

Detección de movimiento

Extracto de “Visión por Computador. Fundamentos y métodos”. Arturo de la Escalera. Prentice Hall, 2001

Copia para el alumno con fines didácticos

6.3. Detección de movimiento.

Otro tipo de característica es la debida al movimiento. Puede ayudar a la segmentación ya que los puntos del mismo objeto presentan unas velocidades semejantes. Los distintos tipos de movimiento que existen son:

- Movimiento de la cámara.
- Movimiento de los objetos.
- Cambios en la iluminación.
- Cambios en la estructura, forma o tamaño del objeto.

Se aplicarán las siguientes restricciones (aunque luego las dos primeras se eliminarán):

- No existen cambios de iluminación entre una imagen y otra.
- Respecto al movimiento se supondrá que la cámara está fija y que son los objetos los que se mueven.
- Todos los objetos son rígidos.

El método más sencillo es aquel que se basa en las diferencias de imágenes. Aunque estos algoritmos son menos precisos, son mucho más simples y por tanto pueden procesar la imagen en poco tiempo, por lo que su utilidad es mayor.

Los pasos a la hora de analizar el movimiento son tres:

- Determinar en que zonas de la imagen existe algún objeto que esté moviéndose. En esta primera etapa no es importante la exactitud, sólo detectar las futuras zonas de interés, en las que se realizará un análisis más detallado.
- En la segunda etapa la atención sólo está en aquellas zonas donde se haya detectado actividad y se extrae información que pueda ser usada para el reconocimiento de objetos, análisis del movimiento, etc.
- La última etapa tiene en cuenta el conocimiento previo del problema en concreto. Así por ejemplo la información obtenida de un objeto pequeño que se mueve lentamente cerca de la cámara de otro con igual forma pero mayor que se mueve rápidamente pero lejos de ella es la misma.

Ahora se tendrá no una imagen sino una secuencia representada por $f(x,y,t)$. Como se indicó el método para determinar el movimiento será la diferencia entre dos imágenes. Luego la imagen que indicará el campo de velocidades entre los instantes i y j será:

$$V_{ij}(x,y) = \begin{cases} 1 & \text{si } |f(x,y,i) - f(x,y,j)| \geq T \\ 0 & \text{en cualquier otro caso} \end{cases}$$

El sentido de aplicar un valor de umbralización es para minimizar la influencia del ruido en ambas imágenes. Obviamente de igual manera movimientos pequeños no serán detectados. También deberá aplicarse un filtro ya que la influencia del ruido provoca pixeles aislados que hay que eliminar. Puede tomarse para ello la condición de considerar solamente aquellos pixeles que este conectados a sus cuatro u ocho vecinos. En la figura 6.17 se observan dos imágenes donde queda reflejado el movimiento de un coche y el de los arbustos. La imagen diferencia sufre una umbralización para después eliminar aquellos pixeles que no estén conectados a sus cuatro o a sus ocho vecinos.

Este método presenta la gran ventaja de su inmunidad a los cambios de iluminación ya que el tiempo transcurrido entre la toma de las dos imágenes suele ser muy pequeño. El gran inconveniente que presenta es que los cambios detectados son de dos tipos: la parte del fondo tapado por el objeto en la segunda imagen y la parte del

fondo que estaba ocluida en la primera imagen y que aparece en la segunda. Puede así verse en la figura 6.17 cómo aparecen grupos de píxeles por parejas representando estas diferencias. No se tiene así información (con solo dos imágenes) de hacia donde se dirige el objeto. Si se tuviese una imagen en la que todos los objetos que apareciesen fueran estáticos se podría tomar como referencia y así estar seguros que todos los píxeles marcados como movimiento fueran del objeto que se mueve. Ahora el problema que surge son los cambios de iluminación que hacen que la imagen de referencia pierda validez.

Se puede seguir entonces un enfoque como el de la figura 6.18. Se toma una imagen base donde se esté seguro de que todos los objetos son estáticos. Cuando llega una nueva imagen se compara con la referencia tal como se vio en el caso anterior. Si la diferencia está por encima de un umbral es que existe movimiento. Si no el cambio es debido a la iluminación y ese valor sustituye al anterior de la imagen base. Resultados pueden verse en la imagen 6.19.

Los métodos presentados hasta ahora tienen como gran ventaja su simplicidad pero también sufren varios inconvenientes:

1. La cámara debe estar fija.
2. La imagen resultado indica en qué zonas existe el movimiento, pero no da ninguna información sobre la magnitud o dirección de la velocidad.

Existen dos grandes grupos de métodos para la obtención del campo de velocidades que tratan de solucionar estos dos problemas. En el primero se buscan una serie de características en una imagen (esquinas, bordes) y se encuentra su correspondencia en la segunda. La dificultad estriba en encontrar qué característica de la segunda imagen corresponde a una dada de la primera. Aunque son métodos rápidos no dan mucha información ya que el número de características elegidas suele ser bajo (si se compara con el número total de píxeles de la imagen).

El segundo grupo de métodos busca obtener un mapa denso del campo de velocidades (flujo óptico u optical flow en la literatura anglosajona). Para ello se tendrán en cuenta todos los píxeles de las dos imágenes. Como idea general se busca en el entorno a un punto cuál es la zona más parecida en la segunda imagen dependiendo de sus variaciones espaciales y temporales. Es un método local y sufre del denominado *problema de la apertura*. Así si se tiene el caso de la figura 6.20, rápidamente se observa que el objeto oscuro se está desplazando arriba a la derecha. Sin embargo este sería un análisis global ya que observamos toda la imagen. El problema de la apertura se ve con claridad en las dos imágenes inferiores. Representan el mismo caso anterior pero ahora la única información es la que rodea a un punto. Solo se podrá concluir que el objeto se está desplazando hacia arriba. Si se podría conocer perfectamente el movimiento de la figura cuadrada si se observase a través del agujero una de las esquinas del mismo, tal como se aprecia en las dos últimas imágenes. Como conclusión no se puede calcular la velocidad verdadera de los puntos que constituyen la imagen, si solamente se tiene información en una pequeña vecindad en torno a un punto, ya que la única información obtenible directamente a partir de las vecindades de cada punto es la componente de la velocidad normal al contorno o borde que aparezca en la imagen (es decir sólo la componente en la dirección de los gradientes presentes en la imagen). Como caso extremo si la parte de la imagen que observamos hubiera estado situada justo en una zona donde no existieran gradientes de intensidad, el centro de cuadrado por ejemplo, no se podría determinar su velocidad. Se deben por tanto imponer restricciones adicionales, que de alguna manera lo que se está haciendo es introducir información adicional al planteamiento del problema.

Dado un punto P en el mundo se proyecta sobre un punto de la imagen $I(x,y)$ tomada en el tiempo t. Este mismo punto P se proyectará posteriormente, para un incremento de tiempo δt , en la posición $(x + \delta x, y + \delta y)$. Matemáticamente se expresa:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$$

La parte derecha de la ecuación puede ser desarrollada usando las series de Taylor en torno al punto (x,y):

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t + \varepsilon$$

Donde ε contiene términos de segundo orden y superior en δx , δy y δt . Restando $I(x,y,t)$ de ambos lados de la ecuación δt se obtiene:

$$\frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t = \theta(\delta t)$$

dividiendo por δt y tomándolo el límite cuando $\delta t \rightarrow 0$ se obtiene la *ecuación de restricción del Flujo óptico* (figura 6.21):

$$I_x u + I_y v + I_t = 0$$

Donde:

- I_x es el gradiente de la imagen respecto al eje horizontal.
- I_y es el gradiente de la imagen respecto al eje vertical
- I_t es el gradiente temporal de las imágenes tomada en tiempos distintos.
- u es la componente horizontal de la velocidad.
- v es la componente vertical de la velocidad.

Puede observarse como se tienen dos incógnitas (u,v) y una sola ecuación por lo que se necesitarían dos puntos del mismo objeto para determinar las componentes de la velocidad. Sin embargo esto no es del todo cierto, como se explicó, debido al problema de la apertura ya que realmente la ecuación anterior indica la componente normal del campo de velocidades, es decir la componente de la velocidad en la dirección del gradiente ya que:

$$I_x u + I_y v + I_t = (I_x \quad I_y) \begin{pmatrix} u \\ v \end{pmatrix} + I_t = 0$$

siendo la componente normal (figura 6.21):

$$V_n = \frac{-I_t}{\sqrt{I_x^2 + I_y^2}}$$

Se debe poner por tanto una nueva restricción, aquí se explicará la propuesta por Horn y Schunk: la suavidad del campo de velocidades (puntos cercanos pertenecientes al objeto poseerán velocidades similares). La estrategia que utilizan Horn y Schunk para obtener las fórmulas del algoritmo es la minimización de una función que contiene la condición de conservación y la condición de suavidad.

Se define el error de suavidad en el punto (i,j) como:

$$s(i, j) = \left(\frac{\partial u(i, j)}{\partial x} \right)^2 + \left(\frac{\partial u(i, j)}{\partial y} \right)^2 + \left(\frac{\partial v(i, j)}{\partial x} \right)^2 + \left(\frac{\partial v(i, j)}{\partial y} \right)^2$$

Además el error en la ecuación de restricción del flujo óptico es:

$$c(i, j) = (I_x(i, j)u(i, j) + I_y(i, j)v(i, j) + I_t(i, j))^2$$

Constituyendo finalmente el error total, que debe ser minimizado:

$$e = \sum_i \sum_j s(i, j) + \lambda c(i, j)$$

Donde λ es una constante que pondera el error en la ecuación del movimiento de la imagen relativo a la desviación de la condición de suavidad. El valor de este parámetro tiene una gran influencia sobre los resultados obtenidos del algoritmo. Si las imágenes son precisas y están libres de ruido el valor de λ será mayor que para el caso en el que la presencia de ruido es significativa.

Diferenciando el error e respecto de las variables u y v se obtiene:

$$\frac{\partial e}{\partial u} = 2(u - \bar{u}) + 2\lambda(I_x u + I_y v + I_t)I_x$$

$$\frac{\partial e}{\partial v} = 2(v - \bar{v}) + 2\lambda(I_x u + I_y v + I_t)I_y$$

donde \bar{u} y \bar{v} son las medias locales de u y v . Estas ecuaciones tendrán el mínimo cuando sus valores sean nulos, por lo que pueden ser reescritas de la forma:

$$(1 + \lambda I_x^2)u + \lambda I_x I_y v = \bar{u} - \lambda I_x I_t$$

$$(1 + \lambda I_y^2)v + \lambda I_x I_y u = \bar{v} - \lambda I_y I_t$$

Si se expresa este sistema de ecuaciones de forma lineal:

$$(A) \begin{pmatrix} u_{kl} \\ v_{kl} \end{pmatrix} = (B)$$

Donde el determinante de A puede ser escrito como :

$$|A| = 1 + \lambda(I_x^2 + I_y^2)$$

De tal manera que al final queda:

$$(1 + \lambda(I_x^2 + I_y^2))u = (1 + \lambda I_y^2)\bar{u} - \lambda I_x I_y \bar{v} - \lambda I_x I_t$$

$$(1 + \lambda(I_x^2 + I_y^2))v = (1 + \lambda I_x^2)\bar{v} - \lambda I_x I_y \bar{u} - \lambda I_y I_t$$

Resolviendo estas ecuaciones para u y v se obtienen, finalmente, las fórmulas recursivas que constituyen el algoritmo:

$$u^{n+1} = \bar{u}^n - \frac{I_x \bar{u}^n + I_y \bar{v}^n + I_t}{1 + \lambda(I_x^2 + I_y^2)} I_x$$

$$v^{n+1} = \bar{v}^n - \frac{I_x \bar{u}^n + I_y \bar{v}^n + I_t}{1 + \lambda(I_x^2 + I_y^2)} I_y$$

La interpretación geométrica puede verse en la figura 6.22. El valor del campo de velocidades en un punto es igual al valor medio del entorno que le rodea menos un ajuste que se produce en la dirección del gradiente de los niveles de gris. La obtención del flujo óptico para un caso real se presenta en la imagen 6.23.

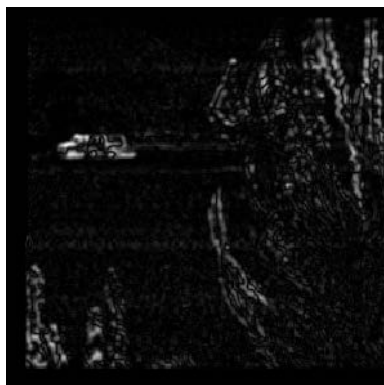


Figura 6.17 Detección de movimiento.

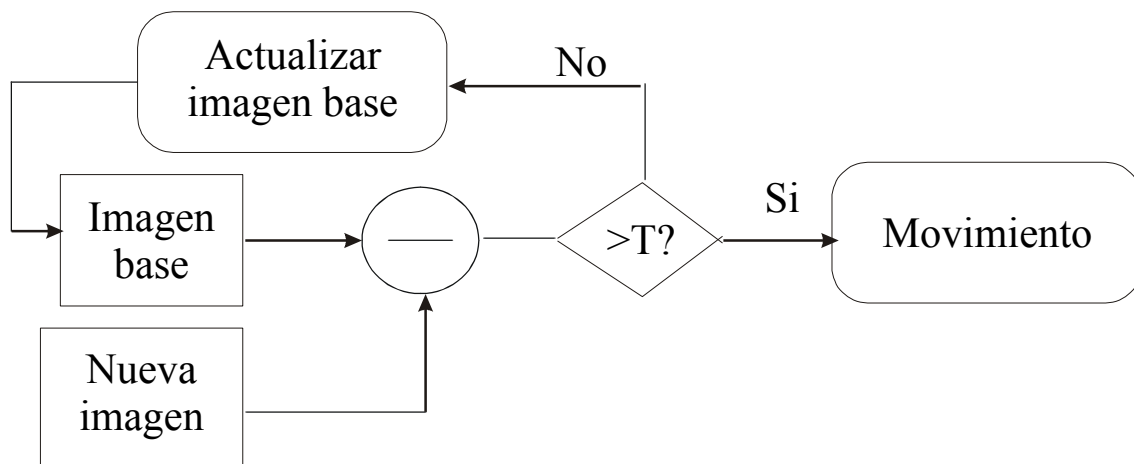


Figura 6.18. Detección de movimiento con actualización del fondo



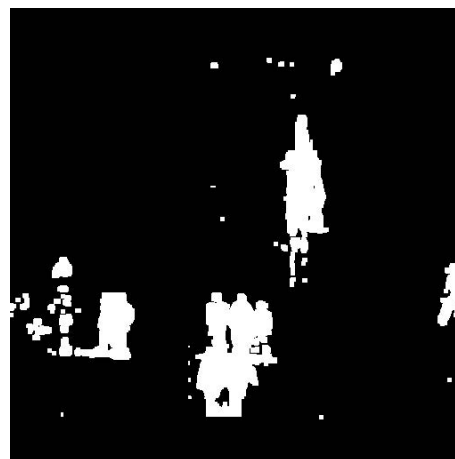
(a)



(b)



(c)



(d)

Figura 6.19. (a) Imagen original (b) Imagen base actualizada (c) Imagen (d) movimiento

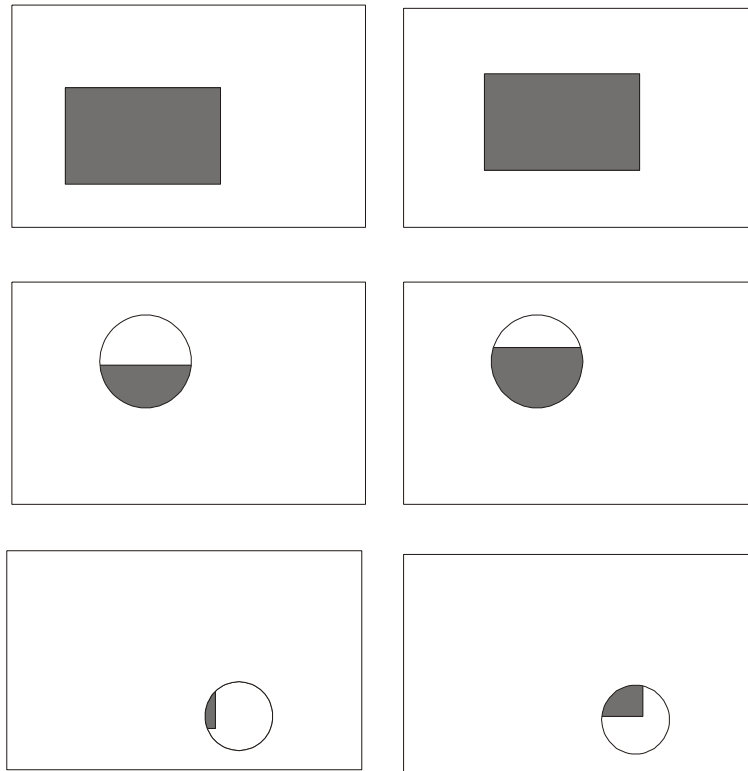


Figura 6.20 Problema de la apertura.

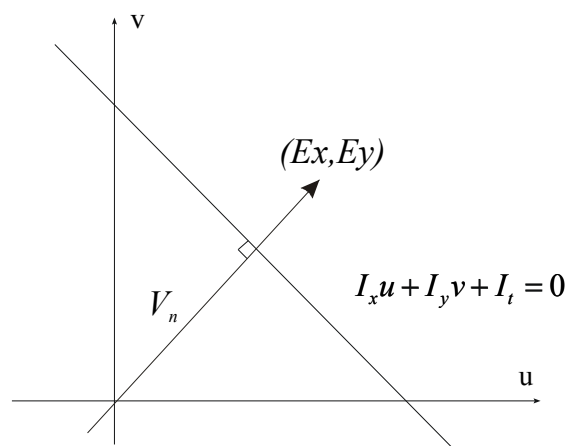


Figura 6.21. Ecuación de restricción del optical flow

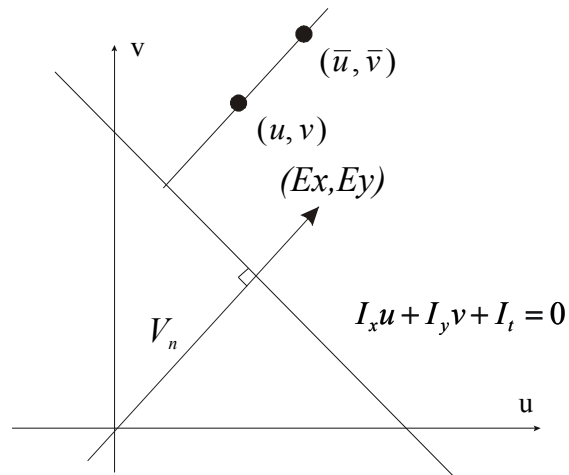


Figura 6.22. Obtención de las dos componentes de la velocidad

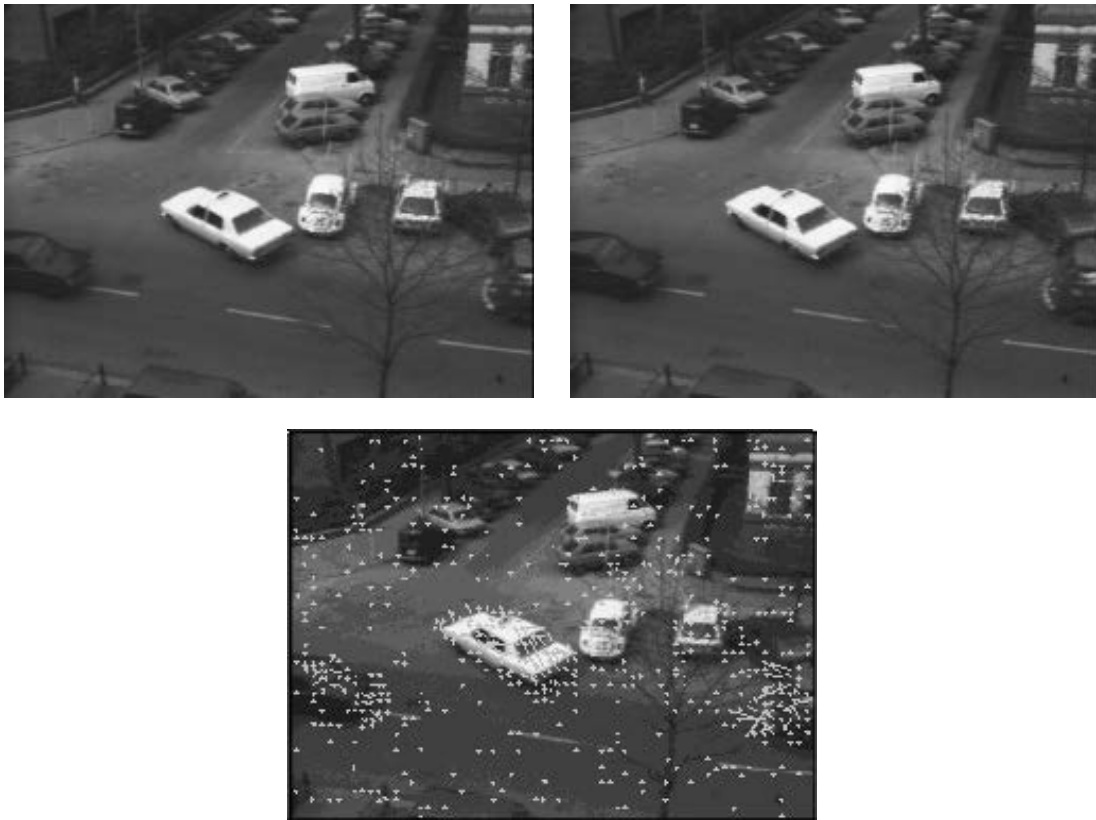


Figura 6.23. Obtención del optical flow.