

Capítulo I.

Introducción.

La visión humana juega un gran papel en la vida del hombre. Su desarrollo ha pasado por un largo proceso evolutivo hasta llegar al estado de hoy. La evolución en sí es algo dinámico y cambiante, entonces la visión no podría existir sin diferentes movimientos que hay alrededor del ser humano. Como dice Jain [Jain, 1995], la mayoría de los sistemas biológicos de visión se han desarrollado con el mundo cambiante. Muchos científicos tratan de descubrir la naturaleza de la visión humana investigando cómo se perciben y procesan los objetos. Por ejemplo en una teoría se dice que durante la percepción o reconocimiento visual los ojos se mueven y se fijan en las partes más informativas de la imagen. Ellos realizan la selección orientada al problema y procesan la información del mundo visual bajo el control de la atención visual [Rybak]. Basándose en varias teorías de visión humana los investigadores construyen diferentes técnicas que con el tiempo se vuelven más y más sofisticadas al encontrar nuevas ideas, desarrollar nuevos métodos. Todas estas técnicas entran a un área llamada visión por computadora.

La visión por computadora incluye muchos campos de desarrollo. Pero en general es un área que usa las técnicas para estimar los detalles en imágenes, relacionar las mediciones de detalles a la geometría de objetos en espacio e interpretar la información geométrica. La meta del sistema de visión por computadora es crear un modelo del mundo real desde las imágenes. El sistema cubre la información útil sobre la escena desde su proyección en dos dimensiones [Jain, 1995]. Una parte importante de la visión por computadora ocupa el procesamiento de imágenes. Es el campo donde sus técnicas transforman la imagen a otra para obtener la información necesaria. Sus algoritmos son útiles en las primeras etapas del área de visión por computadora porque normalmente se usan para mejorar la información en particular y suprimir el ruido. Las imágenes usadas se

obtienen de una cámara y después se aplica algún método para obtener la información necesaria. En general la ingeniería de imagen incluye [Schalkoff, 1989]:

- Análisis de señal para el preprocesamiento.
- Geometría para extraer los detalles.
- Álgebra lineal para representar la geometría de imagen.
- Estimación de movimiento, orientación.
- Reconocimiento de patrones estadístico para la clasificación de objetos.
- Reconocimiento de patrones sintáctico para la descripción de estructura de imagen.

Como el hombre vive en el mundo cambiante, el sistema de visión por computadora tiene que trabajar con los objetos en movimiento o cambio, iluminación cambiante, y el cambio de puntos de vista que es esencial para procesar diferentes tareas [Jain, 1995]. Todo lo que nos rodea y está en la vista forma la escena dinámica. Este concepto también se usa en el área. La entrada al análisis de escena dinámica es la secuencia de imágenes tomada del mundo cambiante. La cámara que se usa para obtener esta secuencia también puede estar en movimiento. Cada imagen representa la escena en una instancia particular del tiempo. Los cambios en la escena pueden proceder del movimiento de la cámara, de objetos, cambios en iluminaciones, o cambios en la estructura, tamaño, forma de objeto. El sistema tiene que detectar los cambios, determinar las características del observador y los objetos en movimiento, caracterizar los mismos usando la abstracción de nivel alto, explorar la estructura de objetos y reconocerlos [Jain, 1995].

El análisis de secuencia de imágenes ha recibido más y más atención desde los años setenta. Intentos de investigación fueron motivados por las necesidades metodológicas y por el crecimiento de aplicaciones basadas en imágenes espacio temporales. Estas imágenes contienen la información importante del ambiente, del mundo físico incluyendo la profundidad, orientación de superficie, estructura y movimiento de objetos en espacio, y movimiento de sensores [Bouthemy, 1989]. En el contexto del análisis de la escena dinámica los esquemas del entendimiento de la visión tienen que relacionarse con sus propias herramientas para la segmentación espacio temporal y el análisis de movimiento.

Como existen diferentes tipos de movimiento, hay varias clasificaciones del mismo. En seguida se describe una, presentada por Zhou [Zhou, 1992]. Él enfatiza el movimiento

como estéreo que es el método para derivar la profundidad de información de cámara en movimiento u objetos en movimiento a través del ambiente estacionario 3D. Puede ser dividida en tres categorías: giratorio, lateral y longitudinal. El movimiento giratorio describe la información de profundidad desde la cámara u objetos en rotación. El movimiento lateral - desde el movimiento lateral de la cámara u objetos. El movimiento longitudinal - desde el movimiento de acercamiento o alejamiento de la cámara u objetos.

Junto con los tipos de movimiento, se destacan tres diferentes campos de vectores de velocidad [Laplante, 1996]: campo de velocidad de objetos, campo de movimiento de imagen, campo de flujo óptico. Cada punto de objeto en la escena tiene su vector de tres dimensiones de movimiento - es el campo de velocidad. Proyectando este campo en el plano da el campo de movimiento de la imagen. El campo de flujo óptico representa los cambios en la intensidad de la luz en la imagen. Normalmente el flujo óptico y el campo de velocidad son diferentes, pero si en dado caso son suficientemente similares entonces sus propiedades cualitativas son iguales [Verri, 1989]. Puede existir más de una velocidad legítima en un marco [Fleet, 1992]:

- Superficies virtuales causadas por el reflejo.
- Sobreposiciones de una superficie con la otra.
- Transmisión de luz a través de superficie.
- Sombras y luminosidad.
- Efectos atmosféricos.

Bajo diferentes tipos de movimiento y velocidades existentes en la escena la estimación del campo de movimiento resulta afectada por tres dificultades principales [Laplante, 1996]:

- Discontinuidades generadas por el ruido en el brillo de la imagen.
- Oclusiones entre diferentes objetos en movimiento que normalmente tienen diferentes velocidades entre los objetos en movimiento y el fondo.
- Problema de apertura. Se refiere a la ambigüedad en determinar el campo de movimiento en los bordes rectos y monótonos. Cualquier componente del campo de velocidad de objeto en la escena que es paralelo a la dirección del borde no puede ser determinado como el flujo óptico. Sólo el componente de velocidad perpendicular al borde se puede detectar.

También el problema podría presentarse en la variabilidad de cámaras existentes por el brillo y luminosidad [Russ, 1992].

Para resolver estos problemas u otros que podría presentar la estimación de movimiento se desarrollan las diferentes técnicas y métodos que se basan principalmente en gradiente y correspondencia. Aunque estos dos campos son muy usados también existen otras técnicas basadas en redes neuronales artificiales y en la obtención de fases utilizando filtros. La mayoría de las aproximaciones construyen el campo de flujo óptico estimando las propiedades cualitativas y cuantitativas del movimiento.

La idea principal de la tesis es describir y analizar las aproximaciones existentes con sus aplicaciones referentes en el área de visión dinámica por computadora para después según la investigación, implementar y probar dos métodos de utilidad par aplicar en otras aproximaciones o en alguna aplicación en específico. El primer método se basa en gradientes donde se multiplican gradiente espacial y temporal [Jain, 1995] y el otro en el filtro SUSAN que encuentra las esquinas principales [Smith, 1997] para detectar los bordes de objetos en movimiento dentro de la escena. Estos dos métodos reducen la cantidad de datos si después se va a construir el campo de flujo óptico en los bordes de objetos en movimiento o en las esquinas significativas.

La tesis se estructura en cinco capítulo. El primer capítulo es la parte introductoria; el segundo habla de las técnicas relacionadas con el área de visión dinámica por computadora; en el tercer capítulo se describen las aplicaciones existentes del área; en cuarto se habla sobre los métodos implementados; y el quinto es de conclusiones de la tesis y trabajos futuros.

1.1. Objetivos, alcances y limitaciones.

Objetivo principal.

Describir y analizar las aproximaciones existentes con sus aplicaciones referentes en el área de visión dinámica por computadora para después, según la investigación, implementar y probar dos métodos de utilidad para aplicar en otras aproximaciones o en alguna aplicación en específico.

Objetivos específicos.

- Describir y analizar las aproximaciones de los grupos basados en gradiente, de correspondencias, de filtración espacio temporal y basado en redes neuronales.
- Mostrar aplicaciones del área que utilizan los métodos adoptados para cada una en específico y proponer otras soluciones a los problemas.
- En base de investigación implementar dos métodos basados en gradiente [Jain, 1995] y en algoritmo SUSAN [Smith, 1997] que encuentran los bordes de objetos en movimiento con su función principal de reducir la información a procesar en etapas posteriores.

Alcances.

- Investigación de las aproximaciones de cada grupo en particular.
- Aplicaciones del área con sus modos adoptados y comentarios sobre las soluciones alternativas.
- Implementación de dos métodos basados en gradiente [Jain, 1995] y en algoritmo SUSAN [Smith, 1997] que detectan los bordes de objetos en movimiento:
 - trabajo con las imágenes reales;
 - secuencias de tiempo distinto entre imágenes;
 - forma de objetos arbitraria.

Limitaciones.

- Manejo de imágenes usando niveles de gris.
- A lo más dos objetos móviles en la escena.

Hardware.

PC Pentium 1.0, sistema operativo Windows NT 4.0

Software.

Halcon 5.0, Visual C++ 5.0 Enterprise Edition.