultados de los algoritmos de seguimiento de persona implementados. Para realizar los experimentos se realizarán varias pruebas, probando diferentes trayectorias dentro de la escena, de tal forma que se obtenga los resultados de estas trayectorias y también se conozca la trayectoria real, para validar si son seguidos correctamente.

Muchas veces el funcionamiento de un algoritmo depende en gran medida de la calidad de la resolución de la imagen sobre la que se trabaja, debido a que facilita y acelera el proceso computacional de detección de los algoritmos. Al tener imágenes con una resolución aceptable, no existiría demasiado ruido en las mismas. En este proyecto se ha trabajado con una cámara web de bajo costo Logitech C170, que presenta una frecuencia de 30 FPS y una resolución de video 640x480.

## Seguimiento simple de persona

En este experimento se desea conocer los errores de estimación del filtro de Kalman. Se realizó una comparación de los resultados de las estimaciones del filtro de Kalman con las mediciones visuales (localización real de la persona). Se consideraron 10 trayectorias diferentes de personas, donde se obtuvo el promedio del *error* de toda la trayectoria, el error en cada fotograma es la norma al cuadrado, la cual es definida como la sustracción de la *estimación* obtenida  $\hat{Q}_k$  y la *medición visual*  $Q_k$ , como muestra la Ecuación 1.

Ecuación 1. Error en un fotograma de la estimación del Filtro de Kalman

$$error = \|\hat{Q}_k - Q_k\|^2$$

En la Tabla 1 se muestra el promedio de la norma al cuadrado del error para las 10 trayectorias, en el cual se evidencia una variación entre 6,64 y 56,50, lo que significa bastante buena para el seguimiento. En general los errores obtenidos disminuyen en el tiempo como resultado de la fase de corrección de los datos predicho en la fase de predicción con la medición visual, por lo que siempre es posible continuar el seguimiento en fotogramas sucesivos correctamente.

Tabla 1. Promedio de la norma al cuadrado del error para 10 trayectorias diferentes

#Trayectoria	Valor
1	14,10
2	6,64
3	55,36
4	15,57
5	54,63
6	44,71
7	17,11
8	56,50
9	41,67
10	13,11

## Seguimiento de múltiples persona

En este experimento se tomó una secuencia de video con la cámara web Logitech C170 descrita anteriormente, con una duración de 2:25:35 horas, situando la cámara en el exterior de la Facultad de Matemática Computación de la Universidad de Oriente. Este video presenta vibraciones y variaciones de la iluminación. Los resultados obtenidos se muestran en la Tabla 2, teniendo en cuenta los siguientes criterios: se considera un *error* de seguimiento cuando se pierde la persona por completo o en otros términos cuando la cantidad de muestras pérdidas del total que se toman

durante su trayectoria en la escena es mayor al 75 %, en el caso de haberse perdido la persona menos del 25% de las muestras se considera *muy bueno*, menor al 50% *bueno* y menor de 75% *regular*. Durante la secuencia de video de detectó 136 personas, en el cual se obtuvo resultado muy satisfactorio, donde se cometió *error* en el seguimiento solamente en el 11% de las personas detectadas. En tanto, el 50 por ciento de las personas se le dio un seguimiento *muy bueno*, el 26 por ciento fue *bueno* y el 13 por ciento *regular*. Los mayores problemas para las pérdidas momentáneas de las personas que se le da seguimiento en la escena fueron las oclusiones personas-persona o persona-objetos.

Tabla 2. Resultados del seguimiento

Pérdidas de muestras	<25%	<50%	<75%	>75%	Total
Cantidad de personas	68	36	17	15	136

< 25%: *Muy Bueno*; < 50%: *Bueno*; < 75%: *Regular*; >75%: *Error*.

## Conclusiones

En el presente trabajo se presentó una aplicación potente, eficiente y flexible, para la realización de tareas de videovigilancia con cámaras de baja resolución. Esta aplicación permite la detección y seguimiento de personas de manera automática. La experimentación realizada, demuestra la validez de los métodos implementados. En la etapa de seguimiento se obtuvieron satisfactorios resultados, donde solo se cometió error en el 11% de las personas detectadas, motivadas fundamentalmente por la ocurrencia de oclusiones persona-persona o persona-objetos.

El sistema propuesto, en este trabajo, puede ser utilizado en una importante clase de aplicaciones para la videovigilancia. Sin embargo, los resultados obtenidos en la experimentación realizada, ponen de manifiesto algunos aspectos en los que se debe seguir trabajando para mejorar su comportamiento en condiciones más complejas. El principal problema en que debe profundizarse, es aquel relacionado con las oclusiones que pueden presentarse en las escenas. Estas introducen mayores complejidades en el seguimiento de objetos.

## Referencias

ABE, S. (2010). *Support Vector Machines for Pattern Classification*. London: Springer.

AGGARWAL, J., & RYOO, M. (2011). Human activity analysis: A review. *ACM Computing Surveys (CSUR)*, 43(3), 16.

- ATHANESIOUS, J., & SURESH, P. (March 2013). Implementation and Comparison of Kernel and Silhouette Based Object Tracking. *International Journal of Advanced Research in Computer Engineering & Technology*, pp 1298-1303.
- ATHANESIOUS, J., & SURESH, P. (October 2012). Systematic Survey on Object Tracking Methods in Video. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 242-247.
- COLE, L., AUSTIN, D., & COLE, L. (2004). Visual Object Recognition using Template Matching. *Australasian Conference on Robotics and Automation*.
- DALAL, N., & TRIGGS, B. (2005). Histogram of oriented gradients for human detection. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (pp. 886-893).
- DOLLAR, P., WOJEK, C., SCHIELE, B., & PERONA, P. (2012). Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4).
- GAIGLITZ, S., HÖLLERER, T., & TURK, M. (2011). Evaluation of interest point detectors and feature descriptors for visual tracking. *International Journal of Computer Vision (IJCV)*, 94(3), pp. 335-360.
- KINJAL, A., & THAKORE, D. (2012). A Survey on Moving Object Detection and Tracking in Video Surveillance System. *International Journal of Soft Computing and Engineering (IJSCE)*, 2(3).
- LOWE, D. (2004). Distinctive image features from scale-invariant keypoints. *International J. Comput. Vision*, 60(2).
- MISHRA, R., MAHESH, K., & NITNA, D. (2012). Multiple Object Tracking by Kernel Based Centroid Method for Improve Localization. *International Journal of Advanced Research in Computer Science and Software Engineering*, pp 137-140.
- PAREKH, H., THAKORE, D., & JALIYA, U. (2014). A Survey on Object Detection and Tracking Methods. *International Journal of Innovative Research in Computer and Communication Engineering*.
- PATEL, H., & THAKORE, D. (April 2013). Moving Object Tracking Using Kalman Filter. *International Journal of Computer Science and Mobile Computing*, 326 – 332.
- PAUL, M., HAQUE, S., & CHAKROBORTY, S. (2013). Human detection in surveillance videos and its applications - a review. *EURASIP Journal on Advances in Signal Processing*(1), 1.

- SANTHANAM, T., SUMATHI, C., & GOMATHI, S. (2012). A survey of techniques for human detection in static images. *Proceeding of the International Conference on Computational Science, Engineering and Information Technology(CCSEIT), ACM*, (pp. 328-336). New York, NY, USA.
- SHI, J., & MALIK, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- SHLENS, J. (2005). Tutorial on Principal Component Analysis. *Systems Neurobiology Laboratory*.
- SMEULDERS, A., CHU, D. M., CUCCHIARA, R., CALDERARA, S., DEGHAN, A., & SHAH, M. (2014). Visual tracking: An experimental survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7), 1442-1468.
- STAUFFER, C., & GRIMSON, W. (2000). Learning patterns of activity using real time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8).
- SUN, Y., & MENG, M. (2015). Multiple moving object tracking for automated visual surveillance. *IEEE International Conference on Information and Automation*, (pp. 1617-1621).
- UKINKAR, V., & SAMVATSAR, M. (2012). Object detection in dynamic background using image segmentation: A review. *International Journal of Engineering Research and Applications (IJERA)*, 2(3), 232-236.
- VIOLA, P., & JONES, M. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2), 137-154.
- WELCH, G., & BISHOP, G. (2006). An Introduction to the Kalman Filter. *Technical report, TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill*.
- YIN, S., NA, J., CHOI, J., & OH, S. (2011). Hierarchical Kalman-particle filter with adaptation to motion changes for object tracking. *Computer Vision and Image Understanding*, 115(6), 885-900.
- ZHANG, R., & DING, J. (2012). Object Tracking and Detecting Based on Adaptive Background Subtraction. *International Workshop on Information and Electronics Engineering (IWIEE)*, 1351–1355.