

Filtro mediana recursivo para la estimación de fondo y segmentación de primer plano en videos de vigilancia

Freddy Alexander Díaz González, David Alejandro Arévalo Suárez

Universidad Sergio Arboleda,
Colombia

{freddy.diaz, david.arevalo}@correo.usa.edu.co

Resumen. El uso de cámaras de video es ampliamente usado en los sistemas de vigilancia, y ofrece la posibilidad de realizar el procesamiento de las imágenes capturadas para la detección automática de eventos de interés que se puedan presentar en la escena. El siguiente trabajo propone un método de estimación del fondo y segmentación del primer plano en videos de vigilancia, mediante el uso de un filtro mediana recursivo, con la aplicación de una ventana móvil temporal en la cantidad de fotogramas a analizar, que ofrezcan una mayor robustez frente al ruido causado por los cambios de iluminación y vibraciones de la cámara, limitando el incremento del costo computacional durante el procesamiento.

Palabras clave. Mediana temporal, estimación de fondo, primer plano, recurrencia.

Recursive Median Filter for Background Estimation and Foreground Segmentation in Surveillance Videos

Abstract. Video cameras are widely used in surveillance systems; this offers the possibility of processing the captured images for automatic detection of events of interest that may arise in the scene. The present paper proposes a method for estimating the background and foreground segmentation in video surveillance using a recursive median filter and applying a temporal moving window in the number of frames to be analyzed, which provide more robustness against noise caused by changes in illumination and camera shake, limiting the increase in the computational cost of processing.

Keywords. Temporal median, background subtraction, foreground, recurrence.

1. Introducción

En años recientes se popularizaron los Sistemas de Vigilancia con Cámaras de Video (SVCV), debido a su bajo costo, facilidad de instalación y mantenimiento. Este fenómeno contribuyó en el incremento del uso de cámaras y dio un gran impulso al campo de investigación de procesamiento de imágenes y visión por computador [1, 2, 3]. Investigaciones realizadas en el año 2011, aproximaron el número de cámaras instaladas en el Reino Unido a unas 1.8 millones [4].

A pesar de toda la información generada por las cámaras, el uso de ésta no es eficiente sino es analizada y debido al alto volumen de videos capturados, esta información debería ser procesada de forma automática. El procesamiento automático de las secuencias de video, es un área de investigación que tiene como objetivos captar, analizar, identificar y alertar, sobre patrones o conductas no deseadas, extraídas de los videos tomados por los SVCV [5, 6].

Generalmente los SVCV se instalan de forma permanente en zonas determinadas, en donde la información útil que se espera capturar se desarrolla en regiones dinámicas, tales como peatones, paquetes abandonados y demás elementos que no pertenezcan a la zona vigilada. Estos se conocen como Primer Plano (PP) [7]. Por otro lado, las imágenes que permanecen constantes con el paso del tiempo en las secuencias de video son denominadas Fondo del Video (FV) [8].

Los objetos de interés en los videos poseen información que puede ser extraída mediante segmentación, representación de objetos y medición de características [9]. La idea más común para la segmentación del PP es contar con un modelo de FV y dado que un video es una secuencia de fotogramas o cuadros, el PP resultará de la diferencia entre el cuadro del video que se esté analizando y el modelo del FV con el que se cuente en ese instante [8]. La segmentación de objetos en videos de vigilancia es el primer paso y también una de las tareas más críticas en el análisis de éstos.

En este trabajo se propone un método recursivo para la estimación del Modelo del Fondo de un Video (MFV), que posteriormente se utiliza para la segmentación del PP en videos de vigilancia. Los resultados se validan contra una estimación del PP hecha manualmente y la modificación de los parámetros necesarios para el cálculo del MFV, con el fin de realizar análisis de errores en la estimación de MFV y PP; y la complejidad del MFV. En la sección 2 de este documento, se analizan los diferentes métodos para la estimación de FV y segmentación del PP, en la sección 3 se formula el modelo desarrollado, en la sección 4 se exponen los resultados de las pruebas realizados contra la segmentación manual del PP, los cuales son analizados y presentados en la sección 5, finalmente la sección 6 concluye el artículo y presenta las líneas de trabajo futuro.

2. Revisión de métodos para la estimación de modelos de fondo en videos

Muchos de los métodos utilizados para el análisis de videos, que tienen como objetivo segmentar el PP, se basan en un conocimiento a priori del MFV, que puede ser estimado a partir de una secuencia de cuadros de la escena que no contenga elementos dinámicos [10, 11]. Comparado a la naturaleza espacial de las imágenes (estática), los videos tienen características espaciales y temporales. La segmentación de un video en el dominio espacial es justo como segmentar una imagen estática. La segmentación de una secuencia de frames de

video en el dominio temporal es llamada segmentación temporal o detección de tomas [12].

Usar secuencias de frames de un video en la segmentación temporal permite que el MFV se actualice y ajuste a los diferentes cambios de iluminación o ruido generado por las vibraciones en la cámara. De esta manera se reducen los falsos positivos generados en la segmentación del PP [13]. A partir de esto se clasifican los métodos de estimación de MFV. La primera categoría se compone de los métodos recursivos, la segunda de los métodos no recursivos.

2.1. Métodos no recursivos

Los métodos no recursivos son los que estiman un MFV basados en " n " cuadros del video previamente almacenados en un buffer [14, 15], los modelos estimados no son adaptables a variaciones del ambiente y ruidos de entrada. En los métodos no recursivos hay modelos que calculan la media para estimar el valor de cada pixel del MFV [16], otros métodos utilizan la mediana que es más robusta ante los ruidos del ambiente [17]. Debido al tipo de almacenamiento de cuadros que requieren los métodos no recursivos, el sistema en el que se implementa presenta grandes requerimientos de memoria [18].

2.2. Métodos recursivos

Los métodos recursivos son aquellos que actualizan el MFV con cuadros de entrada [8, 19]. Dentro de estos métodos se encuentran los modelos estadísticos. Por ejemplo en [20, 21] analizan la densidad de la distribución Gaussiana del comportamiento cuadro a cuadro de un pixel, para determinar la pertenencia o no de este al MFV. También se utilizan modelos Gaussianos para estimar fondos multimodales, como los presentados en [22, 23, 24], en los que se combinan " n " modelos Gaussianos para estimar " n " MFV para cada canal de pixel. Modelos estadísticos más complejos se presentan en [25, 26, 27] en los que se integran métodos como las cadenas de Markov y la redes Bayesianas, en una unión basada en el modelo general de Ising.

Los métodos recursivos estadísticos actualizan constantemente el MFV, lo que incrementa la robustez en los algoritmos de segmentación de PP, debido a que el MFV se ajusta a las variaciones en los cambios de iluminación y otro tipo de ruidos de entrada [28]. Sin embargo, dependiendo de la frecuencia con la que se ingrese un nuevo cuadro al modelo, este puede mantener errores de estimación por largos periodos de tiempo [18]. Del mismo modo, los cálculos de las distribuciones Gaussianas tienen un costo computacional alto [13].

Otra variación de un método recursivo es el Modelo Kernel, que es típicamente un Gaussiano. Este método estima la función de densidad de la probabilidad para cada pixel denominada "Kernel Density Estimation" (KDE) [29, 30]. A diferencia de los Modelos de Mezcla de Gaussianos (GMM del inglés Gaussian Mixture Model), el modelo KDE es una aproximación más general que no asume alguna forma específica de la función de densidad [31, 32]. Este método es muy utilizado por sus propiedades de continuidad, diferenciabilidad y localización. Sin embargo, el mayor problema de usar esta técnica es su costo computacional [33].

La característica de actualizar el MFV que tienen los métodos aumenta su robustez ante los factores de ruido, disminuyendo los falsos positivos en el proceso de segmentación de PP. Sin embargo esa misma característica ocasiona que objetos no pertenecientes al fondo del video, sean asumidos como parte del fondo después de permanecer estáticos en la escena, durante una secuencia de video determinada [17]. Una forma de evitar este efecto, puede basarse en la incorporación de un parámetro de memoria en la estimación del MFV, para que cada nuevo MFV generado recuerde modelos anteriores, evitando que elementos no pertenecientes al fondo sean asumidos como parte de él.

3. Formulación del método

Las principales ventajas de los métodos no recursivos que se basan en el filtro de la mediana, son la alta eficiencia computacional y su robustez al ruido [10, 11, 28, 34]. Adicionalmente, los errores obtenidos en las

estimaciones de MFV basadas en la mediana, son altamente comparables con los resultados obtenidos por métodos más complejos como los basados en GMM [8].

En la actualidad la mayoría de secuencias de video capturadas por los SVCV se almacenan en fotogramas a color. Sin embargo, el procesamiento para la segmentación del PP se puede hacer más rápido en secuencias monocromáticas o escala de grises [35]. Esto debido a que en una imagen gris, el componente rojo, verde y azul de un pixel (RGB del inglés Red Green Blue color model), tienen el mismo valor [36], por lo cual un procesamiento de un cuadro perteneciente a una secuencia de video, se puede llevar a cabo trabajando solo con la información de la matriz generada luego de realizar la conversión a la tonalidad de grises de las tres matrices RGB [10].

A continuación se presenta un Método recursivo de estimación MFV y el Modelo de Segmentación del Primer Plano (MSPP), basado en el filtro de la mediana. El MSPP estima el MFV sobre un número consecutivo y limitado de cuadros de un video, en adelante definido como ventana temporal móvil. Adicionalmente, el modelo propuesto contempla parámetros de memoria de MFV anteriores, un análisis para seleccionar el rango del umbral utilizado en el proceso de segmentación de PP y la determinación del grado de complejidad del modelo dado por el tamaño de la mediana temporal aplicada.

3.1. Modelo de estimación MF

El MFV se construye según la siguiente ecuación:

$$FC_i = \begin{cases} FA \times mediana(VC_0:VC_i) + (FM \times FC_{i-1}), & i < x, \\ FA \times mediana(VC_{i-x}:VC_i) + (FM \times FC_{i-1}), & i \geq x, \end{cases} \quad (1)$$

donde:

- FC_i , cuadro del video del fondo que se está estimando.
- FA , factor de ganancia que tendrá la mediana móvil en el cuadro de fondo actual.
- VC_i , cuadro del video que ingresa a la estimación del cuadro del fondo actual.

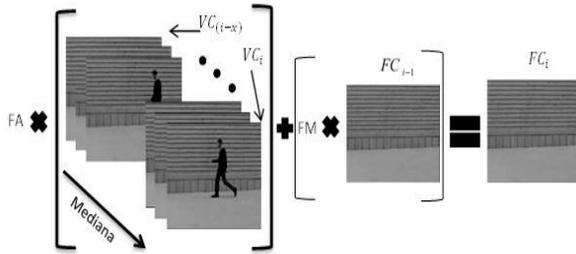


Fig. 1. Representación gráfica de la ecuación (1)

- FM , factor de ganancia que tendrá el cuadro de fondo.
- x , cantidad de cuadros sobre los cuales se aplica el filtro mediana.
- $(VC_0:VC_i)$, cantidad de cuadros que se utilizan para estimar los i primeros cuadros cuando $i < x$

En el método propuesto, la estimación del MFV se compone de la ponderación entre el MFV calculado recursivamente aplicando el filtro mediana ($mediana(VC_{i-x}:VC_i)$), y el MFV calculado inmediatamente anterior (FC_{i-1}). La proporción entre ellos en el MFV FC_i es controlada por las constantes FA y FM , las cuales deben cumplir $FA+FM=1$.

El aporte fundamental del método es la inclusión en el cálculo del MFV, de una componente de memoria que disminuye el error en la estimación del MFV, causado por elementos no pertenecientes al fondo, que tienen una baja dinámica en la escena analizada.

3.2. Método de segmentación de PP

El comportamiento de un pixel en la secuencia de fotogramas que componen el video de vigilancia, se muestra en la figura 2. Se puede observar en esta gráfica, que el valor de pixel no es constante en cada cuadro, debido al ruido inducido por el ambiente y las vibraciones del sistema de captura (cámara de video). Sin embargo, el valor del pixel se mantiene dentro de un rango determinable, siempre y cuando el pixel pertenezca solo al fondo de la escena. Por ejemplo en la figura 2, el valor del pixel entre los cuadros 0 y 35, varía entre 0.4 y 0.45, a priori se conoce que en ese rango de cuadro el pixel solamente pertenece al fondo de la escena. Entre



Fig. 2. Comportamiento de un pixel

los cuadros 36 y 47 el pixel pertenece al PP de la escena, lo cual se refleja en un cambio brusco del valor.

La segmentación del primer plano se basa en un MFV generado por MSPP, el cual sirve como referencia en un clasificador simple, el cual determina si el valor de un pixel perteneciente al cuadro nuevo en la secuencia, se encuentra o no, fuera de un rango de valor. El rango de comparación del clasificador está acotado por el valor del mismo pixel en el MFV estimado, más o menos un umbral.

Los pixeles cuyo valor no se encuentre dentro del umbral son considerados como parte de PP y los restantes son clasificados como fondo de la escena analizada. El modelo para segmentación de PP propuesto está dado por la siguiente ecuación:

$$PC_i = \begin{cases} 1, & (FC_i - r) < VC_i < (FC_i + r), \\ 0, & \dots, \end{cases} \quad (2)$$

donde:

- PC_i , cuadro del primer plano que se está construyendo.
- r , Valor del umbral para determinar pertenencia o no al PP.

Como se observa en la ecuación (2), el valor de un pixel en el cuadro analizado debe mantenerse por encima de $(FC_i - r)$ pero no superar el valor de $(FC_i + r)$, para que sea identificado como parte del fondo, que en este caso es clasificado con un valor de 1, el cual en valores de color representa el blanco. Los valores que no cumplan la condición de mantenerse en el

rango promedio de los pixeles de fondo, se clasifican como PP con un valor de 0, que equivale al color negro.

El valor de r se determina de la siguiente manera: Primero se crea una matriz A de tamaño $(i/3) \times (j/3)$, donde i y j son las dimensiones en pixel de los frame que entrega el SVCV, cada elemento de A es un arreglo de tamaño x , que contiene el comportamiento de un pixel en los x primeros cuadros. Luego se calcula la varianza de cada elemento de A y se almacena en la matriz Var . También se calcula la media de cada elemento de A y se almacena en la matriz MA . MA y Var son de tamaño $(i/3) \times (j/3)$. Luego a cada elemento de los arreglos contenidos en A se les resta su respectivo valor medio almacenado en MA , y se almacenan en B , que es del mismo tipo de A . Se calcula la media de los elementos que contiene Var y se almacena en H ; finalmente se obtiene el máximo del valor absoluto de B y se almacena en C . Para obtener r como $r=C+H$.

3.3. Validación de MSPP (pruebas)

El modelo MSPP fue implementado en la herramienta software Matlab y el video de vigilancia utilizado para las pruebas tiene las siguientes características:

- 84 cuadros en escalas de grises,
- Cuadros de 320 x 160 pixeles,
- Condiciones de iluminación variables,
- Primer plano con movimiento continuo en el video.

Para la validación del MSPP se establecieron las siguientes condiciones:

Contrastar la segmentación del PP entregada por el MSPP para un caso específico, contra la misma segmentación elaborada de forma manual.

Contrastar la estimación del MFV entregada por el MSPP para un caso específico, en el cual no hay presencia actual de PP, pero sí lo hubo en cuadros anteriores.

Para segmentar el PP manualmente se toma un cuadro de análisis de la secuencia de video de prueba. En la escena se presenta cambios de iluminación y vibraciones en la cámara, que inducen ruido al SVCV. En la escena de análisis

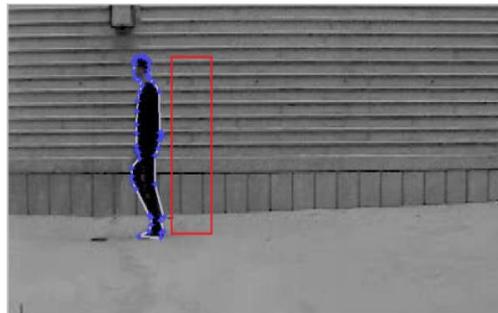


Fig. 3. Regiones de interés

el elemento dinámico avanza de derecha a izquierda. El recuadro rojo de la figura 3 representa una zona que solo debería clasificarse como fondo del video, pero por la cual ya pasó el elemento dinámico, este recuadro tiene una dimensión de 97x46 pixeles.

El contorno azul del elemento dinámico, es segmentado manualmente en un recuadro de 97x46 pixeles. Como resultado se obtiene el PP que se muestra en la gráfica 4.

Ahora es posible realizar la comparación entre la segmentación de PP realizada con la ecuación (2) y la segmentación de PP manual en el cuadro de video escogido para la prueba y determinar el error en el cálculo.

Usando la validación anterior, se procede a realizar pruebas del MSPP en la estimación de MFV y segmentación de PP para cada uno de los cuadros de la secuencia de video. A su vez se realiza una variaron en los siguientes parámetros del modelo:

FA , FM y x de la ecuación (1)

La validación permite evaluar el método con diferentes valores en el parámetro de la ecuación (1) y el error que se genera en la segmentación del PP. El parámetro FA inicia con valor de 0.1, e incrementa en 0.1, hasta alcanzar el valor de 0.9. Dado que se debe cumplir $FA+FM=1$, el valor de FM inicia en 0.9 y termina en 0.1.

El parámetro x que da el tamaño de la ventana temporal móvil sobre la cual se aplica el filtro de la mediana de manera concurrente, inicia con valor de 3, se incrementa en 1, hasta alcanzar un valor de 15 para analizar el comportamiento del modelo en diferentes tamaños de ventana temporal móvil.

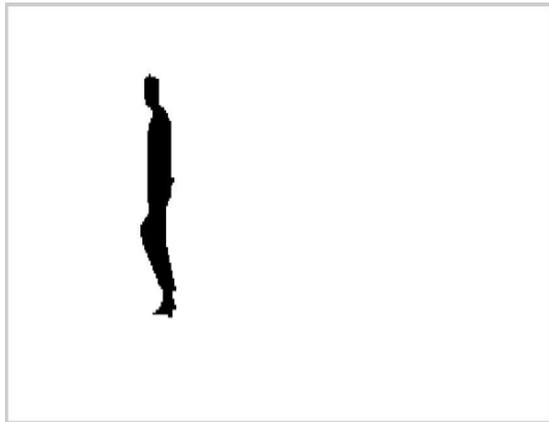


Fig. 4. Segmentación manual

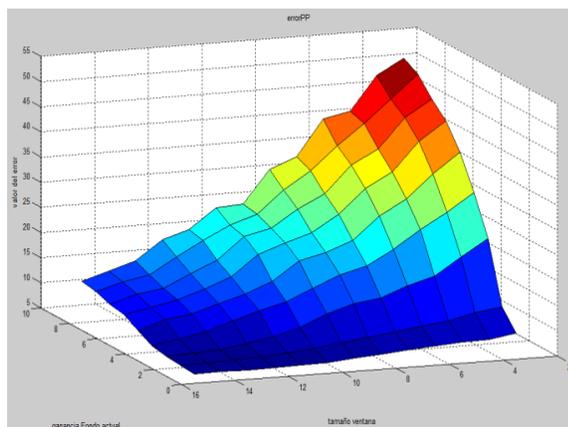


Fig. 5. Error modelo primer plano

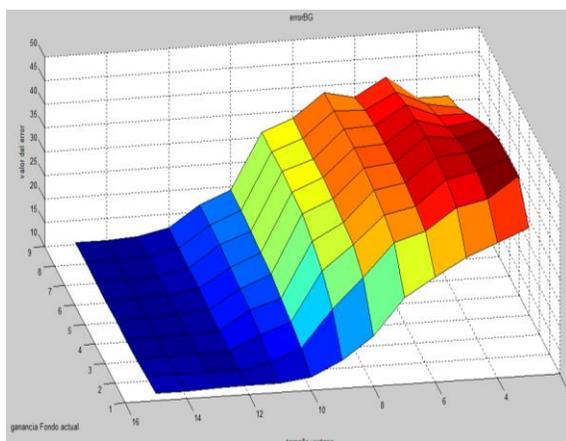


Fig. 6. Error modelo de fondo

4. Resultados

Luego de probar el modelo con distintos valores para los parámetros FA , FM y x , ya mencionados anteriormente, tomar datos de los errores generados en la estimación de MFV y la segmentación de PP, se presenta el comportamiento del error en las figuras 5 y 6.

En la figura 6 se muestra el valor del error en la estimación del fondo generado por MSPP, con respecto al tamaño de la ventana y el factor de ganancia FA .

En la figura 5 se muestra el valor del error en la segmentación del PP generado por MSPP, con respecto al tamaño de la ventana y el factor de ganancia FA .

En la figura 5:

- El error en la segmentación del PP disminuye al aumentar el tamaño del parámetro x .
- El mayor error en la segmentación del PP se presenta con valor de x igual a 3, y valor de FA de 0.9.
- En la segmentación del PP el error aumenta si crece FA .
- Los menores errores en la segmentación del PP se presentan con FA de 0.1 sin importar el tamaño de x .

En la figura 6:

- El error en la estimación del MFV disminuye al aumenta el tamaño del parámetro x .
- En la estimación del MFV los mayores errores se presentan con valores de x entre 3 y 6, con valores de FA entre 0.2 y 0.8.
- El error al estimar el MFV no se ve afectado considerablemente al variar FA .
- En la estimación del MFV los menores errores se presentan con valores de x entre 12 y 15, sin importar el valor de FA .

La complejidad del modelo MSPP depende solamente del parámetro Tamaño ventana temporal (x), que es la cantidad de cuadros en la ventana temporal móvil para cualquier secuencia de video. Esto se debe a que el modelo MSPP se basa en el filtro de la mediana. En la implementación de este filtro se debe organizar

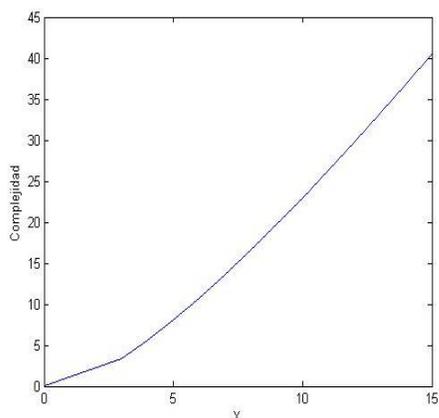


Fig. 7. Complejidad del método

un arreglo y tomar el valor de la posición media, donde el costo computacional se centra en la organización del arreglo y está dado por $f \cdot \log(f)$ [37], donde f es el tamaño del arreglo, para el método MSPP $f=x$.

El comportamiento de la complejidad del modelo MSPP en función de la variación del parámetro x , se muestra en la figura 7, donde la complejidad crece si x aumenta.

El método fue evaluado con distintos tipos de escenas en las secuencias de video. Para una escena de prueba 2, en la que el PP no tiene un buen contraste con el FV se presenta un alto porcentaje de error, como se evidencia en la figura 8 y 9.

Para una tercera escena de prueba en la que hay elementos de fondo en movimiento que generan error en la segmentación del PP. Las figuras 10 y 11 muestran el resultado.

En la escena se encuentra una persona caminando que hace parte del PP. La vegetación hace parte del fondo del video, y está expuesta a una corriente de viento que hace que se mueva y por lo tanto la segmentación del PP detecta ese movimiento y lo clasifique como primer plano.

En una cuarta escena de prueba se presenta un PP con dos personas en movimiento. Las figuras 12 y 13 muestran la escena y la segmentación de PP realizada.

Como se puede ver en la segmentación del PP de la escena de prueba 4, los dos objetos de interés fueron identificados. A su vez las sombras generadas por el PP fueron detectadas y el ruido



Fig. 8. Escena de prueba 2



Fig. 9. Primer Plano escena de prueba 2

generado por los cambios de iluminación y movimientos del FV también fueron detectados.

5. Análisis de resultados

El modelo implementado para estimar el MFV y segmentar el PP, trabaja bien para valores de x mayores a 9, donde se ven menores errores en las pruebas, dado que el filtro de la mediana al tener mayor cantidad de datos a comparar puede eliminar fácilmente valores extremos como lo son los pixeles no pertenecientes al fondo.

Si el valor del parámetro FA aumenta, el error de la segmentación del PP crece, dado que está amplificando, o dándole mayor peso a los valores obtenidos en la mediana; por lo cual, los pixeles al ser la mayoría parte del fondo, algunos



Fig. 10. Escena de prueba 3



Fig. 12. Escena de prueba 4

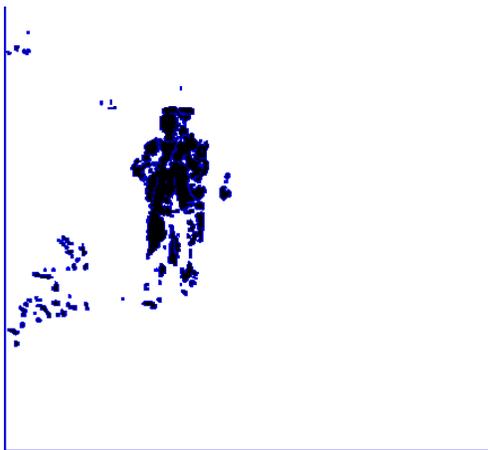


Fig. 11. Primer Plano escena de prueba 3



Fig. 13. Primer Plano escena de prueba 4

alcanzan a sobrepasar el umbral luego de darles ganancia con FA , y son clasificados como parte del PP. Los pixeles que deberían clasificarse como parte del MFV, pero que se incluyen en la segmentación del PP son errores de falsos positivos.

La correcta selección del valor para los parámetros r y FA del modelo MSPP, permite obtener la mejor clasificación de los pixeles en la segmentación del PP evitando los falsos positivos.

Dado que al aumentar el valor de x crece la complejidad del método propuesto "MSPP", se debe establecer un valor de x que disminuya el costo computacional y no genere gran cantidad

de errores en la estimación del MFV y la segmentación del PP. Con valores de x entre 9 y 15 el error del modelo es menor.

En la modificación del parámetro FA , los valores entre 0.1 y 0.3 generaron poco error en la estimación del MFV y la segmentación del PP.

De la combinación de los parámetros x y FA para obtener el menor error en el método MSPP, se consideran valores de x entre 9 y 11, y de FA entre 0.1 y 0.2.

El modelo MSPP obtuvo el menor error al trabajar x con valores entre 9 y 11, y la combinación de FA entre 0.1 y 0.2.

A partir de los resultados obtenidos en la segmentación del primer plano de los distintos

videos de prueba, es importante mencionar el porcentaje de error de ésta disminuye cuando el video cumple con las siguientes características:

- Fondo del video estático.
- Alto contraste del primer plano con el fondo del video.
- Diferenciación entre los colores del objeto del primer plano y el fondo del video.

6. Conclusiones

La implementación del filtro de mediana temporal de manera concurrente sobre una ventana móvil temporal, redujo el error en la estimación del MFV. Adicionalmente, disminuyó el costo computacional y requerimientos de memoria, al limitar el número de frames incorporados en el método para la estimación del MFV.

El modelo propuesto realiza una segmentación del primer plano con muy baja cantidad de falsos positivos, debido al bajo error que presenta la estimación del MFV y el rango utilizado en el clasificador simple propuesto.

La incorporación de un factor de memoria en proceso de estimación del MFV, le brinda una robustez al método MSPP frente a elementos de baja dinámica que aparezcan en la escena, reforzándolos en la segmentación del primer plano.

El método funciona adecuadamente en escenas en las que el fondo es constante y no tiene objetos dinámicos. La segmentación del primer plano mejora cuando el objeto de interés tiene un alto contraste con respecto al fondo del video.

Para trabajos futuros se propone procesar los videos en tiempo real, con la implementación del método MSPP en un lenguaje de bajo nivel, para ser utilizado en sistemas embebidos, con el fin de desarrollar aplicaciones comerciales de sistemas de vigilancia automatizados a bajo costo.

Referencias

1. Guan, Y.P., Du, J.H., & Zhang, C.Q. (2012). *Improved HSV-Based Gaussian Mixture Modeling for Moving Foreground Segmentation*. pp. 52–58.
2. Pava, D.F. (2011). Object Detection and Motion Analysis in a Low Resolution 3-D Model. *Ninth LACCEI Latin American and Caribbean Conference (LACCEI'2011), Engineering for a Smart Planet, Innovation, Information Technology and Computational Tools for Sustainable Development*, Medellín, Colombia.
3. Kim, Z. (2008). Real Time Object Tracking based on Dynamic Feature Grouping. *Computer Vision and Pattern Recognition, IEEE Conference on*, Anchorage, AK, 2008, pp. 1–8.
4. CCTV User Group (2011). Two million cameras in the UK. *CCTV Image*, pp. 10–12.
5. Ferrando, S. (2006). Classification of Unattended and Stolen Objects in Video-Surveillance System. *Video and Signal Based Surveillance*, Sydney, Australia.
6. Plataniotis, K.N. (2005). Visual-centric Surveillance Networks and Services. *IEEE Signal Processing Magazine*, Vol. 22, No. 2, pp. 12–15.
7. Mahadevan, V. (2008). Background Subtraction in Highly Dynamic Scenes. *Conf. on Computer Vision and Pattern Recognition* (p. 1). Anchorage, AK: IEEE.
8. Parks, D.H. (2008). Evaluation of Background Subtraction Algorithms with Post-processing. *Advanced Video and Signal Based Surveillance, (AVSS '08), IEEE Fifth International Conference on*, pp. 192–199, Santa Fe, NM.
9. Zhang, Y.J. (2006). *Advances in Image and Video Segmentation*. IRM Press.
10. González, R.C., Woods, R.E., & Eddins, S.L. (2002). *Digital image processing* (2nd Edition). Englewood Cliffs: Prentice Hall.
11. Kamijo, S. (2000). Occlusion robust tracking utilizing spatio-temporal Markov random field model. *ICPR*, Vol.1, pp. 1140.
12. Zhang, Y.J., & Lu, H.B. (2002). A hierarchical organization scheme for video data. *Pattern Recognition*, Vol. 35, No.11, pp. 2381–2387.
13. McHugh, J.M. (2009). Foreground-Adaptive Background Subtraction. *Signal Processing Letters*, IEEE, pp. 390–393. DOI: 10.1109/LSP.2009.2016447
14. Toyama, K. (1999). Wallflower: principles and practice of background maintenance. *Computer Vision, Proceedings of the Seventh IEEE International Conference on*, Vol. 1, pp. 255–261.
15. Lo, B. (2001). Automatic congestion detection system for underground platforms. *Intelligent Multimedia, Video and Speech Processing, Proceedings of 2001 International Symposium on*, pp. 158–161. DOI: 10.1109/ISIMP.2001.925356

16. **Cucchiara, R. (2003).** Detecting moving objects, ghosts, and shadows in video streams. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 25, No. 10, pp. 1337–1342. DOI: 10.1109/TPAMI.2003.1233909
17. **Caldera, S. (2006).** Reliable background suppression for complex scenes. *Video surveillance and sensor networks*, pp. 211–214. DOI: 10.1109/TPAMI.2003.1233909
18. **Maddalena, L. (2008).** Background Subtraction in Highly Dynamic Scenes. *Conf. on Computer Vision and Pattern Recognition (p. 1)*, Anchorage, AK: IEEE.
19. **Cheung, S.C.S. (2004).** Robust techniques for background subtraction in urban traffic video. *Visual Communications and Image Processing*, 5308, pp. 881–892. DOI: 10.1117/12.526886
20. **Stauffer, C. (2000).** Learning patterns of activity using real-time tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 22, No. 8, pp. 747–757. DOI: 10.1109/34.868677
21. **Tsai, D.M. (2009).** Independent Component Analysis-Based Background. *IEEE Transactions on Image Processing*, Vol. 18, No. 1, pp. 158–167. DOI: 10.1109/TIP.2008.2007558
22. **Power, P. (2002).** Understanding background mixture models for foreground segmentation. *Proc. of the Image and Vision Computing*, pp. 266–271.
23. **Zivkovic, Z. (2006).** Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, Vol. 27, No. 7, pp. 773–780. DOI: 10.1016/j.patrec.2005.11.005
24. **Heikkila, J. (1999).** A real-time system for monitoring of cyclists and pedestrians. *Visual Surveillance, Second IEEE Workshop on*, pp. 74–81. DOI: 10.1016/j.imavis.2003.09.010
25. **Aach, T. (1995).** Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. *Signal Process.* Vol. 7, pp. 147–160. DOI: 10.1016/0923-5965(95)00003-F
26. **Paragios, N. (2001).** A MRF-based approach for real-time subway monitoring. *Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference on, Volume 1*, pp. 1034–1040. DOI: 10.1109/CVPR.2001.990644
27. **Migdal, J. (2005).** Background Subtraction Using Markov Thresholds. *Motion and Video Computing, (WACV/MOTIONS'05), Volume 2, IEEE Workshop on*, pp. 58–65. DOI: 10.1109/ACVMOT.2005.33
28. **Marco Cristani. (2010).** Background Subtraction for Automated Multisensor Surveillance. *EURASIP Journal on Advances in Signal Processing*. DOI: 10.1155/2010/343057
29. **Ahmed, E. (2002).** Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance. *Proc. IEEE*, Vol. 90, pp. 1151–1163. DOI: 10.1109/JPROC.2002.801448
30. **Ahmed, E. (2000).** Non-parametric Model for Background Subtraction. *ECCV '00 Proceedings of the 6th European Conference on Computer Vision*, pp. 751–767. DOI: 10.1007/3-540-45053-X_48
31. **Friedman, N. (1997).** Image segmentation in video sequences: a probabilistic approach. *13th Conference on Uncertainty in Artificial Intelligence*, pp. 175–181.
32. **Stauffer, C. (2000).** Learning patterns of activity using real-time tracking. *Patterns Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 22, No. 8, pp. 747–757. DOI: 10.1109/34.868677
33. **Ahmed, E. (2001).** Efficient non-parametric adaptive color modeling using fast Gauss transform. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 563–570. DOI: 10.1109/CVPR.2001.991012
34. **McFarlane, N.J. (1995).** Segmentation and tracking of piglets in images. *Machine Vision and Applications*, Vol. 8, No. 3, pp. 187–193.
35. **Jung, C.R. (2009).** Efficient Background Subtraction and Shadow Removal. *Multimedia, IEEE Transactions on*, pp. 571–577. DOI: 10.1109/TMM.2009.2012924
36. **Zhu, F. (2009).** A Video-based Traffic Congestion Monitoring System Using Adaptive Background Subtraction. *Electronic Commerce and Security, (ISECS '09), Second International Symposium on*, pp. 73–77. DOI: 10.1109/ISECS.2009.64
37. **Knuth, Donald E. (2005).** *Art of Computer Programming*. Addison-Wesley, United States.

Freddy Alexander Díaz González (MSc) es director del grupo de investigación ESINUSA de la Escuela de Ciencias Exactas e Ingeniería de la Universidad Sergio Arboleda y director del proyecto de desarrollo del Payload óptico para la Misión Satelital Libertad 2.

David Alejandro Arévalo Suárez es ingeniero electrónico de la Universidad Sergio Arboleda, estudiante de maestría en ingeniería electrónica de la Pontificia Universidad Javeriana. Joven Investigador participante del diseño de los

prototipos de los subsistemas OBC, C&DH, y ADCS de la Misión Satelital Libertad 2 de la Universidad Sergio Arboleda.

*Artículo recibido el 13/08/2014; aceptado el 17/04/2015.
Autor de correspondencia es Freddy A. Díaz González.*