Universidad Rey Juan Carlos
Escuela de Ingeniería de Fuenlabrada

Visión Artificial

9. Flujo óptico

JOSÉ MIGUEL GUERRERO HERNÁNDEZ

EMAIL: JOSEMIGUEL.GUERRERO@URJC.ES

Índice de contenidos

Universidad Rey Juan Carlos | Escuela de Ingeniería de Fuenlabrada

Índice de contenidos

1. Introducción

- La detección de movimiento es un tema fundamental en visión artificial y la robótica
- Múltiples aplicaciones:
  - Seguimiento de objetos
  - Cálculo de la estructura a partir del movimiento
  - Reconstrucción 3D
  - Representación y compresión de vídeo

1. Introducción

- El principal objetivo del análisis del movimiento es calcular el movimiento en la escena 3D
- Pueden darse varias posibilidades:
  - 3D-3D: registro de dos conjuntos 3D (ICP)
  - 2D-3D: características 2D sobre datos 3D (RANSAC)
  - 2D-2D: inferir movimiento 3D en base a movimientos detectados en imágenes 2D
- Nos centraremos en este último punto: inferir movimiento 3D en base a imágenes 2D

1. Introducción

- Un aspecto clave para detectar movimiento es el emparejamiento de datos en dos imágenes consecutivas
- Atendiendo a cómo realizamos este emparejamiento agrupamos los métodos en:
  - **Dispersos**: cuando utilizan alguna característica extraída de las imágenes
  - **Densos**: cuando tienden a emparejar todos los puntos de las imágenes

Índice de contenidos

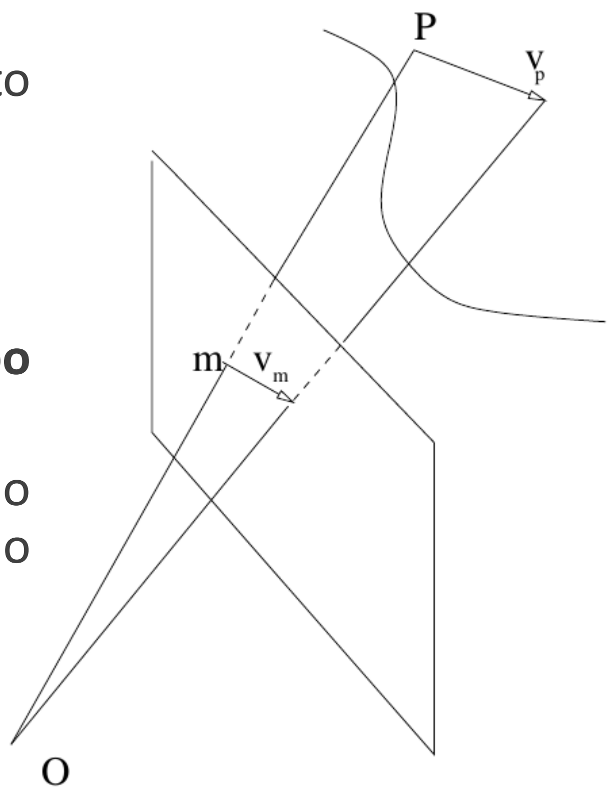
2. Campo de movimiento

- En una escena 3D, un punto P se mueve siguiendo un vector V_p
- La proyección de P en la imagen es un punto m que tiene asociado un vector v_m

$$P = [X, Y, Z]^T \qquad m = [x, y]^T$$

- El conjunto de estos vectores define el **campo de movimiento**
- La posición de m en función del tiempo depende tanto del desplazamiento de P como de su profundidad Z relativa a la cámara

$$\frac{dP}{dt} = \frac{dZ}{dt}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + Z\begin{bmatrix} dx/dt \\ dy/dt \\ 0 \end{bmatrix}$$



2. Campo de movimiento

$$\frac{dP}{dt} = \frac{dZ}{dt}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + Z\begin{bmatrix} dx/dt \\ dy/dt \\ 0 \end{bmatrix}$$

- El primer término representa el efecto del cambio de profundidad Z en el movimiento de la imagen
- El segundo término representa la velocidad del punto en la imagen, escalada por la profundidad
- Desarrollando la ecuación tenemos que $V_p = \left(V_p^T k\right)\widehat{m} + Zv_m$:
  - k es un vector unitario en la dirección de la profundidad
  - V_p es la velocidad del punto en 2D
  - $\left(V_p^T k\right)\widehat{m}$ es la proyección de V_p en la dirección k
  - Zv_m relaciona el campo de movimiento en la imagen con la profundidad

2. Campo de movimiento

$$V_p = \left(V_p^T k\right)\hat{m} + Z v_m$$

- Despejando obtenemos que $v_m = \frac{1}{Z}\left(V_p - \left(V_p^T k\right)\hat{m}\right)$, lo que implica que el campo de movimiento v_m está en función de V_p y Z
- Así para una determinada V_p :
  - Si Z es grande (el objeto está lejos), v_m es pequeño: el movimiento aparente en la imagen es menor
  - Si Z es pequeño (el objeto está cerca), v_m es grande: el objeto se desplaza más en la imagen para la misma velocidad

Índice de contenidos

3. Flujo óptico

- Los vectores de **campo de movimiento** no pueden ser **observados** directamente
- Lo que tenemos son imágenes y puntos de éstas
- Podremos decir que un punto se ha movido de un lugar a otro de la imagen, lo que implica su **flujo óptico**

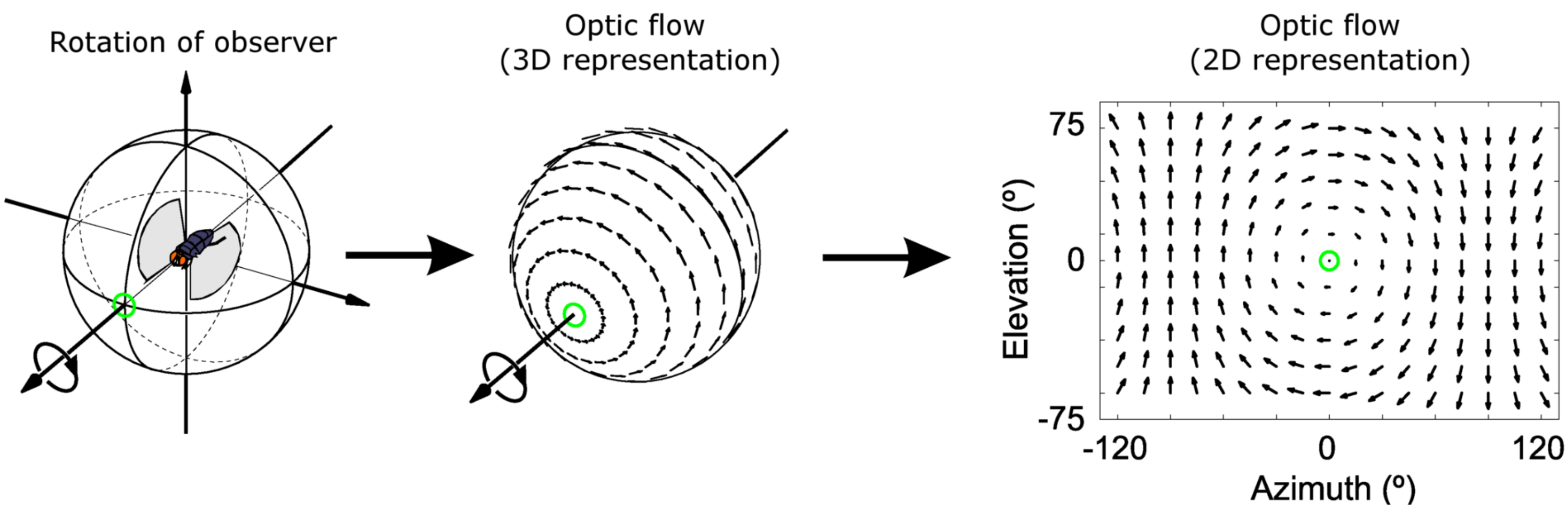
Flujo óptico ≠ Campo de movimiento

3. Flujo óptico

- Flujo óptico (wikipedia): es el patrón del movimiento aparente de los objetos, superficies y bordes en una escena causado por **el movimiento relativo entre un observador** (un ojo o una cámara) **y la escena**
- Aunque, en general, el **flujo óptico** no es igual al **campo de movimiento**, es lo único que tenemos para estimarlo
- Asumiremos que el **campo de movimiento** no estará muy lejos del **flujo óptico** que detectemos en una sucesión de imágenes

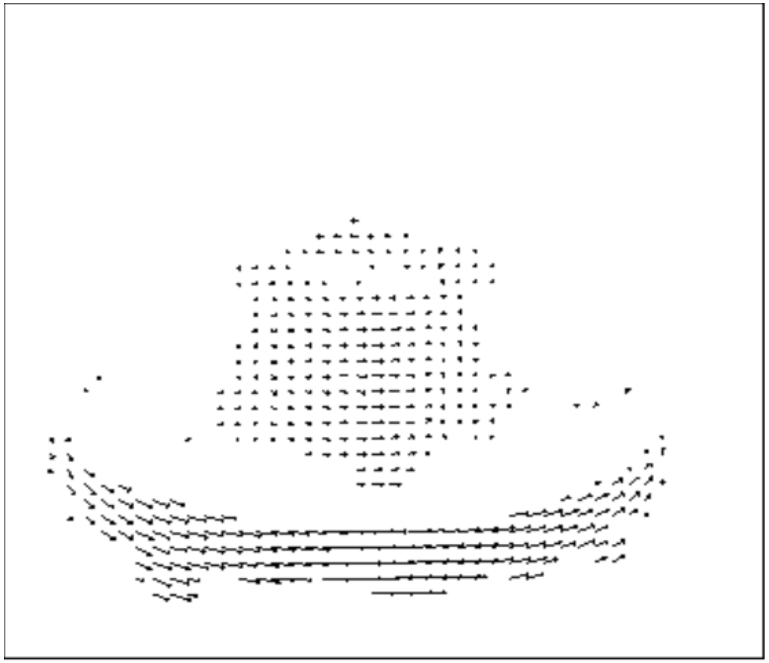
3. Flujo óptico

- Ejemplos:



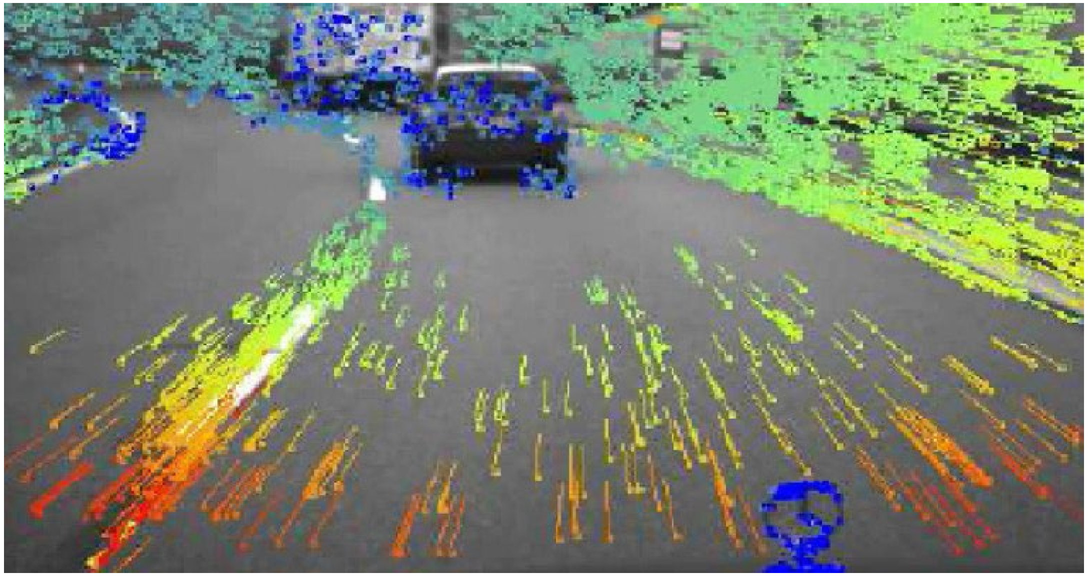
3. Flujo óptico

- Ejemplos:



3. Flujo óptico

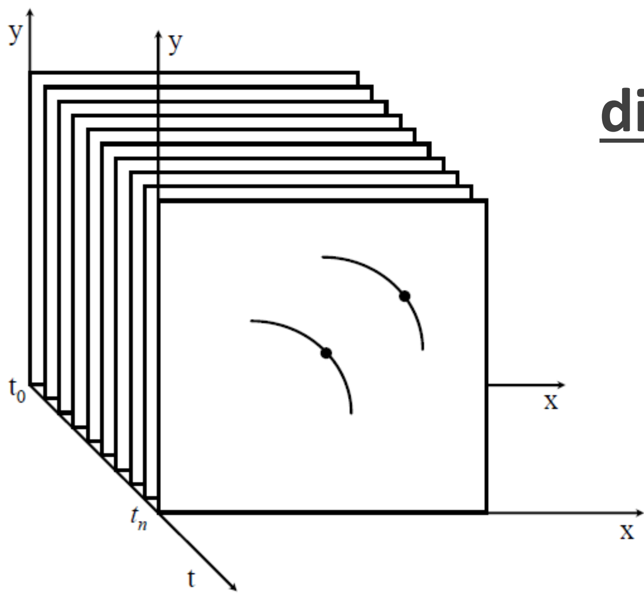
- Ejemplos:



Índice de contenidos

4. Cálculo del flujo óptico

- Dada una distribución espacio temporal de intensidad procedente de la secuencia de imágenes $I(x, y, t)$, el objetivo es obtener un vector de velocidad de la forma $v = (u, v)$



<u>dimensión espacio-temporal</u>

dimensión espacial: x, y

dimensión temporal: t

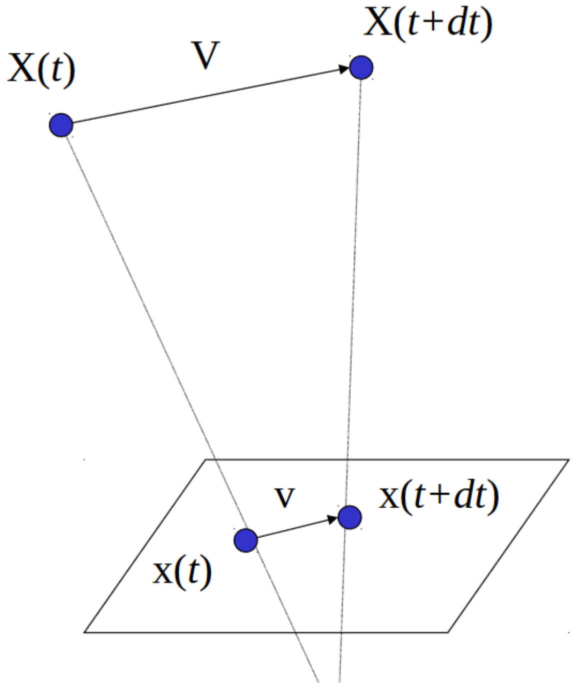
Flujo óptico

4. Cálculo del flujo óptico

- $\mathrm{X(t)}$ es un punto 3D en movimiento
- $\mathrm{x(t)} = \big(\mathrm{x(t)}, \mathrm{y(t)}\big)$ es la proyección de X en la imagen $I(\mathrm{x, y, t})$
- Velocidad aparente del píxel en la imagen v:

$$m = [x, y]^T$$

$$v_m = [v_x, v_y]^T = \begin{bmatrix} dx/dt \\ dy/dt \end{bmatrix}$$



4. Cálculo del flujo óptico

- Si asumimos que la intensidad de m no cambia en el tiempo:

$$I(x + v_x dt, y + v_y dt, t + dt) = I(x, y, t)$$

- Si la intensidad cambia levemente (series de Taylor)

$$I(x, y, t) + \frac{\partial I}{\partial x} v_x dt + \frac{\partial I}{\partial y} v_y dt + \frac{\partial I}{\partial t} dt + O\left(dt^2\right) = I(x, y, t)$$

$$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0$$

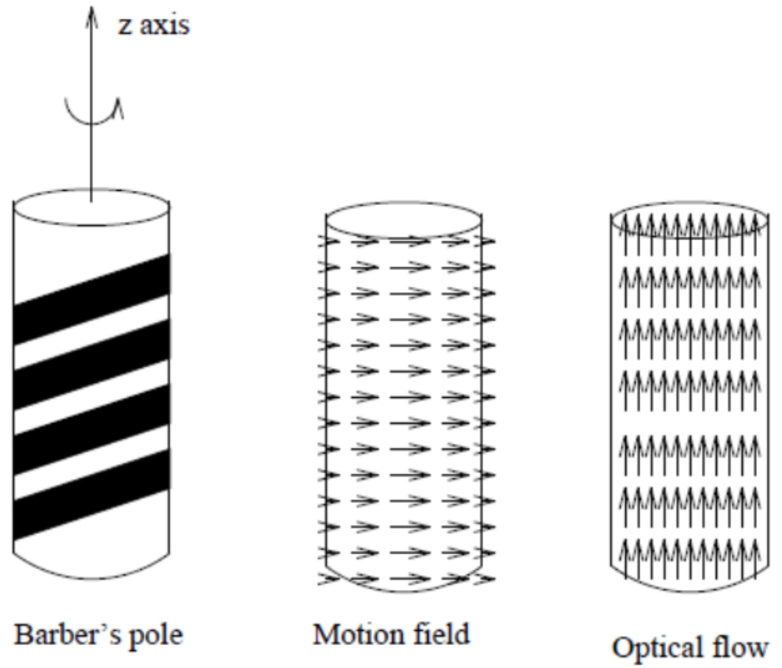
$$\nabla I \cdot v_m + \frac{\partial I}{\partial t} = 0$$

Índice de contenidos

Universidad Rey Juan Carlos | Escuela de Ingeniería de Fuenlabrada

5. Problema de la apertura

- El campo de movimiento y el flujo óptico pueden no coincidir

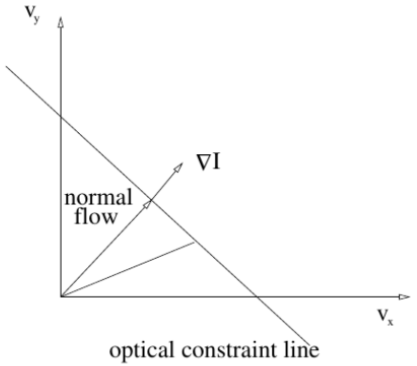


5. Problema de la apertura

- Tenemos una ecuación y dos incógnitas: no podemos determinar unívocamente el flujo óptico

$$\left(\frac{\partial I}{\partial x},\frac{\partial I}{\partial y}\right)\cdot\left(V_x,V_y\right)=-\frac{\partial I}{\partial t}\qquad\rightarrow\qquad I_xV_x+I_yV_y+I_t=0$$

  - I_x y I_y son los gradientes de intensidad en las direcciones x e y
  - I_t es el cambio temporal de intensidad
  - V_x,V_y es el flujo óptico buscado

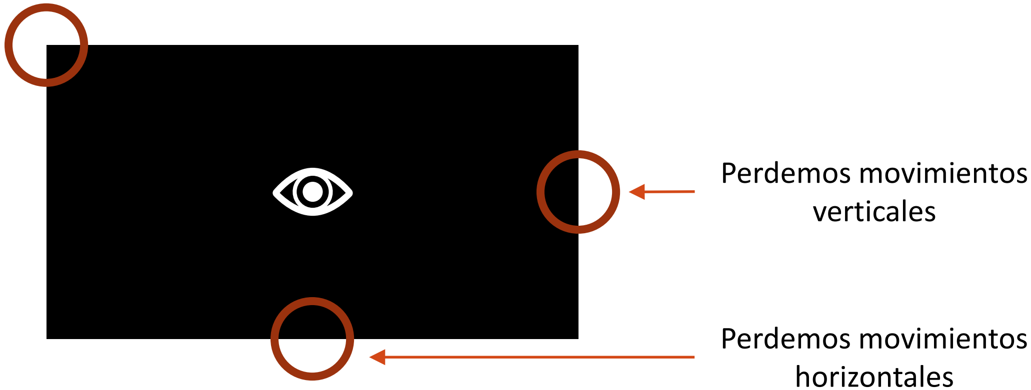


- Sólo podremos calcular el flujo óptico en la dirección normal del gradiente: **problema de la apertura**

Universidad Rey Juan Carlos | Escuela de Ingeniería de Fuenlabrada

5. Problema de la apertura

- Cada neurona del sistema visual es sensible a la entrada en una pequeña parte del campo de visión, ya que la está mirando a través de una pequeña ventana o apertura
- Una variedad de contornos con diferentes orientaciones moviéndose a diferentes velocidades pueden causar respuestas visuales idénticas



https://www.youtube.com/watch?v=vVGorOxMh8w

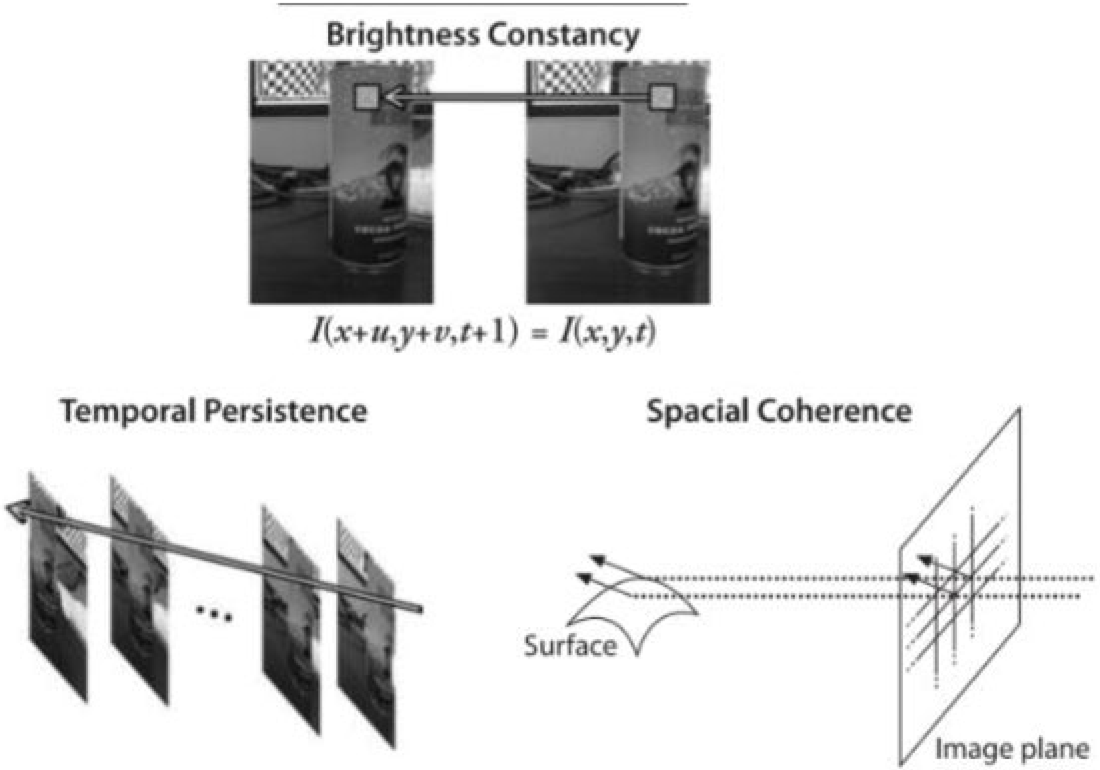
Índice de contenidos

6. Método de Lucas-Kanade

- La solución al problema de la apertura es la propuesta por **Lucas-Kanade**
- Consideraciones previas:
  - **Consistencia luminosa**: se asume que los keypoints cambian su posición, pero no su intensidad. Cualquier cambio en la intensidad de la imagen se debe al movimiento, la iluminación de la escena permanece constante y las superficies de los objetos son opacas (superficies Lambertianas)
  - **Coherencia espacial**: puntos vecinos pertenecen a la misma superficie y comparten misma velocidad
  - **Persistencia temporal**: los movimientos de un objeto en la escena son muy lentos en cuanto al tiempo Δt. Los incrementos temporales son más rápidos que los movimientos de los objetos

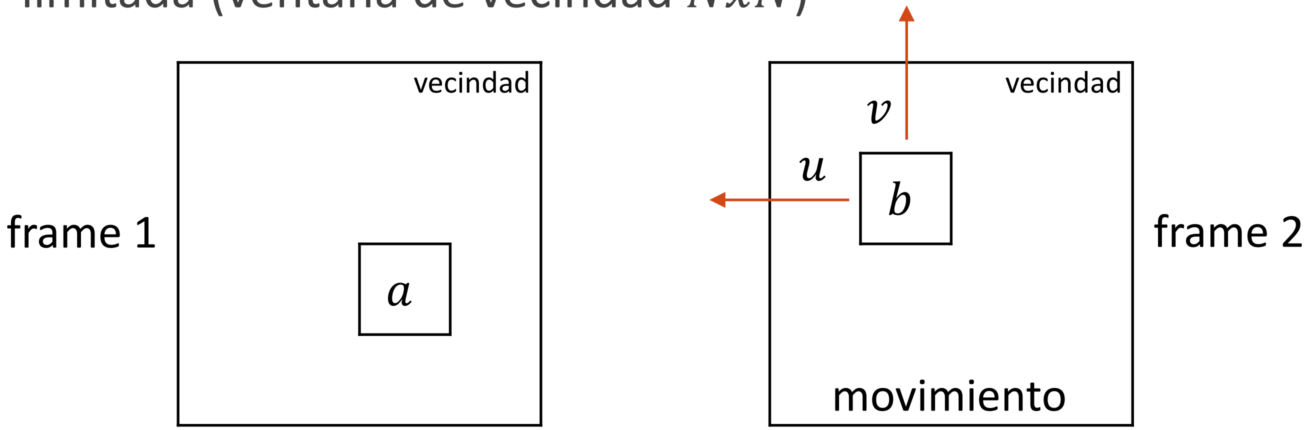
6. Método de Lucas-Kanade

- Consideraciones previas:



6. Método de Lucas-Kanade

- Método local que analiza la vecindad de cada punto:
  - Se asume que nosotros vemos la escena a través de una zona limitada (ventana de vecindad NxN)



  - La intensidad a dentro de esa ventana es variable
  - Del frame 1 al frame 2, a se ha movido a b
  - De este modo, se asume que a se ha desplazado arriba y a la izquierda a una nueva intensidad b, visible dentro de la ventana

6. Método de Lucas-Kanade

- Después de un movimiento de un píxel (x, y), u píxeles en la dirección x, y v en la dirección y, tenemos:

$$I_x(x, y) \cdot u + I_y(x, y) \cdot v$$

- Esto hace que la diferencia local en intensidad sea la ecuación general de flujo óptico:

$$I_x(x, y) \cdot u + I_y(x, y) \cdot v = -I_t(x, y)$$

- El signo negativo es necesario para garantizar que el flujo óptico estimado apunta en la dirección correcta en relación con el cambio de intensidad entre las dos imágenes

6. Método de Lucas-Kanade

- Dentro de una vecindad, por ejemplo, $3x3$, tenemos 9 ecuaciones lineales:

$$I_x(x+\Delta x, y+\Delta y)\cdot u + I_y(x+\Delta x, y+\Delta y)\cdot v = -I_t(x+\Delta x, y+\Delta y)$$

para $\Delta x = -1, 0, 1$ y $\Delta y = -1, 0, 1$

- Las ecuaciones se pueden resumir como:

$$S\binom{u}{v} = \vec{t}$$

donde:

- S es la matriz $9x2$ con derivadas parciales de intensidad $\left(I_x, I_y\right)$
- $\vec{t}$ es un vector que contiene los 9 términos $-I_t$
- $(u, v)^T$ es el vector de flujo óptico que buscamos

6. Método de Lucas-Kanade

- En el caso general, la siguiente ecuación no tiene solución exacta, por lo que la solución por mínimos cuadrados se consigue multiplicando la ecuación por S^T

$$S^T S \binom{u}{v} = S^T \vec{t}$$

- E invirtiendo $S^T S$:

$$\binom{u}{v} = \left(S^T S\right)^{-1} S^T \vec{t}$$

6. Método de Lucas-Kanade

- Finalmente, todo se reduce a la obtener la siguiente solución:

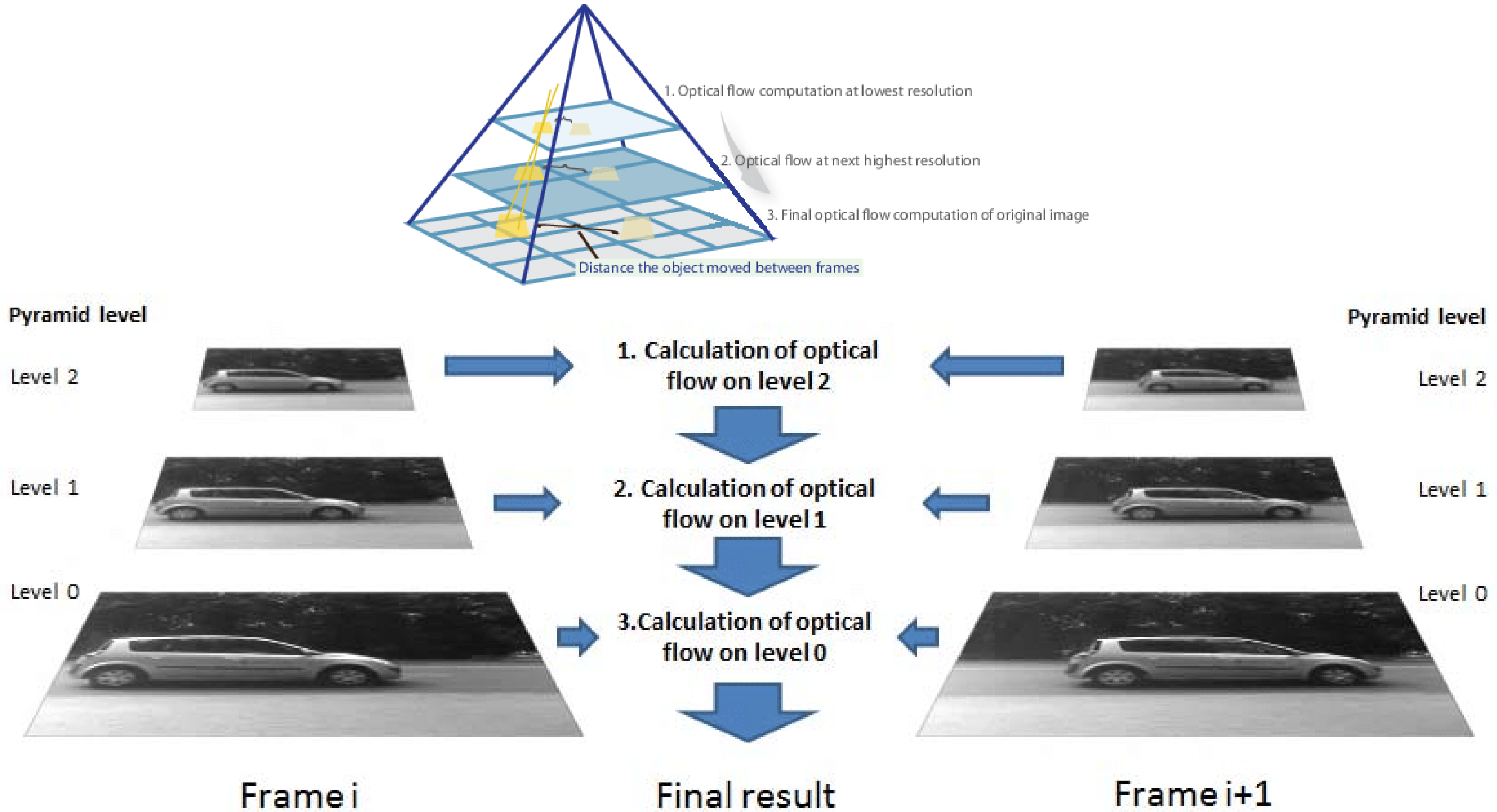
$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}^{-1} \begin{pmatrix} -\sum I_x I_t \\ -\sum I_y I_t \end{pmatrix}$$

- Puede verse la similitud de la matriz inversa con el detector de esquina de Harris, esto nos indica que las esquinas son los mejores puntos para ser rastreados

6. Método de Lucas-Kanade

- La idea después de todo es sencilla:
  - Se obtienen unos puntos para rastrear
  - Se obtienen los vectores de flujo óptico de esos puntos
- Problemas:
  - No funciona bien en zonas homogéneas (problema de apertura)
  - No trabaja bien con movimientos grandes
- Solución:
  - Utilizar una pirámide
  - Cuando se sube en la pirámide, se eliminan los movimientos pequeños y los movimientos grandes se convierten en pequeños
  - Obtenemos el flujo óptico con la escala

6. Método de Lucas-Kanade



Índice de contenidos

7. Método de diferencias

- Sean $f_1(i,j,t_1)$ y $f_2(i,j,t_2)$ dos imágenes consecutivas separadas por un intervalo de tiempo
- Un elemento $d(i,j)$ de la diferencia de imágenes entre f_1 y f_2 puede tener un valor 1, debido a alguna de las siguientes razones:
    1. $f_1(i,j)$ es un píxel de un objeto en movimiento y $f_2(i,j)$ es un píxel estático del fondo o viceversa
    2. $f_1(i,j)$ es un píxel de un objeto en movimiento y $f_2(i,j)$ es un píxel de otro objeto en movimiento
    3. $f_1(i,j)$ es un píxel de un objeto en movimiento y $f_2(i,j)$ es un píxel de una parte diferente del mismo objeto
    4. Ruido o imprecisiones en el posicionamiento de la cámara estacionaria

7. Método de diferencias

- El método de diferencias acumuladas se lleva acabo de la siguiente manera:

$$d_{acum}(i,j) = \sum_{k=1}^{n} a_k |f_1(i,j) - f_k(i,j)|$$

- a_k proporciona la importancia de las imágenes en la secuencia de las n imágenes

$$a_k = \frac{k-1}{n-1} \rightarrow n = 4; a_k = 0, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}$$

7. Método de diferencias

$$f_1 = \begin{bmatrix} 30 & 30 & 0 & 0 & 0 \\ 30 & 30 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad f_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 30 & 30 & 0 & 0 \\ 0 & 30 & 30 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$f_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 30 & 30 & 0 \\ 0 & 0 & 30 & 30 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad f_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & 30 \\ 0 & 0 & 0 & 30 & 30 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$d_{acum} = \begin{bmatrix} 60 & 60 & 0 & 0 & 0 \\ 60 & 50 & 30 & 20 & 0 \\ 0 & 10 & 30 & 50 & 30 \\ 0 & 0 & 0 & 30 & 30 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$