DETECCIÓN DE CAMBIOS EN UN VIDEO

Francisco J. Hernández López

fcoj23@cimat.mx

MOVIMIENTO

- Una simple imagen provee un instante de lo que sucede en una escena
- Un video registra la dinámica de la escena, entonces es posible reconocer los objetos que están en movimiento
- El movimiento lleva la información espacio-temporal entre los objetos que se encuentran en el campo de vista de la cámara, esta información puede ser utilizada en aplicaciones como:
  - Monitoreo de trafico
  - Video-vigilancia
- La intensidad o color de la imagen tienen alta correlación en la dirección del movimiento, esto puede ser utilizado para remover la redundancia temporal en la codificación del video (compresión y filtrado)

Alan C. Bovik. 2009. *The Essential Guide to Video Processing* (2nd ed.). Academic Press

NOTACIÓN

- $I: \Omega \times \mathcal{T} \rightarrow \mathbb{R}^+$: Intensidad de una secuencia de imágenes
- Ω : Dominio espacial
- $\mathcal{T}$: Dominio temporal
- $\vec{x} = (x_1, x_2)^T \in \Omega$: Posición en el espacio
- $t \in \mathcal{T}$: Posición en el tiempo
- $I(\vec{x}, t)$: denota la intensidad en $(\vec{x}, t)$
- I_t : denota una imagen al tiempo t
- $\vec{v} = (v_1, v_2)^T$: Velocidad

DETECCIÓN DE MOVIMIENTO

- Tareas relacionadas con el movimiento:
  - Detección
  - Estimación
  - Segmentación
- La Detección es la más simple de estas tres tareas
- El objetivo es identificar las regiones de la imagen que están en movimiento, ya sea que la cámara esté fija o en movimiento

SIMPLE UMBRAL (THRESHOLD)

- Algoritmo de detección de movimiento más simple

- Asumiendo $I_k(\vec{x}) = I_{k-1}(\vec{x}) + q$

  con q un termino de ruido Gaussiano con media cero y varianza σ^2, $\mathcal{N}(q;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{q^2}{2\sigma^2}}$

- Sea $r_k(\vec{x}) = I_k(\vec{x}) - I_{k-1}(\vec{x})$, entonces un simple detector de cambios podría ser el siguiente:

$$b(\vec{x}) = \begin{cases} 1 & si \;\; \mathcal{N}(r_k(\vec{x});\sigma^2) > \theta \\ 0 & otro\ caso \end{cases}$$

Nota: Podemos considerar que σ^2 es constante en toda la secuencia.

SIMPLE UMBRAL (EJEMPLOS)

1. $r_k^2(\vec{x}) \underset{S}{\overset{M}{\gtrless}} \theta$

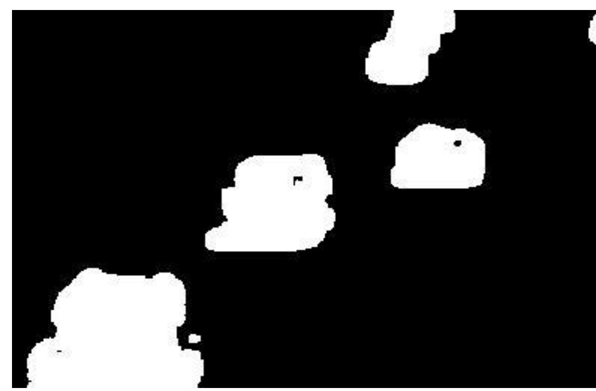
2. $\frac{1}{N} \sum_{\vec{m} \in W_{\vec{x}}} r_k^2(\vec{m}) \underset{S}{\overset{M}{\gtrless}} \theta$

No es robusto al ruido

Para θ muy pequeña, obtenemos mucho ruido

Para θ muy grande, obtenemos solo el contorno de los objetos y las partes que tienen mucha textura

Promediando las observaciones sobre una ventana espacial $W_{\vec{x}}$ centrada en $\vec{x}$ ayuda a disminuir el ruido



ANALIZANDO LAS DIFERENCIAS

- Sea $r_k(\vec{x}) = I_k(\vec{x}) - I_{k-1}(\vec{x})$, otro simple detector de cambios podría ser el siguiente:

$$b(\vec{x}) = \begin{cases} 1 & si \;\; |r_k(\vec{x})| > \theta \\ 0 & otro\ caso \end{cases}$$

con θ un valor pequeño positivo.

- $b(\vec{x}) = 1$, debido a alguna de las siguientes razones:

1. $I_k(\vec{x})$ es un pixel que pertenece a un objeto moviéndose.
   $I_{k-1}(\vec{x})$ es un pixel que pertenece al fondo estático.
   O viceversa
2. $I_k(\vec{x})$ es un pixel que pertenece a un objeto moviéndose.
   $I_{k-1}(\vec{x})$ es un pixel que pertenece a otro objeto moviéndose.
3. $I_k(\vec{x})$ es un pixel que pertenece a un objeto moviéndose.
   $I_{k-1}(\vec{x})$ es un pixel que esta en otra parte del mismo objeto moviéndose.
4. Ruido, pequeños movimientos de la cámara.

Sonka, Milan, Vaclav Hlavac, and Roger Boyle. Image processing, analysis, and machine vision. Cengage Learning, 2014.

ESTIMACIÓN DE MOVIMIENTO

- Puede ser necesario considerar los siguientes elementos:
  - Modelo de movimiento
  - Criterio de estimación
  - Estrategias de búsqueda

- Existen dos modelos fundamentales en la estimación del movimiento:
  1. Representando el movimiento en una secuencia de imágenes
     - Modelos de movimiento espacial
     - Modelos de movimiento temporal
  2. Modelo de observación → Relacionando los parámetros del movimiento con las intensidades de la imagen

MODELOS DE MOVIMIENTO ESPACIAL

- El objetivo es estimar el movimiento de los puntos en la imagen, este tipo de movimiento depende de:
  - Modelo de formación de la imagen (perspectiva, proyección ortográfica, etc.)
  - Modelo de movimiento de los objetos 3D (movimiento afine 3D, etc.)
  - Modelo de la superficie del objeto 3D (plano, parabólico, etc.)

- La velocidad instantánea 2D en la posición $\vec{x}$ en el plano de la imagen puede ser definida por

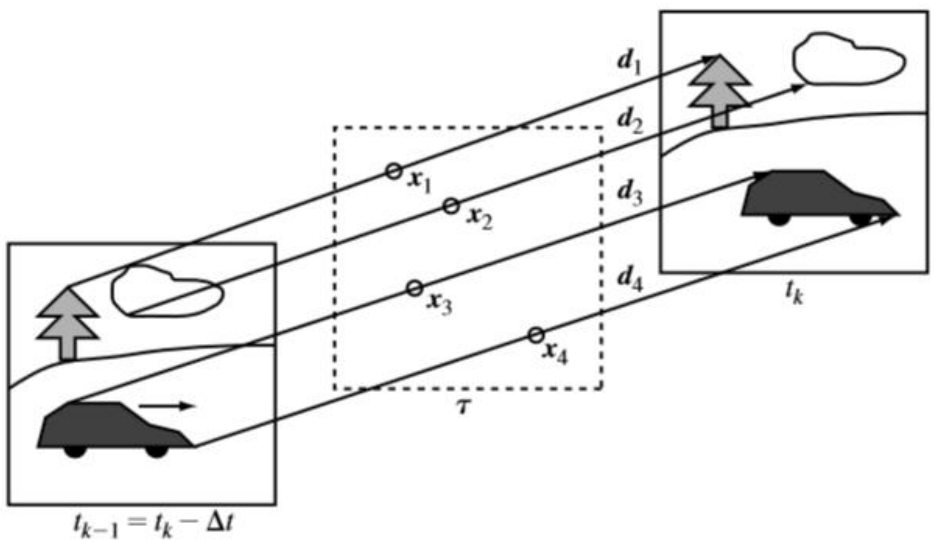
1. $\vec{v}(\vec{x}) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$
2. $\vec{v}(\vec{x}) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} b_3 & b_4 \\ b_5 & b_6 \end{pmatrix} \vec{x},$

donde $(b_1, b_2)^T = (v_1, v_2)^T$, b_3, b_4, b_5 y b_6 dependen de la geometría de la cámara y los parámetros de traslación 3D.

MODELOS DE MOVIMIENTO TEMPORAL

- Asumiendo que la velocidad $\vec{v}_t(\vec{x})$ es constante entre $t = t_{k-1}$ y τ $(\tau > t)$, entonces una trayectoria lineal puede ser:

$$\vec{x}(\tau) = \vec{x}(t) + \vec{v}_t(\vec{x})(\tau - t) = \vec{x}(t) + \vec{d}_{t,\tau}(\vec{x})$$



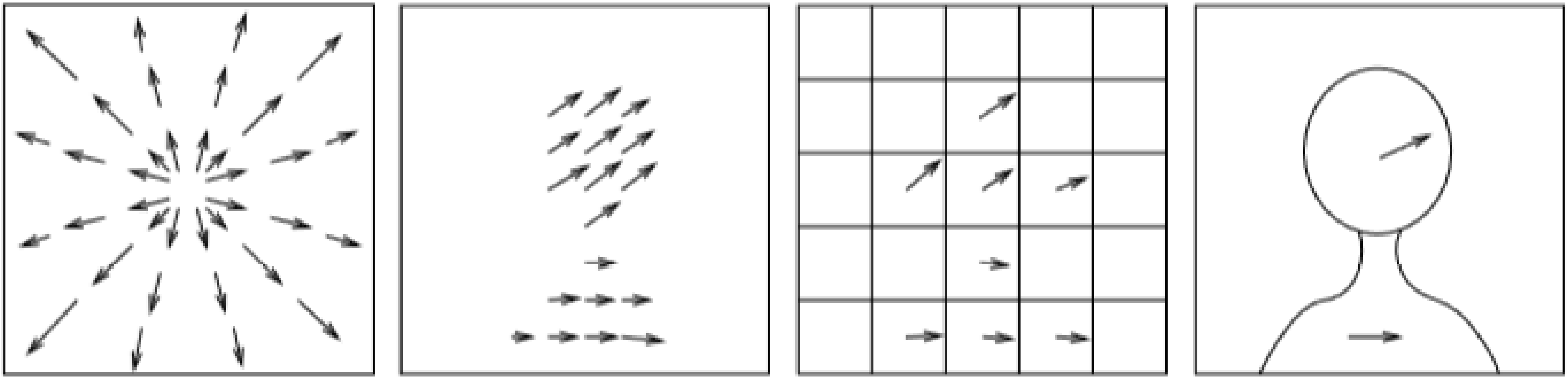
Alan C. Bovik. 2009. *The Essential Guide to Video Processing* (2nd ed.). Academic Press

- Una extensión puede ser el modelo cuadrático:

$$\vec{x}(\tau) = \vec{x}(t) + \vec{v}_t(\vec{x})(\tau - t) + \frac{1}{2}\vec{a}_t(\vec{x})(\tau - t)^2$$

REGIÓN DE SOPORTE

- Es el conjunto de puntos en donde se aplican los modelos de movimiento, denotado por $\mathcal{R}$



Por imagen Por pixel Por bloques Por región irregular

Alan C. Bovik. 2009. *The Essential Guide to Video Processing* (2nd ed.). Academic Press

MODELOS DE OBSERVACIÓN

- El objetivo es estimar el movimiento con base en las variaciones de intensidad en el tiempo
- Suposición: Los objetos no cambian su apariencia al moverse, es decir las intensidades permanecen constantes a lo largo de la trayectoria del movimiento
- Entonces, suponiendo que el movimiento es pequeño

$$I(\vec{x}, t) = I(x_1 + \Delta x_1, x_2 + \Delta x_2, t + \Delta t),$$

por series de Taylor

$$I(x_1 + \Delta x_1, x_2 + \Delta x_2, t + \Delta t) = I(\vec{x}, t) + \frac{\partial I}{\partial x_1}\Delta x_1 + \frac{\partial I}{\partial x_2}\Delta x_2 + \frac{\partial I}{\partial t}\Delta \mathrm{t} + \mathrm{O}(\partial^2),$$

tenemos la ecuación de restricción de movimiento

$$\left(\frac{\partial I}{\partial x_1}\Delta x_1 + \frac{\partial I}{\partial x_2}\Delta x_2 + \frac{\partial I}{\partial t}\Delta \mathrm{t} = 0\right) / \Delta \mathrm{t} \quad \rightarrow (\nabla I)^T \vec{v} + \frac{\partial I}{\partial t} = 0$$

con $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\right)^T$ y $\vec{v} = \left(\frac{\Delta x_1}{\Delta \mathrm{t}}, \frac{\Delta x_2}{\Delta \mathrm{t}}\right)^T$

CRITERIO DE ESTIMACIÓN

- Los modelos anteriores necesitan incorporarse en un criterio de estimación el cual será optimizado

- No existe un criterio único, la selección del criterio depende del problema que vamos a resolver, además, también depende de las capacidades del procesador

- Una elección común es la siguiente:

$$\mathcal{E}(\vec{d}) = \sum_{\vec{x} \in \mathcal{R}} \Phi\left[I_k(\vec{x}) - I_{k-1}\left(\vec{x} - \vec{d}(\vec{x})\right)\right]$$

con Φ una función no-negativa que puede ser:

- $\Phi(\mathrm{z}) = \mathrm{z}^2$ cuadrática
- $\Phi(\mathrm{z}) = \alpha|z|$ valor absoluto
- $\Phi(\mathrm{z}) = \log\left(1 + \frac{\mathrm{z}^2}{2\omega^2}\right)$ Lorentziana con parámetro ω

Estrategias de búsqueda

- Ya que los modelos han sido incorporados en un criterio de estimación, el último paso es desarrollar una eficiente y efectiva estrategia para encontrar los parámetros del movimiento
- La estrategia más común para minimizar un error de predicción $\left[I_k(\vec{x}) - I_{k-1}\left(\vec{x} - \vec{d}(\vec{x})\right)\right]$ es el emparejamiento (matching):
  1. Para varios candidatos $\vec{d}$ se calcula $\mathcal{E}\left(\vec{d}\right)$
  2. El óptimo será el candidato con el mejor match (el mínimo $\mathcal{E}$)
- Otra estrategia es utilizar el gradiente, el cual requiere un criterio de estimación $\mathcal{E}$ que sea diferenciable
  - Para evitar la optimización no-lineal, comúnmente usamos Taylor
  - Con Taylor, asumimos que el movimiento es pequeño
  - Si queremos obtener desplazamientos grandes, usualmente se utilizan estrategias jerárquicas o multi-resolución