

Capítulo III.

Aplicaciones.

Los resultados de estimación de movimiento se usan para diferentes aplicaciones como navegación de robot, seguimiento de objetos, entre otras. Uno de los requerimientos principales para los algoritmos es ser efectivos y adecuados en la velocidad sin olvidar la precisión [Liu, Camus, 1998]. Por un lado el algoritmo puede ser muy preciso pero con muy baja velocidad no aceptada por la aplicación. Por el otro lado el algoritmo puede correr muy rápido pero su precisión no es suficiente para dar los resultados esperados. Los dos precisión y eficiencia son importantes para el desarrollo de algoritmos aplicados al mundo real. A veces, las investigaciones recientes de movimiento toman estas dos aproximaciones en direcciones opuestas. Unas rechazan todas las consideraciones de eficiencia para la precisión más alta posible, y otras tratan de encontrar las soluciones rápidas olvidando de la precisión.

En este capítulo presentaré algunas aplicaciones del área de visión dinámica por computadora en tiempo real tomando en cuenta la precisión necesaria. La primera aplicación se trata de ayudar al conductor a manejar dándole los mensajes de advertencia o haciendo unas tareas asignadas. La segunda aplicación se trata del sistema de clasificación de imágenes de objetos motivado en biología. En la tercera aplicación se presenta el sistema visual bajo el agua aplicada a detección de los tubos. Y en la cuarta se trata de detectar la orientación de la vista del ojo humano.

3.1. Aplicaciones automovilísticas.

En los últimos 25 años los cambios profundos han estado en la transportación de los bienes y la gente como el incremento del tráfico continuo. Diferentes países se lanzaron a la investigación de poder reducir el tráfico congestionado e incrementar la seguridad de conductores. El programa PROMETHEUS (PROgram for a European Traffic with Highest Efficiency and Unprecedented Safety) [Laplante, 1996] es desarrollada por la industria automovilística en Europa en cooperación con la industria electrónica y los componentes de vehículos, con los consultantes de administración de tráfico, y con las investigaciones básicas en universidades y institutos. PROMETHEUS es el programa orientada a los vehículos, sobre todo a su infraestructura. La meta principal de PROMETHEUS es, usando las tecnologías informativas y telecomunicativas, soportar a los conductores, optimizar la efectividad de herramientas de navegación, y ligar el tráfico del camino con los otros sistemas de transporte al beneficio de todos. Entre las técnicas que han explorado para alcanzar la meta, la mayor atención fue prestada a la visión por computadora. El interés en la visión por computadora en aplicaciones automovilísticas es el objeto de las muchas discusiones y soluciones contradictorias. La meta principal es hacer un asistente para el conductor que le diera los mensajes de advertencia, y no excluya del manejo. Las razones de esta opción son técnico, legal y funcional.

Interés especial presenta el sistema de visión que informa al conductor de posibles situaciones peligrosas o da la información útil para la navegación. Este tipo de advertencias es el resultado de datos integrados como el dispositivo del circuito cerrado de la cámara con diferentes fuentes de información (sensores infrarrojos, las bases de conocimiento del mapa de camino). Como ejemplos de información útil que puede darse al conductor son [Laplante, 1996]:

- advertencias sobre las reglas del tráfico;
- advertencias de obstáculos enfrente de un vehículo;
- información sobre la posición del vehículo en el carril;
- información sobre la segmentación del camino;
- recomendaciones sobre la velocidad antes de dar la vuelta o rebasar.

Esta información se puede computarizar por un conjunto de heurísticas y algoritmos que usan la cantidad diferente del tiempo computacional. Como al conductor hay que informar a tiempo es importante decidir qué clase de información puede ser computarizada. Por ejemplo la información sobre los segmentos del camino y de la posición del vehículo en el carril puede ser determinada usando el sistema de detección de bordes.

El hardware que se usa para el sistema se llama PAPRICA que se basa en el concepto de correspondencia derivado de la transformación ganar - perder. También incluye otros operadores morfológicos como retraso, erosión, complemento, translación y transposición.

3.1.1. Detección de bordes en el camino.

Existen dos algoritmos principales al detectar bordes de la secuencia de imágenes: extracción de detalles basada en la calibración de cámara y en el modelo de manejo [Laplante, 1996].

3.1.1.1. Extracción de detalles basada en la calibración de cámara.

Aquí entra el conjunto de los filtros clásicos incluyendo el agrupamiento, detección de bordes y correlación. Como la aplicación requiere realizarse en tiempo real, es necesario mejorar la velocidad de desempeño. Los detalles a extraer son las líneas rectas determinadas por las discontinuidades del brillo y apuntadas al enfoque de expansión de la imagen original. Para extraer este tipo de información se usan los siguientes filtros:

- extraer las discontinuidades del brillo;
- retener sólo las líneas rectas que apuntan al enfoque de expansión;
- descartar los segmentos cortos, normalmente generados por el ruido.

La secuencia primero pasa al agrupamiento, después para detectar los bordes por el filtro basado en gradiente, y al final se aplica la umbralización y la afinación para obtener la imagen binaria compuesta por segmentos con el ancho de un pixel. Después de terminar la

primera parte la secuencia pasa a otra etapa donde primero se determinan las curvaturas locales de líneas siguiendo por el operador de umbralización para dejar sólo las líneas rectas. El algoritmo de detección de detalles se obtiene a través de correlación con la imagen sintética que contiene los conocimientos sobre la posición del enfoque de expansión.

3.1.1.2. Extracción de detalles basada en el modelo de manejo.

La imagen obtenida del vehículo en movimiento puede pertenecer a las categorías como: el conjunto de las calles, obstáculos, otros vehículos y sombras del fondo específico dependiendo del ambiente. Para procesar información tan diversa se aplica el algoritmo de segmentación que restringe el área de búsqueda empezando desde los conjuntos estándares. El modelo que contiene los conocimientos previos sobre los detalles a extraer está codificado en la estructura de datos tradicional y se usa en el procesamiento del nivel bajo. Este modelo se llama la imagen sintética. Las imágenes naturales y la sintética no pueden ser comparadas directo debido a muchos detalles adicionales de la primera. Para detectar los detalles necesarios se usa la estructura piramidal de diferentes resoluciones de la misma imagen. En el primer paso se usa filtración para reducir la influencia del ruido, los detalles redundantes y la distorsión. La imagen se parte en el conjunto de cuadrados de cuatro pixeles cada uno. El filtro consiste en sacar la media entre los valores de pixeles pertenecientes al mismo cuadrado. El conjunto de los valores resultantes forma una imagen muestra. Cuando la resolución mínima es alcanzada, se mejoran los bordes de la imagen original usando agrupamiento. Después se comparan la imagen obtenida y de la muestra. Los detalles codificados en la imagen sintética se adoptan a la imagen natural usando el algoritmo iterativo. Después los resultados pasan a la resolución más alta y el proceso se repite hasta llegar a la resolución final de la imagen original.

En detectar los bordes de objetos en movimiento, a parte del método que se utiliza aquí, también se podría aplicar el método basado en gradiente [Jain, 1995] que hace un AND entre dos gradientes: espacial y temporal (ver 2.1.3). El método no requiere ningún

preprocesamiento de la imagen y puede detectar los bordes de objetos con desplazamiento grande entre dos marcos de la secuencia.

3.1.2. Computación de campos de flujo óptico.

Como el flujo óptico asocia la velocidad de dos dimensiones a cada punto del plano de la imagen, probablemente es mejor determinar primero el flujo óptico de la imagen y después construir las formas basadas en ello [Laplante, 1996]. Entre dos marcos de la secuencia el flujo óptico representa el desplazamiento del mismo detalle entre dos imágenes. Los detalles se obtienen de los patrones locales o se determinan en el nivel de píxeles. El resultado es el campo donde cada posición de la imagen contiene el desplazamiento de detalles entre dos marcos. La correspondencia de detalles se busca entre los vecinos de cada píxel. Si hay más de una correspondencia se aplica el criterio que remueve la ambigüedad.

En computación de campos de flujo óptico también se podrían utilizarse las técnicas de correspondencia de bloques (ver 2.2.1). Son fáciles de implementar, y como para las aplicaciones automovilísticas no se requiere el procesamiento de nivel bajo entonces no hay posibilidad de errores grandes de estas técnicas.

Aquí en las aplicaciones automovilísticas los datos se procesan en dos alternativas. En la primera se detectan los bordes en movimiento en el camino y en la segunda se obtiene el campo de flujo óptico. Estas dos alternativas se podría combinar en una que usaría el método "Extracción de los bordes en movimiento e implementación en el tiempo real" propuesto por Liu [Liu, 1998] (ver 2.1.2.1). Entonces el resultado sería el campo de flujo óptico de los bordes en movimiento que es más preciso y eficiente en comparación con alternativas separadas.

El manejo de los coches es la tarea compleja en la cual el conductor actúa como el punto central. Los errores no pueden ser totalmente excluidos y a veces pueden llevar a situaciones peligrosas. El sistema asistente para el conductor ayuda a reducir los errores

dando mensajes de advertencia al conductor. Existen diferentes niveles de funcionamiento para este tipo de sistemas [Laplante, 1996]:

- Sistemas informativos que proveen a los conductores mensajes de advertencia sobre las situaciones específicas, las cuáles el conductor interpreta rápido, procesa y transforma a una acción apropiada.
- Sistemas de servicio que alivian la realización de las acciones iniciadas por el conductor.
- Sistemas de intervención automática que limitan las consecuencias de alguna irregularidad de mecanismos, controlados para mantener su reacción entre las esperas del conductor.
- Sistemas automáticos que ejecutan las acciones del conductor según sus reglas establecidas.

Las limitaciones de tiempo real son fundamentales para las aplicaciones automovilísticas para poder responder a tiempo al conductor.

3.2. Sistema de clasificación de imágenes de objetos motivado en biología.

El entendimiento de la visión natural y la cognición es el problema que se reconoce como un gran desafío para la ciencia y tecnología. Para su solución se juntan los científicos de diversos áreas como neurofisiología/neuroanatomía, psicología, biofísica y ciencia computacional. La simulación por computadora toma un papel importante en las investigaciones como encontrar las explicaciones y poder predecir los fenómenos de la mente humana. Este tipo de simulaciones puede proveer la base para las pruebas de teorías competitivas de percepción, desarrollo, aprendizaje y cognición entre otras cosas.

La organización de la red neuronal es la propiedad importante del sistema visual. El sistema artificial, que se espera a tener las propiedades perceptibles y cognitivas similares al sistema natural, tiene que ser capaz de organizarse propiamente. Los modelos de la red neuronal artificial (ANN) incorporan el aprendizaje no supervisado, como por ejemplo los de Kohonen [Laplante, 1996]. Su red consiste en dos capas. En la primera capa entran los

canales de filtración de la imagen. La salida de primera capa se usa para la entrada de la segunda en la cual los agrupamientos de la salida corresponden a las clases de objetos. En la fase de aprendizaje el conjunto de las imágenes generadas por los filtros se presenta como el conjunto de las redes de primera capa que son iteradas a través del procesamiento de aprendizaje hasta la convergencia y organización propia. Este sistema es capaz de aprender diferentes modelos visuales y clasificarlos. El efecto de los filtros es mejorar las transiciones de luminosidad y escala dadas. Al calcular las imágenes con los filtros existe cierta redundancia que puede ser omitida en usar las orientaciones cruzadas e inhibición lateral. Este tipo de mecanismos es bastante útil y corresponde a las acciones naturales en la visión, pero su gran desventaja es el requerimiento de gran tiempo computacional.

Una vez que el sistema aprende por los modelos es capaz de construir sus propios representaciones [Laplante, 1996]. El detalle importante de este sistema es no incorporar ningún modelo de conocimiento sobre los objetos que tienen que ser aprendidos y después clasificados. Las unidades de salida del sistema de clasificación corresponden a las clases de objetos y no a los objetos particulares. El aspecto útil de la red de Kohonen es poder asociar cada nodo de salida con el patrón particular de entrada.

En base de este método [Laplante, 1996] el sistema tiene que estar puesto en un lugar donde el objeto se puede capturar bien por la cámara y después decir al sistema que lo memorice. La acción empezará en grabar diferentes imágenes del objeto, calcular los conjuntos correspondientes de los filtros y después aprender los patrones para la clasificación. En la tarea de clasificación, cuando el mismo objeto o similar está puesto enfrente de la cámara y se requiere que el sistema lo clasifique, lo va a grabar en diferentes imágenes, construir los conjuntos de los filtros y usándolos tratar de asignar al objeto su clase correspondiente.

Aquí en la clasificación de imágenes de objetos se podría usar el método de Zhou [Zhou, 1992] que primero obtiene las tendencias de datos de entrada y después las procesa en una red neuronal (ver 2.4), en este caso clasificando la imagen de objeto.

3.3. Sistema de visión bajo el agua aplicada a detección de tubos.

Se desarrolla el PVS (Pipeline Vision System) [Laplante, 1996] para encontrar la posición y los límites de todos los tubos alrededor del sistema con la cámara. Se usa el ciclo de control para mover el vehículo al tubo escogido y después mantenerlo durante la inspección. PVS está diseñada para trabajar bajo el agua, en los lugares incontrolables por el hombre. Existen seis problemas básicos relacionados con éste:

- El fondo es no uniforme y puede dar los resultados falsos en el procesamiento de imágenes.
- La intensidad de la luz va a variar lentamente en las áreas cubiertas y va a cambiarse muy brusco en los bordes.
- Los vehículos tienen que traer su propia luz debido a la profundidad del agua.
- En usar la fuente de luz la calidad de imágenes se degrada por las reflexiones del agua.
- Los modelos de tubos tienen que ser tomados del mapa y estar disponibles para el sistema.
- Existencia de objetos que son difíciles de interpretar, sobreposiciones de los objetos.

El sistema PVS tiene que tratar de resolver estos problemas.

3.3.1. Segmentación.

Segmentación es la base para el análisis de imagen [Laplante, 1996]. Los bordes de tubos tienen que ser usados para el reconocimiento. Ellos pueden ser distorsionados pero se asume que tienen que ser distinguidos por su contraste con el lecho marino. La ventaja de detectar los bordes es la rapidez y el análisis de nivel alto porque el número de los puntos significativos es al menos dos ordenes de la magnitud menor que el número original de los píxeles de la imagen. La detección de bordes empieza al encontrar los puntos significativos en la imagen que pueden ser agrupados. Bajo este algoritmo hay dos problemas:

- El contraste en la imagen puede variar lo que dificulta al encontrar el umbral adecuado para discriminar entre los píxeles significativos y los que no son.
- El algoritmo puede parar al perder los puntos significativos.

Para resolver estos problemas se propone el siguiente proceso [Laplante, 1996]: etapa de preprocesamiento, detección de los contornos (puntos significativos), detección de la línea por la transformada de Hough modificada.

3.3.1.1. Preprocesamiento.

Se usa la lente de auto-iris para preservar la intensidad de la luz en la escena. La dificultad con la lente es capturar la forma y la imagen detalladamente en la apertura alta. El segundo problema es la apertura demasiado pequeña cuando hay los puntos de la intensidad de la luz alta en la escena. Este puede resultar en la imagen con bajo contraste. El filtro promediado va a reducir este tipo de problemas. Los puntos brillantes van a ser suavizados.

3.3.1.2. Detección de contornos.

Hay diferentes técnicas de calcular los contornos. El operador de Sobel es muy popular por su simplicidad computacional. Pero el problema de umbralización de contornos está presente con este método. Para resolver este problema se propone a descartar los puntos, la magnitud de intensidad de los cuales es más baja, tomándolos como el ruido. El porcentaje que se queda de los puntos es de 60 a 70 que es razonable para este ambiente de trabajo. También se asume que la magnitud del ruido tenga una distribución de Gauss con el resultado m y la desviación σ . Si e_i es el pixel de la magnitud de Sobel, el número total de pixeles es n y el umbral es t , entonces

$$\begin{aligned} e_i > t = m + 3\sigma \\ \text{donde } m = 1/n \sum_i e_i \quad \sigma^2 = 1/n \sum_i (e_i - m)^2 \end{aligned} \tag{3.1}$$

El otro problema que aparece aquí es el grosor de los bordes. Entonces después del paso uno descrito arriba los resultados pasan a la detección del borde de un pixel. El método de detección de los contornos es robusto pero también puede encontrar sus desventajas si las imágenes tienen el contraste muy bajo.

Este método detecta los bordes en una imagen sin utilizar la secuencia de marcos. Pero como el sistema de visión se mueve entonces es mejor aplicar los métodos basados en movimiento. En detección de los contornos de tubos bajo el agua como en detección de los bordes de objetos en el camino se puede usar el método de Liu [Liu, 1998] "Extracción de los bordes en movimiento e implementación en el tiempo real" (ver 2.1.2.1) o el método de gradiente [Jain, 1995] (ver 2.1.3).

3.3.1.3. Detección de línea.

El segmento de tubos aparecerá como dos líneas paralelas separadas por el diámetro del mismo. También van a haber algunas distorsiones. La transformada de Hough da la habilidad de mantener los datos de la imagen necesarios y rechazar los extraños. El uso directo de la transformada de Hough tiene tres problemas:

- Las líneas están indicadas por diferentes valores cercanos y no por uno sólo.
- No hay la indicación de longitud y posición de la línea.
- La transformada básica busca todas las líneas posibles que consume el tiempo y no es necesario cuando las direcciones de líneas se conocen.

La transformada modificada resuelve estos problemas. El algoritmo escoge una línea apropiada entre todas basándose en la longitud, el punto medio y el ángulo.

El algoritmo PVS puede interpretar las configuraciones de tubos complejos, más aún cuando la posición de vehículo no es exacta. PVS es robusto contra errores de escala, rotación, contrastes variantes de imagen a imagen, sombras y distribución de la luz brusca. Como los otros algoritmos, PVS tiene varios problemas en el número de parámetros usados (reducir) y el mantenimiento de errores (mejorar).

3.4. Detección de la orientación de la vista del ojo humano.

La gente usa sus ojos para diferentes propósitos: leer, ver, juntar la información para planear las acciones, percibir y aprender nuevas cosas, evaluar sus acciones, etc. Nosotros usamos nuestro sistema visual en forma transparente concentrándose en el proceso de operar los conceptos que nos rodean, dejando el proceso básico de información óptica a nuestros ojos y el sistema visual. El hecho importante - los ojos pueden observar todo lo que los rodea indicando la dirección del punto de interés [Glenstrup, 1995]. Esta habilidad fue usada para las técnicas de comunicación entre el hombre y la máquina. Al usar el ojo humano como el medio de interacción abre nuevos caminos para la tecnología. El usuario podrá combinar los dispositivos estándares y sus ojos para ejecutar diferentes tareas. La comunicación entre el usuario y la computadora será más transparente.

En el día de hoy existen diferentes algoritmos para detectar la orientación de la vista. Las características que debe tener un sistema para la detección ideal de la orientación de la vista del ojo humano son [Glenstrup, 1995]:

- Evitar los equipos para la cara y la cabeza del usuario.
- No tener ningún contacto con el sujeto.
- Ser capaz de estabilizar artificialmente la imagen de retina.
- Tener la exactitud de, por lo menos, algunos minutos del arco. La exactitud es limitada por efectos no lineales, distorsión, ruido, etc.
- Ser capaz de detectar los pequeños cambios de la posición del ojo.
- Extender dinámicamente el rango desde 1 minuto hasta 45 para la posición del ojo.
- Buena dinámica temporal y la velocidad de respuesta.
- Tener la respuesta en el tiempo real.
- Medir las tres dimensiones de rotación y ser insensible a translación ocular.
- Extenderse fácil a grabaciones binoculares.
- Poder grabar los movimientos de la cabeza y el cuerpo.
- Ser fácil de usar en diversos sujetos.

Básicamente existen tres técnicas de detección [Glenstrup, 1995]:

- Medición de reflejo de la luz en el ojo. Normalmente se usa la luz infrarroja para no distraer mucho al usuario y evitar interferencias de otras frecuencias como de la lámpara.
- Medición de potencia eléctrica de la piel cerca del ojo.
- Aplicar las lentes de contacto especiales que faciliten la detección de la posición.

En la técnica uno el contacto con el usuario es menor que en la tercera. Todas estas técnicas requieren algún tipo de calibración antes de usarlas, y después alguna recalibración en la hora de usarlas.

3.4.1. Técnica basada en el reflejo de la luz.

Dentro de esta categoría existen cinco técnicas: sendero de iris, sendero de pupila, relación de reflexión entre córnea y pupila y reflexión de córnea en la imagen de ojo usando redes neuronales artificiales [Glenstrup, 1995].

3.4.1.1. Sendero de iris.

Se estima es la frontera entre el blanco del ojo y el iris. Como el primero es normalmente blanco y el segundo es oscuro, es relativamente fácil detectar el movimiento. Esta técnica se basa en la posición y la forma de iris en relación con la cabeza que tiene que estar bien sujeta. Puede detectar precisamente sólo en sendero horizontal.

3.4.1.2. Sendero de pupila.

Esta técnica es similar a la primera sólo ahora se considera la frontera entre la pupila y el iris. Otra vez la cabeza tiene que estar bien sujeta. Las ventajas de esta técnica en comparación con la de iris son:

- La pupila está menos cubierta por los párpados que iris. Esto abre el sendero vertical.

- La frontera de pupila está formada mejor que la de iris. Esto lleva a la resolución más alta de solución.

La desventaja de este método es que el contraste entre la pupila y el iris es más bajo que entre el iris y el blanco del ojo haciendo la detección de fronteras más difícil.

3.4.1.3. Relación de reflexión entre córnea y pupila.

Cuando la luz ilumina los ojos del usuario, diferentes reflexiones ocurren en las fronteras de pupila y córnea. Las imágenes de este tipo se llaman imágenes de Purkinje. Una de las imágenes se llama fulgor que se obtiene con la reflexión de la luz de la retina. La graban en la cámara que soporta la luz infrarroja, donde se va a verse el punto muy claro del centro y el disco menos claro alrededor. Cuando el ojo se mueve horizontalmente o vertical la posición relativa del fulgor y el centro del ojo cambian respectivamente, y así se puede calcular la dirección de la vista. Los problemas asociados con esta técnica son:

- No siempre se puede obtener la buena imagen del ojo.
- El movimiento de la cabeza puede poner la imagen fuera del enfoque de la cámara.

En las tres técnicas se podría utilizar el algoritmo SUSAN [Smith, 1997] (ver 2.3.2) para detectar la orientación de la vista del ojo humano. Como este algoritmo encuentra las esquinas de la imagen utilizando un umbral, su valor se puede poner de tal modo que se detecten los objetos de nivel de gris necesario.

3.4.1.4. Reflexión de córnea en la imagen de ojo usando redes neuronales artificiales (ANN).

En esta técnica se digitaliza la imagen de usuario implicando la cabeza completa. La luz estacionaria se coloca enfrente del usuario y el sistema empieza a buscar el ojo derecho en la imagen por la reflexión del punto más brillante rodeado por la región oscura. Después se hace la segmentación de imagen y se aplica la red neuronal artificial. La salida de la red es el conjunto de las coordenadas del monitor según la orientación de la vista. ANN tiene que estar entrenada por las diferentes imágenes de los ojos de usuario y

su cabeza. La precisión de resultados de ANN no es tan buena como de otras técnicas. Podría ser mejorada por agregar los cálculos basados en córnea/pupila y la posición del fulgor de ojo.

3.4.2. Mediciones de potencia eléctrica de la piel cerca del ojo.

Se basa en el hecho que existe el campo electrostático alrededor del ojo. Al grabar las diferencias pequeñas de la potencia de la piel la posición del ojo puede ser detectada. Esto se hace con los electrodos ubicados en la piel. No requiere la imagen precisa del ojo. El problema que podría tener esta técnica es el contacto muy cercano de usuario con los electrodos [Glenstrup, 1995].

3.4.3. Técnica basada en las lentes de contacto.

Dentro de esta técnica existen dos métodos [Glenstrup, 1995]:

- Poner varios planos de la superficie espejada para la lente. Reflexión de los rayos de luz se usa para calcular la posición del ojo.
- Implantar un material inductivo en la lente. La posición exacta puede ser determinada por la frecuencia alta de los campos electromagnéticos alrededor de la cabeza del usuario.

Las últimas dos técnicas presentadas parecen ser difíciles de aplicar en el uso diario por las razones de salud y el hecho de poner las lentes o electrodos en la cabeza.

3.4.4. Movimiento de los ojos.

Después de determinar la posición del ojo, hay que seguir su trayectoria. Los movimientos que puede hacer el ojo son los siguientes [Glenstrup, 1995]:

- Convergencia es el movimiento de dos ojos relacionando uno con el otro que asegura la distancia del objeto a los ojos. Entre más cerca está el objeto, más los ojos apuntan uno al otro. Este movimiento puede ser controlado, pero normalmente es el resultado del estímulo de movimiento.
- Ondulante de los ojos, es la rotación alrededor del eje óptico que pasa a través de la pupila. Es el movimiento no controlado y es la influencia del ángulo de cuello, entre otras cosas.
- "Saccades" son los métodos principales para mover los ojos a diferentes partes de la escena visual, es un movimiento muy rápido. Para iniciar "saccade" se toma como 100-300 ms, cuando se presenta el estímulo los ojos empiezan a moverse. Después se toma como 30-120 ms para completar "saccade". Una vez iniciado el "saccade" la ruta y el destino no pueden cambiarse. Esto indica que la atención visual escoge la posición siguiente a dónde van a moverse los ojos.
- El movimiento de búsqueda es suave y lento. Actúa para mantener el objeto en movimiento en la vista.
- "Nystagmus" es el patrón de movimientos de los ojos que ocurren como respuesta a la vuelta de cabeza y el movimiento de la vista. Consiste en movimiento suave en una dirección para seguir a una posición de la escena.
- El flujo y micro "saccades" ocurren durante las fijaciones de ojos y consisten en los flujos lentos seguidos por los "saccades" muy pequeños.
- "Nystagmus" fisiológico es la oscilación de alta frecuencia del ojo que sirve al desplazamiento continuo de la imagen en retina.

Analizando estos movimientos se ve que el sendero de los ojos resultará de los datos que consisten de los movimientos de diferentes orígenes. Esto implica que los datos necesarios tendrán que ser extraídos de todo el conjunto que va a afectar la precisión de los resultados.

3.4.5. Aplicabilidad de las técnicas.

Con la posibilidad de controlar una máquina usando los ojos aparece el nuevo tipo de interfaz que no podrá ser más rápido de los dispositivos tradicionales. El tiempo de respuesta es [Glenstrup, 1995]:

$$T_{\text{exe}} = T_H + T_P + T_K + T_D + T_M + T_R \quad (3.2)$$

H es mover la mano a diferentes dispositivos.

P es apuntar a la posición correcta.

K es hacer clic del ratón.

D es dibujar las líneas de movimiento.

M es para la preparación mental.

R es el tiempo de respuesta del sistema.

La ventaja principal del ratón es el hecho que se basa en la habilidad humana. Y no es tan natural trabajar con los objetos usando sólo los ojos. Las desventajas de esta técnica con el ratón son [Glenstrup, 1995]:

- No se puede sólo cerrar los ojos cerca del objeto si no quiere que haga una acción.
- No se puede hacer los movimientos vagos con los ojos por la pantalla.
- La vista del hombre no es suave y por eso es más difícil de controlarla.
- Es difícil de apuntar a un área monótona porque los ojos normalmente buscan los objetos que atraen por su contenido de información.

Pero la evidencia empírica muestra que el manejo con los ojos podría resultar más rápido y útil. Por eso los investigadores tratan de encontrar las soluciones posibles al problema. Ahora existen diferentes software y hardware relacionados con la detección de la orientación de la vista del ojo humano. Una de las herramientas es el dispositivo llamado "Express Eye" [Express Eye, 1999] que es fiable, preciso y capaz de extraer los datos con la resolución alta del tiempo. Enfrente de los ojos hay un detector sostenido por las bandas. El detector puede ser movido en tres dimensiones. La afinación horizontal asegura la posición óptima que se encuentra fácil y rápido. Otro [EyeTech, 1999] es un sistema digital que provee el hardware necesario y el software que permite usar los ojos como una

alternativa al ratón. El sistema mueve el cursor según el movimiento de los ojos y hace clic con el parpadeo suave o fuerte, o con algún dispositivo. El sistema puede beneficiar a cualquiera persona que no puede operar con las manos o que quiere tenerlas desocupadas.

El sistema de detección de orientación de la vista del ojo humano puede ser muy útil en diferentes situaciones. Por ejemplo en las fábricas donde puede ocurrir el momento de no tener las manos desocupadas para apretar el botón o una tecla. También este podría ayudar a la gente discapacitada. Los ojos servirían como una alternativa a los dispositivos estándares o un ayudante.

En el día de hoy se desarrollan varias técnicas que se puede clasificar en dos grupos principales: los que usan el contacto estrecho con el usuario y los que no lo hacen. Las primeras dan los resultados más precisos que las segundas. Pero para el usuario común el contacto estrecho no es tan deseable. Por eso aparecen los sistemas como de EyeTech que tratan de beneficiar el usuario normal, pero la pregunta es qué tan exactos son los resultados que da el sistema. Para los futuros desarrollos es importante encontrar la solución óptima entre dos grupos de desarrollo; uno que usa el contacto cercano con el usuario y el otro que trata de darle la libertad total.

Existe una gran variedad de aplicaciones que usan la estimación de movimiento. Aquí presenté las cuatro aplicaciones sobre el manejo automovilístico, clasificación de objetos, detección de tubos bajo el agua, y detección de la orientación de la vista del ojo humano. En la primera para dar los mensajes al conductor es necesario analizar el ambiente alrededor del vehículo sacando la secuencia de imágenes. En la segunda la clasificación de imágenes tiene que trabajar en el tiempo real tomando diferentes imágenes del objeto para asignarlo a su clase correspondiente. En la tercera hay que detectar los límites del tubo como bordes y puntos extremos. Para encontrar y después detectar el tubo completo puede resultar útil sacar la secuencia de imágenes trabajando en el tiempo real. En la última aplicación para poder seguir el movimiento de los ojos también se necesita el procesamiento de la secuencia de datos obtenida por las técnicas descritas en el párrafo de detección de la orientación de la vista del ojo humano.

Al investigar las aplicaciones del área y asignar los métodos, analizados en el capítulo de modelos de aproximación, a cada una descrita en este capítulo sirvió para escoger dos métodos que se implementarán en el siguiente capítulo.