

Sistemas de monitoreo de tránsito vehicular basados en técnicas de segmentación de imágenes

Traffic surveillance systems base on image segmentation techniques

Nallig Leal. M. Sc.¹, Esmeide Leal. M. Sc², John Willian Branch Ph. D³

1. Facultad de Ingeniería de Sistemas, Universidad Simón Bolívar, Barranquilla, Colombia

2. Escuela de Sistemas, Universidad Nacional de Colombia - sede Medellín

nleal@unisimonbolivar.edu.co; eleall@unisimonbolivar.edu.co; jwbranch@unal.edu.co

Recibido para revisión 08 de julio de 2010, aceptado 4 de octubre de 2010, versión final 26 de noviembre de 2010

Abstract—Traffic surveillance systems promise to improve traffic safety and alleviate traffic jams, among other things. The continuous increase in processing speed of computers in the last decade has led to an increasing emphasis on research and development of automatic traffic surveillance systems in different countries. However, traffic surveillance systems are still an open research area because they are not in a state of maturity that allows them to operate with complete autonomy in the task of regulating traffic. This article presents a review of automatic traffic surveillance systems based on vehicular image segmentation techniques. We present a classification of these systems according to their purpose and nature, and a classification of the most commonly used techniques in such systems.

Keywords— Traffic Surveillance Systems, Image Segmentation, Computer Vision.

Resumen— Los sistemas de supervisión automática del tránsito vehicular prometen, entre otras cosas, mejorar las condiciones de seguridad del tráfico y aliviar las congestiones vehiculares. El incremento continuo de la velocidad de procesamiento de los computadores en la última década ha propiciado un creciente esfuerzo en investigación y desarrollo de sistemas de supervisión automática del tránsito vehicular en diferentes países; sin embargo, los sistemas de supervisión automática del tránsito vehicular aún son un tema abierto de investigación dado que no se encuentran en un estado de madurez que les permita operar con total autonomía en la tarea de regulación del tránsito vehicular. Este artículo presenta una revisión de sistemas de monitoreo de tránsito vehicular basados en técnicas de segmentación de imágenes. Se realiza una clasificación de dichos sistemas de acuerdo a su propósito y naturaleza; de igual manera, se realiza una clasificación de las técnicas más comúnmente utilizadas en la implementación de dichos sistemas.

Palabras Clave—Sistemas de Monitoreo de tránsito, Segmentación de imágenes, Visión por Computador.

I. INTRODUCCIÓN

La supervisión automática del tránsito ha sido de interés por muchos años; desde los 80s, los ingenieros de tránsito han venido investigando métodos para la implementación de sistemas de supervisión automática del tránsito. Uno de los propósitos de tales sistemas es obtener información del uso de las vías y determinar áreas con necesidad de expansión o con alteración de los requerimientos de los patrones de tránsito existentes [12].

El control automático del tráfico vehicular es un campo que ofrece retos interesantes en la aplicación de técnicas de visión artificial. Es así como recientes investigaciones se han dirigido a la solución del problema del tráfico vehicular [38, 32, 27, 28], aplicando una variedad de técnicas que van desde el procesamiento digital de imágenes clásico y de imágenes en movimiento [13], hasta métodos heurísticos como las redes neuronales artificiales y los algoritmos genéticos.

En la actualidad, las soluciones informáticas a los problemas de tránsito vehicular, no se encuentran en un estado de madurez que les permita operar con total autonomía en la tarea de regulación del tránsito vehicular. Lo anterior se debe principalmente, a problemas de sobresegmentación, subsegmentación y pérdida de información presentados por las técnicas y métodos de segmentación, base primordial de los sistemas de supervisión del tránsito vehicular.

Los problemas en la segmentación de imágenes de tránsito vehicular se presentan, principalmente, por las características de bajo contraste, alto nivel de ruido y distribución no uniforme de la intensidad de los niveles de grises que presentan estas imágenes; dichas características, se deben a factores ambientales como: contaminación, baja iluminación, nubosidad y sobreexposición, entre otros, los cuales afectan la captura de las imágenes.

Los problemas antes mencionados, sugieren el empleo o desarrollo de nuevas técnicas de segmentación que puedan ser aplicadas en la construcción de sistemas de monitoreo del tránsito vehicular.

En este artículo se presenta una inspección de los diferentes trabajos que se han realizado y se vienen realizando en el área de los sistemas de monitoreo de tránsito vehicular (SMTV) basados en técnicas de segmentación de imágenes. Lo que resta de este artículo está organizado de la siguiente manera: La sección 2 presenta una revisión histórica de los SMTV y una taxonomía de clasificación de dichos sistemas. En la sección 3 se abordan las técnicas de segmentación más comunes empleadas por los SMTV y se presenta una taxonomía de clasificación de dichas técnicas. Las conclusiones y tendencias futuras son abordadas en la sección 4.

II. SISTEMAS DE MONITOREO DE TRANSITO VEHICULAR

Aunque los primeros SMTV desarrollados, de los que se tiene registro, tenían como objetivo la supervisión del tránsito en pequeñas zonas urbanas y tramos cortos de carreteras, en años recientes, se han desarrollado sistemas capaces de realizar supervisión en áreas más amplias y entregar información de carácter global del comportamiento del tránsito vehicular a través de fotografía aérea o satelital. A continuación se presentarán algunos aspectos de los sistemas de supervisión del tránsito basado en video cámaras de corto alcance; luego, se presentarán algunos aspectos de los sistemas de supervisión del tránsito mediante fotografía aérea y satelital.

2.1 Sistemas de monitoreo de cobertura local

Uno de los propósitos de los sistemas de monitoreo de tránsito vehicular es obtener información del uso de las vías y determinar áreas con necesidad de expansión o con alteración de los requerimientos de los patrones de tránsito existentes [12].

Un SMTV típico (Figura 1) emplea una cámara CCD que provee video en vivo, el cual se fracciona en marcos (frames), se digitaliza y se almacena en un computador, para luego aplicarle técnicas de visión artificial que permitan realizar la detección de los objetos de interés y poder realizar un análisis de la escena.

Los SMTV basados en visión, además de ser más económicos, son potencialmente más poderosos que aquellos basados en sensores de captura diferentes a las cámaras de video, dado

que tienen la capacidad de extraer características como: la forma, la textura, el tamaño, la ubicación y el color de los vehículos, entre otras. Dichas características le permiten a los SMTV entregar información tanto de carácter general como particular, de una escena de tránsito; como la cantidad de vehículos en la vía, la dirección de mayor flujo vehicular en un instante dado, el número de matrícula y modelo del vehículo, entre otras. La información anterior, junto con una adecuada ubicación de las cámaras, permite a los SMTV basados en visión hacer el seguimiento de vehículos en escenas de tránsito simples o complejas, lo cual difícilmente se puede lograr con otros tipos de sensores. Teóricamente, un SMTV podría tener el mismo poder de observación que un observador humano, pero sin los efectos perjudiciales del cansancio o el aburrimiento.



Figura 1. Sistema típico de visión para monitoreo del tránsito¹ vehicular

Generalmente, los SMTV basados en visión (en adelante referidos simplemente como SMTV) trabajan sobre imágenes de niveles de gris, puesto que esencialmente estos buscan los cambios repentinos de la posición de los píxeles de un marco a otro, para lo cual es suficiente el nivel de intensidad del píxel, siendo irrelevante su color. Otra razón para la utilización de dichas imágenes es que los SMTV, al ser sistemas de procesamiento en tiempo real, deben responder rápidamente a los cambios presentados en la escena, lo cual es complicado cuando se procesan varios canales de color.

El proyecto “Traffic Research using Image Processing (TRIP)”, desarrollado en conjunto por la Universidad de Manchester y el instituto de ciencia y tecnología de la Universidad de Sheffield en los 80s, fue uno de los primeros intentos de usar visión por computador para el monitoreo del tráfico de carreteras, aunque sus resultados iniciales no fueron particularmente impresionantes [31].

1. Adaptado de [Setchell97]

Takaba y su grupo [34], desarrollaron en Tokio en los años 80's un sistema que empleaba un método denominado "puntos muestreados"; en el cual, mediante cada punto muestreado se podía detectar la presencia de un vehículo, y por el posicionamiento de los puntos muestreados en cada carril de tráfico, se podía realizar el conteo de vehículos. La velocidad del vehículo podía ser medida por la ubicación de dos puntos muestreados en cada carril de tráfico, y la medición del tiempo de retardo entre la captura del primer y el segundo punto muestreado. El sistema era dependiente de las líneas de marcación de los carriles de tránsito, por lo que podía presentar fallas en el conteo cuando éstas eran ocluidas por los vehículos, o cuando cambiaba su nivel de gris debido al desgaste natural en el tiempo.

En 1984 el Instituto Real de Tecnología de Suecia desarrolló un SMTV en el cual una línea de detección de 3 píxeles de ancho era ubicada a lo largo de cada uno de los carriles de la carretera. Para cada línea de detección, las intensidades de referencia eran almacenadas. Los valores actuales en las líneas de detección eran restados de los valores de referencia y su resultado era umbralizado. Este proceso segmentaba vehículos, mientras se movían a lo largo de las líneas de detección, permitiendo su seguimiento [1]. La detección de vehículos lograda por este sistema era más rápida que la del sistema de Takaba, aunque adolecía de los mismos inconvenientes.

En 1986, Dick-Manns probó el sistema VITA II de conducción automática en carretera a una velocidad cercana a los 100 Km/h en el proyecto europeo PROMETHEUS, el sistema mostró completa autonomía en conducción y un buen desempeño realizando planificación y maniobras. Los buenos resultados alentaron a los creadores a continuar adelante y entrar en ambientes urbanos mucho más complejos para las aplicaciones de visión por computador. Aunque éste no era propiamente un SMTV, las técnicas de visión por computador empleadas en su construcción contribuyeron al desarrollo de nuevos SMTV [16].

En 1989 el Instituto de tecnología de California, Caltech, y el Laboratorio de Propulsión a Chorro (JPL), desarrollaron un sistema capaz de detectar, contar y medir la velocidad de los vehículos en movimiento [20]. Los algoritmos de detección de los vehículos fueron desarrollados para utilizarse en aplicaciones en tiempo real, por lo que solo operaban sobre pequeñas porciones de las imágenes (líneas de 50 píxeles de ancho). La identificación de los vehículos se hacía mediante diferenciación de marcos. Si la diferencia sobrepasaba un umbral, un vehículo era detectado. El sistema podía realizar la detección de vehículos sobre un máximo de tres carriles de tránsito; aunque, a diferencia de los sistemas antes mencionados, no dependía de marcaciones realizadas en la escena.

Entre 1989 y 1992 en la Universidad College London, N. Hoose [18, 19] desarrolló un sistema para el monitoreo del tráfico. El sistema operaba a un nivel macroscópico, en el que no se identificaban o rastreaban vehículos individuales, sino que se

proveía una descripción cualitativa de la distribución espacial estacionaria y de movimiento en una escena de tránsito. Para detectar un vehículo, el sistema dividía la imagen en celdas, las cuales se etiquetaban en uno de tres posibles estados, celda vacía, celda con vehículo estacionario o celda con vehículo en movimiento; luego, las celdas adyacentes con estados iguales se agrupaban. La detección de vehículos estacionarios se realizaba encontrando los bordes de un objeto completo en las celdas agrupadas, mientras que la detección de los vehículos en movimiento se realizaba por diferenciación de las celdas en dos instantes de tiempo.

El sistema desarrollado por la Universidad de Ciencia y Tecnología de IRAN en 1995, liderado por el profesor Fathy, era capaz de medir los parámetros de las colas de tráfico como por ejemplo: la longitud de la cola, el periodo y la frecuencia de ocurrencia. El sistema operaba en pequeñas regiones de imagen llamadas perfiles, con lo cual se lograba un control en tiempo real. Cada perfil era de solo un píxel de ancho y se extendía a lo largo del centro de cada carril de tránsito. Las colas eran detectadas primordialmente por que no existía movimiento a lo largo del carril. La presencia de vehículos era confirmada con una técnica basada en detección de bordes de vehículos. La presencia de vehículos juntos sin movimiento significaba que había una cola en la escena [11].

En 1995 Koller y su grupo de investigación en la Universidad de California, Berkeley, desarrollaron un sistema capaz de detectar y seguir múltiples vehículos y al mismo tiempo, capaz de suministrar información sobre la forma, esto último le permitía al sistema ser utilizado para propósitos de clasificación [31].

En 1997, Cornish y Wakefield de la Universidad de Huddersfield, desarrollaron un sistema capaz de rastrear vehículos en escenas en movimiento, usando técnicas de diferenciación de marcos (frames) para segmentar los vehículos en la escena. Junto con el sistema también se desarrolló un método para la generación automática de dichas escenas, mientras que simultáneamente se crea una región objetivo de interés, la cual reduce la porción de la imagen que será analizada para encontrar los vehículos, reduciendo así la demanda computacional [6].

Muchos de los sistemas antes mencionados utilizan arquitecturas que fueron diseñadas a la medida de una aplicación específica. La funcionalidad es llevada a cabo por unos cuantos módulos de control de vehículos y visión por computador fuertemente interconectados. Aunque esta clase de arquitectura es adecuada para muchas aplicaciones, presenta como desventajas la falta de escalabilidad, de extensibilidad y de uniformidad de la cooperación entre módulos [16].

Las técnicas de detección de vehículos empleadas por los sistemas hasta aquí mencionados, se basan principalmente en la diferenciación de marcos (Frames) para la detección tanto de vehículos en escenas estacionarias, como de vehículos en escenas en movimiento. Aunque estas técnicas han mostrado

ser adecuadas para la detección de vehículos en dichos sistemas, también es cierto que son restrictivas, por lo que los sistemas basados en estas técnicas resultan ser muy sensibles al cambio o variaciones que se puedan producir en las condiciones para las cuales fueron diseñados; como por ejemplo: solo procesamiento de pequeñas áreas o segmentos de las imágenes, líneas y colores de referencia y marcas realizadas en las escenas, entre otras.

Más recientemente, el interés se ha enfocado en sistemas de "detección temprana", los cuales alertarían a los conductores sobre posibles congestiones vehiculares o accidentes lejanos en la misma dirección que estos llevan. Tales sistemas requieren operar en tiempo real, por lo que sistemas de procesamiento por lotes no serían adecuados. Otro requerimiento de los sistemas diseñados para operar en ambientes reales de tránsito, es que estos no deben ser sensibles a las vibraciones de las cámaras. Normalmente, los sistemas se ubican sobre los puentes o puntos por encima del área de tráfico monitoreada. Ambas ubicaciones están sujetas a gran cantidad de vibraciones ocasionadas por el viento e incluso el mismo transitar de los vehículos [12].

En 1996 W. Kasprzak y H. Niemann, desarrollaron un esquema adaptativo de reconocimiento de objetos en movimiento en una secuencia de escenas. El esquema es aplicado al reconocimiento de objetos en el tráfico. La estimación recursiva del estado del objeto es ejecutada por un filtro extendido de Kalman con una modificación en la estimación del error de seguimiento de los objetos. Dicho error es estimado por medio de un proceso de aprendizaje de una red neuronal. La estimación del error mediante la red neuronal permitió el diseño de un estimador recursivo, que sigue muy bien las señales no estacionarias (los objetos) [21].

En 1997 N. Friedman y S. Russel, desarrollaron una aproximación probabilística para segmentar imágenes en una secuencia de video, basándose en la substracción del fondo (background). La idea era que al restar del marco actual un fondo promediado en el tiempo, quedarían solo los objetos no estacionarios. Este método fallaba cuando existían objetos moviéndose lentamente, lo cual ocasionaba que no se diferenciaban sombras de objetos en movimiento. Este era un enfoque muy simple de la tarea de clasificación de cada píxel del marco actual. Lo que se planteaba, era realizar un modelo probabilístico para la clasificación de cada uno de los píxeles que conformaban la imagen, dicho modelo indicaba que tanto un píxel pertenecía o no a una determinada clase. Esencialmente, los modelos de los píxeles describían la distribución de probabilidad de la apariencia de un píxel en las diferentes clases. La distribución de probabilidad de la apariencia de los píxeles era modelada como mezclas Gaussianas [13].

En 1998 G. Salgian y D.H. Ballard, desarrollaron y probaron rutinas visuales para el control de vehículos. El trabajo se enfocó en la generación de rutinas visuales a partir de imágenes, usando modelos de color y forma. Las rutinas visuales presentadas en

este trabajo son la componente principal de un subsistema de percepción de un vehículo inteligente. Estas rutinas fueron probadas en un sistema de simulación bajo un amplio rango de condiciones controlables, dado que algunas de ellas hubiesen sido peligrosas probarlas en un vehículo real [30].

En 1999 R. Cucchiara y M. Piccardi, desarrollaron algoritmos para la detección automática de vehículos en escenas urbanas, bajo diferentes tipos de condiciones, como es el caso del día y de la noche. Describieron dos tipos de algoritmos de análisis de imágenes, los cuales han sido usados para la extracción de vehículos en secuencias de imágenes adquiridas de día y de noche. En este trabajo se propuso un esquema de propósito general para el monitoreo del tráfico. El esquema propuesto se componía de un sistema de dos niveles; el nivel superior, estaba compuesto de un sistema de razonamiento basado en reglas de producción que supervisaba los diferentes módulos de procesamiento de las imágenes. En el nivel inferior se encontraban los módulos que detectaban los vehículos en movimiento [7].

En 2002, S. Grupte; O Masoud, R.F.K. Martin; N.P. Papanikolopoulos, desarrollaron un algoritmo de detección basado en visión y clasificación de vehículos en secuencias de imágenes monoculares de escenas de tráfico, grabadas por cámaras estacionarias. El método propuesto por este grupo es llevado a cabo en tres niveles: nivel de imágenes sin procesar, nivel de región y nivel de vehículos. Los vehículos son modelados como figuras rectangulares con un cierto comportamiento dinámico. El método propuesto se basa en el establecimiento de correspondencias entre regiones y vehículos, como lo son los vehículos moviéndose a través de la secuencia de imágenes. Resultados experimentales a partir de escenas de autopistas demostraron la efectividad del método que se propuso [17].

En 2002, Alessandro Bevilacqua desarrolló un SMTV el cual emplea un nuevo método de segmentación basado en morfología matemática y substracción de las escenas a partir de un marco de referencia. La operación morfológica propuesta, trabaja muy bien en imágenes muy ruidosas, las cuales es capaz de filtrar, al tiempo que detecta y preserva la estructura de los objetos en movimiento [3].

Recientemente, en 2004, Anurag Mittal y Nikos Paragios propusieron un nuevo método para el modelado y substracción de vehículos en escenas en movimiento. Se propuso una técnica para el modelado de escenas dinámicas con el propósito de abstraer el fondo de la escena y detectar cambios. El método utiliza la técnica de flujo óptico para la detección de cambios; luego, para tratar apropiadamente la incertidumbre de dichas características, se propone una nueva técnica de estimación multivariada de densidad basada en funciones kernel, la cual es de tipo no predictiva, y adapta un ancho de banda o ventana de acuerdo a la incertidumbre en las pruebas y medidas del muestreo [25].

En el 2004, Melo et al [24], proponen un sistema de bajo nivel para el seguimiento de vehículos, el cual produce trayectorias precisas de los vehículos y estima la geometría de los carriles usando cámaras de vigilancia vehicular, las cuales no están calibradas. Ellos desarrollaron un algoritmo llamado predicción de trayectorias por mezcla y división (Predictive Trajectory Merge and Split - PTMS), el cual es capaz de detectar oclusiones parciales o completas durante el movimiento de los objetos, actualizando el número de objetos en cada segmento rastreado. El algoritmo es un híbrido entre el filtro de Kalman y un conjunto de heurísticas para el análisis temporal. En resultados preliminares demostraron la viabilidad de usar cámaras comunes y corrientes las cuales no han sido calibradas, mostrando de esta manera buenos resultados en el análisis del tráfico en tiempo real. Ellos argumentan que su algoritmo es independiente del punto de vista y no hace suposiciones a priori en cuanto a la geometría de los carriles. Este sistema puede utilizarse para monitoreo de alto nivel del tráfico, la velocidad, la frecuencia de cambio de carriles, detección de accidentes y clasificación de comportamientos extraños de los conductores.

En Atkociunas et al 2005 [2], se hace un enfoque con algoritmos de procesamiento de imágenes, reconocimiento de patrones y visión por computador para ser aplicados al análisis y monitoreo del tráfico. Uno de los principales aspectos fue la modificación de estos algoritmos para que fuesen operativos en tiempo real para el proceso de monitoreo, dando como resultado un prototipo de un sistema para el análisis del tráfico. El sistema se compone de video cámaras estacionarias, así como de una red de computadores, este sistema tiene la capacidad de seguimiento de vehículos, medición de sus velocidades, reconocimiento de los números de placas de los vehículos en movimiento y cambio de carriles.

Recientemente Cheung y Kamath [5], proponen un algoritmo de validación del foreground, el cual primero construye una máscara para el foreground usando un filtro de Kalman de retardo adaptativo, validando así píxeles individuales del foreground construyendo un modelo simple del objeto en movimiento, utilizando para ello las estadísticas del foreground y el background así como la diferenciación de marcos (frames). Los experimentos mostraron la utilización del algoritmo en secuencias de tráfico urbano, mostrando significativas mejoras sobre resultado usando únicamente el filtro de Kalman o solamente la diferenciación de marcos (frames) y similares desempeños en técnicas basadas en mezclas Gaussianas, filtros de mediana y filtros de mediana aproximados. Aun con estos resultados el algoritmo muestra limitaciones como la causada por la diferenciación de marcos (frames), la cual produce pequeños caminos falsos en el foreground detrás del objeto. También existen problemas si no se hace un buen muestreo para el cálculo de la densidad de los píxeles del foreground, causando que no se detecten bien los objetos.

También en el 2005 Yoneyama et al [36], desarrollaron un método para resolver tres problemas cruciales encontrados en

los SMTV, los cuales son la eliminación de sombras cuando en vehículo esta en movimiento, la detección de oclusiones y la detección de vehículos en la noche, este método se desarrolló para vigilancia de vehículos en autopistas. Para la eliminación de las sombras ellos utilizan un modelo al que llama juntura 2D entre vehículo y sombra el cual consiste en encerrar el vehículo en la escena 2D en un cubo y luego separar el vehículo de la sombra, a través de proyecciones de puntos de fuga. La detección de oclusiones la solucionan utilizando múltiples cámaras, y de esta manera pueden extraer la localización exacta de cada vehículo. Para la detección de los vehículos en la noche se propone una nueva técnica la cual llaman vista posterior, y lo que hacen es realizar un seguimiento del vehículo desde la parte posterior, la cual combinan con un seguimiento delantero; por ultimo también introducen una mejora en la extracción del background de las escenas para una mejor segmentación y detección del vehículo.

En el 2006 Zheng et al [37], presentan un nuevo algoritmo para la extracción del background de las imágenes de una secuencia de video de tráfico. El algoritmo propuesto analiza cada uno de los valores del color de los píxeles, en una serie de marcos (frames) capturados durante un periodo de tiempo, luego utilizan lo que ellos llaman el modo de las series como el valor adecuado del color para background de la imagen. Para manejar las pequeñas perturbaciones del color del background sobre un corto periodo de tiempo, una función es usada para almacenar los valores de los colores del vecindario para de esta manera calcular el modo. Con esta función, el algoritmo es más robusto y reduce el impacto de las perturbaciones resultantes de factores ambientales. El algoritmo no requiere de ordenamientos y de aquí que su implementación no es dificultosa. Las pruebas preliminares mostraron que el algoritmo es efectivo para generar claros y cercanamente perfectos backgrounds en intersecciones urbanas señalizadas.

2.2 Sistemas de monitoreo de cobertura amplia

Aunque el procesamiento de imágenes ha avanzando rápidamente con el transcurrir de los años, y se ha extendido con éxito hacia muchas áreas como, por ejemplo, el monitoreo del tránsito vehicular; hoy en día, son pocos los trabajos y experiencias que se tienen en aplicaciones de monitoreo de tránsito a través de fotografía satelital. Los pocos esfuerzos que se han hecho en esta área, se han realizado en proyectos de investigación en la unión europea y EE.UU, los cuales analizan el potencial de estos trabajos hacia el futuro.

En 1998 Mark R. McCord, Carolyn J. Merry y Prem Goel, Desarrollaron un sistema de simulación de monitoreo satelital de vehículos para tráfico en autopistas. EL sistema realizaba conteo de vehículos para identificar el flujo vehicular en una ventana de tiempo. Utilizaron seguimiento de píxeles para determinar la velocidad de los vehículos, así como substracción de imágenes y transformaciones de reflectancia. Los datos eran analizados e interpretados automáticamente. Para simular las fotografías satelitales se hicieron tomas aéreas [23].

En 2000 Mario Pent, Fabio Dosis y Paolo Mulassano, desarrollaron la plataforma “HeliNet Network of Stratospheric Platforms for Traffic Monitoring, Environmental Surveillance and Broadband Services”. La plataforma presenta información cuantitativa, precisa y detallada del movimiento del tráfico. La plataforma HeliNet es capaz de monitorear vía satelital pequeñas poblaciones de tráfico, registrar la velocidad de los vehículos y detectar congestiones, al tiempo que reporta vía satélite la información procesada a un sistema central [26].

Entre 2001 y 2003 mediante una serie de trabajos basados en la utilización del dispositivo Three Line Scanner (TLS), se desarrolló un método de detección de vehículos en ambientes urbanos. El trabajo propuso un framework de tres etapas, a saber: detección de vehículos, discriminación entre vehículos parqueados y en movimiento, identificación de vehículos

parqueados. La etapa 1 se refiere a la obtención de la imagen por medio del TLS y su segmentación en regiones con presencia de vehículos y con ausencia de éstos, descartando la segunda. En la etapa 2, se detectan los objetos que dan indicios de vehículos y se generan los candidatos a éstos a partir de las regiones segmentadas por medio de la técnica de expansión de detección de marcos. Como última etapa, los vehículos en la imagen son verificados y detectados por coincidencia de modelado de vehículos [27], [28], [29].

Los diferentes trabajos en el desarrollo de sistemas de monitoreo tránsito vehicular hasta aquí mencionados, muestran la variedad de aplicaciones y servicios que dichos sistemas pueden tener. De acuerdo al propósito general o tipo de servicio que estos sistemas presentan, podemos clasificarlos según la taxonomía presentada en la figura 2.

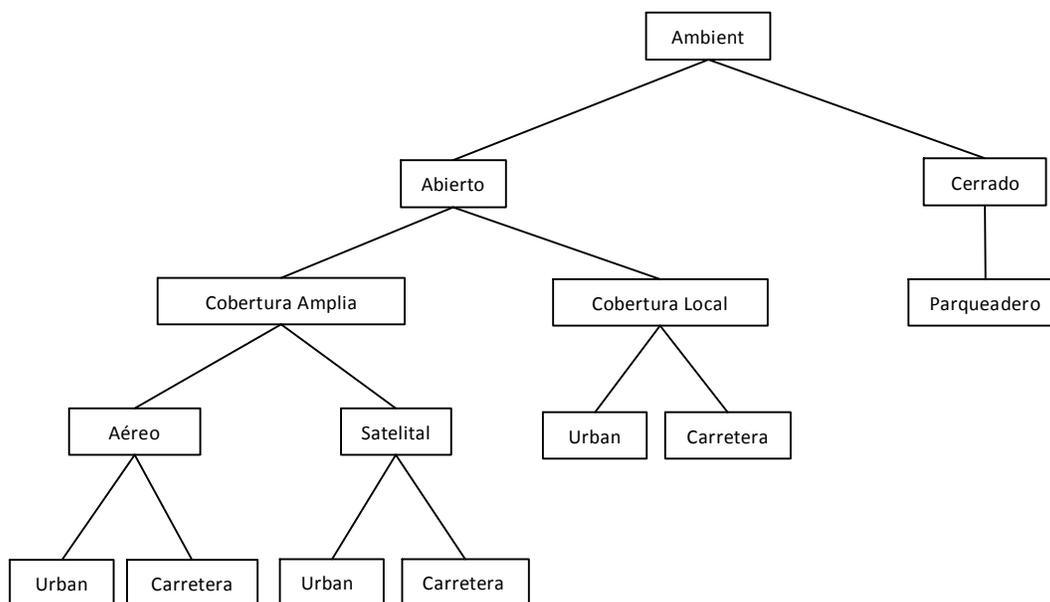


Figura 2. Clasificación de sistemas de monitoreo de tránsito vehicular

Sin importar cual sea el propósito o servicio ofrecido por los SMTV basados en visión, la efectividad en la prestación de éste dependerá, principalmente, de las técnicas de segmentación de imágenes que emplee. Dada la importancia de estas técnicas, la siguiente sección presenta las principales técnicas de segmentación empleadas por estos sistemas.

III. TÉCNICAS DE SEGMENTACIÓN EN LOS SISTEMAS DE MONITOREO DE TRÁNSITO VEHICULAR

Una variedad de técnicas han sido empleadas en el tratamiento digital de imágenes de tránsito vehicular, estas comprenden desde los filtros clásicos de procesamiento de imágenes, hasta técnicas de aprendizaje estadístico.

Dada la amplia variedad de técnicas de segmentación de imágenes empleadas en los sistemas de tránsito vehicular, en la tabla 1 se presenta un cuadro resumen que realiza una clasificación de éstas de acuerdo a su naturaleza.

A continuación, se describen las principales técnicas o metodologías utilizadas en los diferentes proyectos e investigaciones en el procesamiento digital de imágenes de tránsito vehicular, que se han venido utilizando a lo largo de las dos últimas décadas en este campo.

Técnicas de Modelado Espacial

Entre las técnicas de segmentación utilizadas en los SMTV, las cuales tienen su apoyo en el dominio espacial podemos encontrar:

Diferenciación de frames (marcos)

Uno de los primeros intentos en el desarrollo de algoritmos y técnicas de detección de vehículos, fue el de detección de variaciones sobre marcaciones realizadas en la escena. La técnica consistía en trazar líneas a lo ancho de los carriles de una autopista, y detectar si existía una variación en los niveles de gris sobre las líneas trazadas; en caso afirmativo, un vehículo había sido detectado [31].

La diferenciación de frames en una secuencia de imágenes, es una de las técnicas más utilizadas en el desarrollo de algoritmos de tratamiento de imágenes de tránsito vehicular. La diferenciación de frames, como su nombre lo indica, consiste en restar dos frames capturados en tiempos diferentes, para detectar regiones en movimiento. En [31] se establece una diferencia de imágenes, $d(i, j)$ como la generada por el cálculo de la diferencia absoluta entre dos frames f_1, f_2 y el resultado de la umbralización:

$$d(i, j) = \begin{cases} 0 & \text{si } |f_1(i, j) - f_2(i, j)| \leq \mathbf{T}, \\ 1 & \text{en otro caso} \end{cases}$$

donde, \mathbf{T} es el umbral establecido. En el caso del monitoreo del tráfico, usualmente se toma f_1 como el frame entrante, f_2 será una referencia o frame de fondo. El frame de referencia es simplemente una imagen de la escena, la cual no contiene vehículos. El propósito del umbral \mathbf{T} , es reducir el efecto del ruido y los cambios en la iluminación de la escena.

Flujo Óptico

Una técnica reciente empleada en los SMTV es la llamada flujo óptico. El flujo óptico busca correspondencias entre los puntos de una imagen y los de una localidad de la misma en la correspondiente al siguiente cuadro o frame; obteniéndose un campo de movimiento denso que pretende corresponder a la velocidad instantánea de los puntos en el plano de la imagen. El flujo óptico de una secuencia de imágenes es un conjunto de campos de vectores, que relaciona cada imagen con la próxima en la escena. Cada campo de vectores representa el desplazamiento aparente de cada píxel desde una imagen a otra. Si se asume que el píxel conserva su intensidad, se lleva a la “ecuación de conservación del brillo”.

$$I(x, y, t) = I(x + dx, y + dy, t + dt)$$

donde, I es una secuencia de imágenes, $[dx, dy]$ es el vector de desplazamiento para el píxel en las coordenadas $[x, y]$ y t, dt son el frame y el desplazamiento temporal de la secuencia de imágenes respectivamente.

El flujo óptico viene dado por la siguiente expresión.

$$\nabla^T g \cdot f + gt = 0$$

donde, ∇^T y gt son el gradiente espacial y la derivada temporal, de la imagen en una ubicación espacial y tiempos dados y f es el vector de velocidad bidimensional.

Un método de segmentación de objetos en movimiento que una variante del flujo óptico, se presenta en [4], este método segmenta los objetos de interés a partir de secuencia de imágenes que poseen un fondo o background complejo; para ello emplean el cálculo de la componente normal del flujo óptico de las imágenes. Como la estimación precisa de movimiento visual es computacionalmente costoso, se aplica una técnica de filtrado la cual produce una imagen de ceros-cruzados (Zero-Crossings) como salida. Esta salida es dada por la convolución del historial de intensidad da la imagen $I(x, y, t)$ con el filtro espacial-temporal gaussiano de d’Alembertian; el cual es descrito por la siguiente ecuación.

$$S(x, y, t) = \left(\nabla^2 + \left(\frac{1}{u^2} \right) \frac{\partial^2}{\partial t^2} \right) F(x, y, t) \otimes I(x, y, t)$$

$$F(x, y, t) = u \left(\frac{1}{2\sigma^2\pi} \right)^{\frac{3}{2}} \exp \left[-\frac{1}{2\sigma^2} (x^2 + y^2 + u^2 t^2) \right]$$

donde, u es un factor de escalamiento de tiempo.

El próximo paso es detectar los ceros-cruzados en $S(x, y, t)$ y calcular la magnitud de la normal de velocidad en cada uno de los píxeles ceros-cruzados. La normal de velocidad es calculada junto a un vecindario 3×3 ; así que se necesitará al menos tres imágenes sucesivas o frames para calcular los espacios temporales ceros-cruzados. La magnitud de la normal del movimiento visual es estimada por la ecuación:

$$v_n = u \frac{\frac{\partial S}{\partial t}}{\sqrt{\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2}}$$

donde, $\partial S/\partial x, \partial S/\partial y$ y $\partial S/\partial t$ son calculados usando el método numérico de diferencias finitas.

Sensado Tridimensional

Otro tipo de técnica utilizada en los SMTV, es la llamada sensado tridimensional, la cual se basa en el principio de triangulación. Esta técnica se apoya en la localización de dos cámaras que realizan tomas sobre un mismo punto, en el mismo instante de tiempo, lo cual permite tomar medidas de coordenadas tridimensionales de la altura del vehículo, como la

de los bordes de este; permitiendo así, la representación de los vehículos como cubos. La representación en cubos permite diferenciar vehículos traslapados y evitar que se confundan los vehículos con sus sombras o las de otros objetos. La ventaja de esta técnica, es su independencia de la hora del día, el estado del tiempo y la iluminación de la escena. La desventaja es que es un método costoso debido a la tecnología que se emplea como lo es el sensado con láser.

Técnicas de Modelado Estadístico

Son variadas las técnicas de segmentación utilizadas en los SMTV, las cuales tienen su apoyo en la estadística entre las más representativas podemos encontrar:

Filtro de Kalman

El filtro de Kalman, es quizás de las técnicas más utilizadas no solo en los SMTV, si no en otras áreas como economía robótica ingeniería aeroespacial etc. El filtro de Kalman, es un conjunto de ecuaciones matemáticas que proveen un eficiente cálculo de la media en forma recursiva, para estimar el estado de un proceso, de tal forma que minimice la media del error cuadrado. El filtro es poderoso en varios aspectos: este soporta estimaciones de estados pasados, presentes y futuros; haciéndolo aún cuando la naturaleza del sistema modelado es desconocida [35]. El filtro es descrito por las siguientes ecuaciones.

$$\mathbf{S}(k+1) = \mathbf{A}(k+1)\mathbf{S}(k) + \mathbf{W}(k+1)$$

$$\mathbf{Z}(k) = \mathbf{H}(k)\mathbf{S}(k) + \mathbf{U}(k)$$

donde, $\mathbf{S}(k)$ es el vector de estado y contiene los valores de los M parámetros que definen el estado del sistema en el instante k . $\mathbf{A}(k)$ es la matriz de transición de estado de dimensiones $M \times M$. $\mathbf{H}(k)$ es una matriz de dimensiones $L \times M$, que se conoce como la matriz de observaciones. w y u son dos vectores de M y L componentes respectivamente; y representan a dos procesos blancos no correlacionados y de media nula de covarianzas \mathbf{W} y \mathbf{U} , de forma que:

$$E[w(k)w^T(k)] = \mathbf{W}(k)$$

$$E[u(k)u^T(k)] = \mathbf{U}(k)$$

Se supone que las matrices \mathbf{A} , \mathbf{W} , \mathbf{H} , \mathbf{U} , aunque pueden variar con el tiempo, son conocidas a priori [14].

Las características mencionadas, permiten la utilización del filtro Kalman, para tratar secuencias de imágenes y de esta forma detectar vehículos u objetos en movimiento, el filtro es capaz de estimar a partir de estados anteriores de los frames, los estados siguientes de estos. De esta forma el filtro Kalman es usado para hacer seguimiento de la posición y la forma del contorno de los objetos a través de los frames sucesivos [21], [22].

Núcleos de Densidad (Kernel de densidad)

Otras técnicas empleadas, que han tomado relevancia, son las relacionadas con la estadística no paramétrica. Estas técnicas

tienen como base, hacer un modelamiento de los píxeles de una escena determinada, para ello usan una función de densidad de probabilidad que describe el comportamiento de la distribución de los píxeles de una escena a otra. Para poder crear la función de densidad, se toman los píxeles que componen el fondo de una escena actual de referencia (una escena estática y sin objetos en movimiento) y una escena siguiente (con objetos en movimiento), con base en la función de densidad se hace el modelado de ambas escenas, y se analizan que píxeles cambiaron de una escena a otra con base en la función de densidad que los modela, tomando para ello los valores apropiados de la varianza que representa los niveles de intensidad de cada píxel, en cada escena en particular, y distinguiendo las varianzas de los píxeles de una escena a otra, identificando de esa manera que píxeles cambian y cuales no.

En [25] [8] [9] se propone un nuevo método para el modelado y substracción de vehículos en escenas en movimiento. Se modelan escenas dinámicas con el propósito de abstraer el fondo de ésta y detectar cambios. El método utiliza un flujo óptico como característica de detección de cambios. Para tratar apropiadamente la incertidumbre en las características, se propone una nueva técnica de estimación multivariada de densidad basada en kernel, el cual es un método no predictivo, que adapta un ancho de banda o ventana de acuerdo a la incertidumbre en las pruebas y medidas del muestreo.

La idea central en [25], es utilizar las propiedades de la escena para la construcción de una apropiada representación que pueda ser utilizada para la clasificación de una nueva observación del fondo. La información provista por esta representación puede ser considerada como una visualización de bajo nivel, la cual da inicio a la realización de la tarea de alto nivel del análisis de objetos como: detección, seguimiento y clasificación.

El modelamiento del background se realiza mediante un estimador kernel. Este construye una representación probabilística de las observaciones de un píxel en particular. En otras palabras, modela la distribución probabilística de los píxeles del fondo. Este método kernel también llamado "ventanas de parzen" en el reconocimiento de patrones; se representa por la siguiente expresión.

$$\hat{f}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \mathbf{K}_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)$$

donde, n es un conjunto de puntos de una muestra, $\mathbf{K}(\cdot)$ es la función Kernel, en este caso de tipo Gaussiana, la cual está centrada en cada uno de los datos del muestreo y en función de una matriz \mathbf{H} , la cual contiene los parámetros de suavidad y especifica el ancho de la ventana, ya que puede adaptarse a los datos de la muestra.

Una vez establecida la aproximación de la densidad, se procede a determinar un mecanismo de clasificación para los datos observados. Esta puede realizarse por umbralización de

la probabilidad de que una nueva observación pertenezca al fondo. De esta forma, se puede extraer el fondo de las escenas.

El siguiente paso es el modelamiento de las características dinámicas, el cual es basado en el cálculo del flujo óptico; dicho cálculo, es utilizado como una característica en un espacio dimensional superior. Aunque muchos algoritmos de flujo óptico y sus extensiones pueden ser considerados, en este trabajo, se utilizó el flujo óptico diferencial multiescala bayesiano. Este tiene la característica deseada de ser capaz de determinar el error característico de la estimación del flujo óptico.

Mezclas o Misturas Gaussianas

En [38] se modela el fondo, mediante la técnica de Modelado de Mezclas Gaussianas (GMM), con el fin de abstraer el fondo de la escena y de esta manera distinguirlo de foreground o plano superior de ésta. Como en la práctica la iluminación de la escena puede cambiar gradualmente (hora del día, condiciones del tiempo, etc.) o de una manera imprevista como en un cambio de luz al interior de la escena. En ese instante un nuevo objeto puede ser ingresado o borrado de la escena. Para ello, se hace un GMM correspondiente al modelado del fondo, además de un modelo para el plano superior; luego, empleando el teorema de Bayes, se realiza la clasificación de los píxeles cambiantes en la escena.

Los GMM, son modelos semiparamétricos, los cuales son alternativos a los métodos no paramétricos, son una generalización de las redes de función de base radial. Su representación matemática es la siguiente.

$$P(x/j) = \sum_{j=1}^M P(x/j)P(j)$$

Para el caso particular del modelamiento del background, tendríamos que una densidad condicional para el píxel x perteneciente al background j ; será una Mezcla (Mixture) con M componentes de densidad. Donde el coeficiente $P(j)$ es llamado parámetro de Mezcla (Mixing) y

$$\sum_{j=1}^M P(j) = 1$$

es el j -ésimo componente Gaussiano con media de intensidad $\hat{\mu}$ y desviación estándar $\hat{\sigma}$.

Un modelo alternativo al GMM es propuesto en [10], en el cual se combina un modelo robusto de formas, con los histogramas de las características de las imágenes, arrojando buenos resultados en las pruebas.

La fortaleza de este método reside en la búsqueda y detección de propiedades tales como: forma, color y textura. Para ello

modelan la densidad conjunta de dichas propiedades pertenecientes a los objetos presentes en las imágenes. Denotando un vector de características (color, textura), en un punto perteneciente a un objeto f ; se puede usar una distribución conjunta de la forma

$$p(f, x, y) = p(f)p(x, y)$$

Como esta distribución no puede modelar propiedades espaciales variantes en la imagen; una solución es adicionar una copia de las coordenadas espaciales al vector del observador y usar una distribución de la forma:

$$p(f, x, y, x', y') = p(f, x, y)p(x', y')$$

Una posible forma de $p(f, x, y)$, la cual es elegante por su simplicidad es construir un histograma, para lo cual el espacio $\langle x, y, f \rangle$ es dividido en una maya bidimensional, asumiendo que los componentes de f son mutuamente independientes.

Una forma de calcular esta distribución es entonces teniendo un producto de los histogramas unidimensionales de cada una de las componentes independientes, de la forma:

$$p(f, x, y) = \prod_{i=1}^F p_{xy}(f_i)$$

donde, cada f es un histograma unidimensional.

Sumando dicha productoria, se obtiene una “mezcla de histogramas”. O se puede calcular dicha distribución de forma bidimensional de la siguiente manera.

$$p(f, x, y) = \sum_{i=1}^B \sum_{j=1}^B p(f/b_{ij})p(x, y/b_{ij})P(b_{ij})$$

donde B , es la dimensión, y b_{ij} son las subdivisiones del histograma. Lo cual es similar a las mezclas Gaussianas continuas que se presenta en [38].

Filtro de Gabor

Una combinación de wavelets y un filtro Gabor es empleada en [33] con el fin de reducir el error en la detección de objetos de interés en secuencias de imágenes. La combinación utiliza los wavelets para codificar la información de los bordes y el filtro Gabor para determinar la orientación y la realizar un afinamiento de la detección de bordes y líneas

La función general de la familia de filtros Gabor bidimensional, puede ser representada como una función Gaussiana modulada por una función compleja sinusoidal. Específicamente el filtro bidimensional de Gabor es:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x}{\sigma_x} + \frac{y}{\sigma_y}\right)^2\right] \exp[2\pi j\mathbf{W}x]$$

$$\bar{x} = x\cos\theta + y\sin\theta$$

$$\bar{y} = -x\sin\theta + y\cos\theta$$

donde, σ_x y σ_y son los parámetros de escalamiento del filtro que determinan el tamaño efectivo del vecindario de un píxel mediante la suma de los pesos. El parámetro $\theta(\theta \in [0, \pi))$, especifica la orientación del Filtro [32], [33].

Filtrado de partículas

El Filtro de Partículas es un método secuencial de Monte Carlo aplicable a cualquier transición de estados o modelo de medida. En la Visión Artificial se utiliza para el seguimiento de objetos en secuencias de imágenes. Este método lo propusieron N. Gordon, D. Salmond y A. Smith en 1993 [15]. El problema de la estimación secuencial, desde un punto de vista Bayesiano, consiste en el cálculo recursivo, con un cierto grado de confianza, del estado del sistema X_t en el instante t , dadas las

observaciones $Z_{1:t} = \{Z_1, \dots, Z_t\}$. Para ello, es necesario calcular la función de distribución de probabilidad (pdf) $p(X_t | Z_{1:t})$. La pdf inicial $p(x_0 | z_0) \equiv p(x_0)$ se asume conocida. Así, la pdf a posteriori $p(X_t | Z_{1:t})$ se puede calcular en dos etapas:

Tabla 1. Clasificación de técnicas de segmentación de imágenes

		Dominio de aplicación	
		Espacial	Frecuencia
Naturaleza de la técnica	Determinística	<ul style="list-style-type: none"> • Umbralización • Diferenciación de imágenes • Detección de bordes • Morfológicos • Por movimiento • Orientados a regiones • Fusión de regiones [17] [3]	<ul style="list-style-type: none"> • Umbralización • Filtro de Gabor • Wavelets [32] [33]
	Heurística	<ul style="list-style-type: none"> • Redes Neuronales • Algoritmos Genéticos • Lógica Difusa • Clusterización [21]	
	Estadística	<ul style="list-style-type: none"> • Bayesianos • Histograma de Frecuencias • Kernels de Densidad • Mezclas Gaussianas • Filtro de Kalman [21] [8] [9] [17]	

1. **Predicción:** supóngase que se dispone de la pdf $p(X_{t-1} | Z_{1:t-1})$ en el instante $t-1$. La etapa de predicción implica el uso del modelo del sistema para obtener de manera recursiva la pdf a priori $p(x_t | Z_{1:t-1})$ en el instante siguiente t , mediante la ecuación:

$$p(x_t | Z_{1:t-1}) = \int p(x_t | z_{t-1}) p(x_{t-1} | Z_{1:t-1}) dx_{t-1}$$

2. **Evaluación:** En el instante t , se dispone de una nueva medida (Z_t) que se puede utilizar durante la actualización del estado del sistema mediante el uso del teorema de Bayes:

$$p(x_t | Z_{1:t}) = \frac{p(z_t | x_t) p(x_t | Z_{1:t-1})}{p(z_t | Z_{1:t-1})}$$

Las dos ecuaciones de predicción y evaluación, conforman la base de la solución Bayesiana óptima. Esta propagación

recursiva de la densidad a posteriori es únicamente un resultado conceptual, puesto que, en general, no se puede determinar analíticamente. Por esta razón, se encuentra en la literatura un número significativo de trabajos que presentan modelos aproximados de esas funciones de distribución.

IV. CONCLUSIONES Y TENDENCIAS FUTURAS

Hacia el futuro, se espera que los sistemas inteligentes de visión sean aplicados, cada vez más, en tareas más complejas. La creciente complejidad de estos sistemas, requerirá arquitecturas de software

que puedan tratar con varias aplicaciones y módulos de percepción, simultáneamente. Los sistemas de visión inteligente de control de tránsito, evolucionarán a sistemas que entreguen información más precisa y granular sobre las escenas capturadas en tiempo real.

Se espera que éstos, entreguen detalles como cantidad de vehículos, velocidad de desplazamiento de los mismos, modelo, color, identificación del número de matrícula, aún cuando estos se encuentren en movimiento; y en general, detalles minuciosos que los sistemas actuales no están en capacidad de proporcionar y que contribuirían al mejoramiento de las condiciones de seguridad y de regulación del tránsito vehicular.

El avance en las tecnologías de las telecomunicaciones permitirá desarrollar sistemas de tránsito vehicular en red, que cubran y entreguen información detallada de áreas amplias, las cuales son cubiertas en la actualidad por sistemas de cobertura amplia que no entregan detalles de los vehículos de la escena.

REFERENCIAS

- [1] Abramczuk T., 1984. A microcomputed based TV detector for road traffic. In Symposium on Road Research Program, Tokyo, Japan, October.
- [2] Atkojiunas E.; Blake R.; Juozapavicius A. and Kazimianec M., 2005. Image Processing in Road Traffic Analysis. Nonlinear Analysis: Modelling and Control, Vol. 10, No. 4, 315–332.
- [3] Bevilacqua A., 2002. Effective Object Segmentation in a Traffic Monitoring Application. 3rd IAPR Indian Conference on Computer Vision, Graphics and Image Processing.
- [4] Camapum J. and Fisher M., 1998. Segmentation using Spatial-Feature Clustering from Image Sequences. International Conference on Image Processing (ICIP '98) October 04 - 07, Chicago, Illinois.
- [5] Chandrika K; Cheung S. and Ching S., 2005. Robust Background Subtraction with Foreground Validation for Urban Traffic Video. EURASIP Journal on Applied Signal Processing, 2330–2340.
- [6] Cornish M.T. and Wakefield J.P., 1996. Automatically Locating an Area of Interest and Maintaining a Reference Image to Aid the Real Time Tracking of Objects. In Proceedings of British Machine Vision Conference, pages 475-484, September.
- [7] Cucchiara R. and Piccardi M., 1999. Vehicle Detection under Day and Night Illumination'. Proceedings of 3^a International ICSC Symposium on Intelligent Industrial Automation.
- [8] Davis L.; Elgammal A.; and Harwood D., 2000. Non-parametric Model for Background Subtraction, 6th European Conference on Computer Vision, Dublin.
- [9] Davis L.S. ; Duraiswami R.; Elgammal A. ; and Harwood D., 2002. Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance. Proceedings of the IEEE, vol. 90, no. 7, July.
- [10] Everingham M.R.; and Thomas B.T., 2001. Supervised segmentation and tracking of nonrigid objects using a 'mixture of histograms' model. In Proceedings of the 8th IEEE International Conference on Image Processing (ICIP2001), pages 62-65, October.
- [11] Fathy M.; and Siyal M.Y., 1995. A real-time image processing approach to measure traffic queue parameters. IEEE proceedings. Vision, image and Signal processing 142(5): 297-303, October.
- [12] Blake A.; Ferrier N.; and Rowe S., 1994. Real Traffic Monitoring?. IEEE Workshop on Applications of Computer Vision, Sarasota, FL.
- [13] Friedman N.; and Russel S., 1997. Image Segmentation in Video Sequences a Probabilistic Approach. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence.
- [14] Fumero A. M., 2000. El Filtro de Kalman. Temas Especiales de Análisis Numérico. Madrid.
- [15] Gordon N.J.; Salmond D.J.; and Smith A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. Radar and Signal Processing, IEE Proceedings F. 140:2 107-113.
- [16] Franke U.; and Gorzic S., 1998. ANTS - Intelligent Vision in Urban Traffic'. IEEE International Conference on Intelligent Vehicles.
- [17] Grupte S.; Martin R.; Masoud O.; and Papanikolopoulos N., 2002. Detection and Classification of Vehicles. IEEE Transactions on Intelligent Transportation Systems, vol. 3, no. 1, March.
- [18] Hoose N., 1991. Computer Image Processing in Traffic Engineering. Research Studies Press.
- [19] Hoose N., 1989. Queue detection using computer image processing. In 2nd. Int. Conf. Road Traffic Monitoring, pages 94-98, London, February. IEEE.
- [20] Iñigo R. M., 1989. Application of Machine Vision to traffic Monitoring and Control'. IEEE Transaction on Vehicular Technology, 38(3):112-122, August.
- [21] Kasprzak W.; and Niemann H., 1996. Applying a Dynamic Recognition Scheme for Vehicle Recognition in Many Object Traffic Scenes. Workshop on Machine Vision Applications. Tokyo Japan.
- [22] Kwon E. ; Masoud O. ; and Papanikolopoulos N. P., 1998. Vision-Based Monitoring of Weaving Sections. Intelligent Transportation Systems. Proceedings. IEEE/IEE/JSAI International Conference on. 770 – 775.
- [23] Bronzini M.; Goel P.; Gomez R.; Hickman M.; McCord M.; Merry C.; Mirchandani P.; and Morrison J., 2001. Remote Sensing of Transportation Flows, Remote Sensing for Transportation-Product and Results: Foundations for the Future'. Conference Proceedings 29, Committee for Conferences on Remote Sensing and Spatial Information Technologies for Transportation, Transportation Research Board, Washington, DC, December 10-12, pp. 21-27.
- [24] Bernardino A.; Melo J.; Naftel A.; and Santos-Victor J., 2004. Viewpoint Independent Detection of Vehicle Trajectories and Lane Geometry from Uncalibrated Traffic Surveillance Cameras. ICIAR

- Int'l Conference on Image Analysis and Recognition, Porto, Portugal, Sep.29 - Oct.1.
- [25] Mittal A.; and Paragios N., 2004. Motion Based Background Subtraction Using Adaptive Kernel Density Estimation, Siemens Corporate Research. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC, July.
- [26] Dosis F.; Pent M.; and Mulassano P., 2000. HeliNet: Network of Stratospheric Platforms for Traffic Monitoring, Environmental Surveillance and Broadband Services, GNSS2000, Edimburgh, Scotland, UK, 1-4 May.
- [27] Puntavungkour S.; and Shibasaki R., 2003. Vehicle Detection from Three Line Scanner Image'. IEEE 6th International Conference on intelligence transportation Systems, Shanghai, China, 12-15 October.
- [28] Puntavungkour S.; and Shibasaki R., 2003. Feature Object Detection on the urban road surface by The Application of Three Line Scanner Imagery. Poster Session 1, The 23th Asian Conference on Remote Sensing, Kathmandu, Nepal, 25-29 November.
- [29] Puntavungkour S.; and Shibasaki R., 2002. A Study on Possibility of TLS image for Application to Road Monitoring on the street in the urban. Poster session No.1, Geoinformation Forum Japan 2002, Tokyo, Japan, 3-5 July.
- [30] Ballard D.H.; and Salgian G., 1998. Visual Routines for Vehicle Control. In D. Kriegman, G. Hager, S. Morse editors, The Confluence of Vision and Control. Springer Verlag.
- [31] Setchell C.J., 1997. Application of Computer Vision to road-traffic Monitoring. Ph.D Thesis University of Bristol.
- [32] Bebis G.; Miller R.; and Sun Z., 2003. Evolutionary Gabor Filter Optimization with Application to Vehicle Detection. IEEE International Conference on Data Mining, Melbourne FL.
- [33] Bebis G.; Miller R.; and Sun Z., 2002. Improving the Performance of On-Road Vehicle Detection by Combining Gabor and Wavelet Features. IEEE 5th International Conference on Intelligent Transportation Systems, Singapore, Sep.
- [34] Byong W.H.; Kaneko T.; Sakauchi M.; Sekine T.; and Takaba, S., 1982. Measurement of traffic flow using real-time processing of moving pictures. In 32nd. Conf. on Vehicular Technology, pages 488-494, San Diego, C.A.
- [35] Bishop G.; and Welch G., 2006. An Introduction to the Kalman Filter. Department of Computer Science University of North Carolina at Chapel Hill. UNC-Chapel Hill, TR 95-041, July 2.
- [36] Jay K.C.C.; Yeh C.H; and Yoneyama A., 2005. Robust Vehicle and Traffic Information Extraction for Highway Surveillance. EURASIP Journal on Applied Signal Processing, 2305-2321.
- [37] Nihan N.L.; Hallenbeck M.E.; Wang Y.; and Zheng J., 2006. Extracting Roadway Background Image: a Mode-Based Approach. TRB 2006 Annual Meeting.
- [38] Zivkovic Z., 2004. Improved Adaptive Gaussian Mixture Model for Background Subtraction. In Proc. ICPR.