

CAPÍTULO 2. MARCO TEÓRICO

Para lograr un mayor entendimiento sobre el objetivo de esta tesis, es necesario comprender algunos de los conceptos mayormente utilizados en el desarrollo de la misma, este capítulo los expone y también presenta la teoría básica sobre el rastreo de movimiento, lo que fundamenta en conjunto los procedimientos por los que pasa la secuencia de imágenes desde su conformación hasta su rastreo.

2.1 Objeto Rígido Bidimensional

Un objeto rígido bidimensional puede ser definido como aquél que está compuesto por un conjunto de puntos descritos en el plano 'x y', y que se encuentran limitados por un contorno finito y constante, incapaz de sufrir ninguna deformación.

2.2 Movimiento

Una gran cantidad de información puede ser extraída de una secuencia de imágenes variantes en el tiempo, en la mayoría de las veces, mucha mayor que aquella adquirida por imágenes estáticas. Por ejemplo, elementos camuflajeados sólo pueden ser vistos con facilidad cuando se mueven. Además, los tamaños y posiciones relativos de los objetos son más fácilmente determinados cuando éstos se mueven. Incluso las diferencias entre dos imágenes simples permiten determinar las siluetas de objetos libres de texturas que se mueven sobre un fondo estático.

El movimiento es una característica poderosa de las secuencias de imagen, revela el dinamismo de escenas por medio de relaciones espaciales y cambios en el tiempo.

El análisis del movimiento visual se divide en dos etapas principales:

- La medida del movimiento y
- El uso de la información del movimiento para segmentar la escena en distintos objetos y extraer información sobre la forma y el movimiento de éstos.

Existen dos tipos de movimientos a considerar, el movimiento de un plano para un objeto estático, o el movimiento del objeto sobre el plano. Sin embargo, como el movimiento es finalmente relativo, ambos tipos de movimientos podrían ser considerados iguales, aunque no en todos los casos es así, ya que si se consideraran efectos de iluminación y sombras, así como otros aspectos, el movimiento deja de ser relativo. En los casos analizados para este trabajo, tales complicaciones son ignoradas.

2.2.1 Análisis del movimiento

La tarea del análisis del movimiento permanece como un desafío y es un problema de la visión por computadora. Desde secuencias de imágenes bidimensionales, el único parámetro accesible de movimiento es el flujo óptico, una aproximación del campo de movimiento bidimensional en un sensor de imagen. El campo de movimiento es dado como una proyección de un movimiento tridimensional de puntos en la escena en un sensor de imagen. El campo de flujo óptico estimado puede ser usado como una entrada para una variedad de pasos procesadores subsecuentes, incluyendo la detección del movimiento, su compensación, el movimiento basado en la compresión de datos,

reconstrucción de escenas tridimensionales, navegación autónoma y el análisis de procesos dinámicos en aplicaciones científicas.

Las dificultades de la estimación del movimiento son múltiples y se originan en diferencias inherentes entre el flujo óptico y el campo de movimiento verdadero. Únicamente el movimiento aparente en la secuencia puede ser extraído, aún suponiendo *a priori* cambios de brillo, propiedades del objeto, y la relación entre movimientos relativos de escena tridimensionales y la proyección en un sensor de imagen bidimensional son necesarios para cuantificar el análisis de la escena.

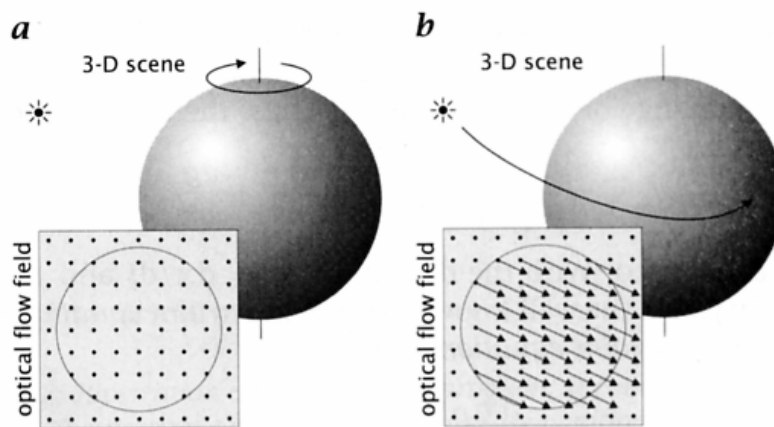


Figura 2.1 Correspondencia física contra visual. **a)** Una esfera que gira con iluminación fija se traduce en flujo óptico cero; **b)** una fuente de iluminación en movimiento causa un campo de flujo óptico aparente son movimiento de la esfera. [4]

Horn propuso una visión optimista de la posibilidad de la reconstrucción de un campo tridimensional desde campos de movimiento. Mostró que el campo de movimiento puede estar casi siempre sin ambigüedades, relacionado a velocidades traslacionales y rotacionales de superficies rígidas. Sin embargo, el campo de movimiento por si mismo es a menudo inaccesible. Esto puede quedar bien demostrado en un simple ejemplo (Figura 2.1) que considera una esfera rígida con reflejo de superficie homogéneo, girando alrededor en un eje por el centro de la esfera. Si la superficie no está texturizada y la

iluminación permanece constante, el campo de flujo óptico aparente podría equivaler a cero en toda la esfera. Si un cambio direccional de luz ocurre alrededor de la misma esfera, la iluminación aparente cambia y puede ser falsamente atribuida al movimiento de la superficie de la esfera. Este problema bastante complejo, muestra que incluso un experimento muy simple bajo perfectas condiciones puede rendir una estimación de flujo imposible. Los problemas encontrados con mayor frecuencia en unas secuencias del mundo real, incluyen una cubierta transparente de movimientos múltiples, oclusiones, cambios de iluminación, movimiento no rígido, movimientos de parar-disparar, niveles de baja señal de ruido (SNR), problemas de apertura y problemas de correspondencia (por mencionar solo algunos de ellos). Por esta razón, el verdadero campo de movimiento es difícilmente accesible y se ha determinado que sólo propiedades cualitativas del campo de movimiento deben ser computadas.

Estos problemas, sin embargo, no siempre están presentes y usualmente no se esparcen por toda el área de la imagen. Así, existen muchas aplicaciones en las que el análisis del movimiento se vuelve factible. Al mismo tiempo, se han colocado una serie de limitaciones de la computación del flujo óptico, que a menudo no se toman en cuenta: los errores deben ser detectados y cuantificados. Esto es especialmente importante para tareas cuantitativas de medición. En contraste con los requerimientos más cuantitativos de las aplicaciones de la visión estándar por computadora, tales como la detección del movimiento o evasión de colisiones, las medidas cuantitativas de procesos dinámicos, requieren precisar y resaltar los campos de flujo óptico en orden para reducir la propagación de errores en los pasos de procesos subsecuentes. En adición al campo de flujo óptico, mediciones confiables deben ser proveídas para descartar puntos de datos erróneos y cuantificar con precisión las medidas.

2.2.2 Movimiento de Objetos Rígidos Bidimensionales

El movimiento de un objeto rígido bidimensional sin escalamiento consiste en aplicarle tres transformaciones al objeto: desplazamiento en 'x', desplazamiento en 'y' y rotación, de manera sistemática en diferentes instantes de tiempo. (Figuras 2.2 y 2.3)

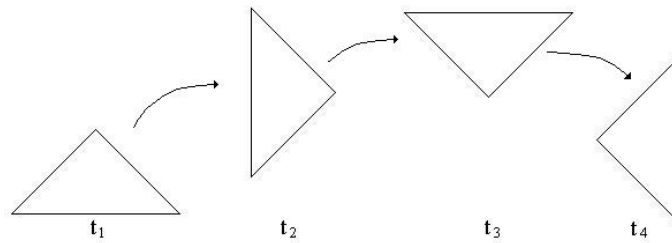


Figura 2.2 Objeto rígido bidimensional en movimiento sobre un plano

$$\begin{aligned} \text{En } t_2 : T\{\triangle, \Delta x_1, \Delta y_1, \theta_1\} &= \triangleright \\ \text{En } t_3 : T\{\triangle, \Delta x_3, \Delta y_3, \theta_3\} &= \nabla \\ \text{En } t_4 : T\{\triangle, \Delta x_4, \Delta y_4, \theta_4\} &= \triangleleft \end{aligned}$$

Figura 2.3 Objeto rígido bidimensional en su posición inicial y el mismo objeto en las posiciones siguientes, después de aplicar transformaciones para cada instante de tiempo

El rastreo de un objeto rígido consiste en encontrar los parámetros $(\Delta x, \Delta y, \theta)$ de la transformación que se aplica al objeto en su posición inicial, los cuales permiten describir las posiciones subsecuentes en la que se encuentra el objeto.

2.2.3 El campo de movimiento

Cuando un objeto se mueve frente a un espectador, hay un cambio correspondiente en la imagen. Así, si un punto p_o en un objeto se mueve a una velocidad v_o , entonces al punto visualizado como p_i se le puede asignar un vector v_i que indica su movimiento en el plano de la imagen. La colección de todos estos vectores forma el campo de movimiento. (Figura 2.4)

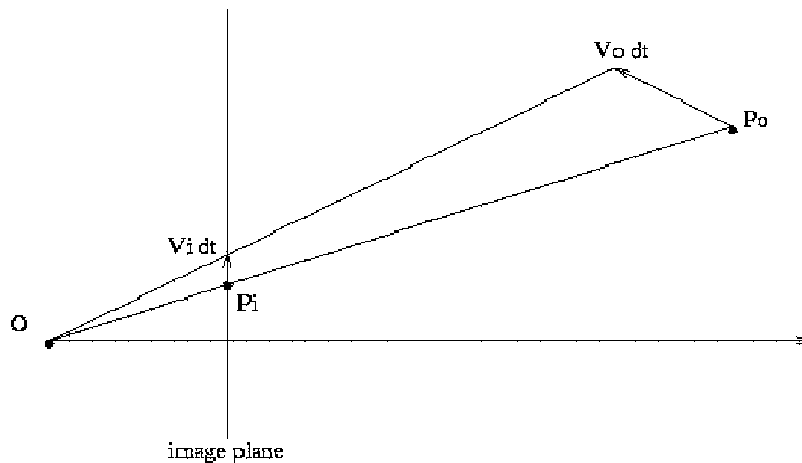


Figura 2.4 Campo de movimiento [2]

En el caso de tratarse del análisis del movimiento y rotación de objetos rígidos bidimensionales, entonces el campo de movimiento será constante, excepto en los límites de la silueta de los objetos. (Figura 2.5)

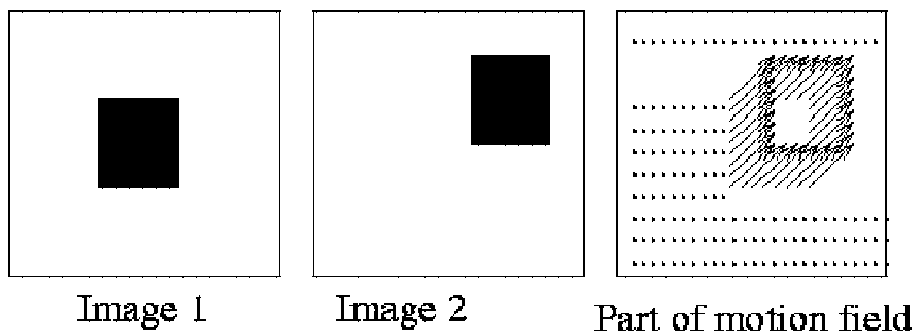


Figura 2.5 El campo de movimiento de un cuadrado móvil [2]

2.3 Técnicas de Rastreo de Movimiento

El flujo óptico es una representación conveniente y útil del movimiento de imágenes. Sin embargo, existen otros métodos que describen el movimiento, algunas veces un poco más generales que el flujo óptico, tales como modelos paramétricos de movimiento o descriptores adaptados a contextos determinados, como elementos de geometría de la escena o del movimiento del sensor visual que son parcial o completamente predeterminados.

Distintas metodologías han sido propuestas para el rastreo del movimiento de objetos en el espacio y han sido clasificadas de diferentes maneras. Algunas de las comúnmente encontradas en la literatura son los siguientes:

2.3.1 Diferenciación de Imágenes

Está basado en la comparación de diferencias entre imágenes sucesivas. Sin embargo, este algoritmo no es útil cuando se busca rastrear áreas en movimiento, pues prácticamente se limita a detectar cambios en las imágenes, sin ser muy exacto en sus resultados.

Sólo requiere calcular los valores absolutos de la diferencia entre los píxeles correspondientes de dos imágenes, y los resultados de esta operación que sean mayores en el mapa indicarán las zonas de cambio. Algunas aplicaciones comunes de este método, incluyen sistemas de detección de intrusos, sistemas de supervisión de vehículos y compresión de datos de cuadros internos. También se ha usado como parte del análisis

de imágenes satelitales, en la medición de la erosión terrestre, la deforestación, y el crecimiento urbano, etc., así como en análisis médicos para calcular la distribución celular.

2.3.2 Substracción del fondo

Su funcionamiento esencial se basa en identificar los objetos móviles de un conjunto de marcos al calcular la diferencia entre una imagen de la secuencia con una imagen de fondo (previamente obtenida).

La tarea de detección de movimiento con este método se hace un poco más complicada debido a distintos factores, pues de no tener una imagen de fondo preestablecida de manera adecuada, se pueden obtener errores recurrentes. Además, el proceso de diferenciación entre los objetos en movimiento que realmente interesan al sistema de aquellos que son irrelevantes es complicado, por lo que se requieren algoritmos muy robustos para su implementación en sistemas reales.

2.3.3. Estimación por movimiento de bloques

Calcula el vector de movimiento entre pequeños bloques que forman parte de distintas imágenes sucesivas. El problema principal de este método recae en la determinación del área que cada bloque cubre dentro de la imagen, por lo que se debe ser bastante específico en los criterios de decisión. Por otro lado, la búsqueda del bloque obtenido en la imagen original sobre el nuevo marco analizado es un tanto complicada, sin embargo, se han desarrollado algoritmos rápidos de detección del bloque bastante robustos pero que requieren procesos computacionales intensos.

2.3.4 Flujo Óptico (*Optical Flow*)

Es el campo de desplazamiento para cada uno de los píxeles en una secuencia de imágenes. Es decir, para cada píxel un vector de velocidad $\left(\frac{dx}{dt}, \frac{dy}{dt}\right)$, o bien $V = (u, v)$,

es encontrado, el cual nos dice:

- La velocidad a la que se mueve cada píxel
- La dirección de su movimiento

Supongamos que la intensidad de la imagen está dada por $I(x, y, t)$, donde la intensidad es ahora una función del tiempo t , así como de x y y . En un punto a una distancia corta poco tiempo después ($I(x + dx, y + dy, t + dt)$), la intensidad varía para cada variable, es decir:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \dots (1)$$

Donde los puntos suspensivos hacen alusión a términos de mayor orden que pueden ser obviados pues no representan variaciones considerables.

Ahora, supongamos que parte de un objeto se encuentra en una posición (x, y) en la imagen para un tiempo t , y que tras un tiempo después dt , éste se ha movido una cierta distancia (dx, dy) a lo largo de la imagen. También, supongamos que la intensidad de esa parte del objeto es justo la misma en la imagen original y en la imagen transformada, entonces:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) \quad (2)$$

Por lo que al sustituir (2) en (1), podemos observar que:

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \dots = 0$$

Sin embargo, al dividir estos términos entre dt tenemos lo siguiente:

$$\frac{dx}{dt} = u, \quad \frac{dy}{dt} = v$$

Las cuales son las velocidades a las cuales se mueve el objeto en las direcciones 'x' y 'y' respectivamente. Así en el límite en el que dt tiende a cero, tenemos:

$$-\frac{\partial I}{\partial t} = \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v \quad (3)$$

La cual es conocida como la ecuación de flujo. Es importante hacer notar que $\frac{\partial I}{\partial t}$ en un píxel dado, sólo muestra qué tan rápido cambia la intensidad en el tiempo, mientras que $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ son las tasas espaciales de cambio de intensidad, de forma que estos tres valores pueden ser calculados para cada píxel de acuerdo a la secuencia de imágenes.

Sin embargo, éste análisis no es suficiente, pues dado el caso de que un píxel de una intensidad determinada, Δt tiempo después, se encuentren dos o más píxeles con la misma intensidad (por ejemplo, una línea de contorno), entonces se requiere más información para la determinación de u y de v .

Una observación crítica en este aspecto es que, excepto de los límites de los objetos móviles, el movimiento observado en el resto de los píxeles adyacentes será muy similar.

Una medida de qué tanto el flujo óptico se desvía suavemente de este ideal, puede ser calculada por medio de la siguiente integral:

$$S = \iint_{\text{imagen}} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right] dx dy \quad [15]$$

Sobre toda la imagen donde las derivaciones como $\frac{\partial u}{\partial x}$ miden qué tan rápido la velocidad cambia a lo largo de la secuencia de imágenes.

La anterior aproximación con suavidad no es necesariamente consistente con la de flujo óptico. Es posible expresar en qué medida una solución para u y v se desvía de la condición requerida por la ecuación de flujo óptico al evaluar:

$$C = \iint_{\text{imagen}} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right)^2 dx dy \quad [15]$$

Para combinar estas dos soluciones de ecuaciones se usa la técnica de multiplicadores de Lagrange:

- Buscan una solución para u y v que minimiza $S + C\lambda$, donde λ es un escalar.
- Típicamente, se da un valor de λ grande si las medidas de intensidad son precisas, donde esperamos que el valor de C sea pequeño.
- Por otro lado, si la información original es ruidosa, entonces el valor de λ debe ser pequeño.

- Un ajuste interactivo debe ser realizado para ofrecer el valor más adecuado de λ .

El proceso de minimización de la integral resultante, puede llevarse a cabo por medio de técnicas del cálculo de variaciones, las cuales muestran que las funciones u y v que son requeridas satisfacen las siguientes ecuaciones diferenciales:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \lambda \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial x}, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = \lambda \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial y} \quad [15]$$

Las derivadas de I para cada píxel son obtenidas de la imagen original, y λ es elegida como ya se ha explicado. Un método iterativo puede entonces ser usado para resolver estas ecuaciones para u y v en cada píxel.

2.3.4.1 Características

La hipótesis inicial en medición de movimiento de imágenes es que la intensidad de la estructura de variaciones locales de tiempo en regiones de imagen, son aproximadamente constantes bajo el movimiento de al menos una corta duración. Formalmente, si $I(x, t)$ es la función de intensidad de la imagen, entonces:

$$I(x, t) \approx I(x + dx, t + dt), \quad (4)$$

donde dx es el desplazamiento de una región determinada de la imagen en (x, t) después de tiempo dt . Extendiendo el lado izquierdo de la mano en esta ecuación en una serie de producción Taylor tenemos:

$$I(x, t) = I(x, t) + \nabla I \cdot dx + dtI_t + O^2 \quad (5)$$

donde $\nabla I = (I_x, I_y)$ e I_t son las derivativas parciales de primer orden de $I(x,t)$ y O^2 el segundo y el de más alto orden, el cual se supone insignificante. Extrayendo $I(x,t)$ en ambos lados, ignorando O^2 y dividiendo por dt obtenemos:

$$\nabla I \cdot v + I_t = 0 \quad (6)$$

donde $\nabla I = (I_x, I_y)$ es el gradiente de intensidad espacial y $v = (u, v)$ es la velocidad de la imagen. Dicha ecuación se conoce como la ecuación de límite del flujo óptico y se determina como una fuerza única y local en el movimiento de la imagen. (Figura 2.6)

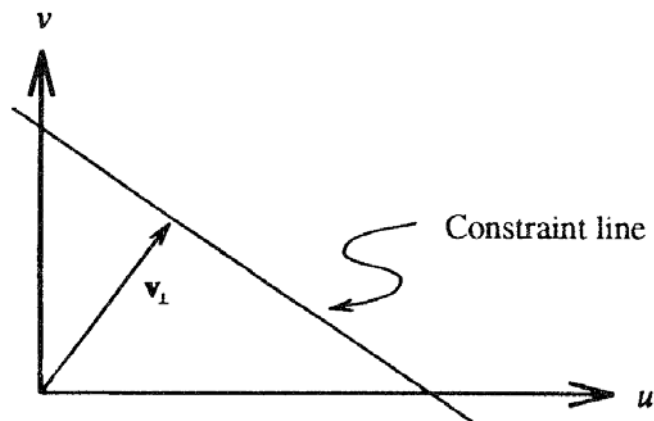


Figura 2.6 La ecuación limitante del flujo óptico define una línea en la velocidad espacial. [9]

En la figura 2.6 la velocidad normal v_{\perp} se define como el vector perpendicular de una línea límite, es decir, la velocidad con la menor magnitud en el límite del flujo óptico. En este límite no es posible calcular los dos componentes de v ya que la ecuación limitante del flujo óptico resulta incorrecta. Es decir, sólo v_{\perp} , el componente del movimiento en la dirección del gradiente de la función de intensidad de la imagen, puede ser estimado.

Ese fenómeno es conocido como el Problema de Apertura y sólo en ciertos puntos de la imagen, donde la estructura de intensidad es suficiente (o en la curvatura Gaussiana), puede ser que el movimiento sea completamente estimado con el uso de la ecuación limitante de flujo óptico (Figura 2.7).

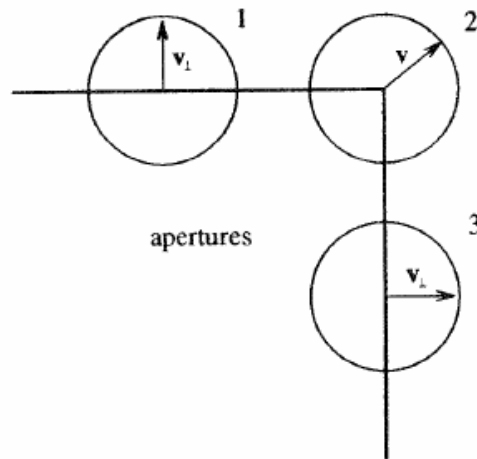


Figura 2.7 El problema de apertura. En las aperturas 1 y 3, sólo movimientos normales de las orillas que forman el cuadrado pueden ser estimados debido a una falta en la estructura local. Dentro de la apertura 2, en el punto de la esquina, el movimiento puede ser medido completamente cuando ahí hay suficiente estructura local, los dos movimientos normales, son visibles. [9]

Por ejemplo, la velocidad de una superficie que es homogénea o que contiene textura en una misma orientación, no puede ser recuperada óptimamente. Debido a que la velocidad normal está en la dirección del gradiente espacial ∇I la ecuación sólo se traduce en:

$$v_{\perp} = \frac{-I_r \nabla I}{\|\nabla I\|_2^2} \quad [9]$$

Así, la medición espacio-temporal de las derivadas, permite la recuperación de una velocidad normal de imagen.

Desde esta definición, se torna claro que para que el flujo óptico obtenga un movimiento de imagen exacto, es necesario satisfacer un número de condiciones, éstas son:

- a) Iluminación uniforme.
- b) Reflejos específicos en superficies y
- c) Movimiento exclusivamente paralelo al plano de la imagen.

Objetivamente, estas condiciones nunca son completamente satisfechas en un paisaje real. Sin embargo, se puede asumir que estas condiciones son retenidas localmente dentro de la escena, y por lo tanto, localmente en el plano de la imagen.

El grado en el que estas condiciones son parcialmente satisfechas, determina la precisión en la que el flujo óptico se aproxima al movimiento de imagen. Alternativamente, se puede medir el desplazamiento de pequeños parches de una imagen, por ejemplo por correlación, en secuencias cortas de imágenes (usualmente de dos o tres pasos). Dicho desplazamientos de la imagen constituyen una aproximación considerable a la velocidad de la imagen, en donde ciertas condiciones son cumplidas. En particular, la proporción de un sensor de velocidad traslacional, la rotación vertical y horizontal tridimensional del sensor, y el intervalo de tiempo entre pasos, deben ser de pequeñas cantidades.

El flujo óptico puede ser también calculado como un campo de disparidad donde, dadas dos imágenes estéreo o dos imágenes contiguas de alguna secuencia, se extraen las características de interés en las imágenes y son colocadas mediante un proceso de correspondencia.

Esencialmente la ejecución de detección de movimiento bidimensional, involucra el procesamiento de escenas donde el sensor se mueve junto el medio ambiente que lo contiene junto a objetos estacionarios y no estacionarios. Además los eventos visuales tales como oclusión, movimiento transparente y objetos no rígidos, aumentan la complejidad inherente de la medición del flujo óptico.

2.3.5 Técnicas de Flujo Óptico

Un método fundamental en el procesamiento de secuencias de imágenes, es el cálculo del flujo óptico, una aproximación del movimiento de imagen, definido como la proyección de velocidades en una superficie de puntos en un plano de la imagen de un sensor visual.

La mayoría de las aplicaciones reales de estimación del movimiento emplean un acercamiento bidimensional, en el que se obtienen estimaciones del movimiento en el tiempo. En este tipo de métodos encontramos la condición de que el brillo de la imagen a lo largo de la trayectoria de movimiento es constante, es decir que cualquier cambio en el tiempo de la intensidad de un punto es debido únicamente al movimiento.

Existen dos ideas del mismo problema divididas en dos teorías diferentes sobre la estimación del movimiento bidimensional: la utilización directa de la ecuación de flujo

óptico, y el cálculo del mínimo de la diferencia de desplazamiento entre tramas en un conjunto de regiones locales, buscando vectores de movimiento.

En base a esto, Barron et. al [1] clasificó los algoritmos de flujo óptico por sus etapas de extracción de la señal. Esto proporciona cuatro grupos: métodos basados en el gradiente, métodos basados en la correlación de regiones, métodos basados en el dominio de frecuencias y métodos bayesianos.

2.3.5.1 Métodos basados en el gradiente

La estimación basada en el gradiente se ha convertido en el principal acercamiento en las aplicaciones de visión por computadora. Ello se debe principalmente a que es eficaz y a que se produce una estimación acertada del campo de movimiento. Estos métodos para poder usarse en la estimación de movimiento requieren de algunas restricciones.

El método basado en el gradiente más conocido es el de Horn y Schunck [17] en el que impuso una restricción de suavizado global en el campo de movimiento. Se han realizado igualmente algunos esfuerzos para obtener modelos de alto nivel a partir del campo de flujo óptico por medio de técnicas de regresión lineal. Las desventajas comunes de todos los métodos basados en el gradiente provienen de los obvios cambios de iluminación y en la dificultad de abordar numéricamente la diferenciación de la función de intensidad de la imagen.

Las técnicas diferenciales están caracterizadas por la búsqueda de un gradiente extraído en primer lugar de derivados espaciales y en segundo de derivados temporales.

Lucas y Kanade [7] también construyeron una técnica de estimación de flujo basada en un primer orden de derivados de la secuencia de imagen. En contraste con la regularización post-uniforme de Horn y Schunk, ellos eligen el pre-suavisado de los datos antes de usar la ecuación de restricción de pendiente.

Generalmente, los métodos diferenciales se pueden considerar como una extracción de la señal sobrepasada. Por lo tanto, ellos sólo proveen una representación local en un espacio de frecuencia y ésta está restringida al perfeccionamiento de un intervalo de velocidades caracterizadas por una pre-uniformidad de la señal espacio-temporal previa a la diferenciación numérica.

La información de flujo puede no siempre estar limitada a una estrecha frecuencia de banda.

2.3.5.2 Métodos basados en la correlación de regiones

Una forma de evitar las limitaciones de los métodos basados en el gradiente está en considerar regiones de la imagen en lugar de estimar el movimiento píxel a píxel. En general los métodos basados en la correlación de regiones son menos sensibles al ruido al tomar más datos de la imagen en el proceso de estimación del movimiento.

Las técnicas de correlación se basan en minimizar la diferencia en el desplazamiento entre tramas en un bloque de píxeles. Representan el acercamiento a la estimación de movimiento mas ampliamente usado hoy en día. En su forma básica, el

algoritmo de correlación de bloques divide una imagen en una serie de regiones de igual tamaño. Para cada una de las regiones se busca, en la siguiente trama, la posible correlación en su vecindad, minimizando un criterio de error como la diferencia en el desplazamiento entre tramas, u otra medida relacionada, sobre un conjunto de vectores de movimiento.

Al trabajar con un tamaño fijo de bloque, se está limitando el campo del movimiento estimado, ya que no hay posibilidad de manejar bloques que contengan varios movimientos.

Una desventaja compartida de los esquemas basados en la correlación de regiones es su dependencia del modelo de movimiento traslacional. Las regiones deben permanecer bastante pequeñas para que el modelo traslacional permanezca válido.

2.3.5.3 Métodos basados en el dominio de frecuencias

Los métodos basados en el dominio de frecuencias se clasifican en dos clases: los métodos basados en las fases, y los métodos basados en la energía.

Los métodos basados en las fases se aprovechan del hecho de que un cambio en el dominio espacial produce un cambio de la fase en el dominio de frecuencias, mientras que los métodos basados en la energía trabajan con la distribución de energía en el espacio de frecuencias al considerar el movimiento en sus orientaciones espacio-temporales.

Se han usado técnicas basadas en las fases para la estimación de la disparidad. En general, la disparidad se define en términos de la diferencia de fase en un punto dado entre sus valores filtro paso-banda de dos tramas.

Fleet y Jepson (1990) aplican una técnica basada en el gradiente a la componente fase de las salidas de un conjunto de filtros direccionales sintonizados a diferentes velocidades. [13]

Los métodos en el dominio de frecuencias presentan algunas ventajas notables respecto de muchos métodos en el dominio espacial. En primer lugar, son menos sensibles a los cambios en la iluminación global, y en segundo lugar, son bastante robustos con respecto al ruido.

a) Métodos basados en energía

La ventaja de los métodos basados en energía es la descomposición jerárquica de la secuencia de imagen en un dominio de frecuencia. Las técnicas basadas en energía, obtienen velocidades usando familias de filtros sobrepasados afinados por la velocidad y orientación. La transformada de Fourier de un patrón 2-D en movimiento es:

$$\hat{I}(\vec{k}, \omega) = \hat{I}_0(\vec{k}) \delta(\omega + \vec{v}^T \cdot \vec{k}) \quad [1]$$

donde $\hat{I}_0(\vec{k})$ es la transformada de Fourier de $I(\vec{x}, 0)$, $\delta(k)$ es una función delta de Dirac, ω denota la frecuencia temporal y $\vec{k} = (k_x, k_y)$ significa frecuencia espacial. Esto implica que todas las potencias asociadas con un patrón 2-D en movimiento serán mapeadas en un plano que atraviese el origen en el espacio de la frecuencia.

Por otro lado, Heeger [16] utilizó una familia de doce filtros Gabor [14] de diferentes resoluciones de espacio para extraer la información de velocidad desde secuencias de imágenes. El filtro de Gabor es un filtro lineal cuya respuesta al impulso es una función armónica modulada por una distribución gaussiana. Utilizando filtros Gabor que provee espacio-temporal simultáneo y localización de frecuencia, se obtiene una representación sobrepasada clara. Queda un ajuste del cuadrado que se aplica a la distribución resultante en la frecuencia-espacio.

b) Métodos basados en las fases

Las estimaciones de flujo más acertadas son producidas utilizando el acercamiento basado en fases de Fleet y Jepson [13]. Los métodos basados en fases, también utilizan una familia de filtros de velocidad que extraen una representación de la frecuencia local de la secuencia de imágenes. El flujo estimado es proporcionado por la búsqueda de un declive en el espacio de fase extraído.

La motivación de este acercamiento, está basado en el argumento de que la evolución del contorno de fase provee una buena aproximación a un campo de movimiento proyectado. La fase de salida de filtros sobrepasados es generalmente más estable que la amplitud cuando son buscados pequeños cambios dentro de la escena.

Aunque el flujo óptico es una medida localizada, también está caracterizada por pequeños desplazamientos. Mediante la derivación de la velocidad a partir de la fase en comparación con la magnitud real se obtienen importantes medidas de eficiencia.

2.3.5.4 Métodos Bayesianos

El problema de la estimación del movimiento puede formularse como un problema de estimación global bayesiano. La idea general consiste en usar un criterio de máximo a posteriori (MAP) para maximizar la probabilidad del campo de movimiento, dada la observación de la intensidad de la imagen en la próxima trama. La estimación del MAP puede lograrse por medio de los métodos de relajación, tales como el temple simulado, que garantizan la convergencia en un máximo global.

Los métodos bayesianos han evolucionado a partir de los métodos recursivos sobre píxeles, y normalmente se formulan usando campos aleatorios específicos, llamados campos de Markov.

Los enfoques bayesianos a la estimación del movimiento son de gran interés al enfrentarse directamente con una de las dificultades fundamentales; cómo segmentar el campo de movimiento en regiones de movimiento coherente trabajando píxel a píxel. Sin embargo, padecen de la lentitud en la convergencia inherente a los métodos de relajación.

2.4 Obtención de parámetros de movimiento

La obtención de los parámetros de movimiento a través del campo de flujo óptico depende del tipo de transformaciones espaciales y geométricas que el objeto dentro de él presente, así como de las características propias del objeto (contorno, estructura rígida o no rígida, textura, etc).

De esta forma, los dos tipos principales de movimientos tienen las siguientes variantes:

- Espaciales: el objeto se mueve linealmente sobre los ejes del plano.
- Geométricos:
 1. Rotación: el objeto gira sobre su centro.
 2. Escalamiento: el objeto se acerca o se aleja del punto de visión.
 3. Esquileo: deformación del objeto en sentidos opuestos.
 4. Presión: deformación del objeto hacia sí mismo.

Para los objetivos que ocupan a esta tesis, sólo se tomarán en cuenta las transformaciones espaciales del objeto en el plano xy , y su transformación geométrica en su variante de rotación.

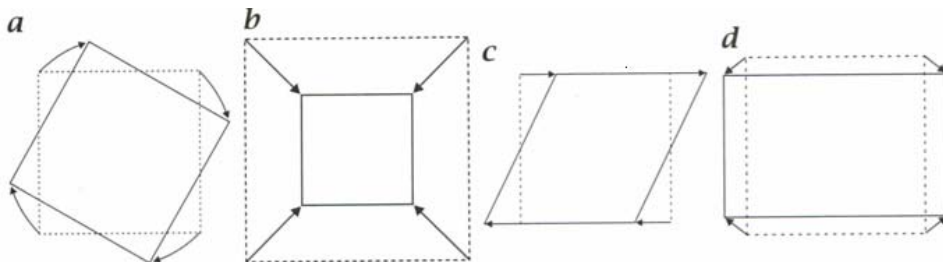


Figura 2.8 Transformaciones geométricas elementales en una superficie plana, el elemento sufre distintas transformaciones: **a)** rotación **b)** escalamiento **c)** esquileo **d)** presión. [4]

2.5 Técnica implementada

Teniendo en cuenta las técnicas que se han analizado y los parámetros que hemos de obtener, se tomó la decisión de utilizar las variantes de flujo óptico basadas en la correlación y en el gradiente, ya que éstas presentan una menor complejidad en su

implementación y al mismo tiempo otorgan resultados satisfactorios de acuerdo a la literatura revisada.

El uso de ambas técnicas se justifica de la siguiente manera: Primeramente se usa la técnica de correlación para estimar el movimiento de un grupo de píxeles dentro de un radio determinado. El resultado de esta búsqueda permite saber cómo se mueve este grupo, sin embargo, no sabemos qué sucede dentro de éste, pues sólo sabremos que un grupo se desplaza en cierta dirección, pero no la manera en que cada píxel se comporta. Por ello, es necesario usar la técnica basada en el gradiente, con la cual obtenemos niveles de precisión para cada píxel dentro de un grupo en movimiento.

Ahora bien, si no se usa únicamente la técnica del gradiente que bien podría darnos precisión de cada píxel sin necesidad de la correlación, es porque el proceso de cálculo sería muy lento, pues se debería hacer el rastreo individual de cada píxel de la imagen, mientras que al aplicar la técnica de correlación, el algoritmo hace el rastreo individual por píxel sólo para aquellos grupos que se han identificado en movimiento, lo que reduce el tiempo de procesamiento.

De esta forma, al conocer la forma en que se mueve cada píxel, es más sencillo identificar el tipo de movimiento que se está presentando en cada transición de imágenes y podremos ofrecer resultados más precisos.