

Adquisición de variables de tráfico peatonal utilizando visión por computador

Pedestrian traffic variables acquisition using computer vision

*Julián Quiroga**, Néstor Romero, Carolina García, Carlos Parra

Grupo de Sistemas Inteligentes, Robótica y Percepción – SIRP, Facultad de Ingeniería, Pontificia Universidad Javeriana. Cra. 7 N.º 40-62 Edificio José Gabriel Maldonado. Bogotá, Colombia.

(Recibido el 14 de abril de 2010. Aceptado el 14 abril de 2011)

Resumen

El problema de tráfico tiene diferentes componentes que pueden ser analizados: los vehículos, los peatones y la interacción entre ellos. En este trabajo se propone un método de adquisición de variables de tráfico peatonal, utilizando técnicas de visión por computador. A partir de una secuencia de vídeo son detectados los peatones aislados, los grupos de peatones y los vehículos de la escena, utilizando un modelo del fondo. Los peatones son seguidos en la imagen por medio de características de su contorno y flujo óptico. El conteo de peatones es realizado sobre cualquier zona de la escena para estimar el flujo y la dirección de movimiento. El método propuesto puede ser configurado bajo diferentes perspectivas a partir de un conjunto de ejemplos. Los resultados experimentales en cruces peatonales demuestran que el método permite estimar las variables de interés en escenas complejas.

----- *Palabras clave:* Detección de peatones, seguimiento de peatones, visión por computador

Abstract

The traffic problem has several components that may be discussed: vehicles, pedestrians and the interaction between them. This paper proposes a method for acquisition of pedestrian traffic variables, using computer vision techniques. Isolated pedestrians, groups of pedestrians and vehicles at the scene are detected from a video sequence, using a background model. Pedestrians are tracked on the image using their shape and optical flow. Counting is done on any area of the scene to estimate the flow and direction of movement. The proposed method can be configured under different perspectives from a set

* Autor de correspondencia: teléfono: + 57 + 1 + 320 83 20 ext. 538, correo electrónico: quiroga.j@javeriana.edu.co (J. Quiroga)

of examples. The experimental results on crosswalks show that this method allows estimating the variables of interest in complex scenes.

----- **Keywords:** Pedestrian detection, pedestrian tracking, computer vision

Introducción

Hoy en día se realizan grandes esfuerzos para obtener información confiable sobre el flujo vehicular y peatonal, con el fin de mejorar la movilidad de vehículos y peatones y disminuir el riesgo de accidentalidad de los mismos. En Colombia los accidentes de tráfico producen 6000 muertes [1] y 40000 heridos [2] cada año. En el año 2006, 27,1% de todos los heridos y 31,5% de las víctimas mortales de accidentes con vehículos automotores en el mundo fueron peatones [3].

El desarrollo de soluciones con miras a disminuir el riesgo al que se ven expuestos los peatones por la interacción con vehículos, requiere la adquisición de información del tráfico vehicular y peatonal. En [4] se presenta un algoritmo confiable y eficiente para el conteo de vehículos en diferentes ambientes. Sin embargo, la adquisición de información del tráfico peatonal es una tarea más compleja debido a la cantidad de variables que deben ser analizadas. Por tal motivo, se hace necesario el desarrollo de herramientas que permitan adquirir automáticamente variables de interés como el flujo de personas, la velocidad promedio y las trayectorias, entre otras.

Diferentes métodos basados en técnicas de visión por computador han sido propuestos para la detección y el seguimiento de personas. Estos métodos pueden dividirse principalmente en dos enfoques, monoculares y de múltiples vistas [5]. El enfoque monocular es de mayor interés pues la mayoría de sistema de monitoreo de tráfico se basan en un sola cámara. Dentro de este enfoque los métodos se diferencian de acuerdo a la ubicación de la cámara.

La utilización de vistas superiores permite simplificar el problema pues se minimizan ocultaciones y variaciones de forma entre

personas. En [6] y [7] se proponen métodos para estimar el flujo y realizar el seguimiento de personas, los cuales son fácilmente segmentados en esta perspectiva. En estos se obtiene una alta precisión con un bajo costo computacional, sin embargo, por la ubicación de la cámara sólo se cubren pequeñas regiones en comparación con las cámaras de vista oblicua.

Por otro lado, la utilización de una vista horizontal facilita la detección de personas incluso en escenas muy complejas, pues es posible extraer una gran cantidad de información para la tarea de reconocimiento. En [8] se propone un método para la detección de personas utilizando máquinas de vectores de soporte (SVM) lineales y un conjunto de características extraídas de la distribución de la orientación de los gradientes de intensidad. Sin embargo, la vista horizontal incrementa en número y en porcentaje las ocultaciones, por lo que el entrenamiento de los clasificadores debe incluir un gran número de casos con los diferentes tipos de traslape. En [9] se realiza una segmentación en varios niveles, en primer lugar se detectan los segmentos de la imagen que corresponden a peatones y posteriormente estos son asignados a un peatón específico utilizando información a priori como el área esperada y el traslape estimado. En [10] se utilizan bordes para detectar personas bajo ocultaciones utilizando un conjunto de clasificadores de las diferentes partes del cuerpo. Cuando el número de ocultaciones se incrementa el seguimiento se convierte en una tarea más demandante y es necesario extraer un conjunto más robusto de características. En [11] cada persona es segmentada en clases de color por medio del algoritmo EM (*Expectation Maximization*) y posteriormente se utiliza un enfoque de maximización de la probabilidad a posteriori para seguir las clases en cada cuadro del vídeo. En general, sobre las vistas horizontales es difícil extraer in-

formación de la ubicación y velocidad del peatón en la escena y se dificulta el proceso de conteo debido a los altos niveles de ocultación.

Muchos de los sistemas de monitoreo de tráfico poseen cámaras en vistas oblicuas al igual que los circuitos cerrados de televisión (CCTV), motivo por el cual la mayoría de trabajos para la adquisición de información del tráfico peatonal se realizan desde este tipo de perspectivas. En [12] se estima el flujo de personas a través de una línea de la escena sin segmentar los peatones. Para tal fin, se detectan los píxeles pertenecientes a peatones y se estima el vector de movimiento, información que permite estimar el número de personas que atraviesan la región de interés a partir de una normalización del tamaño de los píxeles.

El análisis del tráfico peatonal requiere además del conteo de peatones, la estimación de trayectorias y velocidades. En [13] se presenta la adaptación de un algoritmo para la detección y seguimiento de vehículos, basado en un modelo estadístico del fondo. Los peatones son modelados como objetos rectangulares caracterizados por su área y el seguimiento es planteado como un problema de optimización de grafos. Sin embargo, este método asume que todos los objetos de interés en la escena son peatones. En [14] la detección de peatones integra información de intensidad de la imagen con información de movimiento. Se implementa un detector basado en *AdaBoost* sobre dos cuadros consecutivos del vídeo, para aprovechar la información de movimiento e intensidad. No obstante el sistema debe ser exhaustivamente entrenado para cada cambio de perspectiva y presenta problemas con las ocultaciones. En [15] las personas son detectadas por el contorno cabeza-hombro usando contornos activos. A diferencia del cuerpo, el contorno cabeza-cuello parece similar ante diferentes puntos de vista y difícilmente se presenta oculto. Sin embargo, la complejidad computacional del algoritmo requiere un segundo de proceso por cada cuadro de vídeo.

En este trabajo se presenta un método para adquirir información del tráfico peatonal a partir

de vídeo, el cual permite detectar, seguir y contar peatones en un cruce peatonal. El método puede ser utilizado bajo diferentes perspectivas, sólo requiere de una corta configuración previa y tiene un bajo costo computacional. Los objetos de interés de la escena son detectados a partir de un modelo estadístico del fondo basado en [16]. Los peatones aislados son clasificados utilizando información de tamaño y orientación de su contorno. Para realizar una detección robusta, un peatón detectado es confirmado sólo si las propiedades de su contorno se mantienen por un número de cuadros de vídeo. El seguimiento es realizado combinando características del contorno y una versión piramidal de la técnica de flujo óptico de Lukas y Kanade, sobre un conjunto de puntos seleccionados de acuerdo a [17]. Los grupos de peatones son diferenciados de los vehículos utilizando la transformada de Hough. Finalmente el conteo es realizado en cualquier recta sobre la imagen utilizando una estimación del traslape entre peatones. En la figura 1 se ilustra el diagrama de flujo del método propuesto. El método desarrollado ha sido utilizado para la adquisición automática de información en el desarrollo del proyecto “*Determinants of risky behavior in a high-frequency vehicle-pedestrian crash site in Bogotá, Colombia*” financiado por NIH a través de la Universidad de Texas–Houston y la Pontificia Universidad Javeriana.

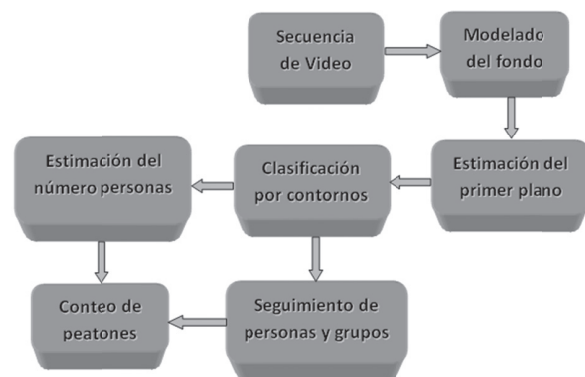


Figura 1 Diagrama general del método

El artículo está organizado como sigue. En la Sección Experimentación se describe la estrategia utilizada para la estimación del primer plano y

los algoritmos para la detección, el seguimiento y el conteo de peatones. Los resultados del método son presentados en la Sección Resultados y Discusión. Finalmente son expuestas las conclusiones sobre el trabajo realizado.

Experimentación

Estimación del primer plano

La estimación del primer plano consiste en la detección de los objetos de interés de la escena y puede realizarse mediante la comparación de un modelo del fondo con cada cuadro del vídeo. El modelo del fondo de la escena es estimado utilizando el algoritmo propuesto por Kim [16], escogido debido a su bajo costo computacional y buen desempeño.

El algoritmo construye para cada píxel un *codebook* conformado por valores (*codewords*) de crominancia y luminancia, cada uno asociado a una frecuencia de aparición. Un píxel es clasificado como parte del fondo si corresponde con alguno de los *codewords*, de lo contrario el píxel es clasificado como perteneciente al primer plano de la escena.

Con el objetivo de mejorar el modelo del fondo fueron implementadas algunas variaciones al algoritmo original. En la etapa de entrenamiento se eliminan todos los *codewords* con baja frecuencia asociada cada cierto tiempo, esto para evitar que ante un eventual tráfico alto de peatones, estos *codewords* se vuelvan “fuertes” y produzcan errores en la estimación del fondo. Adicionalmente, se continúa el entrenamiento del fondo cuando se detectan segmentos del vídeo con bajo flujo peatonal, permitiendo lidiar con posibles cambios de iluminación en la escena.

Un *objeto del primer plano* (OPP) se define como la colección de todos los píxeles que se encuentran conectados, ver figura 2. Para mejorar la calidad de los OPP, se realizan operaciones morfológicas de *closing* y *opening* con el fin que OPP muy cercanos se unan formando un solo objeto y eliminar huecos al interior de los mismos.

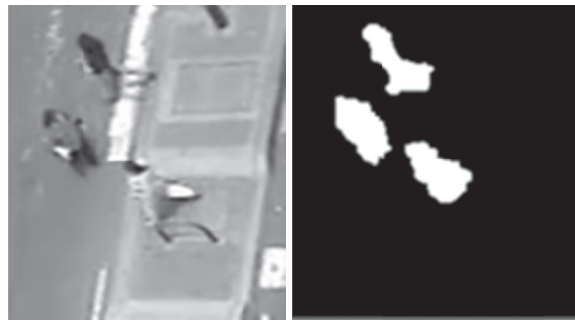


Figura 2 Cuadro de vídeo y objetos del primer plano (OPP)

Detección de peatones

El método propuesto diferencia entre tres clases de objetos de interés: peatón aislado, grupo de peatones y vehículo; los demás objetos detectados son considerados como ruido, ver figura 3. Los OPP muy pequeños, cuyo tamaño se encuentra por debajo de un umbral determinado, son considerados como ruido y no son procesados. Para cada uno de los objetos que superan el umbral de tamaño se obtiene un conjunto de características a partir de su contorno. La clasificación de peatones aislados se realiza utilizando condiciones de tamaño y de orientación.

Con el fin de obtener descriptores que caractericen de forma concisa el contorno se estima su matriz de covarianza, cuyos auto-vectores $\{V_1, V_2\}$ y los auto-valores $\{\lambda_1, \lambda_2\}$ caracterizan el contorno. Los auto-vectores dan información sobre las direcciones principales de dispersión y los auto-valores sobre la dispersión en la dirección de su auto-vector asociado.

Selección de los umbrales de clasificación

Los peatones aislados son modelados asumiendo que su tamaño y orientación siguen una distribución normal. Por tal motivo, se utilizan umbrales de clasificación ajustados en la etapa de configuración sobre un conjunto de cuadros de vídeo. Al seleccionar manualmente los peatones aislados se almacenan los auto-valores y auto-vectores observados de sus contornos. Los

umbrales de clasificación son seleccionados en función de la media estadística y la desviación estándar de las observaciones. Sean M_V y M_λ los promedios de los valores observados de auto-valores y auto-vectores, respectivamente, y sean D_V y D_λ las desviaciones estándar de los mismos. El umbral de tamaño, $U_T = M_\lambda + D_\lambda$, define la cota superior del máximo auto-valor de un contorno

clasificado como peatón aislado. Los umbrales de ángulo, $\{M_V \pm D_V\}$ definen el intervalo en el que debe encontrarse la orientación de una máscara para ser clasificada como peatón aislado. El umbral de longitud, $U_L = 2M_\lambda$ es utilizado para definir la longitud mínima de la recta que debe ser encontrada en una máscara para ser clasificada como vehículo.

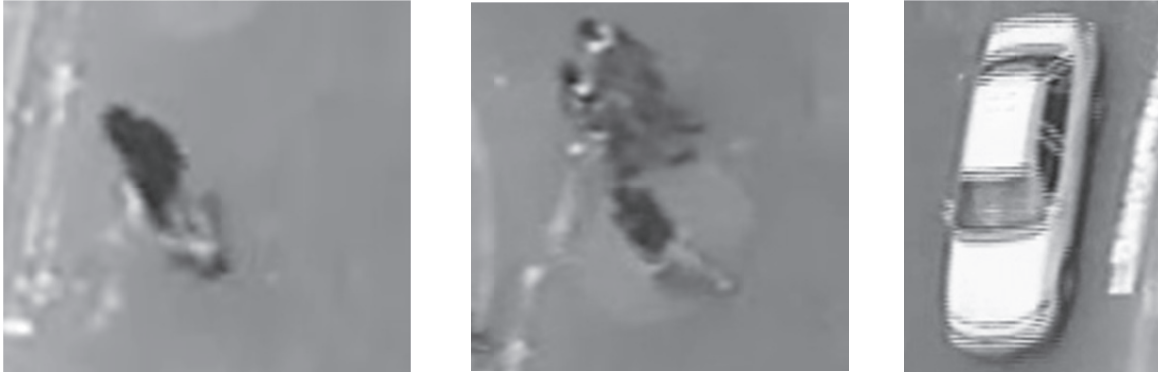


Figura 3 Objetos de interés: peatón aislado, grupo de peatones y vehículo

Detección de peatones aislados

Para la detección se utiliza el auto-vector, V_{\max} , asociado al mayor auto-valor, λ_{\max} , de cada contorno. Se asume que el tamaño del peatón no varía considerablemente sobre la imagen a medida que él se desplaza en la escena. Un objeto es clasificado como peatón aislado si cumple:

- Condición de tamaño: $\lambda_{\max} < U_T$ (se descartan objetos demasiado grandes).
- Condición de orientación: $M_V - D_V < V_{\max} < M_V + D_V$

Por medio de estas dos condiciones se logra clasificar a los peatones aislados, ver figura 4. Los vehículos y grupos de peatones son fácilmente descartados debido a su tamaño.

Detección de grupos de peatones

Para diferenciar entre grupos de peatones y vehículos se utiliza la transformada de Hough [18], con la cual se pueden detectar líneas rectas en una imagen. Conociendo las características

de los vehículos se espera que por su estructura rectangular contengan bordes rectos, mientras que un grupo de peatones debido a las formas arbitrarias que puede tomar, difícilmente presente bordes rectos de longitud considerable. Cada punto del contorno es transformado a una función en el plano paramétrico utilizando la transformada de Hough. Si la intersección de funciones en un punto del plano supera el umbral U_L , una recta es asociada a los píxeles correspondientes y el OPP asociado es clasificado como carro, de lo contrario es clasificado como grupo de peatones.



Figura 4 Contornos de los OPP y peatones detectados delimitados por un rectángulo

Seguimiento de peatones

El método de seguimiento propuesto combina dos estrategias, características del contorno y flujo óptico, con el fin de estimar la posición de cada OPP en el plano imagen, en función del tiempo.

Seguimiento utilizando características del contorno

Se utilizan como características por cada peatón detectado, su centro de masa y su auto-vector λ_{\max} , pues estas dos presentan una baja variación entre 2 cuadros consecutivos de vídeo.

Sobre un cuadro de vídeo n se consideran todos los peatones aislados, los cuales son clasificados como *Posibles Peatones (PP)* y cada uno de ellos es etiquetado con un identificador (ID) único. Un *PP* es seguido y posteriormente confirmado como *Peatón* si su seguimiento es realizado durante N cuadros consecutivos, con el fin de no considerar como peatones a aquellos falsos positivos que existan durante un corto tiempo. Cada *PP* del cuadro n es asociado a un vector de peatón $V_p[n]$ que almacena el ID de peatón, su centro de masa y su ángulo. Posteriormente, sobre el cuadro de vídeo $n+1$ son detectados todos los *PP* y sus respectivas características son almacenadas en $V_p[n+1]$. Sean $P_{i,n}$ y $P_{j,n+1}$ dos *PP* de los cuadros de vídeo n y $n+1$, respectivamente, entonces los dos peatones son asociados (considerados como el mismo peatón) si: la distancia entre sus centros de masa es menor al *umbral de distancia de seguimiento* y la diferencia de orientación en valor absoluto es menor al *umbral de ángulo de seguimiento*. Bajo ninguna circunstancia dos peatones del cuadro n son asociados al mismo peatón del cuadro $n+1$. Si a un peatón $P_{j,n+1}$ no le es asociado ningún peatón del cuadro anterior entonces $P_{j,n+1}$ se considera como un nuevo peatón, al cual se le asigna un ID y sus características son almacenadas en $V_p[n+1]$ para iniciar su seguimiento. Por otro lado, si a un peatón $P_{i,n}$ no se le asocia ningún peatón del cuadro siguiente entonces $P_{i,n}$ se considera como *Peatón Perdido* y se etiqueta de acuerdo a los siguientes casos. Si entró a un grupo, es

etiquetado como *Peatón en Grupo*. Si se perdió en medio de la escena, es etiquetado como *Peatón en Predicción*. Si se perdió en los límites de la escena significa que este salió de escena y es eliminado del vector. El procesamiento de los *Peatones en Grupo* y *Peatones en Predicción* se describe a continuación.

Seguimiento utilizando flujo óptico

El algoritmo de flujo óptico entra en funcionamiento cuando no es posible realizar el seguimiento por características de contorno. El flujo óptico se define como el movimiento aparente de los patrones de intensidad en una imagen y puede relacionarse directamente al movimiento de los peatones en el plano imagen. Para disminuir el tiempo de cómputo, el flujo óptico es estimado sólo sobre un conjunto de puntos de interés, los cuales son seleccionados utilizando el criterio de Shi y Tomasi [17]. El flujo óptico es estimado sobre los puntos de interés utilizando una versión piramidal del algoritmo de Lucas y Kanade [19]. Los puntos son seleccionados en el interior del OPP, como se visualiza en la figura 5 y son actualizados en cada cuadro de vídeo.

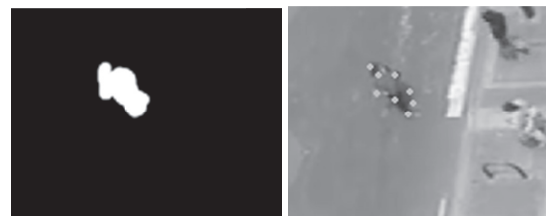


Figura 5 Puntos de interés seleccionados en interior del peatón

Un peatón perdido en el seguimiento es etiquetado como *Peatón en Predicción* y el algoritmo de seguimiento por flujo óptico entra en funcionamiento. A medida que el algoritmo de Lucas y Kanade realiza el seguimiento de los puntos de interés, se calcula el promedio de la posición de éstos para estimar el centro de masa del OPP. Cuando en el cuadro n se encuentra un *Posible Peatón*, $P_{i,n}$, tal que la distancia de su centro de masa y la del centro estimado del *Peatón en*

Predicción es menor al *umbral de recuperación*, el *Peatón en Predicción* es asociado al peatón $P_{i,n}$, tomando sus características de seguimiento pero conservando su ID.

Unión de peatones

Cuando un peatón aislado se une a un grupo, no es posible realizar su seguimiento utilizando características de contorno o flujo óptico. Sin embargo, se plantea una solución basada en el seguimiento de grupos de peatones y en la detección de la separación y unión de OPP.

El seguimiento de un grupo de peatones se realiza utilizando como características el centro de masa y el área de su OPP asociado. Cada grupo es identificado con un ID único. Cuando un peatón se une a un grupo ó cuando peatones aislados se unen entre sí, el OPP de los peatones aislados se pierde y se crea un grupo nuevo. Este grupo es marcado como *Grupo en Predicción* y le son asociados los ID de los peatones que ingresaron, con el fin de poder seguirlos cuando salgan del grupo. Los peatones que ingresan al (o crean el) grupo son marcados como *Peatones en Grupo*, entrando en un estado de espera en el cual no son eliminados del vector de peatones y no se realiza ningún procesamiento sobre ellos.

Para determinar cuáles fueron los peatones que ingresaron al grupo, el OPP asociado al nuevo grupo es encerrado por un rectángulo delimitador (RD) de ancho y alto conocido. Los peatones cuyo último centro de masa se encuentra dentro del RD son asociados a este grupo y marcados como *Peatón en Grupo*. Si no es posible asociar algún peatón al grupo éste se marca como *Grupo Simple*, el cual sólo es seguido y tenido en cuenta para realizar el conteo de peatones.

Cuando se divide un *Grupo en Predicción* los peatones resultantes deben ser asociados a los peatones que crearon o se unieron al grupo. La asociación se realiza recuperando del vector de peatón el centro de masa correspondiente a cada peatón del grupo, luego estos son comparados con los centros de masa de los peatones resultantes en el momento de la separación.

Conteo de peatones

El conteo de peatones se realiza estimando el número de personas que atraviesan una línea virtual (LV) creada sobre la imagen. El centro de masa de los OPP de los peatones y grupos de peatones es utilizado para determinar el paso a través de una LV. Sin embargo, para realizar el conteo de un grupo es necesario estimar el número de personas que conforman el mismo. Por tal motivo, el algoritmo se divide en dos partes, primero la estimación del número de peatones en un grupo y segundo la detección del paso de un OPP a través de una LV.

Estimación del número de peatones en un grupo

Utilizando la información recopilada de los peatones aislados en la etapa de configuración, se calcula el área promedio de un peatón (A_{peaton}) y el área del rectángulo delimitador promedio de un peatón (ARD_{peaton}). Estos valores definen el modelo del peatón aislado.

En un grupo de peatones tanto el número de personas que conforman el grupo como el nivel de traslape entre ellas controlan el área del grupo y el RD del OPP. Con el fin de estimar el nivel de traslape se definen 2 relaciones que utilizan el modelo del peatón aislado para estimar el número de personas que conforman el grupo. La relación de área R_A se define como la relación entre el área del grupo de peatones (A_{grupo}) y A_{peaton} . La relación de rectángulo delimitador R_{RD} se define como la relación entre el área del RD del grupo de peatones (ARD_{grupo}) y ARD_{peaton} . El valor de R_A indica el número esperado de peatones en el grupo utilizando la información del área del OPP, sin embargo, debido al traslape que puede presentarse entre personas, R_A es sólo una cota inferior para la estimación. El valor de R_{RD} indica el número esperado de peatones en el grupo utilizando la información del RD del OPP, no obstante, debido a las diferentes ubicaciones relativas que pueden presentar las personas, el número estimado es sólo una cota superior.

Utilizando estas dos relaciones se define el nivel de ocupación del grupo de acuerdo a:

$$\tau = \frac{R_A}{R_{RD}} = \frac{A_{grupo}/A_{peaton}}{ARD_{grupo}/ARD_{peaton}} \quad (1)$$

Cuando $\tau \approx 1$, los peatones se encuentran poco traslapados y distribuidos uniformemente de forma vertical y/u horizontal. De forma similar, si $\tau < 1$ los peatones se encuentran poco traslapados pero distribuidos arbitrariamente en el RD. En estos dos casos, de bajo traslape, la estimación de número de personas es realizada utilizando R_A . Por otro lado, si $\tau > 1$ los peatones se encuentran traslapados y el valor de τ es un buen indicador de la cantidad de traslape. En este caso R_A es inferior al número real de peatones, debido al traslape. De acuerdo a estas observaciones la estimación de peatones en un grupo se realiza como sigue:

$$N_p = \begin{cases} R_A & \text{si } \tau < 1,25 \\ \tau R_A & \text{si } \tau \geq 1,25 \end{cases} \quad (2)$$

Conteo de peatones

Una *Recta de Conteo* (RC) puede ser ubicada en cualquier región de la escena, como se ilustra en la figura 6. La ubicación de un peatón o de un grupo de peatones es determinada por el algoritmo de seguimiento, definida como el centro de masa del OPP asociado. Para cada OPP se define un punto de inicio (x_0, y_0) , como la ubicación en la imagen en donde éste fue detectado por primera vez. Sean (x_n, y_n) el centro de masa de un OPP en el cuadro n de vídeo y RD la *Recta de Desplazamiento* que une los puntos (x_0, y_0) y (x_n, y_n) . Si (x_i, y_i) es el punto de intersección entre la RC y la RD, entonces el cruce del OPP a través de la RC es detectado cuando (x_i, y_i) se encuentra ubicado entre los puntos (x_0, y_0) y (x_n, y_n) . Posteriormente, se incrementa el contador teniendo en cuenta dos aspectos, la ubicación inicial y final, regiones A y B de la figura 6, y el número de peatones que han atravesado la RC, uno para un peatón aislado y el número de peatones estimados para un grupo de peatones. El conteo es realizado en cada sentido,

con el fin de estimar el flujo en cada dirección. Con el objetivo de evitar múltiples conteos sobre una RC, un OPP es contado sólo si su último conteo se realizó por lo menos 10 cuadros antes.

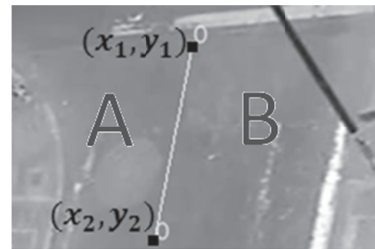


Figura 6 Línea de conteo definida por los puntos (x_1, y_1) y (x_2, y_2)

Resultados y discusión

La evaluación del método se realizó con vídeos de cruces peatonales de la ciudad de Bogotá. Estos fueron capturados entre las 9 a.m. y las 2 p.m., y presentan variaciones de iluminación por nubes. Se utilizaron vistas oblicuas como las presentadas en la figura 7.



Figura 7 Ejemplos de las perspectivas utilizadas para la evaluación del método

Un primer conjunto de vídeos, con bajo flujo peatonal, se utilizó para analizar las etapas de detección y de seguimiento del sistema. Estos tienen una duración total de 29,1 minutos y un flujo promedio 10,46 peatones por minuto. En este experimento se evaluó la detección de los objetos: peatón aislado y grupo de peatones. Se establece que un objeto es detectado por el método, si éste, es localizado y seguido por al menos 10 cuadros de vídeo. En la tabla 1 se presentan los resultados de la etapa de detección, discriminando la tasa de falso rechazo (FR) y la tasa de falsa aceptación del método. La tasa de FR

corresponde al porcentaje de objetos que no son detectados, mientras la tasa de FA corresponde a las falsas detecciones realizadas, ambos con respecto a la cantidad de objetos.

La detección de peatones aislados presenta fallas cuando las condiciones de ruido de la escena

hacen que las propiedades de los contornos cambien drásticamente. Adicionalmente, una mala estimación del primer plano puede causar la división del contorno en varios conjuntos provocando errores en la detección debido a las variaciones en tamaño y forma de los contornos.

Tabla 1 Evaluación de la etapa de detección

<i>Tipo de Objeto</i>	<i>Cantidad</i>	<i>Detecciones</i>	<i>Tasa de FR</i>	<i>Tasa de FA</i>
Peatón aislado	72	70	4,16%	1,38%
Grupo de peatones	93	95	5,37%	7,52%

Los errores de FR en la detección de grupos se presentan debido a problemas en el seguimiento. Las propiedades de los grupos utilizadas en el seguimiento pueden cambiar rápidamente debido al desplazamiento relativo de los peatones en su interior. Por tal motivo, es necesario utilizar características robustas a las deformaciones naturales del grupo. Por otro lado, los errores de FA se presentan cuando no es posible ajustar rectas sobre los vehículos, debido principalmente a la calidad del vídeo, y estos son clasificados como grupos de peatones.

En el segundo experimento se evaluó el conteo de peatones realizado por el método. Para ello se utilizó un segundo conjunto de 9 fragmentos de vídeo de 1 minuto de duración, los cuales fueron obtenidos aleatoriamente sobre 9 vídeos de 10 minutos de duración. Estos fueron capturados en un punto de la ciudad y tienen un flujo promedio 25,44 peatones por minuto. Sobre un total de 229 peatones el método dio como resultado un conteo de 208, presentando un error del 9,17%. El mayor porcentaje de error es aportado por la etapa de estimación de personas en un grupo, en donde la cantidad de traslape entre los integrantes del grupo es difícilmente estimada.

El tiempo de ejecución del sistema depende de la cantidad de grupos y carros que se presentan en la escena, así como del área de la región de interés seleccionada. El método se implementó en C++

en un computador con procesador AMD Turion 64X2 a 1,9 GHz y 2 Gb de memoria RAM, sobre el cual tuvo una tasa promedio de procesamiento de 10,1 cuadros de vídeos por segundo.

El método propuesto es comparable con trabajos previos realizados para el conteo de peatones en vídeo. En [20] se evalúa un sistema comercial para el conteo automático de peatones en exteriores, utilizando 12 horas de vídeo con un flujo muy variado, con un promedio de 6 peatones por minuto, obteniendo una tasa promedio de error del 13,2%. En [15] se presenta un algoritmo para el conteo de peatones, el cual es evaluado utilizando un vídeo en una zona interior de 1 hora de duración y una tasa de 33,3 peatones por minuto. En promedio se obtiene un error del 10% en el conteo y se necesita de 1 segundo para procesar cada cuadro de vídeo. Para vídeos en interiores y después de un ajuste por regresión lineal, los autores obtienen un error cercano al 2% para vídeos con una tasa de 6,6 peatones por minuto.

Conclusiones

En este trabajo se propone un método para la detección, seguimiento y conteo de peatones en diferentes ambientes. A diferencia de la mayoría de métodos que deber ser exhaustivamente entrenados, el método presentado permite una rápida configuración en función de la perspectiva

de la escena, a partir de un pequeño conjunto de ejemplos de peatones aislados. El desempeño del método tiene como etapa fundamental la estimación del primer plano, para lo cual se utiliza un segmento inicial del vídeo para modelar el fondo, disminuyendo así el costo computacional en comparación con algoritmos de modelado adaptativo. Con el fin de considerar los cambios temporales en el modelo del fondo, se implementan intervalos de actualización basados en un conjunto de reglas que disminuyen la cantidad de información ruidosa que es agregada al modelo. La clasificación de los objetos se realiza utilizando información a priori de los objetos de interés, recolectada en la etapa de configuración. Debido a la poca información utilizada para la clasificación de los objetos interés, el proceso de clasificación es más eficiente que otros enfoques. Adicionalmente, a diferencia de otros trabajos previos se permite la presencia de objetos diferentes a peatones en la escena y es posible la detección de grupos de peatones. Para el seguimiento de peatones se combinan dos técnicas ampliamente usadas con el fin de dotar al sistema de la robustez necesaria ante condiciones de ruido y variaciones en la iluminación. A partir de esta información, puede determinarse la trayectoria seguida por el peatón en la escena y otras variables como velocidad promedio e instantánea. Para el conteo de personas se propone una regla para la estimación del número de integrantes de un grupo en función de la información a priori de los peatones aislados. Los resultados experimentales muestran el desempeño del sistema en ambientes con flujo peatonal de bajo a medio y ante la presencia de vehículos y oclusiones parciales. Como trabajo futuro se planea explorar la clasificación estadística de los objetos de primer plano como en [21], utilizando clasificadores de las partes del cuerpo, nuevas estrategias para la estimación del número de personas en un grupo y la utilización de múltiples vistas para superar los problemas debidos a los altos niveles de traslape.

Referencias

1. DANE. "Estadísticas vitales". Disponible en: http://www.dane.gov.co/files/investigaciones/poblacion/defunciones/Defunciones_causa_externa_2008.xls. Recuperado: marzo de 2009.
2. Fondo de Prevención Vial. *Accidentalidad vial en Colombia 2007*. 2007. pp. 13.
3. World Health Organization. *World report on road traffic injury prevention: summary*. Geneva. 2004. pp. 1-9.
4. G. Urrego, F. Calderón, A. Forero, J. Quiroga. "Adquisición de variables de tráfico vehicular usando visión por computador". *Revista de Ingeniería, Universidad de los Andes*. Vol. 30. 2009. pp. 7-15.
5. J. Berclaz, F. Fleuret, P. Fua. "Robust people tracking with global trajectory optimization." *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2006. pp. 744-750.
6. K. Terada, D. Yoshida, S. Oe, J. Yamaguchi. "A method of counting the passing people by using the stereo images." *Proc. of the IEEE International Conference on Image Processing*. Vol. 2. 1999. pp. 338-342.
7. S. Velipasalar, Y. L. Tian, A. Hampapur. "Automatic counting of interacting people by using a single uncalibrated camera." *Proc. of the IEEE International Conference on Multimedia and Expo*. Vol. 1. 2006. pp. 1265-1268.
8. N. Dalal, B. Triggs. "Histograms of oriented gradients for human detection." *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2005. pp. 886-893.
9. B. Leibe, E. Seemann, B. Schiele. "Pedestrian detection in crowded scenes." *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2005. pp. 878-885.
10. B. Wu, R. Nevatia. "Detection of multiple, partially occluded humans in a single image by Bayesian combination of edge part detectors." *Proc. Of the IEEE International Conference on Computer Vision*. Vol. 1. 2005. pp. 90-97.
11. S. Khan, M. Shah. "Tracking people in presence of obstacles." *Proc. of the First Asian Conference on Computer Vision*. Vol. 1. 2000. pp. 1132-1137.
12. G. G. Lee, B. S. Kim, W. Y. Kim. "Automatic estimation of pedestrian flow." *Proc. of the First ACM/IEEE Int. Conference on Distributed Smart Cameras*. Vol. 1. 2007. pp. 291-296.

13. O. Masoud, N. P. Papanikolopoulos. "A novel method for tracking and counting pedestrians in real-time using a single camera." *IEEE Transactions on Vehicular Technology*. Vol. 50. 2001. pp. 1267-1278.
14. P. Viola, M. J. Jones, D. Snow. "Detecting pedestrians using patterns of motion and appearance". *Proc. Of the IEEE Int. Conf. on Computer Vision*. Vol. 2. 2003. pp. 734-741.
15. O. Sidla, Y. Lypetsky, N. Brandle, S. Seer. "Pedestrian detection and tracking for counting applications in crowded situations." *Proc. of the IEEE International Conference on Video and Signal Based Surveillance*. Vol. 1. 2006. pp. 70-75.
16. K. Kim, T. Chalidabhongse, D. Harwood, L. Davis. "Real-time foreground-background segmentation using codebook model." *Real-Time Imaging*. Vol. 11. 2005. pp. 167-256.
17. J. Shi, C. Tomasi. "Good features to track." *Proc. Of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 1. 1994. pp. 593-600.
18. R. O. Duda, P. E. Hart. "Use of the Hough transformation to detect lines and curves in pictures". *Communications of the ACM*. Vol. 15. 1972. pp. 11-15.
19. J. Y. Bouguet. "Pyramidal implementation of the Lucas Kanade feature tracker – Description of the algorithm." *Intel Corporation – Microprocessor Research Labs*. 2000.
20. R. Greene-Roesel, M. C. Diogenes, D. Ragland, L. Lindau. "Effectiveness of a Commercially Available Automated Pedestrian Counting Device in Urban Environments: Comparison with Manual Counts." *Transportation Research Board 87th Annual Meeting*. 2008. pp. 16.
21. P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan. "Object Detection with Discriminatively Trained Part-Based Models." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol 32. 2010. pp. 1627-1645.