

Capítulo I. Introducción.

1.1 Descripción del trabajo.

Cuando se obtiene un número de imágenes tomadas en un determinado tiempo y existe uno o varios objetos en movimiento dentro de las imágenes, o es la misma cámara la que se encuentra en movimiento, se puede obtener información importante como es el caso del número de objetos en movimiento dentro de la imagen, la dirección en que se mueven, si el movimiento es lineal o rotacional, que tan rápido se mueven o en que posición se encontrará en un determinado tiempo, entre otras.

Esta información puede ser utilizada para predicción de movimiento y colisión con obstáculos así como el análisis de comportamiento de objetos en escenas dinámicas.

Los métodos de análisis de movimiento se pueden dividir en dos grupos: métodos basados en el cálculo de gradiente espacio-temporal y basados en correspondencia de bloques en escenas dinámicas [Starostenko, et al., 2001].

La definición de la escena dinámica es posible después de la construcción del flujo óptico o el campo de movimiento, los cuales tienen toda la información necesaria para calcular las características de movimiento como velocidad, vector de desplazamiento, tiempo de propagación del objeto, etc.

En muchas ocasiones se confunde al flujo óptico con el campo de movimiento, sin embargo no siempre es este el caso, ya que el campo de movimiento es como se mueven los objetos y es un concepto puramente geométrico, mientras que el flujo óptico es un concepto basado en el cambio de intensidad del brillo entre imágenes. Ver Figura 1.1. El

campo de movimiento se puede calcular por varios métodos y una vez obtenido, se pueden contestar las preguntas mencionadas en el párrafo anterior [Marshall, 2002].

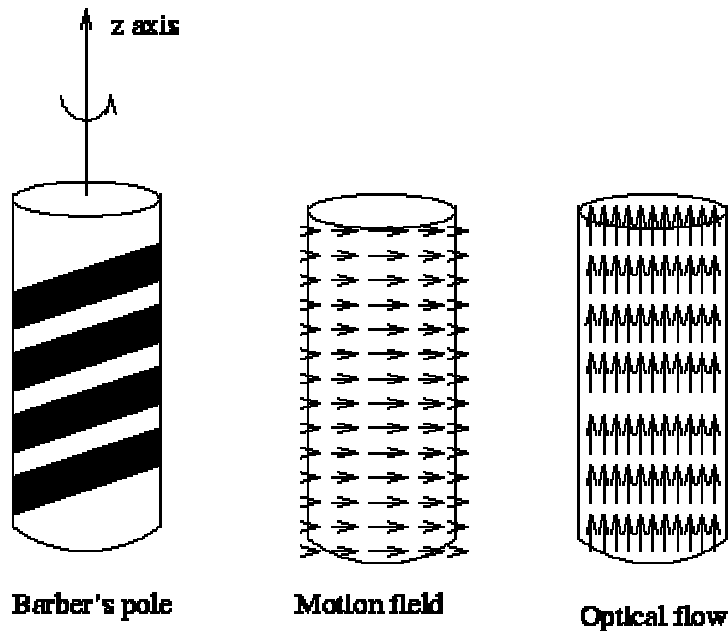


Figura 1.1 Diferencia entre campo de movimiento y flujo óptico.

Existen varios métodos para obtener la estimación de las características de movimiento entre las que se destacan las siguientes: método de acercamiento basado en concordancia (Matching-Based Approaches), método de acercamiento basado en gradiente (Gradient-Based Approaches), método SUSAN de detección de esquinas principales, método de estimación de características del campo de movimiento, método basado en regularización (Regularization-Based Algorithm), método basado en multirestricciones (Multiconstraint-Based Algorithm) [Starostenko, et al., 2001][Laplante, 1996].

Los métodos mencionados anteriormente serán analizados en esta tesis para evaluar su eficiencia y seleccionar aquellos que sean los más adecuados para cumplir con los objetivos de la misma.

1.2 Estado del problema.

1.2.1 Sistemas existentes.

Motion Mining: durante los últimos años un gran número de sistemas de computación han recaudado datos sobre el movimiento humano por medio de diversos dispositivos. Éstos sistemas generan un flujo de trayectorias en 3-D utilizadas para animación por computadora, efectos especiales, análisis biomecánico del hombre, vigilancia de las actividades humanas entre otras. Con el crecimiento de esta colección de datos sobre movimiento humano, existe la posibilidad de crear algoritmos que nos permitan mejorar el entendimiento del movimiento humano y así poder crear métodos asistidos por computadora para rehabilitación de problemas físicos, entrenamiento de deportes, medicina y diagnósticos. Incluso Motion Mining permitirá mejorar los algoritmos existentes de visión por computadora y de reconocimiento de patrones de movimiento. Este proyecto está un poco atrasado porque los algoritmos de indexación y de búsqueda en las bases de datos de movimiento necesitan información proporcionada por humanos, la cual puede llevar a cometer errores o no percibir algún movimiento que un sistema computacional reconocería. Para revisar video y grabar los eventos visuales del contenido de las bases de datos, se necesitan nuevas técnicas de visión, las cuales sean capaces de clasificar y predecir el

movimiento de objetos en una secuencia de imágenes, interpretar el lenguaje por medio de señales y detectar patrones de movimiento que son inusuales [Kollinos, et al.].

Safe and Precise Landing: Este sistema consiste en mantener la ubicación exacta de vehículos tanto móviles como aéreos. Para lograrlo investigadores han trabajado con el Interval Navigation System (INS), el cual integra las señales del Inertial Measurement Unit (IMU) y las actualizaciones periódicas del Global Position System (GPS) para obtener la posición del vehículo autónomo, sin embargo, han encontrado problemas cuando éste se encuentra en lugares cerrados, superficies que no reciben señales satelitales como la superficie de Marte o junto a edificios muy altos donde la señal del GPS no es la correcta. Investigaciones recientes sobre estimación de movimiento basado en visión permite estimar tanto un movimiento lineal como angular por medio de una sola cámara. Al utilizar métodos de estimación de movimiento se han encontrado con pequeños problemas como son el caso de no distinguir entre rotaciones y traslaciones pequeñas, pérdida de información en movimientos bruscos. Sin embargo, éstos problemas se han solucionado utilizando esquemas de estimación de movimiento basado en filtros combinados con la información del IMU [Montgomery, et al., 2003].

Derbis Mine-Clearance Vehicle: este vehículo está diseñado para eliminar minas principalmente en los campos agrícolas o de cultivo. Su funcionamiento se basa en patrones complejos de navegación y estimación de movimiento desde una secuencia de imágenes con el fin de mejorar la información de posición del vehículo. Debido a la complejidad de las imágenes que debe de procesar, ya que los campos de trabajo son de tierra, lodo, pasto, concreto entre otros, utiliza un método de creación de flujo óptico y campo de movimiento basado en correlaciones. Este método producía resultados de gran precisión pero a un muy

alto costo computacional, por lo que se cambió el método por el de correspondencia de bloques con algunos ajustes extras como son el caso de búsquedas en rotaciones, zoom y permitir una mayor cantidad de “frames” sin perder su exactitud. Y después de investigación sobre algunos otros métodos, el de correspondencia de bloque ha sido el que mejor resultados ha dado para ésta aplicación [Howorth, 2001].

Pipeline Vision System (PVS): Este sistema intenta probar la robustez de los sistemas de visión bajo el agua. Como se mencionó este proyecto se basa en un vehículo autónomo que viaje por debajo del agua y sea capaz de distinguir todas y cada una de las tuberías que se encuentran en la profundidad y ser capaz de seguir alguna de ellas. Para éste vehículo la robustez del sistema de visión es sumamente importante debido a que los humanos no pueden interactuar con el PVS y corregir los errores. Este sistema de visión tiene algunos problemas con los que se tiene que enfrentar por el simple hecho de tomar imágenes por debajo del agua, estos problemas son: el fondo del mar no es uniforme, lo cual puede producir errores en el procesamiento de imágenes, el vehículo cuenta con luz propia lo cual tiende a degradarse por las reflexiones del agua, identificación de objetos difíciles de identificar en el caso de sobre posición y la presencia de alga marina en las tuberías [Laplante, 1996].

Traffic Control Systems: últimamente tanto la industria automotriz como la población mundial ha crecido considerablemente a diferencia de la infraestructura necesaria para soportar la cantidad de vehículos, produciendo un incremento en el tráfico vehicular.

Debido a esto, varios países han lanzado programas con el fin de reducir el tráfico vehicular y aumentar la seguridad de los conductores. Algunos de estos programas son: The Mobility 2000 y Automated Highway System en los Estados Unidos, RACS y AMTICS en

Japón y DRIVE y EUREKA en Europa. Otro programa es PROMETHEUS (PROgramM for a European Traffic with Highest Efficiency and Unprecedented Safety) el cual fue iniciado por la industria automotriz en Europa y es un programa orientado al vehículo, a diferencia de DRIVE, el cual es soportado por la Comunidad Europea, que es orientado a la infraestructura.

La meta de PROMETHEUS es usar tecnologías informáticas y de telecomunicaciones para ayudar a los conductores. Por medio de un sistema de visión basado en correspondencia, PROMETHEUS informa al conductor de posibles riesgos, obstáculos frente al vehículo, si el vehículo está fuera de su carril y recomendaciones de velocidad según el sector de la carretera [Laplante, 1996] [Ahlgren, 1989].

1.2.2 Métodos existentes para la detección del campo de movimiento.

Como se mencionó con anterioridad, existen principalmente dos aproximaciones para la estimación del campo de movimiento, la primera es la basada en correspondencia y la segunda es la basada en gradiente. Dentro de esta última se destacan dos técnicas, la de regularización y la de multirestricción. Mientras que en la de correspondencia la más importante es la de correspondencia de bloques.

Existen tres problemas principales al momento del cálculo del campo de movimiento. El primero se refiere a la discontinuidad en el campo de movimiento ocasionado por la presencia de ruido en el brillo de la imagen. Este problema puede ser solucionado o parcialmente solucionado aplicando un operador Gaussiano de alisamiento. El segundo problema es la oclusión entre distintos objetos en la imagen, estos objetos pueden estar en movimiento con distintas velocidades o uno de ellos estático. Para

solucionar este problema se debe tener en cuenta los bordes de las oclusiones ya sea a priori o a posteriori del cálculo. El tercer problema se denomina de apertura y éste se presenta también en la visión humana y se refiere a ver un objeto de tamaño mayor a la apertura con la que se ve. En este problema sólo se puede determinar una aproximación a la velocidad real [Starostenko, et al., 2001] [Laplante, 1996].

El estudio de esta tesis se enfocará en la detección del campo de movimiento de un objeto en movimiento, así como la velocidad con la que se desplaza. Estas características del movimiento son importantes debido a que por medio de ellas se puede calcular la dirección del objeto y a su vez calcular la posición del mismo en un tiempo t determinado.

A continuación se presentan brevemente tanto los métodos más importantes dentro de los mencionados con anterioridad como también el método SUSAN de detección de esquinas principales.

Método basado en correspondencia.

Este tipo de métodos se basan en identificar un conjunto de características bien identificadas de un objeto móvil. Una vez que se tiene el conjunto de características se rastrean en el siguiente frame para determinar el movimiento de ese conjunto entre las imágenes en un determinado transcurso de tiempo.

Las características identificables se les divide en dos tipos: de alto nivel, por ejemplo formas o figuras, y de bajo nivel, por ejemplo puntos característicos como son esquinas, curvaturas o contrastes resaltantes de color.

Este tipo de técnicas son adecuadas para rastrear movimientos en espacios 3-D ya que las características a resaltar son partes importantes del objeto en movimiento. Sin

embargo, para poder rastrear las características de un objeto se necesita de un costo computacional previo, ya que se necesita una fase de procesamiento de la imagen. Dentro de esta técnica la más importante es la de correspondencia de bloques [Laplante, 1996].

- **Método de correspondencia de bloques.**

Este método consiste en tomar un segmento de la imagen como un conjunto de características a ser rastreado. Posteriormente la estimación del movimiento consiste en obtener un punto de coordenadas (x, y) el cual será el centro del conjunto a ser rastreado, posteriormente se busca una nueva coordenada (x', y') correspondiente al mismo centro del conjunto pero en una imagen posterior. Si la nueva coordenada (x', y') es la misma que (x, y) del centro original, esto indica que no hubo desplazamiento. De otra manera, se puede obtener las componentes de x y y del desplazamiento así como la velocidad del píxel en un tiempo Δt entre imágenes. La coordenada (x', y') se toma como el nuevo punto a ser buscado en la siguiente imagen, pero esto puede llevar a un problema llamado acumulación de error.

Este método tiene algunos problemas como es el caso del acumulación de error, el cual se puede reducir tomando como segmento de características puntos en esquinas u orillas. Generalmente el método de correspondencia de bloques utiliza secuencias pequeñas de imágenes, es muy sensible al ruido para lo cual se debe de aumentar el tamaño de la muestra a ser rastreada y funciona bien cuando el desplazamiento es pequeño, es decir, que la velocidad del objeto no sea muy alta. A cambio de esto, el hecho de actualizar en cada imagen la coordenada central de la muestra a ser rastreada le permita al método estimar movimientos de rotación [Laplante, 1996].

Método basado en gradiente.

Este tipo de métodos también proveen solución al problema de estimación de movimiento observando los cambios en el brillo de la secuencia de imágenes en el transcurso del tiempo. Estos cambios de brillo son modelados por ecuaciones diferenciales parciales a las que se les conoce como ecuaciones de restricción. Al campo de los vectores de velocidad obtenido al solucionar estas ecuaciones de restricción se le llama flujo óptico o campo de movimiento.

La ecuación diferencial parcial más importante para determinar el campo de movimiento se obtiene al considerar los cambios en el brillo de las imágenes. Estas ecuaciones también se pueden representar como la ecuación de una línea donde todos y cada uno de los puntos que pasan por la línea recta de la ecuación son posibles soluciones y no ofrece una solución única, por esto se necesita de restricciones de bordes y suavización para obtener una solución computacional.

Este método no está exento de problemas en la estimación del campo de movimiento, con el se presenta la llamada apertura, solamente para los puntos donde se pueden definir dos restricciones lineales que no presentan este problema, tal es el caso de las esquinas. También es sensible al ruido, el cual permite obtener varias soluciones y a la discontinuidad en el brillo de las imágenes [Laplante, 1996]

- **Método basado en regularización.**

En estos algoritmos se define una función donde la restricción de suavización, la cual se altera por medio de un peso (una constante positiva), se utiliza para regularizar la solución de la ecuación diferencial parcial. Posteriormente esta función es minimizada ya sea por medio de cálculo de variaciones o por relajamiento estocástico. Básicamente estos métodos llevan a soluciones iterativas.

Dentro de éste método de estimación del campo de movimiento, la más famosa es la propuesta por Horn y Schunck, la cual propone minimizar una función que consta de dos factores, el primero se refiere a la función del campo de movimiento y la segunda es la función de suavización multiplicada por el peso de la misma. La minimización de la formulación de Horn y Schunck se lleva a cabo por medio de cálculo de variaciones el cual lleva a un sistema de dos ecuaciones diferenciales parciales [].

Existe otro método dentro de regularización que fue propuesto por Fennema (Determinación de velocidad en las escenas con varios objetos en movimiento). Este método se aplica en traslaciones rígidas sin rotación. En este método la variación de intensidad con el tiempo y el gradiente espacial aplicado para cada punto de la imagen genera la velocidad y dirección del objeto en ese punto, pero aun con esta información no se puede obtener el vector único de desplazamiento por lo que se aplica la técnica de agrupamiento por la velocidad común. Este método trabaja bien pero solo cuando los objetos se mueven linealmente [Laplante, 1996].

- **Método basado en Multirestricciones.**

En estos métodos para poder estimar el campo de movimiento por medio de la función general, se tiene que determinar un sistema de ecuaciones de restricción. Este sistema de ecuaciones puede ser resuelto tanto por mínimos cuadrados como por razonamiento probabilístico.

Estos métodos son clasificados en base a la aproximación adoptada para construir el sistema de ecuaciones de restricción y las técnicas matemáticas para la solución de éste. Existen tres aproximaciones principales.

La primera se refiere a la hipótesis de lo estacionario, por lo que se supone que debe ser válida para cualquier función invariante de movimiento. En esta aproximación, se permite definir varias ecuaciones diferenciales parciales de campo de movimiento. Para la segunda aproximación se obtienen nuevas ecuaciones al tomar las derivadas de las ecuaciones fundamentales del campo de movimiento. Por último, la tercera aproximación asume que el cambio del campo de movimiento es lineal, lo cual solo es válido cuando dicho campo es suavizado.

Como consecuencia de lo anterior, la calidad en la estimación del campo de movimiento depende fuertemente de la elección de ecuaciones de restricción [Laplante, 1996].

1.2.3 Método SUSAN de detección de esquinas principales.

Por sus siglas en inglés Smallest Univalued Segment Assimilating Nucleus. Este algoritmo fue propuesto por Smith, [Smith, 1997] y se basa en el método de detección de esquinas

principales de los objetos en una imagen y cubre los problemas como son filtración de ruido en las imágenes, búsqueda de bordes y búsqueda de esquinas. Una vez que se obtienen los puntos principales en la secuencia de imágenes, se pueden sobreponer dos imágenes con sus puntos característicos para determinar si hubo movimiento o no.

El principio del algoritmo SUSAN para detectar los detalles de la imagen es la máscara circular la cual se posa sobre el núcleo o píxel de interés. Si el brillo de cada píxel bajo la máscara tiene el mismo valor del núcleo de la máscara entonces a esa área se le llama USAN (Univalued Segment Assimilating Nucleus) y esta área contiene la información de la región bajo la máscara. El área USAN está en un máximo cuando el núcleo está en la superficie plana de la imagen, está en un medio cuando el núcleo está muy cerca de un borde y en un mínimo cuando el núcleo está en una esquina. El área USAN se utiliza para detectar la presencia de bordes en figuras de dos dimensiones. [Starostenko, et al., 2001].

Dado que este algoritmo no utiliza derivadas de la imagen para la detección de bordes, es un método no tan susceptible al ruido.

El procedimiento de detección de bordes consiste en:

- Poner la máscara circular sobre el píxel de interés y calcular su valor por medio de la función de distribución de Gauss
- Los valores obtenidos bajo la máscara se suman
- Examinar las áreas USAN para detectar los puntos que son bordes.

1.3 Objetivo general.

- Desarrollo de algoritmos de detección de movimiento de objetos en escenas dinámicas invariantes a rotación desde una secuencia limitada de imágenes y construcción del campo de movimiento basándose en métodos de gradiente espacio-temporales y el método de correspondencia de bloques.

1.4 Objetivos específicos.

- Analizar las ventajas y desventajas de métodos existentes, así como una comparación entre ellos.
- Generar al menos dos algoritmos que intenten tomar las ventajas de los métodos existentes para así mejorar el funcionamiento de los mismos.
- Estudiar los métodos más recientes para la construcción de campo de movimiento así como los métodos de estimación de movimiento en una secuencia de imágenes limitada.
- Los algoritmos que se propondrán deberán de ser capaces de detectar cuando más un objeto que se encuentre en movimiento, pero no solamente que siga un movimiento lineal, sino que también sea rotacional.
- Los algoritmos serán capaces de detectar el vector de velocidad de el objeto que se encuentre en movimiento.
- Desarrollo de una aplicación para la estimación del campo de movimiento.
- El sistema será desarrollado como una aplicación web para aprovechar las ventajas de los sistemas distribuidos.

1.5 Alcances y limitaciones.

1.5.1 Alcances.

- Diseño de un software de prueba de métodos existentes para detectar sus fortalezas y debilidades.
- Diseño de un algoritmo de detección de características de movimiento basado en el método de gradiente.
- Diseño de un algoritmo de detección de características de movimiento basado en el algoritmo de detección de puntos esquina SUSAN.
- Implementación de ambos algoritmos.
- Implementación de interfaces accesibles vía Web para interactuar con los métodos implementados.

1.5.2 Limitaciones.

- Se tocarán muchos de los métodos para solucionar el problema, sin embargo, solo se probará el funcionamiento de algunos de los métodos más recientes.
- La calidad de los resultados de los algoritmos dependerá de la calidad de imágenes tomadas desde la cámara de video o las imágenes proporcionadas.
- Los algoritmos implementados deberán detectar a lo más un objeto rígido en movimiento.

1.6 Software y hardware.

1.6.1 Software.

Actualmente existen muchos sistemas para el análisis de imágenes digitales, entre los que destacan ATIC, HALCÓN y NeatVision. Tanto HALCÓN como NeatVision permiten programar subrutinas o incluso modificar operaciones ya implementadas, pero debido a las características que ofrece NeatVision y a su facilidad de uso será el software de análisis que se utilizará en esta tesis. A continuación se presenta una lista del software a utilizar para el desarrollo de la tesis así como la justificación del porque se utilizará.

- NeatVision es un software de visión por computadora basado en Java el cual es muy amigable y al mismo tiempo potente, además que permite al usuario programar sus propias rutinas, es por eso que se utilizará para probar los métodos existentes.
- Muchas de las implementaciones de los métodos de visión están programados en C debido a la rapidez del lenguaje, pero debido a las facilidades que proporciona Java y a que es muy portable, ya que se trabajará en máquinas con diferente sistema operativo, el lenguaje que se utilizará para las implementaciones de los algoritmos será Java 1.4.2.
- Apache Tomcat 4, ya que el sistema se basará en la arquitectura Cliente/Servidor utilizaré éste servidor web que además de ser potente es de código abierto, lo que quiere decir que es un software gratuito.

- Para guardar tanto los datos como las imágenes de prueba y las imágenes resultantes, será utilizado el manejador de bases de datos de código abierto más rápido que existe en la actualidad, que es MySQL.
- Visual Paradigm for UML Community Edition será el software utilizado para la construcción de diagramas UML en la generación del diseño del sistema.
- Excel es un software que permite hacer cálculos fácilmente, genera documentos fácilmente exportables y permite hacer gráficos fácilmente, es por esto que se utilizará para generar gráficas comparativas.
- Word lo usaré para generar los reportes ya que es un sistema que cumple con las necesidades del proyecto y es muy fácil importar documentos de Excel.

1.6.2 Hardware.

- Computadora personal con procesador Intel Celeron 2.0 Ghz, 256 Mb en RAM y 32 Mb de video.
- Computadora Sun con software para servidor Sun Ray, ambiente operativo Solaris 2.6.