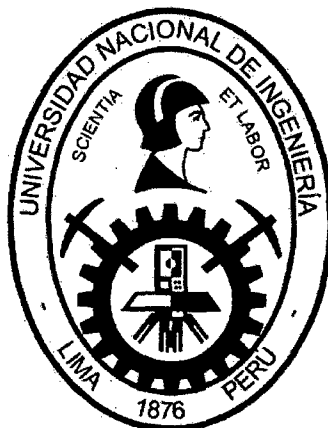


**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS**



**CONTEO DE VEHÍCULOS EN TIEMPO REAL MEDIANTE  
TÉCNICAS DE SEGUIMIENTO DE OBJETOS APLICADO AL  
TRÁNSITO EN LIMA METROPOLITANA EN CONDICIONES DE  
LUZ DIURNA**

**TESIS**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:  
INGENIERO DE SISTEMAS**

**PALACIOS REZZA, ROY DAVID  
VARGAS MENDIETA, JOSE GIANCARLO**

**LIMA – PERÚ**

**2014**

**Digitalizado por:**

**Consortio Digital del  
Conocimiento MebLatam,  
Hemisferio y Dalse**

## **DEDICATORIA**

**A mis padres por enseñarme y motivarme mediante su ejemplo.**

**Roy Palacios Rezza**

**A mis padres, por el gran amor y apoyo incondicional que siempre me han brindado, y a mis hermanos, por los todos los momentos maravillosos que me han regalado.**

**Jose Vargas Mendieta**

## ÍNDICE GENERAL

<b>DESCRIPTORES TEMÁTICOS .....</b>	<b>6</b>
<b>RESUMEN .....</b>	<b>7</b>
<b>ABSTRACT.....</b>	<b>8</b>
<b>INTRODUCCIÓN.....</b>	<b>9</b>
<b>CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA .....</b>	<b>13</b>
1.1. DESCRIPCIÓN DE LA SITUACIÓN PROBLEMÁTICA.....	13
1.2. DESCRIPCIÓN DEL PROBLEMA.....	16
1.3. HIPÓTESIS.....	16
1.4. OBJETIVO DE LA INVESTIGACIÓN .....	17
1.4.1. OBJETIVO SUPERIOR.....	17
1.4.2. OBJETIVO PRINCIPAL .....	17
1.4.3. OBJETIVOS ESPECÍFICOS .....	17
1.5. JUSTIFICACIÓN .....	18
1.6. ALCANCES Y LIMITACIONES.....	19
1.6.1. ALCANCES DE LA INVESTIGACIÓN.....	19
1.6.2. LIMITACIONES DE LA INVESTIGACIÓN .....	19
1.7. METODOLOGÍA.....	20
<b>CAPÍTULO II: ANTECEDENTES .....</b>	<b>22</b>
2.1. SEGUIMIENTO DE PUNTOS.....	23
2.2. SEGUIMIENTO DE NÚCLEOS .....	23
2.3. SEGUIMIENTO DE SILUETAS .....	24
<b>CAPÍTULO III: MARCO TEÓRICO .....</b>	<b>26</b>
3.1. MARCO TEÓRICO CONCEPTUAL .....	26
3.1.1. TECNOLOGÍAS DE MONITOREO DE TRÁFICO .....	26
3.1.2. DURACIÓN DEL MONITOREO DE TRÁFICO.....	51
3.1.3. SELECCIÓN DE DISPOSITIVOS DE VOLUMEN DE TRÁFICO.....	56
3.1.4. SEGUIMIENTO DE OBJETOS .....	61
3.1.5. PROBLEMAS Y TENDENCIAS DEL SEGUIMIENTO DE OBJETOS .....	64
3.2. MARCO TEÓRICO INSTRUMENTAL .....	66
3.2.1. REPRESENTACIÓN DE OBJETOS DE ACUERDO A SU FORMA.....	66

3.2.2.	REPRESENTACIÓN DE OBJETOS DE ACUERDO A SU APARIENCIA.....	68
3.2.3.	SELECCIÓN DE CARACTERÍSTICAS .....	70
3.2.4.	DETECCIÓN DE OBJETOS.....	72
<b>CAPÍTULO IV: DESCRIPCIÓN DE DATOS DE ENTRADA.....</b>		<b>76</b>
4.1.	DESCRIPCIÓN GENERAL .....	76
4.2.	FUENTE DE DATOS.....	76
4.3.	DESCRIPCIÓN DE LOS DATOS.....	77
4.3.1.	SECUENCIAS DE VIDEO .....	77
4.3.2.	MEDICIONES.....	79
4.4.	RESUMEN DEL CAPÍTULO .....	83
<b>CAPÍTULO V: MODELO DE SOLUCIÓN .....</b>		<b>84</b>
5.1.	DESCRIPCIÓN GENERAL .....	84
5.2.	MODELO DE SOLUCIÓN DE ALTO NIVEL .....	84
5.3.	MODELO DE SOLUCIÓN DETALLADO .....	87
5.3.1.	PRE PROCESAMIENTO DEL VIDEO .....	87
5.3.2.	SEGMENTACIÓN DEL FONDO .....	88
5.3.3.	CONTEO DE VEHÍCULOS .....	92
5.4.	RESUMEN DEL CAPÍTULO .....	94
<b>CAPÍTULO VI: PRE PROCESAMIENTO DEL VIDEO .....</b>		<b>95</b>
6.1.	DESCRIPCIÓN GENERAL .....	95
6.2.	OBTENCIÓN DE PROPIEDADES DEL VIDEO .....	95
6.3.	CONVERSIÓN A ESCALA DE GRISES.....	96
6.4.	REDUCCIÓN DE DIMENSIONES .....	96
6.5.	RESUMEN DEL CAPÍTULO .....	97
<b>CAPÍTULO VII: DETERMINACIÓN DEL NIVEL DE MOVIMIENTO.....</b>		<b>98</b>
7.1.	DESCRIPCIÓN GENERAL .....	98
7.2.	PROCEDIMIENTO.....	98
7.3.	RESUMEN DEL CAPÍTULO .....	100
<b>CAPÍTULO VIII: ANÁLISIS DEL NIVEL DE MOVIMIENTO .....</b>		<b>101</b>
8.1.	DESCRIPCIÓN GENERAL .....	101
8.2.	PROCEDIMIENTO.....	101
8.3.	RESUMEN DEL CAPÍTULO .....	103

<b>CAPÍTULO IX: ACTUALIZACIÓN DEL MODELO DE FONDO .....</b>	<b>104</b>
9.1. DESCRIPCIÓN GENERAL .....	104
9.2. PROCEDIMIENTO.....	104
9.3. RESUMEN DEL CAPÍTULO .....	105
<b>CAPÍTULO X: SUBSTRACCIÓN DE FONDO .....</b>	<b>106</b>
10.1. DESCRIPCIÓN GENERAL .....	106
10.2. PROCEDIMIENTO.....	106
10.3. RESUMEN DEL CAPÍTULO .....	107
<b>CAPÍTULO XI: FILTRO MORFOLÓGICO .....</b>	<b>108</b>
11.1. DESCRIPCIÓN GENERAL .....	108
11.2. PROCEDIMIENTO.....	108
11.3. RESUMEN DEL CAPÍTULO .....	109
<b>CAPÍTULO XII: CONTEO DE VEHÍCULOS.....</b>	<b>110</b>
12.1. DESCRIPCIÓN GENERAL .....	110
12.2. TRAZADO DE LÍNEA IMAGINARIA .....	110
12.3. CONTEO DE VEHICULOS .....	111
12.4. RESUMEN DEL CAPÍTULO .....	114
<b>CAPÍTULO XIII: EXPERIMENTACIÓN .....</b>	<b>115</b>
13.1. DESCRIPCIÓN GENERAL .....	115
13.2. DISEÑO DEL EXPERIMENTO.....	115
13.3. VARIABLES INDEPENDIENTES Y DEPENDIENTES .....	116
13.4. RESULTADOS .....	120
13.5. RESUMEN DEL CAPÍTULO .....	122
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>123</b>
<b>GLOSARIO DE TÉRMINOS.....</b>	<b>127</b>
<b>BIBLIOGRAFÍA.....</b>	<b>130</b>

## ÍNDICE DE CUADROS

<b>Cuadro 2.1: Métodos de seguimiento de objetos</b> .....	25
<b>Cuadro 4.1: Características de las secuencias de video</b> .....	79
<b>Cuadro 4.2: Mediciones del volumen vehicular – Video 1</b> .....	80
<b>Cuadro 4.3: Mediciones del volumen vehicular – Video 2</b> .....	80
<b>Cuadro 4.4: Mediciones del volumen vehicular – Video 3</b> .....	81
<b>Cuadro 4.5: Mediciones del volumen vehicular – Video 4</b> .....	81
<b>Cuadro 4.6: Mediciones del volumen vehicular – Video 5</b> .....	82
<b>Cuadro 4.7: Total de vehículos</b> .....	82
<b>Cuadro 5.1: Variables del modelo de solución general</b> .....	86
<b>Cuadro 13.1: Ambiente de experimentación</b> .....	116
<b>Cuadro 13.2: Variables dependientes e independientes del modelo de solución</b> .....	117
<b>Cuadro 13.3: Parámetros del modelo de solución</b> .....	118
<b>Cuadro 13.4: Resultados del experimento – Parte 1</b> .....	120
<b>Cuadro 13.5: Resultados del experimento – Parte 2</b> .....	120
<b>Cuadro 13.6: Resultados del experimento – Parte 3</b> .....	121
<b>Cuadro 13.7: Resultados del experimento – Parte 4</b> .....	121
<b>Cuadro 13.8: Combinaciones de parámetros que minimizan el error</b> ....	122

## ÍNDICE DE FIGURAS

<b>Figura 1.1: Metodología Top-Down</b> .....	21
<b>Figura 2.1: Técnicas de seguimiento de objetos</b> .....	24
<b>Figura 3.1: Representación de objetos de acuerdo a su forma</b> .....	68
<b>Figura 4.1: Secuencias de video tomadas para el estudio</b> .....	78
<b>Figura 5.1: Modelo de solución general</b> .....	85
<b>Figura 5.2: Ejemplo de segmentación de fondo</b> .....	90
<b>Figura 5.3: Modelo de solución detallado</b> .....	93
<b>Figura 6.1: Conversión a escala de grises</b> .....	96
<b>Figura 6.2: Imagen escalada por un factor de 0.5</b> .....	97
<b>Figura 7.1: Determinación del nivel de movimiento</b> .....	99
<b>Figura 8.1: Análisis del movimiento</b> .....	102
<b>Figura 9.1: Resultados de la actualización del modelo de fondo</b> .....	105
<b>Figura 10.1 Substracción de fondo</b> .....	107
<b>Figura 11.1: Filtro morfológico</b> .....	109
<b>Figura 12.1: Línea Imaginaria</b> .....	111
<b>Figura 12.2 Conteo de vehículos</b> .....	112
<b>Figura 12.3 Conteo de vehículos en paralelo</b> .....	113

## **DESCRIPTORES TEMÁTICOS**

- **Monitoreo de Tráfico Vehicular**
- **Conteo de Volumen Vehicular**
- **Sensores Intrusivos y No Intrusivos**
- **Visión Computacional**
- **Representación de Objetos**
- **Seguimiento de Objetos**
- **Segmentación del Fondo**
- **Nivel del Movimiento**
- **Substracción del Fondo**
- **Filtro Morfológico**



## **RESUMEN**

En este trabajo se presenta un algoritmo eficaz para el conteo de vehículos en tiempo real. Este algoritmo está compuesto por tres procedimientos: pre procesamiento del video, en el cual se homogeniza tanto el color como las dimensiones del video; segmentación de fondo, el cual permite identificar los objetos en movimiento; y el seguimiento de objetos, a través del cual se realiza el conteo de vehículos. Asimismo, en base a este algoritmo se desarrolló el prototipo de un Sistema de Conteo Vehículos. De acuerdo a diversos experimentos se determinó que el margen de error promedio (MAPE) de la solución es de 9.9%.

## **ABSTRACT**

This work presents an effective real-time vehicle counting algorithm. The algorithm is composed by three main procedures: preprocessing, in which the video size and color is normalized; background segmentation, which identifies the moving objects; and object tracking, which finally counts and tracks the vehicles. Moreover, we have developed a Vehicle Counting System Prototype based in this algorithm. Several experiments show that the algorithm mean error (MAPE) is around 9.9%.

## INTRODUCCIÓN

La explotación de los recursos naturales y de los mercados, así como la conservación de un margen competitivo de una región sobre otra, están estrechamente vinculados con la calidad del sistema de transporte (N. Garber et al. [12]), entendiéndose como tal al movimiento de personas o bienes de una localidad a otra. La vitalidad económica de un área depende significativamente de la rapidez, el costo y la capacidad del transporte disponible. Así, se observa que las sociedades más desarrolladas e industrializadas, como Estados Unidos, Canadá y Japón, destacan por sus servicios de transporte de alta calidad. Sin un sistema de transporte adecuado, una región disminuye sus posibilidades de maximizar la ventaja comparativa que pueda tener (recursos naturales o humanos).

El equilibrio de un sistema de transporte es el resultado de las fuerzas del mercado (estado de la economía, competencia y precios del servicio), la acción gubernamental (reglamentos, subsidios) y la tecnología del transporte (velocidad, capacidad, confiabilidad). La continua variación de estas fuerzas tiene como consecuencia que el sistema de transporte se encuentre en un estado de constante flujo. Por ejemplo, en épocas de desempleo o de alza

de precios de combustible, la demanda del transporte (vehículos/día) tiende a disminuir.

Además, existen diversos problemas, como choques, dificultades de aparcamiento y congestión. Para reducir el impacto negativo es necesario recolectar información descriptiva del alcance de los problemas de tránsito e identificar su ubicación. Esta información se obtiene a través de diversos estudios y encuestas, entre los cuales están los estudios de volumen de tránsito. Dichos estudios se realizan para recolectar datos del número de vehículos que pasan por un punto determinado en un periodo de tiempo específico.

En la actualidad, los conteos vehiculares se llevan a cabo utilizando métodos intrusivos o no intrusivos, que requieren el uso de dispositivos especializados y de alto precio (revisar sección 3.1.1). La presente tesis de grado propone el uso de técnicas de seguimiento de objetos para el conteo de vehículos en tiempo real (ver capítulo 2). Las técnicas de seguimiento de objetos representan una alternativa no intrusiva que, además de tener un bajo costo, puede ser utilizada simultáneamente para fines adicionales como el relevamiento de tiempos de espera, video vigilancia, entre otros.

En el primer capítulo se realiza una descripción de la situación problemática y el problema, así como los alcances y limitaciones y la metodología empleada para el desarrollo del algoritmo de solución.

El capítulo 2 resume los principales métodos de seguimiento de objetos, así como los trabajos representativos realizados en la actualidad.

El capítulo 3 presenta una revisión de la bibliografía y los conceptos necesarios para entender el tema abordado.

En el capítulo 4 se presenta la fuente de datos y se describe los datos utilizados en esta investigación.

En el capítulo 5 se detalla el modelo de solución, el cual esquematiza las sucesivas etapas utilizadas para abordar el problema planteado.

En el capítulo 6 se hace una explicación del procedimiento de pre procesamiento del video, el cual es el primer paso en el método utilizado.

En los capítulos 7 al 11 se detalla el procedimiento de segmentación de fondo, el cual consta de las etapas de determinación y análisis del nivel de movimiento, actualización del modelo de fondo, sustracción de fondo y filtro morfológico.

En el capítulo 12 se hace una revisión de la última etapa del modelo de solución, la cual realiza el conteo de los vehículos detectados en los procedimientos anteriores.

En el capítulo 13 se hace una descripción de los experimentos llevados

a cabo para calcular el porcentaje promedio de error absoluto.

Finalmente, se presentan las conclusiones y recomendaciones de la presente investigación.

# **CAPÍTULO I**

## **PLANTEAMIENTO DEL PROBLEMA**

### **1.1. DESCRIPCIÓN DE LA SITUACIÓN PROBLEMÁTICA**

Según PROTRANSPORTE [16], la ciudad de Lima Metropolitana ha desarrollado un sistema de movilidad caótico, con exceso de oferta de vehículos, congestión y altos niveles de contaminación y accidentalidad. Entre las principales causas del actual problema de transporte de la ciudad podemos mencionar la baja exigencia de servicios de mejor calidad, inexistencia de una cultura de movilidad urbana, el exceso de oferta que promueve la competencia por pasajeros, importación de vehículos usados que contribuye a la obsolescencia del parque vehicular, incremento de taxis e inexistencia de una adecuada planificación del transporte urbano.

Para lograr alternativas efectivas que solucionen los problemas presentados es necesario el desarrollo de modelos de demanda de transporte o la calibración de modelos ya existentes, es decir, su actualización debido a variaciones en las condiciones del mercado. Existen numerosos estudios y planes viales propuestos para la ciudad de Lima, entre los cuales podemos

mencionar el estudio realizado por el Consorcio Getinsa – Taryet – Geoconsult, por encargo de FONAM y PROTRANSPORTE (FONAM et al. [14]), donde se calibró y validó el modelo de simulación de transporte para Lima, el cual es un modelo de demanda desarrollado por PROINVERSION en el año 2007 para la concesión del Tren Eléctrico (PROTRANSPORTE [15]). Esta actualización fue necesaria debido a la aparición del sistema de transporte público denominado Metropolitano, además de otras intervenciones en la red vial que produjeron variaciones en el tráfico de Lima.

Los modelos de transporte, debidamente calibrados y validados, se pueden tomar como base para simular escenarios en los años siguientes al estudio y poder realizar una adecuada planificación de movilidad urbana que promueva la movilización económica, rápida y masiva de la población, eleve el nivel de confiabilidad del servicio (horarios, frecuencias, paraderos, cumplimiento de la ruta, etc.), mejore las condiciones de prestación del servicio (limpieza, confort, facilidades a discapacitados, etc.), prevenga accidentes de tránsito, reduzca la contaminación ambiental, entre otras cosas.

Existe una variedad de análisis de campo que deben ser llevados a cabo durante el desarrollo o calibración de los modelos mencionados. Por ejemplo, tenemos los estudios de volumen vehicular, frecuencia de paso de rutas de transporte público, encuestas de selección de alternativas basadas en tiempos de espera, caminata y a bordo, encuestas de selección de alternativas basadas en tiempo a bordo, tarifa y número de transbordo, etc. Generalmente, entre los primeros estudios de campo que se realizan están los de volumen o aforo



vehicular, los cuales consisten en determinar cuántos vehículos pasaron por un punto específico de la vía en un intervalo de tiempo dado. Actualmente, este tipo de análisis emplea métodos no automáticos (a través de contómetros o instrumentos para el conteo manual) o métodos automáticos que requieren la instalación de sensores en los lugares donde se obtendrá datos del tráfico vehicular. Los sensores utilizados pueden ser intrusivos (se instalan sobre la superficie de la vía o debajo de la misma) o no intrusivos.

Existen proyectos de movilidad urbana de Lima, como el llevado a cabo por Pait Consultores [17] en el distrito de Chorrillos (Lima), donde el análisis de volumen vehicular se ha realizado de manera manual, método que tiene una mayor probabilidad de presentar errores humanos. En dicho estudio, los encuestadores hicieron uso de contómetros y los resultados fueron registrados en formatos específicos según el tipo de aforo.

Por otro lado, el uso de la mayoría de sensores tiene algunas desventajas. Por ejemplo, tal como indica Pornpanomchai et al. [6], se utilizan sólo para el conteo en tiempo real, es decir, el usuario no tiene la posibilidad de ingresar algún input (como videos) al sensor para que cuente la cantidad de vehículos. Además los dispositivos requeridos tienen un alto costo y es necesaria la presencia de personal capacitado para su uso.

## 1.2. DESCRIPCIÓN DEL PROBLEMA

En Lima Metropolitana, el conteo de volumen vehicular se realiza de manera manual (por lo que se incrementa la posibilidad de ocurrencia de errores humanos) y/o con la utilización de dispositivos especializados, cuya adquisición, instalación, calibración y mantenimiento requieren de una gran cantidad de recursos humanos y tecnológicos. Por ello, es necesario el desarrollo de un método de conteo de vehículos que haga uso eficiente de los recursos humanos y tecnológicos y que tenga un bajo nivel de error.

## 1.3. HIPÓTESIS

Con las técnicas de visión computacional se puede desarrollar un método de conteo de volumen vehicular en tiempo real para Lima Metropolitana, el cual necesite un bajo nivel de inversión tanto en recursos humanos como tecnológicos y que tenga un error promedio absoluto (MAPE) menor o igual a 10%.

## **1.4. OBJETIVO DE LA INVESTIGACIÓN**

### **1.4.1. OBJETIVO SUPERIOR**

Diseñar e implementar un método de conteo de vehículos cuya efectividad contribuya en la solución de los problemas del sistema de transporte de Lima Metropolitana.

### **1.4.2. OBJETIVO PRINCIPAL**

Diseñar e implementar un método efectivo de conteo de vehículos en tiempo real utilizando visión computacional, teniendo en consideración los escenarios más comunes de Lima Metropolitana, tales como cambios ligeros de iluminación o en el segundo plano de las imágenes y cuyo error promedio absoluto (MAPE) sea menor o igual a 10%.

### **1.4.3. OBJETIVOS ESPECÍFICOS**

- Diseñar un método de solución para el conteo de vehículos en tiempo real.
- Determinar los parámetros de los cuales depende el modelo de solución.
- Diseñar experimentos que permitan medir la efectividad del método planteado.
- Determinar la combinación de parámetros que minimice el error medio

absoluto (MAPE) de la solución.

## 1.5. JUSTIFICACIÓN

Los siguientes factores motivan el desarrollo de la presente investigación:

- **Costo estimado.** Según Pornpanomchai et al. [6], el uso de sensores especializados en el monitoreo de tráfico conlleva un alto costo, mientras que la adquisición e instalación de una cámara de video requiere menos inversión económica. Además, muchos usuarios ya cuentan con cámaras de seguridad instaladas en sus oficinas o edificios, por lo que se puede gestionar el uso de dichas cámaras para los estudios de aforo vehicular. Además, el software que procesa los videos no requiere ser ejecutado en un hardware especial, por lo que con una típica cámara de video y una computadora personal se puede implementar el modelo de solución propuesto en la presente tesis.
- **Innovación.** Las técnicas de seguimiento de objetos funcionan en casos de retrocesos y cambios de carril. Además, dichas técnicas se encuentran en estado de investigación activa y continuamente se proponen nuevas soluciones que mejoran los problemas existentes (A. Yilmaz et al. [1]).

## 1.6. ALCANCES Y LIMITACIONES

### 1.6.1. ALCANCES DE LA INVESTIGACIÓN

Si bien el método de solución puede aplicarse en diversos tipos de vías, zonas geográficas y condiciones de visibilidad, se ha delimitado el alcance considerando las siguientes restricciones:

- Zona Geográfica: Lima Metropolitana.
- Tiempo: De 10 a.m. a 3 p.m., en condiciones de máxima visibilidad.

### 1.6.2. LIMITACIONES DE LA INVESTIGACIÓN

El método de solución de esta investigación tiene las siguientes limitaciones:

- Se asume que no hay cambios bruscos en la iluminación durante la captura del video.
- La cámara utilizada para la grabación de videos no presenta movimiento.
- Se considera que los objetos del fondo son estacionarios, es decir, no tienen movimientos.
- Se asume que no hay variaciones bruscas en la intensidad o color de los objetos que forman parte del fondo.
- Se consideran únicamente vías unidireccionales.

- No se contempla obstáculos, como puentes o túneles, que impidan momentáneamente la visibilidad de los vehículos.

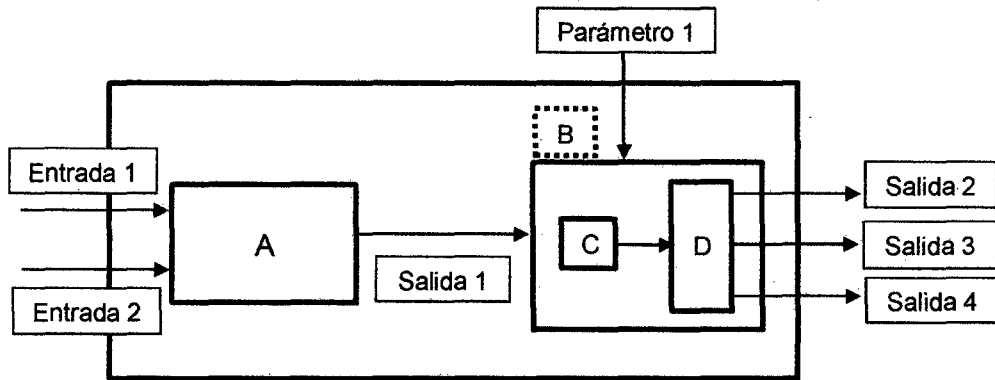
## 1.7. METODOLOGÍA

En el desarrollo de la presente investigación se hizo uso de la metodología Top-Down o Arriba-Abajo planteada por H. D. Mills [20]. En dicho enfoque, el modelo de solución se diseña inicialmente con especificaciones de alto nivel y luego se procede a refinamientos sucesivos a través de especificaciones de menor jerarquía hasta llegar al nivel del código fuente.

Este tipo de metodología enfatiza la planificación y un completo entendimiento del modelo global, el cual se divide en etapas y cada etapa a su vez se divide en sub etapas de menor nivel o jerarquía. No se puede dar inicio a la codificación hasta que se haya llegado al suficiente nivel de detalle.

Cada unidad del modelo se puede esquematizar como una caja negra que toma como entrada ciertas variables y/o parámetros y produce como resultado otras variables que dependen de los datos de entrada (ver Figura 1.1).

**Figura 1.1: Metodología Top-Down**



Fuente: Elaboración propia.

## **CAPÍTULO II**

### **ANTECEDENTES**

En esta sección se describe los principales métodos existentes de seguimiento de objetos, disciplina que se encarga de la estimación de la trayectoria de objetos en un determinado video. El objetivo del seguimiento de objetos es generar la trayectoria de un objeto a lo largo del tiempo localizando su posición en cada fotograma del video.

Las tareas de detectar el objeto y establecer la correspondencia entre las instancias del objeto en los fotogramas pueden ser llevadas a cabo separadas o juntas. En el primer caso, las posibles regiones de objetos en cada fotograma se obtienen mediante un algoritmo de detección de objetos, y luego se realiza la correspondencia entre los objetos en los diferentes fotogramas del video. En el otro caso, la región del objeto y la correspondencia son estimadas conjuntamente mediante actualizaciones iterativas de la ubicación del objeto e información de la región obtenida de fotogramas previos. En cualquiera de los enfoques de seguimiento, los objetos se representan usando modelos de forma o apariencia (revisar sección 3.2).



El modelo seleccionado para representar la forma del objeto limita el tipo de movimiento o deformación que puede existir. Por ejemplo, si un objeto se representa como un punto, entonces sólo se puede usar un modelo traslacional. Si se usa una representación de forma geométrica, como una elipse, se puede usar modelos de movimiento paramétrico. Estas representaciones pueden aproximar el movimiento de objetos rígidos. Para los objetos no rígidos, la representación más descriptiva es la de silueta o contorno, y se puede usar modelos paramétricos o no paramétricos para especificar su movimiento.

Los principales tipos de algoritmos de seguimiento son:

## 2.1. SEGUIMIENTO DE PUNTOS

Los objetos detectados en fotogramas consecutivos son representados por puntos, y la asociación de los puntos se basa en el estado previo de los objetos, que puede incluir tanto posición como movimiento. Este enfoque requiere de un mecanismo externo para detectar los objetos en cada fotograma.

## 2.2. SEGUIMIENTO DE NÚCLEOS

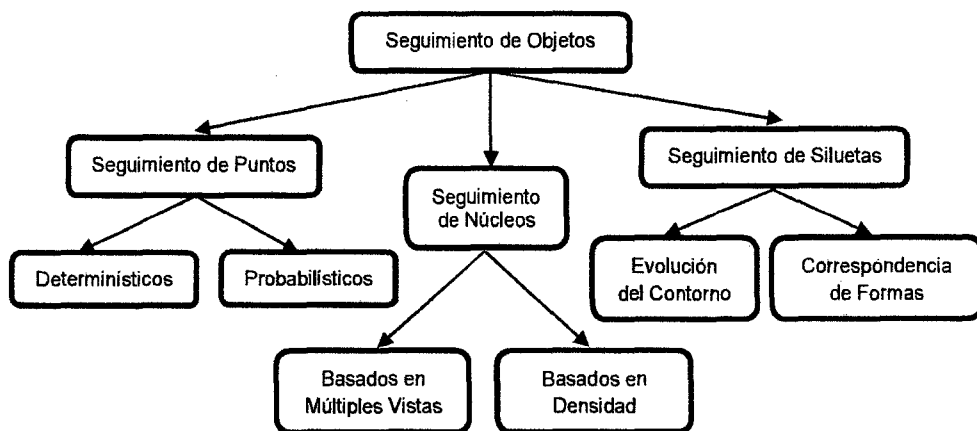
El núcleo se refiere a la forma y apariencia del objeto. Por ejemplo, el núcleo puede ser una plantilla rectangular o una forma elíptica con un histograma asociado. El seguimiento se realiza calculando el movimiento del núcleo en fotogramas consecutivos. Generalmente, el movimiento se encuentra en la forma de transformaciones paramétricas, como traslación o rotación.

### 2.3. SEGUIMIENTO DE SILUETAS

El seguimiento es realizado estimando la región del objeto en cada fotograma y utilizando la información codificada en esta región. Esta información puede ser de densidad de apariencia y modelos de forma. Dados los modelos de objetos, se realiza el seguimiento de las siluetas mediante la correspondencia de formas o evolución del contorno.

En la Figura 2.1 se esquematiza los tipos de seguimiento de objetos mencionados.

**Figura 2.1: Técnicas de seguimiento de objetos.**



Fuente: Yilmaz, Javed y Shah (2006).

En el Cuadro 2.1 se presenta el trabajo representativo para cada categoría de métodos de seguimiento de objetos (A. Yilmaz et al. [1]).

**Cuadro 2.1: Métodos de seguimiento de objetos**

<b>Categoría</b>	<b>Trabajo</b>		
<b>Seguimiento de puntos</b>			
<b>Métodos determinísticos</b>	MGE Tracker	Salari y Sethi	1990
	GOA Tracker	Veenman et al.	2001
<b>Métodos estadísticos</b>	Kalman Filter	Broida y Chellappa	1986
	JPDAF	Bar-Shalom y Foreman	1988
	PMHT	Streit y Luginbuhl	1994
<b>Seguimiento de núcleos</b>			
<b>Modelos basados en densidad</b>	Mean-shift	Comaniciu et al.	2003
	KLT	Shi y Tomasi	1994
	Layering	Tao et al.	2002
<b>Modelos de múltiples vistas</b>	Eigentracking	Black y Jepson	1998
<b>Seguimiento de siluetas</b>			
<b>Evolución de contorno</b>	State space models	Isard y Blake	1998
	Variational methods	Bertalmio et al.	2000
	Heuristic methods	Ronfard	1994
<b>Correspondencia de formas</b>	Hausdorff	Huttenlocher et al.	1993
	Hough transform	Sato y Aggarwal	2004
	Histogram	Kang et al.	2004

Fuente: Yilmaz, Javed y Shah (2006)

## **CAPÍTULO III**

### **MARCO TEÓRICO**

#### **3.1. MARCO TEÓRICO CONCEPTUAL**

##### **3.1.1. TECNOLOGÍAS DE MONITOREO DE TRÁFICO**

Según J. Wu et al. [9], el conteo vehicular y la detección de velocidades se constituyen en los dos indicadores más importantes para describir la condición del tráfico.

Las estadísticas del tráfico que se proveen a los usuarios finales son tan precisas y confiables como los datos recolectados en el campo. En esta sección se presenta diversas alternativas utilizadas para recolectar datos de volumen de tráfico, clasificación, peso y velocidad. Asimismo, se establece la diferencia entre las tecnologías empleadas en el monitoreo permanente y continuo del tráfico, de aquellas utilizadas en sesiones de recolección de datos de corta duración (AASHTO et al. [11]).

No sólo existe una gran variedad de sensores para observar el tráfico,

sino que diferentes vendedores de equipos frecuentemente utilizan diferentes algoritmos para convertir los datos enviados por el sensor en estadísticas de tráfico. El grado de eficacia de una tecnología específica en un esfuerzo de recolección de datos depende de las características y capacidades físicas del sensor, del ambiente donde éste opera y qué tan robusto es el algoritmo empleado por el dispositivo electrónico que interpreta las señales que provee el sensor.

La complejidad de estos factores significa que no existe una simple respuesta a la interrogante de qué tecnología es la que mejor funciona. La diferencia en rendimiento de un dispositivo respecto de otro dentro de la misma tecnología es más significativa que la diferencia entre una tecnología y otra (AASHTO et al. [11]).

Además de las diferencias entre implementaciones específicas de distintos vendedores, cada tecnología tiene sus propias fortalezas y debilidades, cada una funciona mejor bajo circunstancias específicas. Una tecnología puede ser poco confiable en las circunstancias equivocadas o cuando existe una mala instalación, calibrado o mantenimiento. Algunas tecnologías funcionan bien sólo cuando existe libre flujo de tráfico, mientras que otras pueden operar satisfactoriamente en cualquier condición, pero su costo es más alto y requieren un sistema de sensores extenso y permanentemente instalado.

Predecir el rendimiento de los equipos se torna aún más complicado debido a que cada vendedor implementa de manera muy diferente una

tecnología dada. Los equipos basados en la misma tecnología de diferentes vendedores pueden utilizar diferentes algoritmos para procesar las señales y, por ende, tomar los mismos datos del sensor y producir estadísticas de tráfico diferentes. Incluso los distintos modelos de equipos que ofrece un mismo vendedor pueden tener un software interno significativamente diferente, lo que resulta en grandes diferencias en los datos recolectados.

Además de diferir en costos, los diversos equipos ofrecidos por vendedores distintos pueden diferenciarse en la cantidad de datos recolectados que puede almacenar el equipo electrónico, la naturaleza en que dichos datos son almacenados, recuperados y resumidos, cómo se realizan las conexiones con el equipo electrónico y el tipo y calidad de soporte que el vendedor provee. Todos estos factores deben ser tomados en cuenta al momento de decidir qué dispositivo utilizar.

Las acciones más importantes necesarias para recolectar datos confiables son:

- Probar los equipos de recolección de datos bajo las condiciones de estudio esperadas antes de comprarlos. Esto se debe realizar con el fin de asegurarse que los dispositivos funcionarán bajo las condiciones en que serán usados.
- Seguir cuidadosamente las instrucciones de instalación de los equipos.
- Calibrar y probar los dispositivos instalados para asegurar que funcionen

según lo esperado.

- Revisar periódicamente el rendimiento de los equipos.

Estas acciones podrían incrementar el costo de las actividades de recolección de datos, pero asegurarán la calidad de la información obtenida. Además, se reducen gastos en la recolección de datos erróneos, disminuyendo el costo de recolectar datos confiables. Es mejor obtener menos datos en menos ubicaciones, pero con la certeza que sean válidos, que obtener datos de muchas ubicaciones y no utilizarlos debido a que los equipos no funcionaron correctamente.

Las tecnologías de monitoreo de tráfico se dividen en dos grandes categorías: intrusivas, es decir, aquellas que requieren la instalación de un sensor en la superficie de la calzada; y no intrusivas, aquellas que son ubicadas por encima, al lado o debajo de la calzada, de tal manera que no interfieren con el flujo del tráfico (AASHTO et al. [11]).

En algunas instalaciones, los detectores no intrusivos requieren el cierre del carril para su ubicación, aun incluso cuando el sensor en sí mismo o su instalación no interfieren con la calzada o impiden el flujo del tráfico. Esto ocurre frecuentemente cuando los sensores deben ser instalados en ubicaciones por encima de la calzada, por ejemplo, cuando cuelgan de un puente. Las normas de seguridad frecuentemente requieren que el carril debajo de las actividades de instalación o mantenimiento sea cerrado. Esto quiere decir que algunos dispositivos considerados como no intrusivos no deben ser seleccionados si el

carril no puede ser cerrado para su instalación o mantenimiento.

#### 3.1.1.1. TECNOLOGÍAS INTRUSIVAS

La mayoría de estudios tradicionales de tráfico requieren el uso de tecnologías intrusivas. Los diferentes dispositivos de este tipo difieren unos de otros en los atributos que miden y en la manera cómo lo hacen. Asimismo, también existen diferencias en la velocidad con que responden al tráfico. Aquellos con mayor velocidad tienen menos probabilidad de presentar errores en condiciones de alto volumen, densidad o velocidad. Por otro lado, aquellos que tienen tiempo de respuesta más lento (como los tubos de carretera) pueden dejar de contar vehículos cuando estos se mueven rápidamente o están poco espaciados entre sí. Las tecnologías intrusivas más importantes son:

- TUBOS DE CARRETERA

Los tubos de carretera frecuentemente son los sensores más utilizados para conteos de corta duración. Son tubos huecos, tradicionalmente hechos de caucho, unidos a un interruptor de aire. Cuando un eje o rueda pasa sobre un tubo, éste es comprimido forzando el aire a través del interruptor, el cual se cierra y envía una señal eléctrica al dispositivo electrónico de recolección de datos, indicando el paso de un eje.

El tubo de carretera tradicional consiste en un único sensor que puede ser instalado en múltiples carriles. Cuando esto ocurre, el sensor detecta el



paso de ejes en todos los carriles donde es ubicado. La versión tradicional del tubo de carretera no podía diferenciar entre el pulso de aire generado por el tráfico en diferentes carriles, por lo que se ha desarrollado nuevas versiones que sí pueden hacer esta diferencia, generando pulsos de aire sólo cuando el tráfico se mueve a través de una sección específica del sensor. Este diseño permite que los tubos de carretera sean instalados en múltiples carriles pero que observen el tráfico sólo en el carril deseado.

Sin embargo, pueden ocurrir errores si el tubo no está correctamente posicionado en la carretera. Es decir, si la sección activa del sensor se ubica parcialmente en el segundo carril, éste detectará algún tráfico en dicho carril, además del tráfico en el carril para el cual fue instalado.

Cuando los tubos de carretera se utilizan para la clasificación de vehículos, es muy importante que los dos tubos que monitorean cada carril sean de la misma longitud. Si son de diferente longitud, el tiempo requerido para que el pulso de aire alcance el dispositivo de recolección de datos será diferente para cada sensor y las estimaciones serán diferentes.

Por otro lado, los tubos de carretera son lentos debido al caucho, el cual requiere cierto tiempo para recuperarse antes que se produzca un segundo pulso de aire bien definido. Esto significa que los tubos de carretera pueden dejar de contar vehículos en tráfico de alto volumen, cuando los ejes pasan por el sensor en una sucesión rápida.

- **LAZOS INDUCTIVOS**

Los lazos inductivos probablemente son los detectores permanentes de vehículos más usados en los Estados Unidos (AASHTO et al. [11]). Estos dispositivos consisten en lazos de alambres instalados en la superficie de la carretera, cuya forma y tamaño dependen de la aplicación que se les va a dar. Se establece una pequeña corriente eléctrica en el lazo y se monitorea la inductancia del mismo. Cuando un vehículo pasa sobre el lazo, el metal de dicho vehículo produce un cambio en la inductancia del alambre, el cual es detectado por un sensor.

Los lazos inductivos difieren de los tubos de carretera y otros sensores en que miden la cantidad de vehículos, no la cantidad de ejes. Esto aumenta su precisión como contadores de volumen, ya que no es necesaria ninguna conversión del número de ejes observado al número de vehículos. Sin embargo, los detectores de lazos en realidad no miden vehículos, sino la presencia de un metal. Por ende, una limitación de este tipo de tecnología es que su precisión es menor cuando se trata de contar vehículos con poco o nada de metal, por ejemplo motocicletas. Por otro lado, si la sensibilidad del dispositivo se establece lo suficientemente alta para reconocer motocicletas, entonces es probable que vehículos con grandes cantidades de metal activen el sensor aun cuando se encuentren en un carril distinto del carril donde se está realizando el estudio.

La mayoría de instalaciones de lazos inductivos son permanentes. Éstos se colocan en ranuras en la superficie del pavimento y, por lo general, tienen un tiempo de vida largo. Las malas condiciones del pavimento o los filtros de agua pueden afectar el rendimiento de estos dispositivos. Además, para funcionar correctamente, los lazos inductivos necesitan un sensor de inductancia de alta calidad.

- **DETECTORES MAGNÉTICOS**

Los detectores magnéticos utilizan una idea similar a los lazos inductivos para detectar la presencia de vehículos. Estos dispositivos miden el cambio en el flujo magnético creado cuando un vehículo pasa por la zona de detección. Al igual que los lazos, los detectores magnéticos captan la presencia del metal de los vehículos y, por ende, cuentan vehículos en vez de ejes. Asimismo, pueden presentar problemas de precisión cuando se trata de contar vehículos que poseen bajas cantidades de metal.

Este tipo de tecnología se puede instalar sobre o debajo del pavimento. Los detectores magnéticos que se colocan debajo del pavimento son mejor clasificados como “no intrusivos”, a pesar que su instalación puede requerir el cierre de un carril.

- **SENSORES PIEZOELÉCTRICOS**

Algunos materiales presentan una propiedad inusual llamada piezoelectricidad, es decir, presión-electricidad. La aplicación de fuerzas mecánicas externas en estos materiales induce la polarización y establece un campo eléctrico (W. Callister [18]). Si se invierte el signo de la fuerza externa, es decir, de tensión a compresión o viceversa, la dirección del campo se invierte. La propiedad piezoeléctrica, descubierta por Antonie Becquerel en 1819, sólo se manifiesta en algunos minerales como el cuarzo, la turmalina, la blenda, etc. y es reversible (A. Castejón et al. [19]).

Los materiales piezoeléctricos tienen una variedad de aplicaciones. Se utilizan en los transductores, dispositivos que se encargan de convertir la energía eléctrica en deformaciones mecánicas o viceversa. Encontramos elementos piezoeléctricos en encendedores de gas, micrófonos, auriculares de teléfonos móviles, agujas de tocadiscos, entre otros. En estos últimos, a medida que la aguja pasa por el disco, se produce una tensión en el material piezoeléctrico localizado en el cabezal, lo que genera una señal eléctrica que se dirige al altavoz.

Otra de las aplicaciones de los materiales piezoeléctricos son los sensores utilizados en estudios de tráfico. Un sensor piezoeléctrico es un dispositivo de detección de ejes que es instalado, por lo general de forma permanente, sobre o bajo la superficie de la autopista. Puede registrar datos de volumen de tráfico, clasificación, velocidad y peso en movimiento de los

vehículos. Los sensores piezoeléctricos funcionan produciendo una señal (voltaje y corriente) cuando un eje o rueda entra en contacto con ellos en la autopista. El tamaño de señal producida está en función de la fuerza aplicada al sensor (AASHTO et al. [11]). Por ende, los sensores piezoeléctricos proveen una señal más robusta que los tubos de carretera o los interruptores de cierre, y la señal puede ser utilizada para estimar el peso del eje o rueda.

Estos dispositivos tienden a ser más costosos que otros sensores de ejes convencionales, pero son también más robustos y pueden ser utilizados en más que la sola detección de ejes. Además, tienen una rápida velocidad de detección y, por ende, se debe tener especial cuidado en su instalación. Si no son colocados perpendicularmente al flujo del tráfico (de manera que las dos ruedas de un eje golpeen simultáneamente el sensor) son más propensos a registrar el mismo eje más de una vez. Sin embargo, la velocidad de estos dispositivos les da mayor precisión en la recolección de datos en carreteras de alta velocidad.

Otra de las debilidades del sensor piezoeléctrico es su diseño delgado. Si bien este tipo de diseño reduce el impacto en el pavimento y acelera la instalación, también ocasiona que la rueda que está siendo pesada no se apoye totalmente en el sensor. Cuando el neumático está en contacto con el sensor, una porción es soportada por el pavimento que rodea al dispositivo. Como resultado, el peso realmente aplicado al sensor piezoeléctrico depende de la respuesta estructural del pavimento. Mientras más rígido sea, menos peso se aplicará al dispositivo.

De la misma manera, la naturaleza delgada del sensor aunada a patrones inusuales en los neumáticos o presiones anormales origina que algunos ejes no sean medidos con precisión.

- CABLES DE FIBRA ÓPTICA

Los sensores de fibra óptica surgieron como alternativas a las tecnologías piezoeléctricas. En estos dispositivos, la luz pasa a través de un sensor (de la misma manera que la corriente pasa por los lazos inductivos) y cuando un neumático pasa sobre el cable de fibra óptica, lo comprime, ocasionando que la cantidad de luz a través del sensor cambie. Este cambio en la intensidad de la luz es medido por un detector, el cual indica el paso de un eje o neumático.

Los sensores de fibra óptica son sensores de ejes con alta velocidad. Es por ello que requieren ser instalados de manera perpendicular al flujo del tráfico para evitar que registren el mismo eje dos veces. La longitud de estos sensores depende de la cantidad de carriles que se va a monitorear.

Estos sensores pueden ser ubicados sobre del pavimento o en ranuras cortadas dentro del mismo, las cuales son más pequeñas que las ranuras necesarias para los sensores piezoeléctricos.

- **SENSORES DE CAPACITANCIA**

Los sensores de capacitancia consisten en dos o más conductores eléctricos separados por un material dieléctrico (no conductor) y que transportan cargas opuestas. Cuando un eje o neumático pasa sobre el sensor, la fuerza aplicada reduce la distancia entre los dos conductores y la capacitancia del sistema eléctrico aumenta. Mientras mayor sea el peso aplicado al sensor, mayor será el cambio en la capacitancia. El dispositivo electrónico unido al sensor estima el peso del eje en base a la señal producida. El costo de los sensores de capacitancia suele ser mayor que otros sensores.

### 3.1.1.2. TECNOLOGÍAS NO INTRUSIVAS

Las tecnologías intrusivas son las que se usan más frecuentemente en el monitoreo del tráfico, pero requieren que los sensores sean ubicados dentro o sobre el pavimento, lo cual trae desventajas significativas. Entre dichas desventajas tenemos que, con los sensores intrusivos, el personal debe trabajar en situaciones peligrosas, por ejemplo en carriles con alto flujo de tráfico. Otra desventaja es que muchas veces el flujo de tráfico se debe interrumpir para instalar los sensores en el pavimento.

En consecuencia, los investigadores han desarrollado una variedad de nuevas tecnologías de recolección de datos de tráfico que no requieren acceso a los carriles de la carretera. Estos son los llamados sensores no intrusivos, los cuales observan el flujo de tráfico desde arriba o a un lado de la carretera. Esto

permite que la recolección de datos se pueda realizar sin el cierre de los carriles o exponiendo al personal a condiciones peligrosas.

Sin embargo, las tecnologías no intrusivas también tienen sus propias desventajas (AASHTO et al. [11]). La primera es que la mayoría no observa los ejes de los vehículos, lo cual implica que la clasificación de los mismos no puede estar basada en el número y ubicación de ejes. Otra desventaja es que cuando el tráfico se observa desde arriba o a un lado de la carretera, algunos vehículos pueden quedar escondidos de la vista del sensor por otros vehículos. Este problema se conoce como oclusión y es frecuente en las tecnologías intrusivas. Por otro lado, cada tecnología tiene sus propias fortalezas y debilidades, las cuales la hacen trabajar bien bajo algunas condiciones y mal bajo otras. Por ejemplo, los sensores basados en imágenes de video no funcionan bien en condiciones de baja visibilidad, como tormentas; mientras que las tecnologías acústicas tienen dificultades en altos niveles de ruido ambiental.

Entre las principales tecnologías no intrusivas tenemos:

- **SISTEMAS DE DETECCIÓN DE IMÁGENES DE VIDEO**

Se trata de sistemas basados en cámaras de video que utilizan tecnologías de procesamiento digital de imágenes para extraer información de tráfico. Por lo general, los sistemas basados en videos ofrecen una serie de ventajas comparados con otras técnicas, tales como mayor información de tráfico, son fácilmente instalables, etc. Sin embargo, su eficacia se ve



influenciada por factores como lluvias, variaciones en la iluminación, sombras en movimiento, entre otros (M. Lei et al. [7]). Es por ello que se debe incluir alguna protección contra las precipitaciones y la luz solar. Además, se debe establecer la zona de detección para evitar contar los mismos vehículos varias veces. Existen soluciones que ofrecen una variedad de funciones para minimizar el conteo doble o las detecciones falsas, pero estos factores incrementan la complejidad de la configuración. Por otro lado, los principales problemas de mantenimiento son limpiar los lentes de la cámara y asegurar las conexiones de los cables.

Existen tres tecnologías básicas de análisis de imágenes de video (AASHTO et al. [11]): sistemas de línea de viaje, sistemas de seguimiento de video y sistemas de extracción de atributos. Todos estos sistemas comienzan con imágenes de cámara digital e identifican grupos de píxeles que representan un vehículo.

- Sistemas de línea de viaje

Los sistemas de línea de viaje se comportan como lazos inductivos. Producen las mismas estadísticas básicas, es decir, información de volumen y ocupación de carril. En esta tecnología, los "lazos virtuales" o zonas de detección son ubicados en la imagen de la carretera obtenida por la cámara. Cuando se observa que los vehículos cruzan la zona de detección, el "lazo" es "activado" y se realiza el conteo y se hace posible detectar el tiempo que el lazo es "ocupado". Si se ubica dos zonas de detección en serie separadas por una

distancia conocida se logra calcular la velocidad de los vehículos (la distancia entre las líneas de viaje dividida entre el tiempo que le toma al objeto en moverse de la primera a la segunda línea). Esto también hace posible estimar la longitud del objeto en observación (dividiendo la velocidad del vehículo entre el tiempo en que la zona de detección es ocupada).

En la imagen de la cámara se puede establecer múltiples zonas de detección, por lo que los sistemas de imágenes de video basados en líneas de viaje pueden monitorear más de un carril de tráfico desde una misma cámara, es decir, una sola cámara puede realizar la tarea de recolección de datos de varios lazos inductivos reales (AASHTO et al. [11]).

#### – Sistemas de seguimiento de vehículos

Las técnicas de seguimiento de video emplean algoritmos para identificar y seguir vehículos cuando pasan por el campo de observación. Los grupos de píxeles asociados con un vehículo específico son seguidos imagen tras imagen en el video. El seguimiento de objetos específicos junto con el conocimiento de la imagen de fondo (ubicación de los carriles, longitud de la parte de la carretera observada por la imagen, distancias entre puntos de referencia específicos) permite determinar automáticamente atributos del tráfico, como la velocidad de los vehículos y el número y ubicación de cambios en los carriles.

El modelo solución planteado en la presente investigación (véase capítulo 5) se basa en el método de seguimiento de objetos de Tao Yang et al. [2].

- Sistemas de extracción de atributos

Este tipo de técnica examina detalladamente los píxeles de la imagen asociados con un objeto e intenta determinar características específicas de los vehículos, así como lo hace el cerebro humano. Mientras que los sistemas de línea de viaje sólo pueden clasificar vehículos basándose en la longitud total de los mismos, los sistemas de extracción de atributos intentan clasificar los mismos basándose en los atributos obtenidos. Por ejemplo, un bus escolar es observado por un sistema de líneas de viaje y clasificado como un vehículo de “longitud de 40 pies”, mientras que un sistema de extracción de atributos lo puede clasificar como un bus escolar propiamente dicho, dados su altura, longitud, ancho y color.

- Calidad de las imágenes

Los sistemas de procesamiento de imágenes, independientemente de la tecnología de análisis que utilicen, son tan buenos como la calidad de las imágenes, las cuales son el punto de partida para los cálculos. Las imágenes de video, por su parte, dependen de la eficiencia de la cámara para observar el flujo del tráfico. Asimismo, existen muchos factores que influyen en qué tan bien la cámara observa el tráfico, por ejemplo:

- Calidad de los lentes de la cámara.
- Cambios en las condiciones de iluminación, particularmente en el

amanecer y atardecer.

- Falta de iluminación.
- Ángulo, altura y posición de la cámara.
- Condiciones climáticas adversas.
- Dirección del flujo de tráfico relativo a la cámara.
- Movimiento de la cámara.

Generalmente, la diferencia entre fotogramas consecutivos de un video se debe a la combinación del movimiento de la cámara y del movimiento de los objetos. El movimiento de la cámara ha sido un tema ampliamente estudiado y se ha desarrollado modelos que permiten eliminar del video las variaciones producidas por dicho factor (Y. Wang et al. [3]), entre los cuales está el modelo descriptivo/bilineal. Este modelo permite la caracterización de casi cualquier movimiento de cámara, como traslación, rotación, acercamiento, etc. Sin embargo, la implementación de dicho algoritmo conlleva un alto costo computacional (Y. Wang et al. [3]), por lo que en la presente investigación se consideró que la cámara que realiza las capturas de video permanece sin movimiento durante las grabaciones (véase alcances y limitaciones en la sección 1.6).

La calidad de la cámara define el número de píxeles disponibles para el análisis, la claridad de la imagen a analizar y si el sistema de imágenes trabaja con píxeles de colores o escala de grises. Un tema importante a considerar es la limpieza de los lentes de la cámara. Incluso aquellas de alta calidad pueden producir imágenes de baja calidad si los lentes no se mantienen limpios, lo cual

resulta en errores en la recolección de datos.

Por otro lado, la ubicación de la cámara es, quizás, el factor más importante para asegurarse que los sensores basados en video funcionan con precisión. Las cámaras deben ser ubicadas de tal manera que se minimice los factores listados anteriormente. Particularmente, la cámara debería ser posicionada a una altura suficiente para observar el tráfico y limitar los problemas de oclusión. En general, mientras más alta esté la cámara, menor será la oclusión que se produzca en la imagen y, por consiguiente, mayor será la precisión en el conteo. Sin embargo, mientras más alta se ubique la cámara, la distancia entre ésta y los objetos en observación será mayor, lo cual aumenta la posibilidad de que una visión pobre (debido a la lluvia, neblina, nieve o polvo) oscurezca los vehículos en observación.

Los sistemas de cámara pueden operar por encima o a un lado de la carretera. Generalmente, la ubicación por encima de la carretera resulta en datos más precisos, pero los sistemas colocados a un lado de la carretera son más fáciles y menos costosos de instalar y mantener. Algunos sistemas de detección de videos utilizan métodos que requieren que la cámara se ubique a cierta altura del suelo. Por ejemplo, el método desarrollado por J. Wu y C. Gu [9] indica que la cámara se debe situar por lo menos a 6 metros de la carretera para lograr una precisión satisfactoria.

Otros de los factores a tomar en cuenta es el fondo del video. Se considera como fondo a aquellas escenas sin la presencia de los objetos de

interés, tales como humanos o vehículos en movimiento (L. Li et al. [4]). Generalmente, el fondo está formado por objetos que permanecen en estado estacionario durante el video (tales como edificios o puentes). Sin embargo, en muchos casos el fondo puede consistir en un conjunto de objetos estacionarios y objetos en movimiento (como ramas de árboles que se mueven, superficies de agua en movimiento, etc.). Existen métodos para procesar escenas de fondo con objetos en movimiento (L. Li et al. [4]). Sin embargo, en la presente investigación se considera que el fondo sólo contiene objetos estacionarios y que no se producen cambios bruscos en el color o intensidad de los mismos (véase alcances y limitaciones en la sección 1.6). También se debe considerar que la precisión del sistema de detección de video disminuye cuando el color del vehículo en movimiento es igual al color de la carretera o cuando existen sombras de edificios o de los propios vehículos (H.S. Mohana et al. [8]).

- **RADARES DOPPLER Y MICROONDAS**

Los radares Doppler y microondas transmiten radiación microondas de baja energía a un área en el pavimento y luego analizan la señal que es reflejada al detector. Los sistemas de radares microondas o pulsos transmiten un pulso de energía y miden el tiempo que le toma a dicha energía en ser reflejada al sensor, calculando de esta manera la distancia hacia el vehículo. Por otro lado, los sistemas Doppler transmiten una onda continua de energía y miden el cambio en la frecuencia de la señal causada por el efecto Doppler cuando ésta se refleja en los vehículos en movimiento (AASHTO et al. [11]).

Los radares microondas tienen la habilidad para detectar la presencia de vehículos estacionarios y para detectar múltiples zonas. Los radares Doppler generalmente no pueden detectar con precisión vehículos estacionarios.

Además de detectar la presencia de vehículos, la tecnología de radar puede medir su velocidad. Sin embargo, por lo general no son capaces de clasificarlos. Pueden ser ubicados encima o a un lado de la carretera y observar más de un carril de tráfico. Algunos radares pueden diferenciar en el tráfico en diferentes carriles.

Por ejemplo, cuando se colocan sobre los carriles, los detectores de radares microondas generalmente sólo observan un carril por sensor. Sin embargo, esta ubicación tiende a reducir los problemas de oclusión y, por consiguiente, se logra un aumento en la precisión de los resultados. El mismo sensor puede ser ubicado a un lado de la carretera y detectar todo el flujo del tráfico, sin importar el carril, pero esto reduce el nivel de precisión.

- **SISTEMAS INFRARROJOS**

Los sistemas infrarrojos pueden ser activos o pasivos. Los sensores pasivos detectan cambios en la cantidad de energía que se trasmite en las ondas infrarrojas desde la carretera y sus alrededores. Los sensores activos emiten un rayo láser infrarrojo de baja energía y miden el tiempo que demora en regresar dicho rayo. La presencia de un vehículo es medida por la reducción del tiempo en que la señal regresa.

Los detectores pasivos se pueden ilustrar como sistemas que utilizan la energía infrarroja (calor) y detectar las variaciones de las fuentes de calor, para luego convertirlas en estadísticas acerca de las características de los vehículos, generalmente volumen y velocidad (AASHTO et al. [11]). Pueden ser ubicados encima o a un lado de la carretera, dependiendo de la implementación específica.

Los sistemas infrarrojos activos crean rayos de luz que son “cortados” cuando un vehículo pasa a través de ellos. Estos sistemas pueden medir vehículos o ejes, dependiendo de la altura en que la luz es ubicada y de su dirección. El uso de múltiples sensores infrarrojos en serie permite medir la velocidad del vehículo y, por consiguiente, permite la clasificación de vehículos basada en el número y espacio entre los ejes. Existen sistemas infrarrojos activos que son capaces de recolectar de manera precisa información de volumen, velocidad y clasificación en carreteras de varios carriles.

- **SISTEMAS DE DETECCIÓN LÁSER**

Los sistemas de detección láser son funcionalmente similares a los sistemas infrarrojos activos, pero la luz generada es de una frecuencia diferente. Al igual que muchos sensores no intrusivos, pueden ser instalados encima o a un lado del pavimento. Cuando se instalan por encima, pueden ser utilizados para contar o clasificar vehículos. Cada sensor (láser) provee información de detección (cuando un vehículo rompe el rayo de luz), así como información de altura del vehículo (AASHTO et al. [11]). Si se coloca dos sensores de manera consecutiva se logra obtener información de velocidad y clasificar a los vehículos



en base a la combinación de longitud y altura.

Cuando se instalan a un lado de la carretera, los sistemas láser pueden ser utilizados como detectores de ejes, de manera similar a los sistemas infrarrojos, así como para identificar y clasificar a los vehículos en base a su perfil. Al igual que otras tecnologías que se instalan al costado de la carretera, una preocupación importante en estos casos es la posible oclusión de vehículos.

- **SISTEMAS ACÚSTICOS PASIVOS**

Los dispositivos acústicos pasivos consisten en arreglos verticales de micrófonos dirigidos al flujo del tráfico. Los micrófonos “escuchan” los neumáticos y el ruido de los motores de los vehículos que se aproximan. Los ruidos de los neumáticos son el mecanismo principal mediante el cual se detectan los vehículos a altas velocidades, mientras que para bajas velocidades, los ruidos de los motores son los más utilizados. Este tipo de tecnología es capaz de medir volumen, velocidad y presencia de los vehículos.

Los dispositivos acústicos pasivos operan mejor cuando son instalados a un lado de la carretera y su precisión depende de las condiciones ambientales que dificulten la propagación del sonido, como vientos fuertes o precipitaciones. Además, algunos vehículos que emiten ruidos fuertes pueden causar falsas lecturas.

- **SISTEMAS ULTRASÓNICOS**

Estos sistemas emiten energía de sonido ultrasónico y miden la reflexión de dicha energía. Existen dos tipos básicos, los cuales son análogos a los dos tipos de sistemas de radar descritos anteriormente: los detectores ultrasónicos de pulso, que miden el tiempo que necesita la energía emitida para ser reflejada al sensor; y los dispositivos ultrasónicos Doppler, los cuales emiten un señal ultrasónica continua y utilizan el principio Doppler para medir el cambio en la señal reflejada debido al movimiento de los objetos (AASHTO et al. [11]).

Ambos tipos de detectores pueden medir volumen, ocupación de carril y velocidad de los vehículos. En particular, los sensores Doppler tienen dificultades en medir vehículos detenidos. Por otro lado, los detectores ultrasónicos de pulso, cuando se colocan encima del carril, pueden ser utilizados para clasificar vehículos en base a su ancho y longitud. La técnica utilizada por esta tecnología es similar a la que utilizan los sistemas láser.

- **SISTEMAS MANUALES**

La recolección manual de datos es el método original de monitoreo de tráfico. En este método, una o más personas acuden a la ubicación para observar y registrar las estadísticas de tráfico, generalmente de volumen, aunque también se obtienen clasificaciones de vehículos y otras estadísticas que se pueden observar con el ojo humano. El conteo manual de imágenes de

video es otra opción que se puede usar, principalmente cuando existe un alto nivel de congestión en las carreteras y no hay lugar para que el observador se ubique.

Los conteos manuales son comúnmente usados para llevar a cabo conteos de poca duración donde es difícil ubicar equipos de monitoreo de tráfico o donde los equipos no funcionarían correctamente debido a velocidades variables (altos niveles de aceleración o desaceleración en el flujo de vehículos). También son utilizados para recolectar datos que los dispositivos aún no pueden obtener con confiabilidad, como el número de personas dentro de los vehículos (ocupación de vehículos).

La calidad de los conteos manuales está en función de la calidad del personal que los realiza y en su habilidad para ver claramente el flujo de tráfico. La precisión de este método tiende a disminuir luego de aproximadamente tres horas, debido a que la capacidad de atención de los individuos se reduce al estar concentrados por más de dicho tiempo (AASHTO et al. [11]).

Existen dispositivos que sirven de apoyo para el conteo manual, los cuales registran de manera electrónica los datos de tráfico ingresados manualmente. La entrada de datos puede ser a través de un teclado o algún software de reconocimiento de voz.

A pesar de su flexibilidad, los conteos manuales tienen una serie de limitaciones, como el costo (lo cual hace que sólo se utilice en conteos de corta

duración) o la falta de lugares en algunas carreteras donde el personal puede ubicarse de manera segura para observar el tráfico. Incluso bajo buenas condiciones, las habilidades de las personas que realizan el conteo manual son limitadas. Es por ello que, en carreteras de alto volumen de tráfico, generalmente se necesita una persona por cada carril sólo para realizar el conteo.

### 3.1.2. DURACIÓN DEL MONITOREO DE TRÁFICO

La mayoría de dispositivos de monitoreo de tráfico permiten el almacenamiento de datos estadísticos en amplios rangos de intervalos. Los intervalos de datos son necesarios para un gran porcentaje de análisis. Mientras más pequeño sea el intervalo, mayor será la capacidad para examinar el comportamiento de la carretera. Además, la habilidad para registrar datos de flujo de tráfico en pequeños intervalos permite una variedad de análisis de control de calidad que no son posibles cuando los equipos sólo son capaces de recolectar datos en intervalos grandes.

Por ejemplo, una clase de equipos para el monitoreo de tráfico se llama "registradores de eventos". Estos dispositivos utilizan sensores de eje y registran el tiempo en que cada eje cruza el sensor. Estos registros son analizados posteriormente para producir registros de monitoreo de tráfico en cualquier periodo de tiempo requerido, por ejemplo, intervalos de una hora o intervalos de cinco minutos. Se puede examinar el volumen y clasificación de vehículos (utilizando el sistema de clasificación de ejes) y realizar una variedad de verificaciones de errores. El programa de oficina que se utiliza en conjunto con el dispositivo de recolección de datos permite al analista producir y agregar las estadísticas de intervalos de tiempo requeridas.

La ventaja de utilizar estos equipos es que un sólo dispositivo es capaz de recolectar una variedad de datos que pueden ser convertidos en una serie de estadísticas de flujo de tráfico. La desventaja es que el almacenamiento de

datos requiere de espacio electrónico considerable, lo cual puede limitar el tiempo que el dispositivo de recolección puede ser dejado en el campo sin intervención.

Por otro lado, las características que necesitan tener los equipos para realizar registros en un gran número de ubicaciones que duran sólo unos cuantos días son diferentes de las características necesarias para recolectar datos en periodos de tiempo largos en una sola ubicación.

#### 3.1.2.1. MONITOREO PORTÁTIL

Los sensores portátiles se utilizan para realizar recolecciones de corta duración en un gran número de ubicaciones. Es necesario que cuenten con las siguientes características (independientemente de las características de tráfico a ser monitoreadas):

- Deben ser fácilmente transportables.
- Su instalación debe ser sencilla y requerir poco tiempo.
- Deben contener una fuente de poder lo suficientemente grande para que puedan operar hasta que sean recogidos.
- Deben ser resistentes al vandalismo. Tradicionalmente, esto implica que el dispositivo debe ser protegido y encadenado a un lugar seguro para prevenir robos.
- Necesitan mecanismos robustos de transferencia de datos hacia el

repositorio central.

### 3.1.2.2. MONITOREO SEMI-PERMANENTE

Cuando no es posible instalar de manera fácil o segura sensores portátiles se debe considerar ubicaciones semi-permanentes. Esto implica la presencia permanente de sensores a los que se unen (por una corta duración) dispositivos de recolección de datos.

Este tipo de estudio es utilizado muchas veces en carreteras rurales de alto volumen. Tradicionalmente, los carriles son cerrados y se colocan sensores permanentes (como los detectores de lazo) con cables conductores (líneas de comunicación) que van hacia la cabina a un lado de la carretera. Los dispositivos electrónicos portátiles de recolección de datos son llevados a estas ubicaciones periódicamente y conectados a los cables conductores del sensor. Estos dispositivos se dejan en la cabina a un lado de la carretera y son recogidos cuando la sesión de recolección de datos ha terminado. Con sensores no intrusivos, el mismo proceso se puede llevar a cabo sin necesidad de cerrar el carril.

El método semi-permanente tiene la ventaja de permitir visitas más rápidas y seguras en ubicaciones de alto volumen. Además, permite recolectar datos en zonas donde utilizar otros métodos no es seguro.

La desventaja de este método es que requiere un mayor nivel de capital

(en comparación con el método portátil). Es necesaria una mayor cantidad de sensores que en el método portátil pues cada sensor debe instalarse en cada ubicación, en vez de moverse junto con los dispositivos electrónicos de recolección de datos. Además, los sensores sufren fallas periódicamente debido a que están expuestos al ambiente de la carretera incluso cuando no están siendo activamente utilizados para la recolección de datos. El mantenimiento para este tipo de sensores suele costar más que el mantenimiento de sensores portátiles.

### 3.1.2.3. MONITOREO PERMANENTE

Los sensores permanentes son utilizados para proveer mediciones de flujo de tráfico y atributos que describan como dicho flujo cambia en una ubicación a través del tiempo. Las ubicaciones de monitoreo de tráfico permanente requieren:

- Sensores que resistan las condiciones ambientales de la carretera por largos periodos de tiempo.
- Fuentes de energía (energía eléctrica, solar, etc.).
- Sistemas de comunicación.
- Protección ambiental (de la temperatura, sobrecarga eléctrica, etc.).

Debido a que los sensores permanentes representan una fuerte inversión económica, la selección, instalación y calibración de los equipos es



particularmente importante. Si los recursos son pocos, se debería priorizar la recolección de mayor calidad y volumen de datos en menos ubicaciones permanentes, en lugar de recolectar datos de poca calidad en muchos lugares. Esto significa utilizar más tiempo para seleccionar y probar los equipos antes de comprarlos con el fin asegurarse que trabajen de acuerdo a los niveles de precisión deseados. También significa invertir mayor cantidad de tiempo, dinero y esfuerzo en la instalación y calibración de los equipos para garantizar que reportan satisfactoriamente los atributos de tráfico observados y que tendrán un tiempo de vida largo.

### 3.1.3. SELECCIÓN DE DISPOSITIVOS DE VOLUMEN DE TRÁFICO

El conteo preciso de volumen vehicular permite extraer información importante de tráfico, tal como el nivel de congestión y ocupación de los carriles, así como la estimación de tiempos de viaje y análisis del comportamiento del conductor. Dicha información puede ser utilizada para reducir la congestión y aumentar la seguridad en las carreteras (C. Chun et al. [5]).

En general, se deben tener en cuenta los siguientes factores al momento de elegir el equipo de recolección de datos de tráfico (AASHTO et al. [11]):

- Los atributos de tráfico a recolectar.
- La precisión con la cual dichos atributos deben ser recolectados, y cualquier limitación en el lugar o tiempo donde el dispositivo puede ser usado sin comprometer la precisión.
- El ciclo de vida de los dispositivos.
- Los requerimientos de instalación del sistema (tiempo, costo, personal).
- Los requerimientos operacionales del sistema, incluyendo las necesidades de energía y comunicación.
- El costo del sistema.
- El soporte que provee el vendedor del dispositivo.
- La compatibilidad de los datos y el software que los almacena con las bases de datos y sistemas de gestión de datos existentes.

### 3.1.3.1. CONTEO DE EJES Y DE VEHÍCULOS

Los equipos que miden el volumen de tráfico pueden dividirse en equipos que registran el paso de ejes y equipos que registran el paso de vehículos (se puede llevar ambos conteos con tecnologías intrusivas o no intrusivas). Por ejemplo, los tubos de carreteras (tecnología intrusiva) miden la cantidad de ejes, así como los sensores infrarrojos (tecnología no intrusiva). De manera similar, los lazos inductivos (tecnología intrusiva) miden la cantidad de vehículos, así como los radares microondas (tecnología no intrusiva).

Los contadores de ejes tienen la desventaja de tener que convertir el número de ejes en un estimado del número de vehículos. En la mayoría de casos esto se logra dividiendo el número de ejes contados entre un factor de ajuste (factor de corrección de ejes), el cual es un estimado del número promedio de ejes por vehículo para dicha carretera. Existen contadores de ejes más sofisticados en los cuales la distancia entre los ejes observados se utiliza para determinar si dichos ejes pertenecen a un mismo vehículo. Dados las distancias entre los ejes y el conocimiento de la clasificación de vehículos de acuerdo a dichas distancias, estos contadores son capaces de estimar directamente el número de vehículos en base a los ejes observados.

Por otro lado, las tecnologías que miden el volumen de vehículos directamente tienen la ventaja de no requerir factores de corrección de eje para convertir las mediciones de ejes (las cuales pueden ser muy precisas) en conteos de vehículos (la cual puede no ser tan precisa si el factor de corrección

de ejes no es el apropiado).

### 3.1.3.2. CONTEO DE VOLUMEN PORTÁTIL Y PERMANENTE

Debido a que la elección del mejor contador de volumen varía de una ubicación a otra, muchas veces es necesario contar con más de un tipo de dispositivo de monitoreo de tráfico.

La tecnología de conteo portátil de tráfico más usada es el sensor de tubo de carretera (AASHTO et al. [11]). Estos dispositivos son baratos y fáciles de ubicar en carreteras de bajo o mediano volumen de tráfico. Cuentan el paso de ejes con un alto nivel de precisión.

A medida que el volumen de tráfico, el número de carriles, la duración de la sesión de recolección de datos o la complejidad geométrica de la carretera se incrementa, los tubos de carretera se vuelven menos apropiados. Aunque se ha diseñado tubos de carretera para mejorar la precisión bajo estas condiciones, muchas veces se prefiere la elección de otro tipo de tecnología.

Tradicionalmente, los sensores basados en lazos inductivos han sido los dispositivos que más se han utilizado para el conteo permanente de volumen de tráfico. Sin embargo, en la actualidad se ha incrementado el uso de clasificadores de vehículos basados en ejes y dispositivos de medición de peso en movimiento, los cuales también proveen datos de volumen de tráfico.

Asimismo, también se utiliza detectores no intrusivos para conteo permanente en los casos en que el pavimento está en malas condiciones o debido a que no se desea cerrar los carriles de tráfico para la instalación o mantenimiento de los sensores. La tecnología basada en video se utiliza en estos casos (cuando la iluminación no es un problema y las condiciones ambientales no impiden la visibilidad).

### 3.1.3.3. PROBLEMAS DE LOS SENSORES INTRUSIVOS

Muchos de los problemas que presentan estos tipos de tecnología son resultado directo de una mala instalación, una inspección nula o casi nula y el uso de componentes no adecuados para el pavimento. Típicamente, se producen los siguientes errores:

- Fallas en el sensor.
- Fallas en el sello entre el sensor y el pavimento, lo que resulta en la generación de malas señales.
- Fallas en el pavimento, lo que resulta en la pérdida de unión entre el sensor y el pavimento.
- Intrusión de agua en el sensor o en el cableado que conecta el sensor con los dispositivos de recolección de datos que están a un lado de la carretera.

Muchos de estos problemas se incrementan debido a malas técnicas de

instalación, las cuales aceleran el deterioro producido por las condiciones ambientales y resulta en alguno de los cuatro problemas mencionados líneas arriba.

### 3.1.4. SEGUIMIENTO DE OBJETOS

El seguimiento de objetos es una rama importante dentro del campo de la visión computacional, y se define como la estimación de la trayectoria de objetos conforme estos se mueven en una secuencia de imágenes. En otras palabras, consiste en identificar y etiquetar objetos, de manera única y persistente, a lo largo de varias imágenes de un video. Adicionalmente, dependiendo del escenario, es posible brindar mayor información como la orientación, área o forma de los objetos identificados.

La proliferación de computadoras potentes y cámaras de video de bajo costo y alta calidad, así como la necesidad de analizar automáticamente secuencias de video, han generado un gran interés en el desarrollo de algoritmos para el seguimiento de objetos, los cuales poseen tres pasos generales (A. Yilmaz et al. [1]):

- Detección de objetos de interés en movimiento.
- Seguimiento de dichos objetos de imagen a imagen en el video.
- Análisis de los resultados para reconocer el comportamiento de los objetos.

Las técnicas de seguimiento de objetos se utilizan en actividades como:

- Reconocimiento basado en movimiento. Identificación de humanos

caminando, detección automática de objetos, etc.

- Vigilancia automática. Monitorear una escena para detectar actividades sospechosas.
- Indexado de videos. Registro y recuperación automática de videos en bases de datos multimedia.
- Interacción humano-computador. Reconocimiento de gestos, seguimiento del movimiento de ojos, etc.
- Monitoreo del tráfico. Obtención de estadísticas del tráfico en tiempo real para dirigir el flujo vehicular.
- Ruta de vehículos. Determinación de la ruta a seguir con capacidad para evitar obstáculos.

Algunas de las dificultades que se pueden presentar durante el seguimiento son (A. Yilmaz et al. [1]):

- Pérdida de información debido a la proyección del mundo tridimensional en una imagen bidimensional.
- Ruido en las imágenes, por ejemplo, reflexión de la luz causada por superficies metálicas (A. Gyaourova et al. [10]).
- Movimientos complejos de los objetos.
- Objetos articulados o con estructuras no rígidas.
- Objetos con formas complejas.
- Cambios en la iluminación.
- Movimiento de la cámara.



- Requerimientos de procesamiento en tiempo real.

El seguimiento de objetos se puede simplificar estableciendo restricciones en el movimiento o apariencia de los objetos. Por ejemplo, casi todos los algoritmos de seguimiento asumen que el movimiento de los objetos no presenta cambios abruptos. Otra posible restricción es que el movimiento de los objetos sea de velocidad o aceleración constante. Un conocimiento previo del número y del tamaño de los objetos, así como de su apariencia y forma, podrían ser usados para simplificar el problema.

Se han propuesto numerosas aproximaciones para el seguimiento de objetos, las cuales difieren unas de otras en la manera cómo responden a las siguientes preguntas:

- ¿Cuál es la mejor representación de los objetos para el seguimiento?
- ¿Qué características de las imágenes se deberían usar?
- ¿Cómo se debería modelar el movimiento, apariencia y forma de los objetos?

Las respuestas a estas preguntas dependen del contexto en que es llevado a cabo el seguimiento y el uso final que se dará a la información resultante. Existe un gran número de métodos de seguimiento que intentan responder estas preguntas para diferentes escenarios (véase capítulo 2).

### 3.1.5. PROBLEMAS Y TENDENCIAS DEL SEGUIMIENTO DE OBJETOS

#### 3.1.5.1. OCLUSIÓN

Podemos clasificar la oclusión en tres categorías: auto oclusión, oclusión entre objetos y oclusión por la estructura de fondo de la escena. La auto oclusión ocurre cuando una parte del objeto oculta otra parte. Esta situación generalmente se presenta cuando se hace seguimiento a objetos articulados. La oclusión entre objetos se produce cuando se hace seguimiento a dos objetos que se ocultan entre sí. De manera similar, la oclusión por el fondo ocurre cuando una estructura en el fondo oculta a los objetos en seguimiento (A. Yilmaz et al. [1]).

Un enfoque común para resolver el problema de la oclusión completa durante el seguimiento es representar el movimiento de los objetos mediante modelos dinámicos lineales o no lineales y predecir la ubicación del objeto hasta que reaparezca.

#### 3.1.5.2. SEGUIMIENTO CON MÚLTIPLES CÁMARAS

Existen dos razones para el uso de múltiples cámaras. La primera es para utilizar información detallada en el seguimiento y solución de la oclusión. La segunda razón es para incrementar el área de observación. Para los algoritmos de seguimiento que requieren información detallada, se debe

considerar también el alto costo computacional en que se incurre.

### 3.1.5.3. TENDENCIAS FUTURAS

En los últimos años se han llevado a cabo grandes progresos en el seguimiento de objetos. Sin embargo, muchos de los parámetros asumidos, como la mínima cantidad de oclusión, la iluminación constante, el alto contraste con respecto al fondo, etc. son violados en escenarios realistas y, por ende, se limita la usabilidad de los métodos de seguimiento. Es por ello que el seguimiento y los problemas asociados, como la selección de características, representación de objetos, formas dinámicas y estimación de movimientos son áreas activas de investigación y continuamente se proponen nuevas soluciones.

Unos de los retos que se tiene es desarrollar algoritmos para el seguimiento de objetos en videos sin restricciones, como aquellos obtenidos de redes de noticias, o videos caseros. Estos videos tienen alto nivel de ruido, son comprimidos, sin estructura y usualmente contienen escenas captadas de cámaras en movimiento desde múltiples vistas. Otro tipo de videos son los referidos a reuniones formales o informales, los cuales usualmente contienen varias personas en una vista reducida, por lo que hay gran nivel de oclusión y las personas son visibles sólo parcialmente. Existen varios métodos para estimar el punto de ubicación de la fuente de audio, por ejemplo la boca de la persona, basados en cuatro o seis micrófonos. Esta localización provee información adicional que puede ser usada en conjunto con los métodos de seguimiento para resolver el problema de la oclusión.

## 3.2. MARCO TEÓRICO INSTRUMENTAL

Existen diferentes temas que se deben abordar para el seguimiento de objetos. El primero es definir una representación adecuada del objeto, como puntos o formas geométricas primitivas. El siguiente tema es seleccionar las características de imagen, como el color o borde, que se usarán para el seguimiento (A. Yilmaz et al. [1]). También se debe definir las estrategias para detectar objetos en una escena, así como el algoritmo de seguimiento más adecuado, el cual depende de la apariencia, forma y número de objetos, así como del movimiento de los objetos y de la cámara y de las condiciones de iluminación.

### 3.2.1. REPRESENTACIÓN DE OBJETOS DE ACUERDO A SU FORMA

En un escenario de seguimiento, un objeto puede ser definido como cualquier cosa de la que se tiene interés de análisis. Por ejemplo, vehículos en una vía, botes en un lago, personas caminando, o peces en un acuario son objetos cuyo seguimiento puede ser importante en un dominio específico. Los objetos pueden ser representados por su forma y por su apariencia.

#### 3.2.1.1. PUNTOS

El objeto es representado por un punto, el centroide (Figura 3.1(a)), o por un conjunto de puntos (Figura 3.1(b)). En general, la representación por puntos

es ideal para el seguimiento de objetos que ocupan regiones pequeñas en la imagen.

#### 3.2.1.2. FORMAS GEOMÉTRICAS PRIMITIVAS

La forma del objeto es representada por un rectángulo, elipse (Figura 3.1(c),(d)), etc. Generalmente son usadas para el seguimiento de objetos rígidos simples (A. Yilmaz et al. [1]).

#### 3.2.1.3. SILUETA Y CONTORNO DEL OBJETO

La representación de contorno define los límites del objeto (Figura 3.1(g),(h)). La región dentro del contorno se denomina silueta (Figura 3.1(i)). Este tipo de representación es ideal para el seguimiento de formas complejas no rígidas.

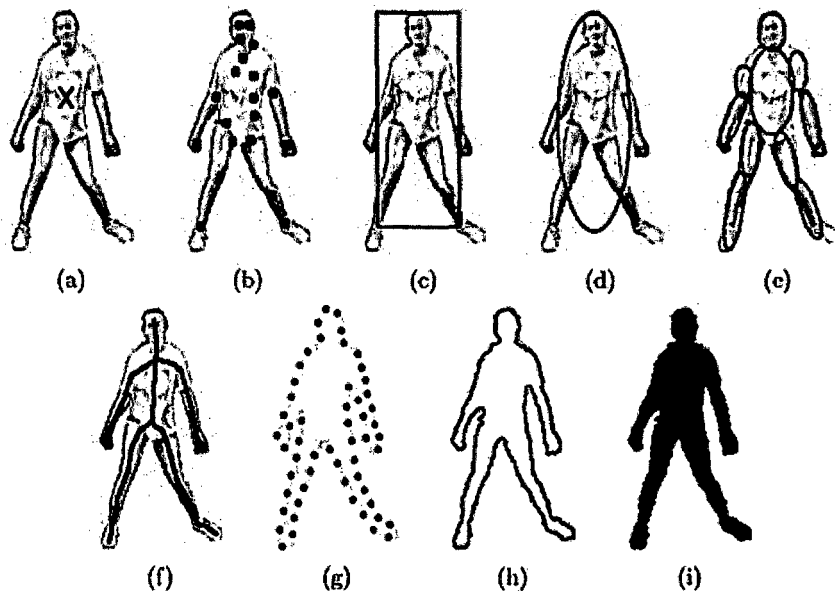
#### 3.2.1.4. MODELOS DE FORMA ARTICULADA

Los objetos articulados están compuestos por partes del cuerpo unidas por articulaciones. Para representar un objeto articulado, se puede modelar las partes constituyentes usando cilindros o elipses, como se muestra en la Figura 3.1(e).

### 3.2.1.5. MODELOS DE ESQUELETO

Es comúnmente usado como una representación de forma para reconocer objetos. Puede ser empleado para modelar objetos articulados o rígidos (Figura 3.1(f)).

**Figura 3.1: Representación de objetos de acuerdo a su forma**



Fuente: Yilmaz, Javed y Shah (2006).

### 3.2.2. REPRESENTACIÓN DE OBJETOS DE ACUERDO A SU APARIENCIA

Existen muchas maneras de representar los objetos de acuerdo a las características de su apariencia. La representación de la forma se puede combinar con la representación de la apariencia para realizar el seguimiento (A. Yilmaz et al. [1]). Las representaciones de apariencia más comunes son:

### 3.2.2.1. DENSIDADES DE PROBABILIDAD DE LA APARIENCIA DE LOS OBJETOS

Los estimados de densidad de probabilidad de la apariencia de los objetos pueden ser paramétricos o no paramétricos. Las densidades de probabilidad de las características de apariencia del objeto (color, textura) pueden ser calculadas a partir de las regiones de la imagen especificadas por los modelos de forma (región interior de un elipse o contorno, etc.).

### 3.2.2.2. PLANTILLAS

Se crean a partir de formas geométricas simples o siluetas. La ventaja de una plantilla es que contiene información tanto espacial como de apariencia. Sin embargo, las plantillas sólo determinan la apariencia del objeto generada a partir de una vista simple, es por ello que se utilizan para realizar el seguimiento de objetos cuya posición no varía considerablemente (A. Yilmaz et al. [1]).

### 3.2.2.3. MODELOS DE APARIENCIA ACTIVA

Son generados mediante un modelado simultáneo de la forma y de la apariencia. La forma se define por un conjunto de marcas, que pueden estar en los límites o dentro de la región del objeto. Para cada marca se almacena un vector de apariencia (color, textura, etc.). Los modelos de apariencia activa requieren una fase de entrenamiento en que la forma y su apariencia asociada sean aprendidas.

#### 3.2.2.4. MODELOS DE APARIENCIA DE MÚLTIPLES VISTAS

Estos modelos codifican diferentes vistas de un mismo objeto (A. Yilmaz et al. [1]).

En general, hay una relación fuerte entre las representaciones de objetos y los algoritmos de seguimiento. Las representaciones de objetos usualmente se eligen de acuerdo al dominio de aplicación. En la siguiente sección veremos la selección de características para el seguimiento.

#### 3.2.3. SELECCIÓN DE CARACTERÍSTICAS

Seleccionar las características indicadas es una etapa crítica en el seguimiento. En general, la propiedad más buscada en una característica visual es su unicidad, de tal manera que los objetos sean fácilmente distinguidos. La selección de las características está estrechamente relacionada a la representación del objeto. Por ejemplo, el color es usado como una característica para las representaciones de apariencia basadas en histogramas, mientras que, para representaciones basadas en el contorno, se utilizan los bordes del objeto. Generalmente, los algoritmos de seguimiento usan una combinación de estas características.

La mayoría de características son seleccionadas manualmente por el usuario, dependiendo del dominio de aplicación. Sin embargo, el problema de la selección automática de características ha recibido una significativa atención



en el reconocimiento de patrones. A continuación se describe las características visuales más comunes:

### 3.2.3.1. COLOR

El color aparente de un objeto está determinado básicamente por dos factores físicos: la distribución espectral de la iluminación y las propiedades de reflectancia de la superficie del objeto. En el procesamiento de imágenes generalmente se utiliza el RGB (rojo, verde, azul) como el espacio para representar los colores. Sin embargo, el RGB no es un espacio de colores perceptualmente uniforme, es decir, las diferencias entre los colores del espacio RGB no corresponden con las diferencias percibidas por los humanos. Otros espacios de colores como el  $L^*u^*v$  y  $L^*a^*b$  son perceptualmente uniformes, mientras que el HSV es aproximadamente uniforme (A. Yilmaz et al. [1]). Sin embargo, estos espacios de colores son sensibles al ruido. No existe, por lo tanto, un acuerdo en qué espacio de color es el más eficiente, por lo que se ha usado una variedad de éstos en las técnicas de seguimiento.

### 3.2.3.2. BORDES

Los límites de un objeto usualmente generan cambios fuertes en las intensidades de la imagen, los cuales son identificados usando la detección de bordes. Una propiedad importante de los bordes es que, comparados con las características de color, son menos sensibles a los cambios en la iluminación. Los algoritmos que realizan seguimiento a los bordes de los objetos usualmente

usan estos bordes como características representativas.

#### 3.2.3.3. FLUJO ÓPTICO

Consiste en vectores de desplazamiento que definen la traslación de cada píxel en la región. Se calcula usando la restricción del brillo, la cual asume que el brillo se mantiene constante en píxeles equivalentes en imágenes consecutivas.

#### 3.2.3.4. TEXTURA

Es una medida de la variación de intensidad de una superficie, la cual cuantifica propiedades como suavidad y regularidad.

#### 3.2.4. DETECCIÓN DE OBJETOS

Los métodos de seguimiento requieren un mecanismo de detección de objetos, ya sea en cada fotograma o cuando el objeto aparece por primera vez en el video. Un enfoque común en la detección de objetos es usar información de un solo fotograma. Sin embargo, algunos métodos hacen uso de información temporal calculada a partir de una secuencia de fotogramas para reducir la cantidad de falsas detecciones. Generalmente, esta información temporal está en la forma de diferencias entre fotogramas, en las que se resalta regiones cambiantes entre fotogramas consecutivos. Dadas las regiones de los objetos en la imagen, se procede a realizar la correspondencia entre un fotograma y el

siguiente para generar el seguimiento.

#### **3.2.4.1. DETECTORES DE PUNTOS**

Son usados para encontrar puntos de interés (que tengan una textura expresiva) en las imágenes (A. Yilmaz et al. [1]). Los puntos de interés han sido ampliamente utilizados en el contexto de seguimiento de objetos. Una propiedad deseable de dichos puntos es que no varíen ante los cambios en la iluminación o en la perspectiva de la cámara.

#### **3.2.4.2. SUBSTRACCIÓN DEL FONDO**

La detección de los objetos se puede lograr construyendo un modelo del fondo y encontrando desviaciones de dicho modelo para cada fotograma del video. Cualquier cambio significativo entre una región de la imagen y el modelo de fondo indica un objeto en movimiento. Los píxeles que constituyen las regiones en cambiantes son marcados para un procesamiento posterior. Usualmente, se aplica un algoritmo para obtener regiones conectadas correspondientes a los objetos. Este proceso se conoce como substracción del fondo.

#### **3.2.4.3. SEGMENTACIÓN**

El objetivo de los algoritmos de segmentación es partir la imagen en regiones perceptualmente similares. Los algoritmos de segmentación tienen

dos problemas, el criterio para una buena partición y el método para realizarla eficientemente.

#### 3.2.4.4. APRENDIZAJE SUPERVISADO

La detección de objetos se puede realizar aprendiendo automáticamente diferentes vistas de objetos a partir de un conjunto de ejemplos. Dado un conjunto de ejemplos de aprendizaje, los métodos de aprendizaje supervisado generan una función que relaciona las entradas con las salidas deseadas. Una formulación estándar de aprendizaje supervisado es el problema de clasificación donde el aprendiz aproxima el comportamiento de una función generando una salida en la forma de valores continuos (regresión) o etiquetas de clase (clasificación). En el contexto de detección de objetos, los ejemplos de aprendizaje están compuestos de pares de características de objetos y su clasificación asociada, siendo ambos definidos manualmente (A. Yilmaz et al. [1]).

La selección de características juega un rol importante en el desempeño de la clasificación, por lo que se debe elegir características que discriminen una clase de otra. Se pueden usar características como el área, orientación y apariencia del objeto. Una vez que las características son seleccionadas, las diferentes apariencias de un objeto se pueden aprender escogiendo un enfoque de aprendizaje supervisado, como redes neuronales, árboles de decisión, etc.

Los métodos de aprendizaje supervisado usualmente requieren una gran

colección de ejemplos de cada clase de objetos. Adicionalmente, esta colección debe ser etiquetada manualmente.

## **CAPÍTULO IV**

### **DESCRIPCIÓN DE DATOS DE ENTRADA**

#### **4.1. DESCRIPCIÓN GENERAL**

En este capítulo se hace una descripción de las fuentes y datos que se utilizarán como entrada para el modelo de solución. Se presenta las propiedades de los videos obtenidos, así como las mediciones manuales realizadas con el fin de medir la eficacia del modelo de solución descrito en el capítulo 5.

#### **4.2. FUENTE DE DATOS**

Con el objetivo de calcular la eficacia del método de solución bajo diversas configuraciones de parámetros, tal como se indica en el capítulo 13, se obtuvo cinco secuencias de video en vías de tránsito de Lima Metropolitana seleccionadas arbitrariamente.

La cámara de video que se utilizó para las grabaciones tiene una resolución de 720p, es decir, 1280 x 720 píxeles, y un ratio de 25fps (fotogramas

por segundo).

### 4.3. DESCRIPCIÓN DE LOS DATOS

El conjunto de datos de entrada, requerido por el modelo de solución y el diseño experimental, comprende tanto las secuencias de videos como las mediciones manuales del volumen de tránsito vehicular.

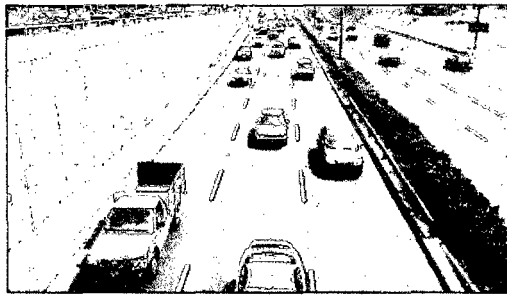
#### 4.3.1. SECUENCIAS DE VIDEO

Cada secuencia de video corresponde al tránsito regular de vehículos en una sola dirección y en condiciones de luz diurna, de acuerdo al alcance y limitaciones de esta investigación, indicados en la sección 1.6.

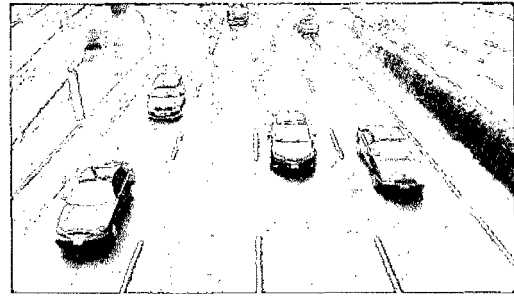
Estas secuencias de video corresponden a los siguientes cinco puentes (véase Figura 4.1):

- Video 1: Av. Javier Prado - Puente Av. Aviación
- Video 2: Av. Javier Prado - Puente Vía Evitamiento
- Video 3: Av. Javier Prado - Puente Universidad de Lima
- Video 4: Av. Paseo de la República - Puente Av. México
- Video 5: Av. Javier Prado - Puente Quiñones

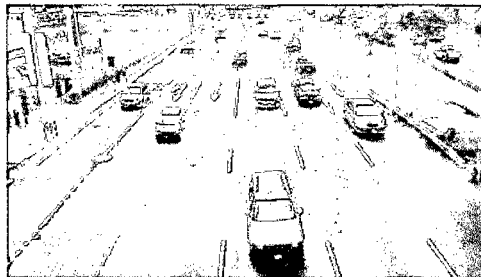
**Figura 4.1: Secuencias de video tomadas para el estudio**



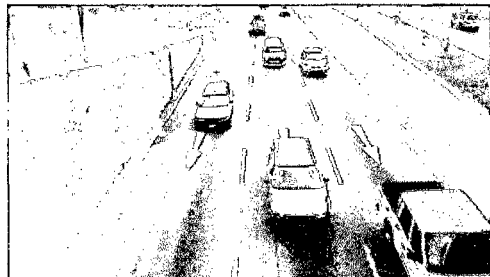
(a) Video-1 Puente Av. Aviación



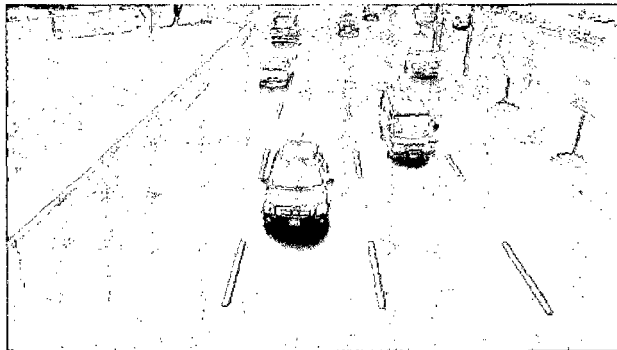
(b) Video-2 Puente Vía Evitamiento



(c) Video-3 Puente Universidad de Lima



(d) Video-4 Puente Av. México



(e) Video-5 Puente Quiñones

Fuente: Elaboración propia.



Cada secuencia de video presenta las características descritas en el Cuadro 4.1:

**Cuadro 4.1: Características de las secuencias de video**

<b>Característica</b>	<b>Descripción</b>
Duración	65 s (aprox.)
Tamaño	100 MB (aprox.)
Ancho de Fotograma	1280
Altura de Fotograma	720
Ratio de Fotograma	25 fps
Formato	.MP4

Fuente: Elaboración propia.

#### 4.3.2. MEDICIONES

Los 5 videos tienen una duración aproximada de 65 segundos cada uno y se dividieron en 7 intervalos (de 10 segundos aproximadamente) para facilitar su análisis. En la medición de cada intervalo se registró los siguientes datos:

- Número del último fotograma correspondiente al intervalo.
- Tiempo correspondiente al fin del intervalo, en segundos.
- Número de vehículos que transitaron en el intervalo, es decir, desde la última medición.
- Número acumulado de vehículos que transitaron desde el inicio de las mediciones.

Las mediciones realizadas se presentan en los cuadros 4.2, 4.3, 4.4, 4.5 y 4.6:

**Cuadro 4.2: Mediciones del volumen vehicular – Video 1**

<b>Video 1</b>				
<b>Nro. de Intervalo</b>	<b>Nro. de Último Fotograma</b>	<b>Tiempo de Fin de Intervalo (s)</b>	<b>Nro. de Vehículos</b>	
			<b>Del Intervalo</b>	<b>Acumulado</b>
0	148	5	0	0
1	376	15	14	14
2	628	25	14	28
3	877	35	13	41
4	1125	45	16	57
5	1385	55	19	76
6	1647	65	21	97
			<b>Total = 97</b>	

Fuente: Elaboración propia.

**Cuadro 4.3: Mediciones del volumen vehicular – Video 2**

<b>Video 2</b>				
<b>Nro. de Intervalo</b>	<b>Nro. de Último Fotograma</b>	<b>Tiempo de Fin de Intervalo (s)</b>	<b>Nro. de Vehículos</b>	
			<b>Del Intervalo</b>	<b>Acumulado</b>
0	147	5	0	0
1	385	15	11	11
2	641	25	13	24
3	891	35	7	31
4	1132	45	10	41
5	1361	55	7	48
6	1628	65	9	57
			<b>Total = 57</b>	

Fuente: Elaboración propia.

**Cuadro 4.4: Mediciones del volumen vehicular – Video 3**

<b>Video 3</b>				
<b>Nro. de Intervalo</b>	<b>Nro. de Último Fotograma</b>	<b>Tiempo de Fin de Intervalo (s)</b>	<b>Nro. de Vehículos</b>	
			<b>Del Intervalo</b>	<b>Acumulado</b>
0	126	5	0	0
1	387	15	15	15
2	643	25	10	25
3	881	35	7	32
4	1147	45	17	49
5	1395	55	13	62
6	1643	65	8	70
			<b>Total = 70</b>	

Fuente: Elaboración propia.

**Cuadro 4.5: Mediciones del volumen vehicular – Video 4**

<b>Video 4</b>				
<b>Nro. de Intervalo</b>	<b>Nro. de Último Fotograma</b>	<b>Tiempo de Fin de Intervalo (s)</b>	<b>Nro. de Vehículos</b>	
			<b>Del Intervalo</b>	<b>Del Intervalo</b>
0	137	5	0	0
1	382	15	4	4
2	644	25	9	13
3	881	35	7	20
4	1135	45	11	31
5	1380	55	2	33
6	1647	65	6	39
			<b>Total = 39</b>	

Fuente: Elaboración propia.

**Cuadro 4.6: Mediciones del volumen vehicular – Video 5**

<b>Video 5</b>				
<b>Nro. de Intervalo</b>	<b>Nro. de Último Fotograma</b>	<b>Tiempo de Fin de Intervalo (s)</b>	<b>Nro. de Vehículos</b>	
			<b>Del Intervalo</b>	<b>Del Intervalo</b>
0	123	10	0	0
1	242	20	7	7
2	361	30	6	13
3	487	40	8	21
4	611	50	6	27
5	728	60	7	34
6	847	70	6	40
			<b>Total = 40</b>	

Fuente: Elaboración propia.

El total de vehículos contados en los cinco videos es:

**Cuadro 4.7: Total de vehículos**

<b>Video</b>	<b># de Vehículos</b>
1	97
2	57
3	70
4	39
5	40
<b>TOTAL GENERAL</b>	<b>303</b>

#### 4.4. RESUMEN DEL CAPÍTULO

- En el presente capítulo se hizo una descripción de las secuencias de video de tráfico vehicular a analizar y de las mediciones manuales de volumen llevadas a cabo.
- En los videos analizados se observa un flujo libre de vehículos, sin congestión ni accidentes de tránsito que impidan el normal paso de los mismos. Asimismo, en el primer intervalo de cada video no se registró el flujo de ningún vehículo.
- En el capítulo 12 se realiza una comparación de las mediciones manuales con los resultados de la implementación del modelo de solución para determinar su margen de error.

## **CAPÍTULO V**

### **MODELO DE SOLUCIÓN**

#### **5.1. DESCRIPCIÓN GENERAL**

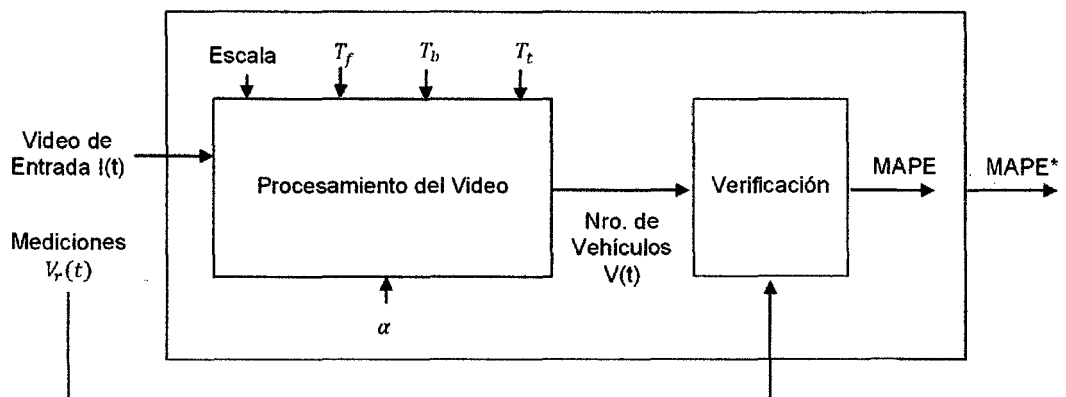
En el presente capítulo se describe el modelo de solución utilizado en la presente investigación para realizar el conteo vehicular. Primero se hace una descripción de alto nivel, es decir, indicando los parámetros globales de la solución; luego se describe detalladamente las etapas a seguir, con las variables obtenidas en cada una de ellas.

#### **5.2. MODELO DE SOLUCIÓN DE ALTO NIVEL**

El modelo de solución planteado en la presente tesis está basado en el método de seguimiento de múltiples objetos en tiempo real presentado por Tao Yang et al. [2]. En el algoritmo utilizado, las variables de entrada son el video a procesar y las mediciones del nivel de tránsito, ambos descritos en el capítulo 4. El procesamiento del video está sujeto a la selección de valores de múltiples parámetros, la cual se realiza de acuerdo a diversos experimentos descritos en el capítulo 13.

Como resultado del algoritmo, se obtiene el número acumulado de vehículos que transitaron hasta un determinado instante de tiempo o, de manera equivalente, el número de vehículos que han transitado en un intervalo de tiempo específico. El esquema general de la solución empleada se presenta en la Figura 5.1:

**Figura 5.1: Modelo de solución general**



Fuente: Elaboración propia.

Los variables y parámetros que intervienen en el modelo de solución, de acuerdo al método de Tao Yang et al. [2], se describen en el Cuadro 5.1:

**Cuadro 5.1: Variables del modelo de solución general**

<b>Variable</b>	<b>Descripción</b>
$I(t)$	Fotograma en el tiempo $t$ (en segundos) del video a procesar.
$V_r(t)$	Número de vehículos, determinado por el conteo manual, que han transitado hasta el tiempo $t$ (en segundos).
Escala	Factor de escalamiento del video.
$T_f$	Valor umbral para detectar cambios en los píxeles de fotogramas consecutivos.
$T_t$	Número de fotogramas a analizar para detectar píxeles correspondientes al fondo.
$T_b$	Valor umbral para decidir si un píxel corresponde al fondo.
$\alpha$	Factor de ponderación para la actualización del modelo de fondo.
$V(t)$	Número de vehículos, determinado por el modelo de solución, que han transitado hasta el tiempo $t$ (en segundos).
MAPE	Porcentaje promedio de error absoluto.
MAPE*	Mínimo porcentaje promedio de error absoluto.

Fuente: Elaboración propia



La selección de los parámetros *Escala*,  $T_f$ ,  $T_b$ ,  $T_t$  y  $\alpha$ , necesarios para el procesamiento del video, se realiza en el capítulo 13. Uno de los objetivos planteados (ver sección 1.4) es determinar la combinación de dichos parámetros que minimice el error absoluto de los resultados obtenidos. Es por ello que, con el fin de medir la eficacia del método utilizado, se emplean los indicadores MAPE y MAPE\*. MAPE (Mean Absolute Percentage Error o Porcentaje de Error Medio Absoluto) es una medida estadística utilizada para determinar la precisión de algún método (R. Yaffee et al. [13]). Para conocer la fórmula de cálculo de dicho indicador, revisar la fórmula 13.1.

### 5.3. MODELO DE SOLUCIÓN DETALLADO

El algoritmo de solución empleado (T. Yang et al. [2]) comprende diversas etapas para la obtención de las mediciones automáticas del volumen de flujo vehicular. A continuación se presenta un resumen de dichas etapas (para mayor detalle revisar los capítulos 6 al 12):

#### 5.3.1. PRE PROCESAMIENTO DEL VIDEO

Este primer procedimiento homogeniza el video de entrada  $I(t)$  (ver Figura 5.3) al reducir sus dimensiones y convertirlo a escala de grises de ser necesario. Además, se determina propiedades como el número de imágenes por segundo que conforman el video (FPS), número de filas y columnas de cada imagen y el número de bits por cada píxel. En el capítulo 6 se hace una

descripción detallada de esta etapa.

### 5.3.2. SEGMENTACIÓN DEL FONDO

Para propósitos de procesamiento de videos, se define el fondo como las escenas sin la presencia de los objetos de interés, tales como vehículos en movimiento (L. Li et al. [4]). En la presente investigación se consideró que los objetos del fondo permanecen en un estado estacionario, es decir, sin movimiento, como es el caso de los edificios. Además, se asume que el color o intensidad de los objetos que pertenecen al fondo no varía de manera brusca (véase los alcances y limitaciones en la sección 1.6).

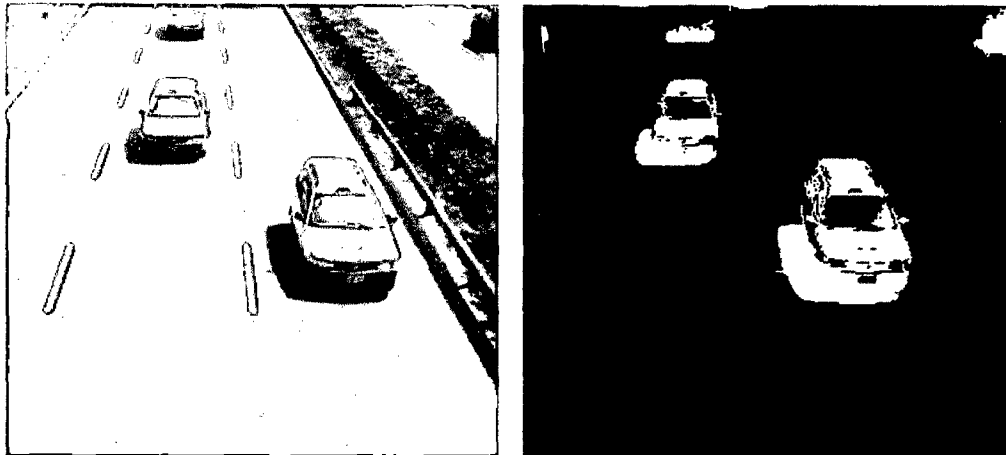
El procedimiento de Segmentación del Fondo comprende varias etapas que tienen por objetivo discriminar entre los píxeles de una imagen que pertenecen al fondo y aquellos que pertenecen a objetos en movimiento.

En el método de Tao Yang et al. [2] se presenta un algoritmo rápido para la segmentación del fondo en tiempo real. En la presenta tesis se utilizó la actualización de fondo a nivel de píxel, cuya idea principal se basa en asumir que el valor del píxel en la posición del objeto en movimiento cambia más rápido que el valor del píxel del fondo. Esto es válido en la mayoría de campos de aplicación y nos permite distinguir el fondo con precisión utilizando un método de diferencia de fotograma a fotograma, el cual puede detectar los cambios rápidos en los píxeles.

Sin embargo, este método falla cuando el color de la superficie del objeto es uniforme y, por ende, los píxeles dentro del objeto no varían. Para resolver este problema, se hace uso de una matriz dinámica que permite analizar los resultados de detección de cambio del método de diferencia de fotograma a fotograma (T. Yang et al. [2]). En dicha matriz se almacena el estado de movimiento de cada píxel y sólo aquellos cuyos valores no cambien un número determinado de veces se actualizarán como parte del fondo.

Este procedimiento toma como entrada una secuencia de video  $I$ , donde  $I(t)$  representa el fotograma en el tiempo  $t$  (en segundos) e  $I_{i,j}(t)$  es el píxel en la posición  $(i, j)$  del fotograma  $I(t)$ . La salida de este algoritmo es la secuencia de video correspondiente al primer plano,  $I'_{i,j}(t)$ . Es decir, se obtiene la secuencia en la cual se han removido todos los píxeles que corresponden a la imagen de fondo del video. En la Figura 5.2 se muestra un ejemplo de entrada y salida del procedimiento.

**Figura 5.2: Ejemplo de segmentación de fondo**



(a) Imagen original en escala de grises

(b) Imagen segmentada

Fuente: Elaboración propia.

El procedimiento de segmentación de fondo comprende las siguientes etapas:

#### 5.3.2.1. DETERMINACIÓN DEL NIVEL DE MOVIMIENTO

En esta etapa se determina los píxeles que han cambiado entre imágenes consecutivas. El resultado de este sub procedimiento es la matriz  $F(t)$ , la cual identifica qué píxeles se encuentran en movimiento en un determinado fotograma. Según se detalla en el capítulo 7, aquellos cuya variación (con respecto al fotograma anterior) ha sido superior al valor umbral  $T_f$  se consideran en movimiento.

### 5.3.2.2. ANÁLISIS DEL NIVEL DE MOVIMIENTO

El nivel de movimiento  $F(t)$  es analizado en intervalos de tiempo  $T_t$  con el fin de identificar aquellos píxeles que han tenido cambios constantes y aquellos que no han tenido cambios en dichos intervalos de tiempo. Los píxeles que no han cambiado en los últimos  $T_t$  fotogramas son considerados como parte del fondo y se identifican en la matriz  $U(t)$ , tal como se indica en el capítulo 8.

### 5.3.2.3. ACTUALIZACIÓN DEL MODELO DE FONDO

Luego de identificar aquellos píxeles que no han tenido cambios constantes y que, por lo tanto, pueden ser considerados como parte del fondo, se actualiza el modelo de fondo del video  $B(t)$  en base a una ecuación lineal que considera el modelo de fondo del fotograma anterior y la imagen actual. Dicha ecuación se presenta en el capítulo 9 de la presente investigación.

### 5.3.2.4. SUBSTRACCIÓN DEL FONDO

Se sustrae el modelo de fondo  $B(t)$  de la imagen del video  $I(t)$ , con lo que se logra obtener los objetos del primer plano de la imagen. Cada fotograma del video es comparado con el modelo de fondo y se obtiene aquellos píxeles cuya variación sea mayor que el valor umbral  $T_b$ . El resultado de este sub procedimiento es la matriz  $P(t)$  que identifica los píxeles que corresponden a la imagen en movimiento. La ecuación utilizada se describe el capítulo 10.

#### 5.3.2.5. FILTRO MORFOLÓGICO

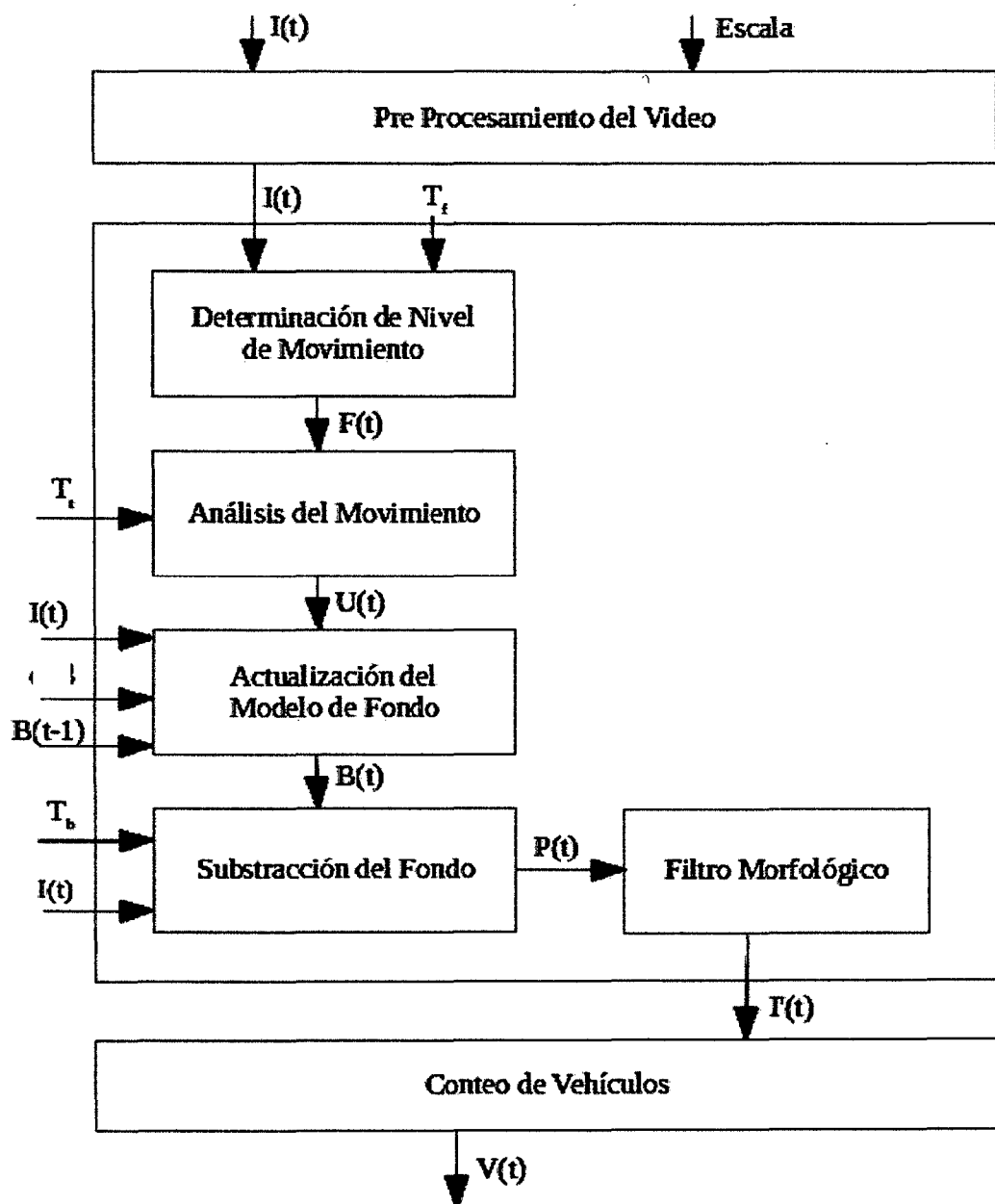
Con el objetivo de reducir el nivel de ruido, se eliminan los componentes del primer plano cuyo tamaño sea menor a un valor umbral. Esta etapa permite mejorar la calidad del resultado final y obtener resultados de mayor precisión. En el capítulo 11 se presenta un ejemplo de aplicación de este procedimiento.

#### 5.3.3. CONTEO DE VEHÍCULOS

La última etapa del modelo de solución es el conteo de vehículos, el cual consiste en el trazado de una línea imaginaria (a través de la cual pasan los vehículos detectados en las etapas anteriores) y el uso de un contador que almacena el volumen de vehículos detectados. En el capítulo 12 se detalla el procedimiento a seguir en esta etapa.

El modelo de solución detallado se esquematiza en la Figura 5.3, donde se puede observar las sucesivas etapas y las variables y parámetros involucrados, los cuales son definidos en el Cuadro 5.1.

**Figura 5.3: Modelo de solución detallado**



Fuente: Elaboración propia.

#### 5.4. RESUMEN DEL CAPÍTULO

- En el presente capítulo se realizó una descripción general del modelo de solución empleado, las variables y parámetros que intervienen y los diferentes procedimientos que conforman el algoritmo para el conteo de flujo vehicular.
- Para conocer el valor que se debe asignar a cada uno de los parámetros es necesario realizar una comparación de los resultados del modelo de solución con los resultados del conteo manual llevado a cabo en el capítulo 4.
- En los capítulos 6 al 12 se presentará el detalle de cada una de las etapas del modelo de solución y en el capítulo 13 se hará una descripción del diseño experimental realizado para determinar la combinación de parámetros que minimice el error medio absoluto del método empleado.



## **CAPÍTULO VI**

### **PRE PROCESAMIENTO DEL VIDEO**

#### **6.1. DESCRIPCIÓN GENERAL**

En el presente capítulo se explica la primera etapa del modelo de solución, el cual toma como entrada el video a procesar y lo homogeniza con el fin de prepararlo para los análisis posteriores.

#### **6.2. OBTENCIÓN DE PROPIEDADES DEL VIDEO**

El primer procedimiento del modelo de solución, descrito en el capítulo 5 (ver Figura 5.3), es el pre procesamiento del video de entrada, el cual puede ser a colores o blanco y negro, o puede tener dimensiones arbitrarias. Esta etapa permite homogenizar las entradas del algoritmo y maximizar su eficacia mediante una adecuada selección de parámetros de conversión.

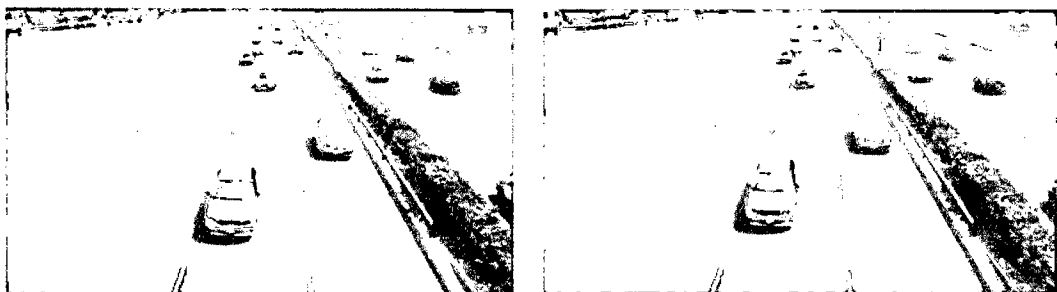
En este sub procedimiento se determinan las siguientes propiedades del video:

- Número de imágenes por segundo que conforman el video (FPS).
- Número de filas de cada imagen.
- Número de columnas de cada imagen.
- Número de bits por píxel.

### 6.3. CONVERSIÓN A ESCALA DE GRISES

Para facilitar el procesamiento del video, se realiza una conversión del espacio de colores de 3 canales (RGB o YUV) a escala de grises con 8 bits por píxel, tal como se indica en la Figura 6.1.

**Figura 6.1: Conversión a escala de grises**



(a) Imagen original en RGB    (b) Imagen convertida a escala de grises

Fuente: Elaboración propia.

### 6.4. REDUCCIÓN DE DIMENSIONES

Dependiendo de las dimensiones del video, este puede ser escalado vertical y horizontalmente por un factor ( $1/2$  o  $1/4$  en la mayoría de casos prácticos). Podemos observar un ejemplo de reducción de dimensiones en la Figura 6.2:

**Figura 6.2: Imagen escalada por un factor de 0.5**



(a) Imagen original en 1280x720 (b) Imagen escalada a 640x360

Fuente: Elaboración propia.

## 6.5. RESUMEN DEL CAPÍTULO

- En esta primera etapa del modelo de solución se prepara el video de flujo vehicular para que sea utilizado como entrada al algoritmo definido por Tao Yang et al. [2].
- La homogenización del video consiste en modificar (de ser necesario) los atributos de color y tamaño con el fin de maximizar la eficacia de la solución.
- En las actividades descritas en el capítulo 7 se toma como entrada el video (pre procesado) con el fin de determinar el fondo de cada fotograma del mismo y realizar el conteo de los vehículos (objetos en movimiento).

## CAPÍTULO VII

### DETERMINACIÓN DEL NIVEL DE MOVIMIENTO

#### 7.1. DESCRIPCIÓN GENERAL

En el presente capítulo se hace una descripción detallada de la primera etapa de la segmentación del fondo. Se presenta los parámetros y las fórmulas utilizadas, así como la matriz que almacena los resultados obtenidos.

#### 7.2. PROCEDIMIENTO

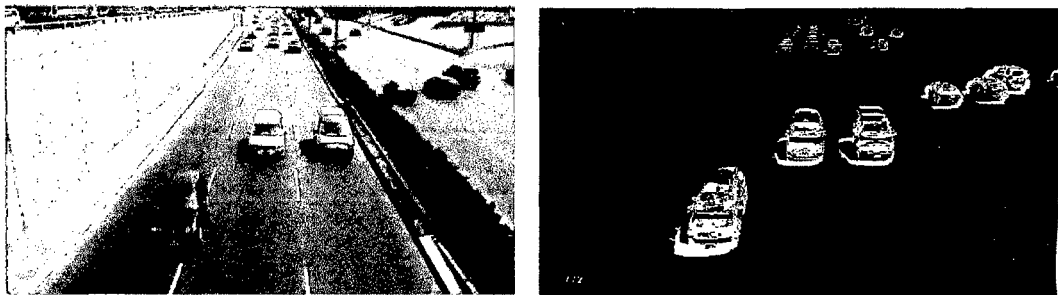
Esta etapa toma como entrada la secuencia de video  $I$ , la cual es analizada para determinar, en base a un valor umbral  $T_f$ , qué píxeles han cambiado entre imágenes consecutivas. De esta manera, es posible detectar secciones en movimiento y secciones que posiblemente correspondan al fondo.

Aquellos píxeles cuyos valores varíen en más de  $T_f$ , serán categorizados como píxeles en movimiento, en caso contrario serán considerados como posibles componentes del fondo, tal como se expresa en la función 7.1.

$$F_{i,j}(t) \begin{cases} 0, & \text{si } |I_{i,j}(t) - I_{i,j}(t-1)| \leq T_f \\ 1, & \text{si } |I_{i,j}(t) - I_{i,j}(t-1)| > T_f \end{cases} \quad (7.1)$$

La salida de este sub procedimiento es la matriz F que identifica a los píxeles en movimiento. En la Figura 7.1 se presentan los resultados en un determinado instante de tiempo. Los píxeles mostrados en blanco son aquellos en los que se ha detectado movimiento, mientras que aquellos en negro no han presentado movimiento con respecto al fotograma anterior.

**Figura 7.1: Determinación del nivel de movimiento.**



(a) Entrada

(b) Salida

Fuente: Elaboración propia.

Sin embargo, este método podría no detectar cambios cuando la superficie de los vehículos es de un color homogéneo. Para superar este inconveniente, se utiliza una matriz dinámica que almacena el número de veces consecutivas en que un píxel ha cambiado (descrita en el siguiente capítulo).

### 7.3. RESUMEN DEL CAPÍTULO

- En el presente capítulo se detalló el sub procedimiento de determinación del nivel de movimiento de cada fotograma del video en análisis. El resultado consta de una matriz F por cada fotograma, donde se indica el estado de movimiento de los píxeles. Si la variación de un píxel (con respecto al píxel correspondiente en el fotograma anterior) es mayor que un valor umbral  $T_f$ , entonces se considera que el píxel está en movimiento.
- En el siguiente capítulo se toma como entrada la matriz F para determinar aquellos píxeles que deben ser actualizados como parte del fondo.

## CAPÍTULO VIII

### ANÁLISIS DEL NIVEL DE MOVIMIENTO

#### 8.1. DESCRIPCIÓN GENERAL

En este capítulo se detalla el análisis matemático llevado a cabo en los píxeles para determinar si corresponden al fondo o si forman parte de los objetos de interés en movimiento. Se presenta además un ejemplo de entrada y salida de esta etapa del algoritmo solución.

#### 8.2. PROCEDIMIENTO

En este paso se determina aquellos píxeles que no han cambiado en los últimos  $T_t$  fotogramas, los cuales son considerados como parte del fondo de la imagen. Se toma como entrada la matriz  $F$  de la etapa anterior y se calcula la matriz dinámica  $D$ , según la fórmula 8.1:

$$D_{i,j}(t) \begin{cases} T_t, & \text{si } F_{i,j}(t) = 1 \\ \max(0, D_{i,j}(t-1) - 1), & \text{si } F_{i,j}(t) = 0 \end{cases} \quad (8.1)$$

Observamos que el valor de un píxel en la matriz  $D$  irá disminuyendo a

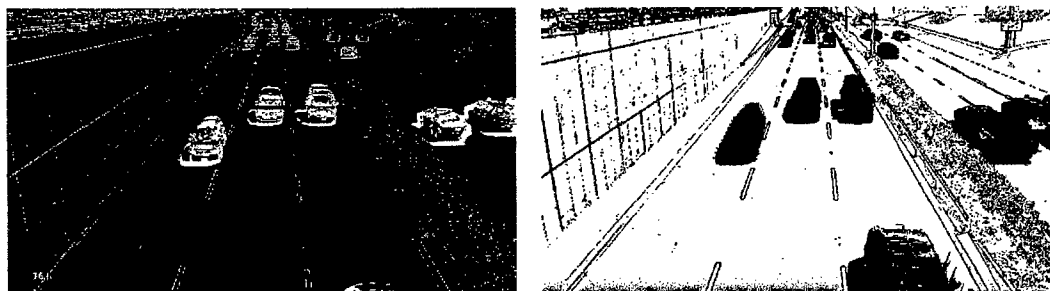
medida que permanezca sin variación en fotogramas sucesivos, hasta hacerse cero. Cuando  $D_{i,j}(t) = 0$  se procederá a actualizar el fondo de la imagen, teniendo en cuenta el fondo anterior y la imagen actual. La matriz  $U_{i,j}(t)$  identifica aquellos píxeles del fondo que serán actualizados.

$$U_{i,j}(t) \begin{cases} 1, & \text{si } D_{i,j}(t) = 0 \\ 0, & \text{si } D_{i,j}(t) \neq 0 \end{cases} \quad (8.2)$$

La fórmula 8.2 indica la lógica de cálculo de la matriz  $U_{i,j}(t)$ . Cuando  $D_{i,j}(t) = 0$ , se establece en 1 el valor de  $U_{i,j}(t)$ ; en caso contrario, toma el valor de 0. De esta manera, los píxeles del fondo que deben ser actualizados se identificarán tomando todos aquellos píxeles de la matriz  $U_{i,j}(t)$  que tienen como valor 1.

Un ejemplo del análisis del movimiento se representa en la Figura 8.1. Los píxeles mostrados en blanco (Figura 8.1(b)) son aquellos que han permanecido sin variación a lo largo de  $T_t$  fotogramas.

**Figura 8.1: Análisis del movimiento**



(a) Entrada

(b) Salida

Fuente: Elaboración propia.



Con este procedimiento se logra evitar errores cuando los vehículos en movimiento tienen superficies homogéneas, las cuales de otra manera serían consideradas como parte del fondo de la imagen.

### 8.3. RESUMEN DEL CAPÍTULO

- En esta etapa del modelo de solución se analiza si cada píxel del fotograma se ha mantenido sin variación considerable una determinada cantidad de veces (lo cual permite resolver el problema de superficies homogéneas en movimiento). De ser así, se procede a la actualización del modelo de fondo (descrita en el siguiente capítulo).

## CAPÍTULO IX

### ACTUALIZACIÓN DEL MODELO DE FONDO

#### 9.1. DESCRIPCIÓN GENERAL

La etapa de actualización de fondo permite modificar el modelo de fondo del video de acuerdo a los resultados obtenidos del análisis del nivel de movimiento de los píxeles (capítulo 8). En esta sección se detalla la fórmula empleada en la presente investigación, de acuerdo a lo indicado por Tao Yang et al. [2], así como un ejemplo de aplicación.

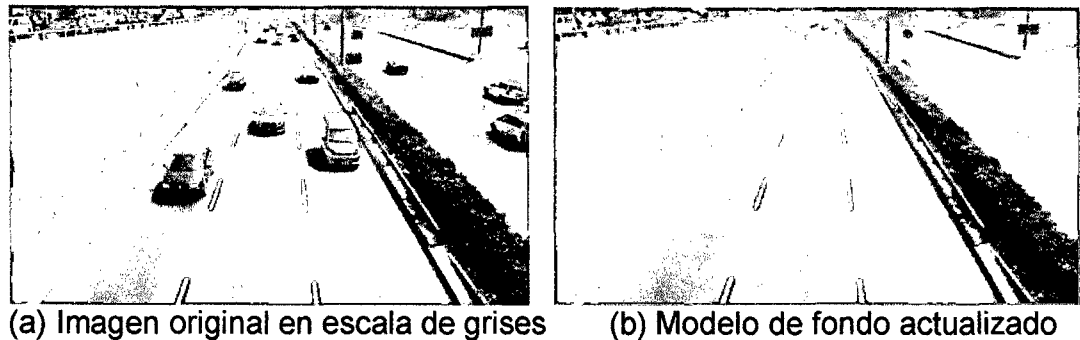
#### 9.2. PROCEDIMIENTO

En este sub procedimiento se actualizan los píxeles del fondo detectados producto del análisis del movimiento del video. Esta actualización se realiza mediante el promedio ponderado del modelo de fondo previo y la imagen actual, tal como se presenta en la ecuación 9.1.

$$B_{i,j}(t) = \alpha * I_{i,j}(t) + (1 - \alpha) * B_{i,j}(t - 1) \quad (9.1)$$

En dicha ecuación, la matriz  $B(t)$  representa la imagen de fondo en el tiempo  $t$ . El parámetro  $\alpha$  puede tomar valores en el rango  $< 0, 1 >$ .

**Figura 9.1: Resultados de la actualización del modelo de fondo**



Fuente: Elaboración propia.

Un ejemplo del resultado de esta etapa se presenta en la Figura 9.1, en la cual observamos el fondo de un determinado fotograma, el cual fue obtenido utilizando la ecuación 9.1.

### 9.3. RESUMEN DEL CAPÍTULO

En esta etapa se realiza la actualización del modelo de fondo en base al análisis del nivel de movimiento de los píxeles (determinado en el capítulo anterior). En el siguiente capítulo se describe la obtención de los fotogramas correspondientes a los objetos en movimiento a partir del modelo de fondo calculado en el presente apartado.

## CAPÍTULO X

### SUBSTRACCIÓN DE FONDO

#### 10.1. DESCRIPCIÓN GENERAL

A continuación se establece las fórmulas a utilizar para la etapa de sustracción de fondo, cuyo objetivo es obtener la imagen correspondiente a los vehículos en movimiento (sin la presencia del fondo).

#### 10.2. PROCEDIMIENTO

Esta etapa permite obtener los objetos del primer plano de la imagen (vehículos) a través de una sustracción entre los valores de los píxeles de la imagen y del fondo (determinado en los pasos anteriores).

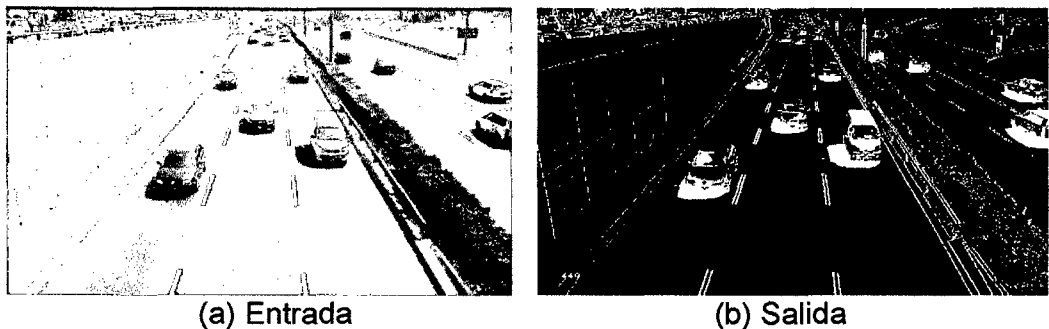
$$P_{i,j}(t) = \begin{cases} 1, & \text{si } |I_{i,j}(t) - B_{i,j}(t)| \geq T_b \\ 0, & \text{si } |I_{i,j}(t) - B_{i,j}(t)| < T_b \end{cases} \quad (10.1)$$

Cada fotograma del video es comparado con el modelo de fondo y se obtiene aquellos píxeles cuya variación sea mayor que el valor umbral  $T_b$ . El

resultado de este sub procedimiento es la matriz  $P(t)$  que identifica los píxeles que corresponden a la imagen en movimiento.

En la Figura 10.1 se observa el resultado de la substracción de fondo (en blanco se muestran los píxeles correspondientes al primer plano del video).

**Figura 10.1 Substracción de fondo**



Fuente: Elaboración propia.

### 10.3. RESUMEN DEL CAPÍTULO

Se detalló el algoritmo para substraer el fondo de cada fotograma del video. De esta manera, el resultado final será una secuencia de imágenes conformadas sólo por los objetos en movimiento, lo cual hace posible el conteo de los vehículos. En la siguiente sección se procederá a realizar el filtro morfológico de las imágenes obtenidas según el procedimiento indicado en este capítulo.

## **CAPÍTULO XI**

### **FILTRO MORFOLÓGICO**

#### **11.1. DESCRIPCIÓN GENERAL**

En este capítulo se describe el filtro morfológico aplicado en la presente investigación. Este paso corresponde a la última etapa de la segmentación del fondo y permite obtener resultados finales de mayor precisión.

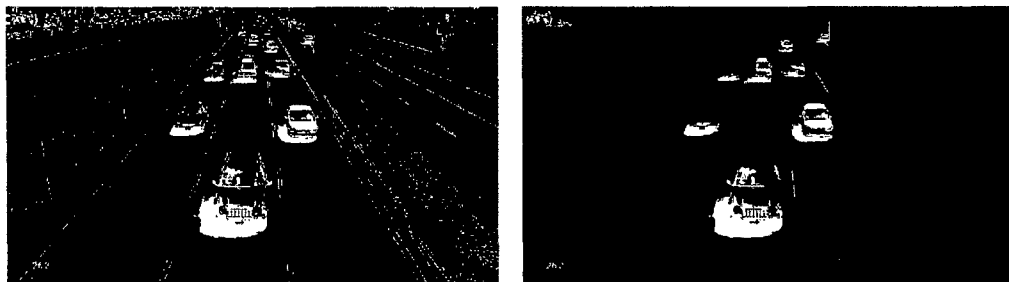
#### **11.2. PROCEDIMIENTO**

Los fotogramas obtenidos en la etapa previa pueden contener numerosas imperfecciones que dificultan el conteo de vehículos. El filtro morfológico consiste en disminuir el ruido de las imágenes mediante la extracción de aquellos componentes que no son de interés para el análisis.

En la presente investigación se utilizó como criterio el tamaño de los objetos, es decir, se extrajo aquellos componentes cuyo tamaño sea menor a un valor umbral. De este modo se elimina gran parte del ruido presente en el video. El resultado de la aplicación de esta técnica se puede observar en la Figura 11.1,

donde la imagen de salida contiene una menor cantidad de elementos que no son vehículos.

**Figura 11.1: Filtro morfológico**



(a) Entrada

(b) Salida

Fuente: Elaboración propia.

### 11.3. RESUMEN DEL CAPÍTULO

- En el presente capítulo se hizo una descripción del último paso correspondiente a la segmentación del fondo del video. En esta etapa se logra eliminar aquellos objetos de menor tamaño que no corresponden a los objetos de interés (vehículos).
- En el siguiente capítulo se toma entrada los resultados obtenidos después del filtro morfológico para realizar el conteo de los vehículos.

## **CAPÍTULO XII**

### **CONTEO DE VEHÍCULOS**

#### **12.1. DESCRIPCIÓN GENERAL**

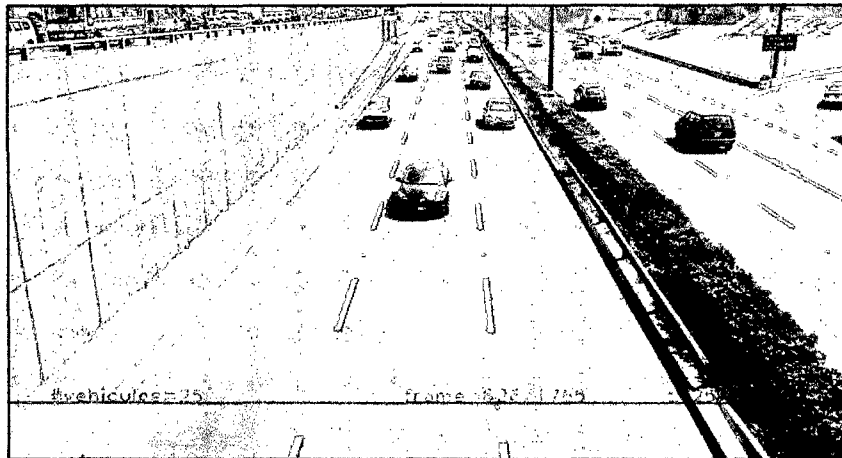
En este capítulo se describe el procedimiento empleado para determinar la cantidad total de vehículos que se desplazan en cada video. Dicha cantidad se comparará con la cantidad real, indicada en el capítulo 4, para obtener el error medio absoluto (MAPE).

#### **12.2. TRAZADO DE LÍNEA IMAGINARIA**

Para definir la cantidad de vehículos es necesario trazar una línea imaginaria que establezca el punto a partir del cual se empezará a contar los objetos en movimiento, tal como se puede observar en la Figura 12.1.



**Figura 12.1: Línea Imaginaria**

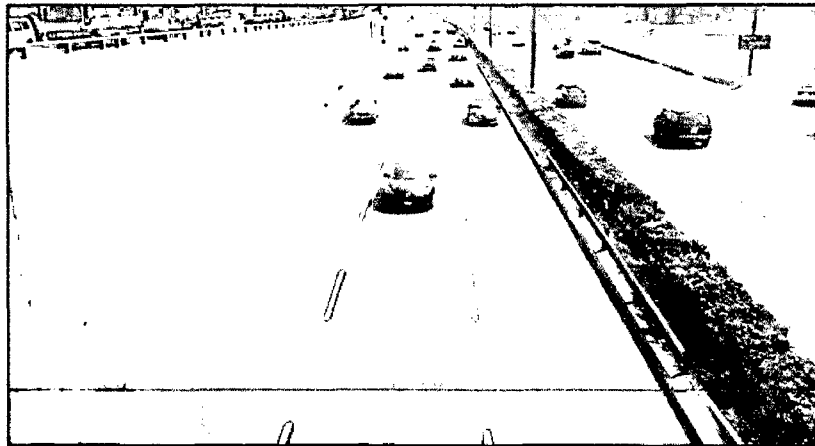


El trazado de esta línea imaginaria permite excluir del conteo aquellos vehículos que hayan sido detectados en la vía de tránsito, pero que no hayan continuado su desplazamiento sobre ella, así como aquellos vehículos que se encuentran en el carril contrario.

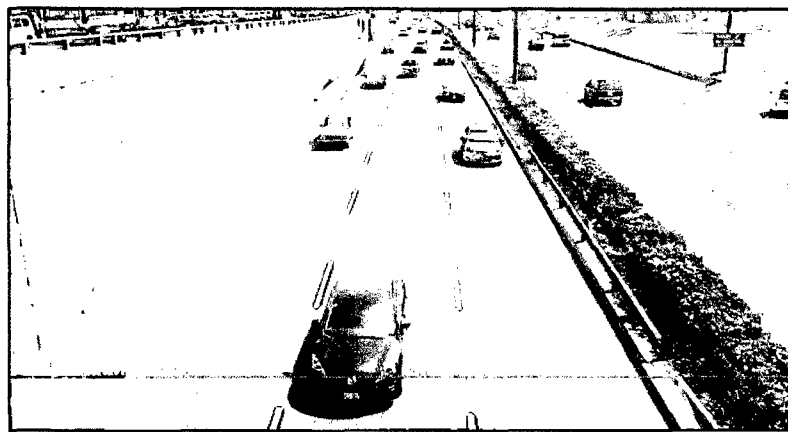
### 12.3. CONTEO DE VEHICULOS

Cuando un vehículo pasa por la línea imaginaria trazada, el contador se incrementa en uno. En la Figura 12.2(a) se puede observar que el contador tiene valor de 25 y que al pasar el vehículo (Figura 12.2(b)) se incrementa en una unidad. Luego del pase del vehículo mostrado, observamos que el contador se mantiene en el valor de 26 (Figura 12.2(c)).

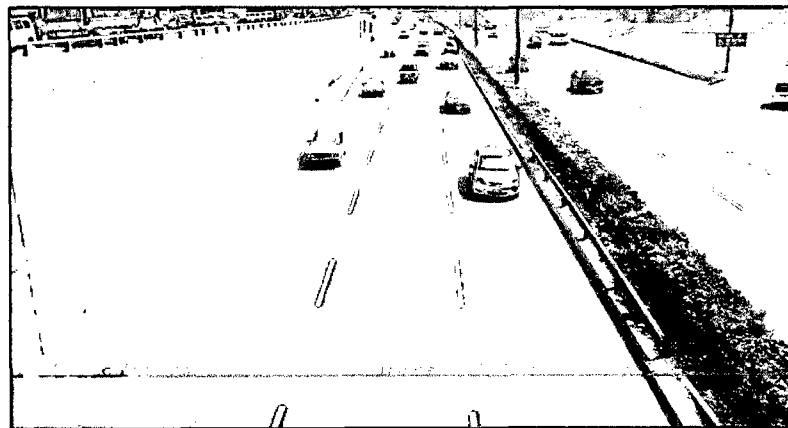
**Figura 12.2** Conteo de vehículos



(a) Antes del paso del vehículo



(b) Durante el paso del vehículo

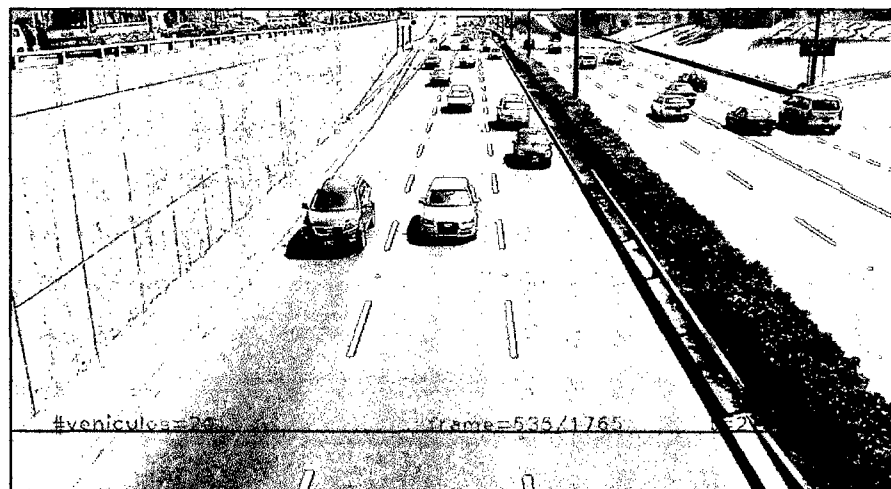


(c) Después del paso del vehículo

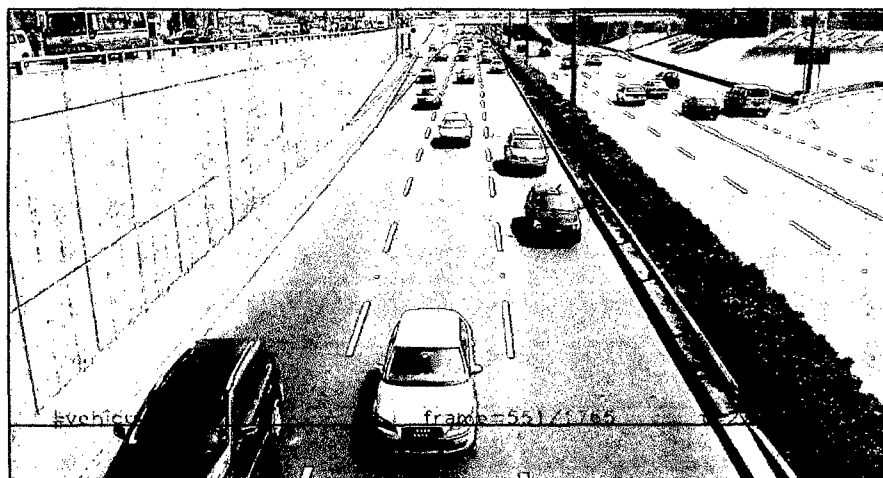
El uso de una línea imaginaria permite contar dos vehículos que pasan al

mismo tiempo de manera paralela. En la Figura 12.3(a) se observa los dos vehículos que se aproximan a la línea imaginaria y en la Figura 12.3(b) podemos ver que el contador se incrementó en dos unidades (de 20 a 22):

**Figura 12.3** Conteo de vehículos en paralelo



(a) Antes del conteo



(b) Después del conteo

Sin embargo, el método de solución empleado no detecta dos vehículos que pasan juntos. En dichos casos, los considera como uno solo.

#### 12.4. RESUMEN DEL CAPÍTULO

En el presente capítulo se hizo una descripción detallada del último paso correspondiente al conteo de vehículos. En la siguiente etapa se realiza el experimento necesario para determinar el error medio absoluto de la solución (MAPE).

## **CAPÍTULO XIII**

### **EXPERIMENTACIÓN**

#### **13.1. DESCRIPCIÓN GENERAL**

A continuación se hace una descripción de los experimentos llevados a cabo para determinar los valores de los parámetros que minimizan el error medio absoluto. Se hace una presentación del diseño del experimento y de los resultados obtenidos.

#### **13.2. DISEÑO DEL EXPERIMENTO**

Para maximizar la eficacia del modelo de solución, se ha diseñado experimentos en los que se calcula el porcentaje promedio de error absoluto (MAPE) bajo distintas configuraciones de parámetros (R. Yaffee et al. [13]). En cada experimento se analiza las variaciones del indicador MAPE y se selecciona los valores de los parámetros que lo minimicen.

La ejecución se llevó a cabo en un ambiente con las siguientes características:

**Cuadro 13.1: Ambiente de experimentación**

<b>Característica</b>	<b>Valor</b>
Sistema Operativo	Distribución de GNU/Linux
Procesador	Intel Core i7 1.6 GHz
Memoria RAM	6 GB

Fuente: Elaboración propia.

Los datos de entrada se describen en el capítulo 4 de la presente tesis (cinco videos de flujo vehicular tomados a un ratio de 25 FPS).

### 13.3. VARIABLES INDEPENDIENTES Y DEPENDIENTES

Para los experimentos llevados a cabo se realizó una clasificación de las variables en independientes y dependientes. Las independientes son el video a procesar y las mediciones manuales de flujo vehicular, cuya metodología de obtención fue descrita en el capítulo 4 de la presente investigación. Por otro lado, se considera variables dependientes a aquellas cuyo valor está en función de las variables independientes y de los parámetros determinados, como se puede observar en los cuadros 13.2 y 13.3.

**Cuadro 13.2: Variables dependientes e independientes del modelo de solución**

<b>Variables Independientes</b>	
<b>Variable</b>	<b>Definición</b>
$I(t)$	Fotograma en el tiempo $t$ (en segundos) del video a procesar.
$V_r(t)$	Número de vehículos, determinado por el conteo manual, que han transitado hasta el tiempo $t$ (en segundos).
<b>Variables Dependientes</b>	
<b>Variable</b>	<b>Definición</b>
$V(t)$	Número de vehículos, determinado por el modelo de solución, que han transitado hasta el tiempo $t$ (en segundos).
$MAPE$	Porcentaje promedio de error absoluto en un determinado video.
$MAPE^*$	Mínimo porcentaje promedio de error absoluto.

Fuente: Elaboración propia.

**Cuadro 13.3: Parámetros del modelo de solución**

Parámetro	Definición	Valores posibles
Escala	Factor de escalamiento del video.	0.5
$T_f$	Valor umbral para detectar cambios en los píxeles de fotogramas consecutivos.	10; 20; 30; 40; 50
$T_t$	Número de fotogramas a analizar para detectar píxeles correspondientes al fondo.	2; 4; 6; 8
$T_b$	Valor umbral para decidir si un píxel corresponde al fondo.	10; 20; 30; 40; 50; 60; 70; 80; 90; 100
$\alpha$	Factor de ponderación para la actualización del modelo de fondo.	0.2

Fuente: Elaboración propia.

En la ecuación 13.1 se detalla el cálculo del MAPE para cada combinación de parámetros en cada video:

$$MAPE = \frac{1}{n} \times \sum_{i=1}^n \frac{|V(i) - V_r(i)|}{V_r(i)} \quad (13.1)$$

Donde:

- $n$  = Número total de intervalos en que se dividió cada video. Según se detalla en el capítulo 4, todos los videos se dividieron en 7 intervalos.
- $V_r(i)$  = Número de vehículos reales que han transitado en el intervalo  $i$ .  
Revisar los cuadros 4.2 al 4.6.
- $V(i)$  = Número de vehículos determinado por el modelo de solución en el intervalo  $i$ . Se obtiene a partir de la variable  $V(t)$ .



Como resultado de la fórmula 13.1 se obtiene un *MAPE* para cada combinación de parámetros en cada video. El siguiente paso es obtener un único *MAPE* ponderado (*MAPE<sup>P</sup>*) para cada combinación de parámetros, para lo cual se aplica la siguiente fórmula:

$$MAPE^P = \frac{1}{\sum_{k=1}^m V_{r,k}(T)} \times \sum_{k=1}^m (MAPE_k * V_{r,k}(T)) \quad (13.2)$$

Donde:

- *m* = Número de videos. Según se indica en el capítulo 4, en la presente investigación se tomó como entrada 5 videos.
- $V_{r,k}(T)$  = Total de vehículos reales del video *k* (véase los cuadros 4.2 al 4.6).
- $MAPE_k$  = *MAPE* de una combinación específica de parámetros en el video *k* (obtenido de la fórmula 13.1).

#### 13.4. RESULTADOS

Los valores  $MAPE^P$  obtenidos para cada combinación de parámetros, según la ecuación 13.2, se presentan en los cuadros 13.4, 13.5, 13.6 y 13.7:

**Cuadro 13.4: Resultados del experimento – Parte 1**

Tt	Tb	Tf				
		10	20	30	40	50
2	10	0.59076	0.59076	0.60726	0.60066	0.61056
2	20	0.19472	0.14191	0.14851	0.14521	0.15182
2	30	0.16502	0.12871	0.11551	0.12871	0.11551
2	40	0.15842	0.11551	0.13531	0.12541	0.11221
2	50	0.13531	0.11881	0.11551	0.12541	0.11221
2	60	0.11221	0.11221	0.11551	0.11881	0.11881
2	70	0.10891	0.09901	0.10891	0.10891	0.13201
2	80	0.13201	0.12211	0.11221	0.13861	0.14191
2	90	0.15512	0.16832	0.18152	0.21452	0.22112
2	100	0.25743	0.27063	0.28383	0.30033	0.31023

Fuente: Elaboración propia.

**Cuadro 13.5: Resultados del experimento – Parte 2**

Tt	Tb	Tf				
		10	20	30	40	50
4	10	0.60396	0.60066	0.58086	0.57756	0.56766
4	20	0.21122	0.16172	0.14521	0.14521	0.13861
4	30	0.17162	0.12211	0.10891	0.11221	0.11221
4	40	0.17162	0.11881	0.12211	0.13201	0.12871
4	50	0.15182	0.11221	0.10231	0.12871	0.12541
4	60	0.10561	0.11221	0.11551	0.11551	0.11221
4	70	0.10231	0.10561	0.09901	0.10561	0.12211
4	80	0.12211	0.10891	0.10231	0.13201	0.14191
4	90	0.15512	0.16502	0.17492	0.21122	0.21782
4	100	0.25082	0.28383	0.27723	0.29043	0.29043

Fuente: Elaboración propia.

**Cuadro 13.6: Resultados del experimento – Parte 3**

Tt	Tb	Tf				
		10	20	30	40	50
6	10	0.55446	0.58746	0.56436	0.55446	0.57426
6	20	0.21782	0.16502	0.15512	0.15182	0.14851
6	30	0.19142	0.12541	0.11221	0.11881	0.10891
6	40	0.17162	0.11881	0.12541	0.12541	0.12211
6	50	0.15842	0.12211	0.10561	0.13201	0.12211
6	60	0.10891	0.12211	0.12211	0.12541	0.11881
6	70	0.10561	0.10561	0.10561	0.10231	0.12541
6	80	0.11881	0.10891	0.10231	0.13201	0.14191
6	90	0.15512	0.16502	0.17492	0.21122	0.22112
6	100	0.25413	0.28713	0.28053	0.29373	0.29043

Fuente: Elaboración propia.

**Cuadro 13.7: Resultados del experimento – Parte 4**

Tt	Tb	Tf				
		10	20	30	40	50
8	10	0.53795	0.56106	0.56436	0.55776	0.55445
8	20	0.23102	0.15842	0.14851	0.14521	0.14851
8	30	0.18812	0.12871	0.11221	0.11881	0.10561
8	40	0.18812	0.11551	0.12541	0.12871	0.11881
8	50	0.16172	0.12541	0.10891	0.13861	0.12541
8	60	0.12871	0.11551	0.11551	0.12211	0.12541
8	70	0.10891	0.10561	0.09901	0.10561	0.12871
8	80	0.12211	0.10891	0.10231	0.13201	0.14521
8	90	0.15512	0.16832	0.17492	0.21122	0.22442
8	100	0.25413	0.28713	0.28053	0.29043	0.29373

Fuente: Elaboración propia.

Observamos que el menor  $MAPE^P$ , al cual denotaremos por  $MAPE^*$ , es 0.09901 y se produce al utilizar tres combinaciones de parámetros (ver Cuadro 13.8).

**Cuadro 13.8: Combinaciones de parámetros que minimizan el error**

<b>Parámetros</b>	<b>Combinación 1</b>	<b>Combinación 2</b>	<b>Combinación 3</b>
Escala	0.5	0.5	0.5
$T_f$	20	30	30
$T_t$	2	4	8
$T_b$	70	70	70
$\alpha$	0.2	0.2	0.2

Fuente: Elaboración propia

### 13.5. RESUMEN DEL CAPÍTULO

- Con el fin de determinar la eficacia del modelo de solución, se realizó una comparación entre los resultados del algoritmo utilizado (para cada combinación de parámetros) y las mediciones manuales descritas en el capítulo 4.
- Se estableció la combinación de valores de los parámetros que intervienen en la solución y que minimizan el error medio absoluto (MAPE), el cual es indicador estadístico ampliamente utilizado para medir la eficacia de métodos cuantitativos (R. Yaffee et al. [13]).

## CONCLUSIONES Y RECOMENDACIONES

### CONCLUSIONES

De acuerdo a los resultados obtenidos en la presente investigación, se formulan las siguientes conclusiones:

- El método confirma la hipótesis planteada en la presente investigación, debido a que ofrece un porcentaje promedio de error absoluto de 9.901%. De esta manera, se constituye en una alternativa confiable al conteo manual y se resuelve el problema descrito en el primer capítulo.
- Existen tres combinaciones de parámetros que minimizan el error medio absoluto: Escala = 0.5, Tf = 20, Tt = 2, Tb = 70 y  $\alpha = 0.2$ ; Escala = 0.5, Tf = 30, Tt = 4, Tb = 70 y  $\alpha = 0.2$ ; Escala = 0.5, Tf = 30, Tt = 8, Tb = 70 y  $\alpha = 0.2$ .
- El pre procesamiento del video facilita los cálculos de las etapas posteriores, disminuyendo el costo computacional requerido. Esto se logra con la reducción de las dimensiones del video y la simplificación del

mismo de RGB a escala de grises.

- Para llevar a cabo los estudios de volumen vehicular a través del método de solución descrito en la presente investigación, no es necesario el uso de algún dispositivo especializado en monitoreo de tráfico. La tecnología a emplear consiste en una cámara digital para la captura del video, el cual es utilizado posteriormente como input para el software que implemente el algoritmo de solución.
- No es necesaria la presencia de personal capacitado en la instalación, uso y mantenimiento de alguna tecnología especializada en monitoreo de tráfico. Con esto se logra disminuir costos tiempos de capacitación durante la ejecución de los estudios de volumen vehicular.
- Debido a la naturaleza del método planteado, el software no requiere ejecutarse en algún computador de grandes capacidades de memoria o procesamiento.

## RECOMENDACIONES

Según lo analizado durante el desarrollo de la presente tesis, se plantean las siguientes recomendaciones:

- Con el fin de tener mayor visibilidad, instalar la cámara de video en ubicaciones altas como semáforos o puentes. Considerar la colocación de algún material protector contra el viento, lluvias o temperaturas extremas.
- Determinar la duración de la batería de la cámara utilizada y realizar el cambio periódico de la misma. Asimismo, por motivos de mantener la calidad de los videos capturados, programar el mantenimiento físico de la cámara durante el periodo de estudio (limpiar lentes, revisar buen estado de funcionamiento, entre otros).
- Debido a que muchas cámaras dividen automáticamente los videos capturados en varios archivos digitales de igual tamaño, durante el diseño del software que implementará el algoritmo descrito se recomienda incluir la capacidad para procesar varios videos en lote, es decir, que al terminar uno, comience el procesamiento del siguiente de manera automática. Asimismo, implementar algún método de compresión de videos con el fin de reducir el tamaño necesario para el almacenamiento de los mismos.
- Con el fin de disminuir costos, utilizar herramientas de desarrollo libres

para la implementación del método de solución. Existen librerías gratuitas de programación para el manejo de videos.



## GLOSARIO DE TÉRMINOS

**BASE DE DATOS:** Es un conjunto de datos almacenados sistemáticamente para su posterior uso.

**CONTEO VEHICULAR:** Es un tipo de estudio en el que se determina cuántos vehículos pasaron por un punto específico de la vía en un intervalo de tiempo dado.

**EJE VEHICULAR:** Los ejes de un vehículo son las líneas transversales respecto de las cuales giran las ruedas cuando el vehículo avanza recto. Cada eje une los centros de un par de ruedas (una a cada lado del vehículo). Los ejes mantienen la posición relativa de las ruedas entre sí y éstas respecto al chasis del vehículo. Además, deben soportar el peso del vehículo y su carga adicional.

**FONAM:** Fondo Nacional del Ambiente. Es una institución de derecho privado sin fines de lucro encargada de promover la inversión pública y privada en proyectos ambientales en el Perú.

**FOTOGRAMA:** Es una imagen específica de una sucesión de imágenes que

forman parte de una animación o video.

**FPS:** Frames per second, o fotogramas por segundo. Es el ratio o frecuencia en el que un dispositivo produce imágenes (fotogramas).

**PÍXEL:** Es la menor unidad que forma parte de una imagen digital. Cada píxel se codifica mediante un conjunto de bits de una determinada longitud. Por ejemplo, si se codifica un píxel con un byte (8 bits), éste podrá admitir 256 ( $2^8$ ) colores (de 0 a 255).

**PROINVERSION:** Agencia de Promoción de la Inversión Privada. Es un organismo público ejecutor, adscrito al Ministerio de Economía y Finanzas, que promueve la incorporación de inversión privada en servicios públicos y obras de infraestructura.

**PROTRANSPORTE:** El Instituto Metropolitano PROTRANSPORTE de Lima es un organismo público descentralizado de la Municipalidad Metropolitana de Lima que se encarga de poner en práctica diversos proyectos de transporte público en dicha ciudad.

**RGB:** Son las siglas en inglés: Red, Green, Blue (rojo, verde, azul). RGB es un modelo de color con el que se puede representar un color a través de la adición de los tres colores primarios mencionados.

**RUIDO DIGITAL:** Es la variación aleatoria (que no tiene correspondencia con la

realidad) de los atributos de las imágenes digitales, como el brillo o color, producido por dispositivo de captura, por ejemplo una cámara.

**SENSOR INTRUSIVO:** Es aquel sensor que requiere la instalación de un dispositivo en la superficie de la calzada.

**SENSOR NO INTRUSIVO:** Es aquel sensor ubicado por encima, al lado o debajo de la calzada, de tal manera que no interfiere con el flujo del tráfico.

**VISIÓN COMPUTACIONAL:** Es una rama de la Inteligencia Artificial cuyo propósito es lograr la interpretación de imágenes digitales, capturadas mediante una cámara por ejemplo, y reconocer los diversos objetos en el ambiente y su posición en el espacio.

**720p:** Es un modo de video que indica 720 líneas horizontales de resolución. Tiene una resolución de 1280 x 720 píxeles.

## BIBLIOGRAFÍA

- [1] Alper Yilmaz, Omar Javed y Mubarak Shah. Object Tracking: A Survey. *ACM Computing Surveys*, 38(13):1-45, ACM, Estados Unidos, 2006. ISSN: 0360-0300.
- [2] Tao Yang, Stan Z. Li, Quan Pan y Jing Li. Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:970-975. IEEE, Estados Unidos, 2005. ISBN: 0-7695-2372-2.
- [3] Yiwei Wang, John F. Doherty y Robert E. Van Dyck. Moving Object Tracking in Video. *Proceedings of the 29th IEEE Applied Imagery Pattern Recognition Workshop*, p.95-101. IEEE, Estados Unidos, 2000. ISBN: 0-7695-0978-9.
- [4] Liyuan Li, Weimin Huang, Irene Y.H. Gu y Qi Tian. Foreground Object Detection from Videos Containing Complex Background. *Proceedings of the 11th ACM International Conference on Multimedia*, p.2-10. ACM, Estados Unidos, 2003. ISBN: 1-58113-722-2.

- [5] Clement Chun Cheong Pang, William Wai Leung Lam y Nelson Hon Ching Yung. A Method for Vehicle Count in the Presence of Multiple-Vehicle Occlusions in Traffic Images. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):1-19. IEEE, Estados Unidos, 2007. ISSN: 1524-9050.
- [6] Chomtip Pornpanomchai, Thitinut Liamsanguan y Vissakorn Vannakosit. Vehicle Detection And Counting From A Video Fotograma. *International Conference on Wavelet Analysis and Pattern Recognition*, 1:356-361. IEEE, Estados Unidos, 2008. ISBN: 978-1-4244-2238-8.
- [7] Manchun Lei, Damien Lefloch, Pierre Gouton y Kadder Madani. A Video-Based Real-Time Vehicle Counting System Using Adaptive Background Method. *IEEE International Conference on Signal Image Technology and Internet Based Systems*, p.523-528. IEEE, Estados Unidos, 2008. ISBN: 978-0-7695-3493-0.
- [8] H.S. Mohana, M. Ashwathakumar y G. Shivakumar. Vehicle Detection and Counting by Using Real Time Traffic Flux Through Differential Technique and Performance Evaluation. *International Conference on Advanced Computer Control*, p.791-795. IEEE, Estados Unidos, 2009. ISBN: 978-1-4244-3330-8.
- [9] Jianping Wu y Caidong Gu. The Design and Implementation of Real-time Automatic Vehicle Detection and Counting System. *International Conference on Information Engineering and Computer Science*, p.1-4. IEEE, Estados Unidos, 2009. ISBN: 978-1-4244-4994-1.

[10] A. Gyaourova, C. Kamath y S.-C. Cheung. Block Matching for Object Tracking. Lawrence Livermore National Laboratory, Estados Unidos, 2003

[11] AASHTO, FHWA y A. VANDERVALK-OSTRANDER. AASHTO Guidelines for Traffic Data Programs. AASHTO, Estados Unidos, 2009. ISBN: 978-1-5605-1054-3.

[12] Nicholas Garber y Lester Hoel. Ingeniería de tránsito y de carreteras, 3a. ed. Thomson., México, 2005. ISBN: 978-9-7068-6364-5.

[13] Robert A. Yaffee y Monnie McGee. Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS. Academic Press, Estados Unidos, 2000. ISBN: 978-0-12-767870-2

[14] FONAM, PROTRANSPORTE. Estudio para la Consolidación del Sistema Integrado de Transporte Público de Lima. PROTRANSPORTE, Perú, 2010.

[15] PROTRANSPORTE. Estudio de Preinversión a Nivel de Perfil Proyecto Construcción Corredor Vial de Transporte Público Masivo Este – Oeste Carretera Central – Av. Grau – Av. Venezuela, Provincia de Lima – Lima. PROTRANSPORTE, Perú, 2011.

[16] PROTRANSPORTE. Términos de Referencia del Estudio de Preinversión a Nivel de Factibilidad del Proyecto: "Construcción Corredor Vial de Transporte Público Masivo Este - Oeste Carretera Central - Av. Grau - Av. Venezuela, Provincia de Lima - Lima". PROTRANSPORTE, Perú, 2011.

[17] Pait Consultores. Levantamiento de Observaciones del Informe Final Estudios Definitivo de Arquitectura e Ingeniería del Patio Taller Sur. PROTRANSPORTE, Perú, 2009.

[18] William Callister. Introducción a la Ciencia e Ingeniería de los Materiales, Volumen 2. Reverte, Estados Unidos, 2002. ISBN: 978-8-42-917253-9.

[19] Agustín Castejón Oliva y German Santamaría Herranz. Electrotecnia Ciclos Formativos. Editex, España, 2009. ISBN: 978-8-49-771536-2.

[20] H. D. Mills. The Management of software Engineering Part I: Principles of Software Engineering. *IBM Systems Journal*, 19(4):1-65. IBM, Estados Unidos, 1980. ISSN: 18-8670.