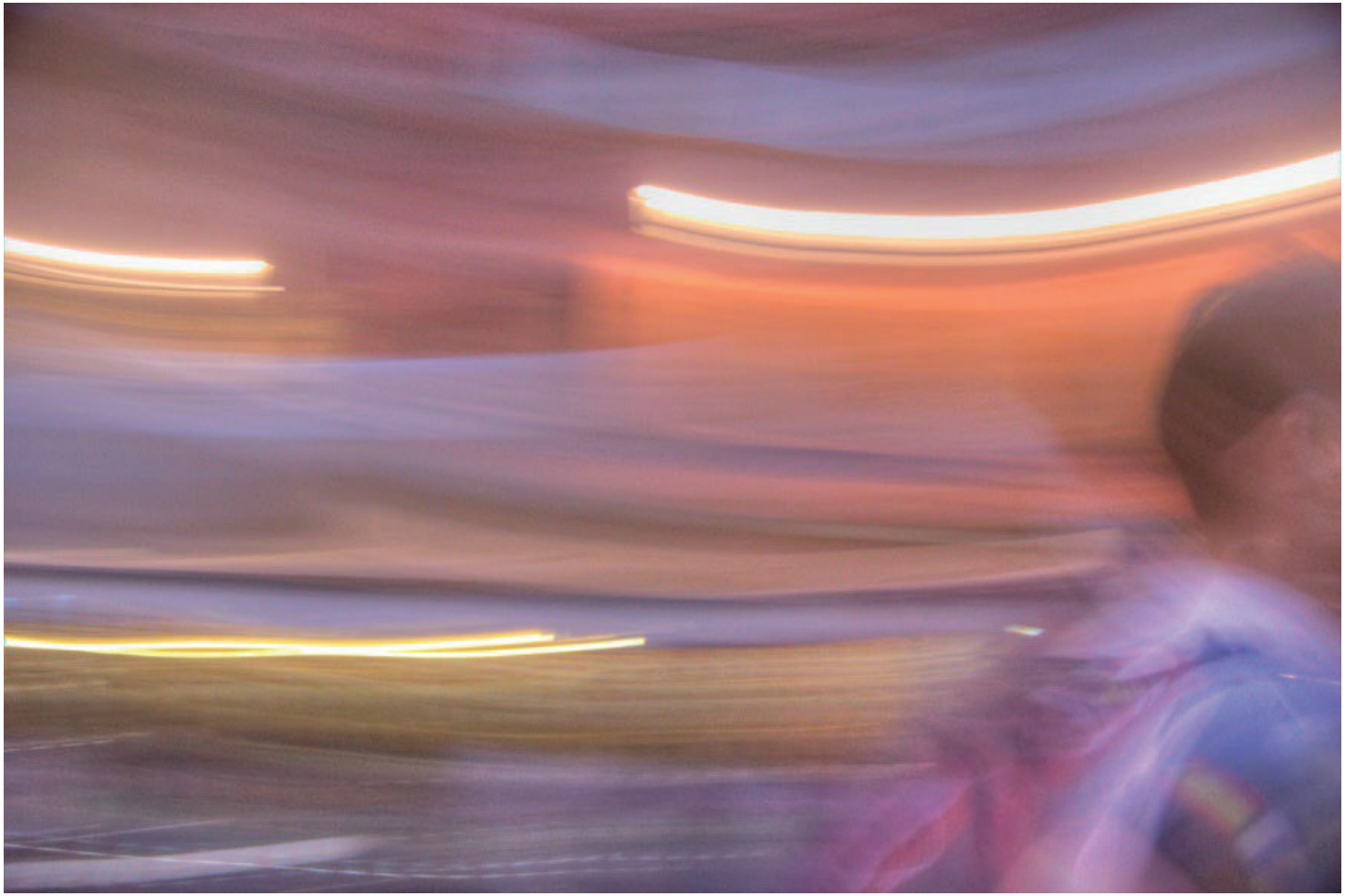
Visión por Computador

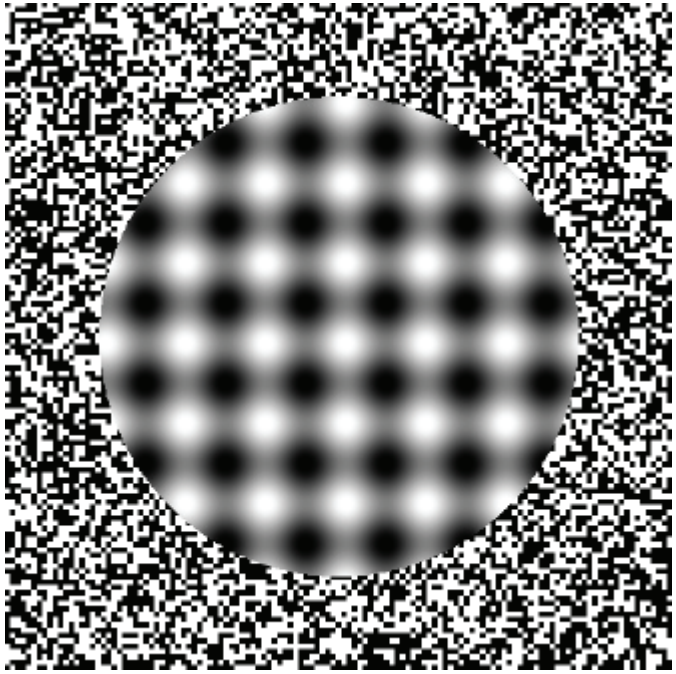
Tema 7. Movimiento



EII, GII. *Modesto Castrillón Santana*

Contenidos

- Introducción
- Traslacional
- Flujo óptico
- Seguimiento
- Filtro de Kalman



Introducción

¿Se mueve?

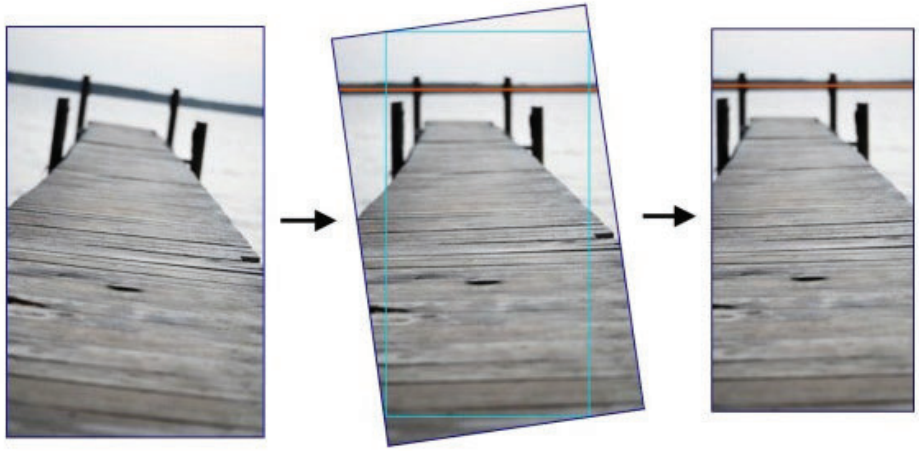
S. Seitz, R. Szeliski, M. Pollefeys

Introducción

¿Se mueve?

Introducción

Estimación de movimiento



Introducción

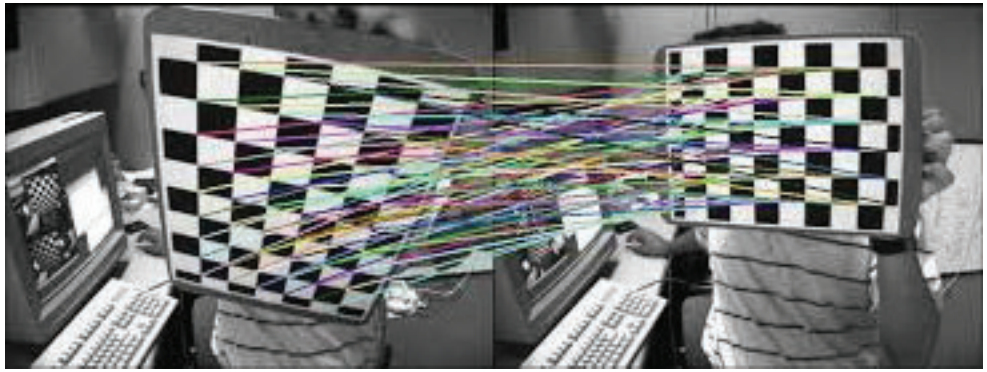
Alineamiento en imágenes

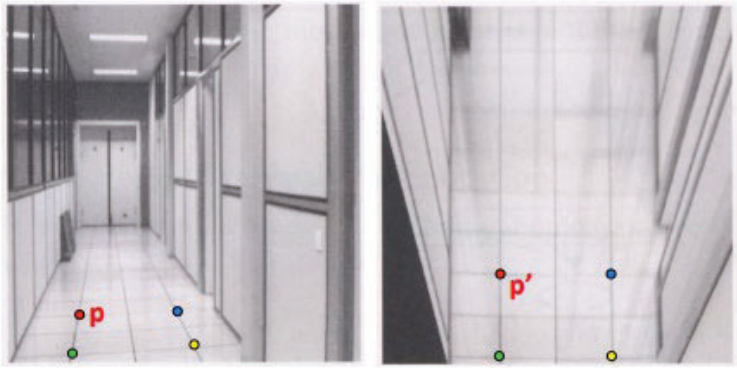
Rafa Afonso

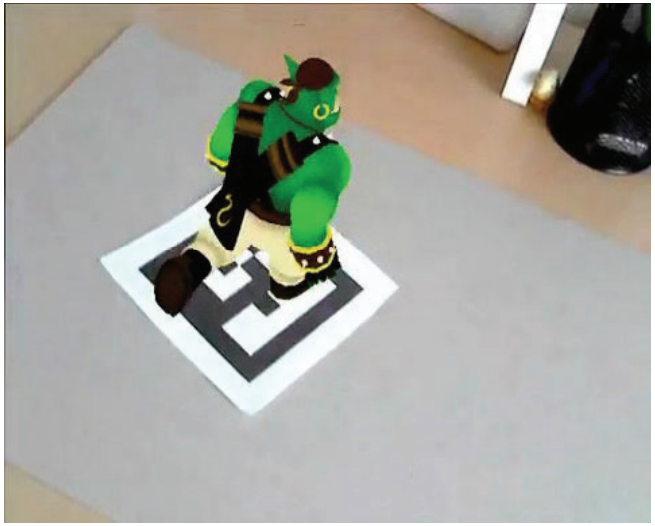
Introducción

Alineamiento en imágenes

Homografías (no necesariamente imágenes consecutivas)





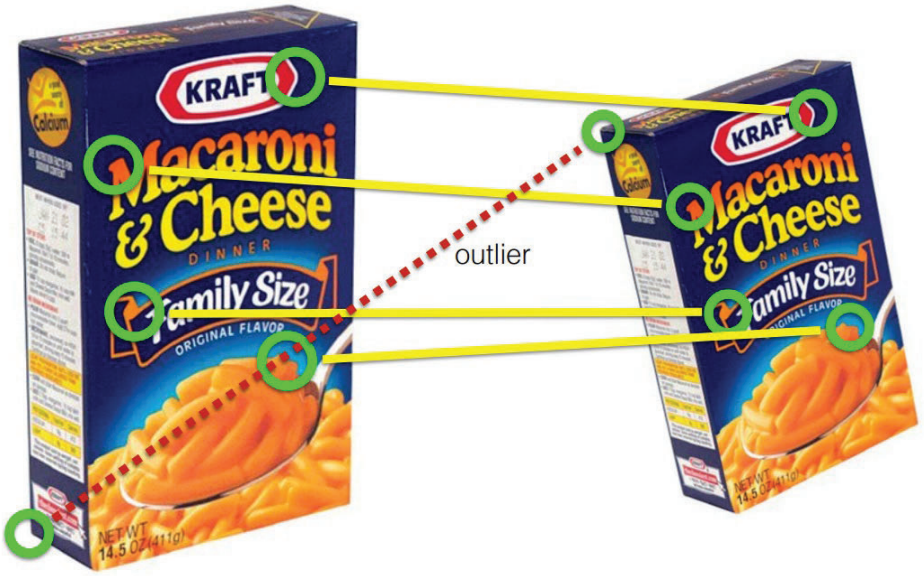


Introducción

Aplicado a *image stitching*

Introducción

¿Siempre buenas correspondencias?



RANSAC (RANdom SAmple Consensus, 1981) para eliminar *outliers*

<50% *outliers*

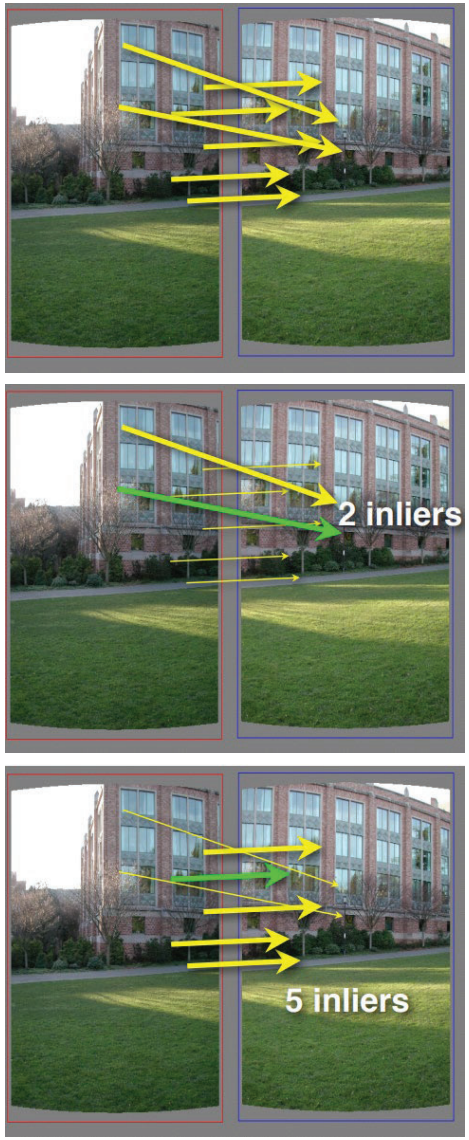
Introducción

RANSAC:

- Escoger par(es) aleatoriamente
- Contar *inliers*
- Repetir el proceso

Traslación con mayor número de *inliers*

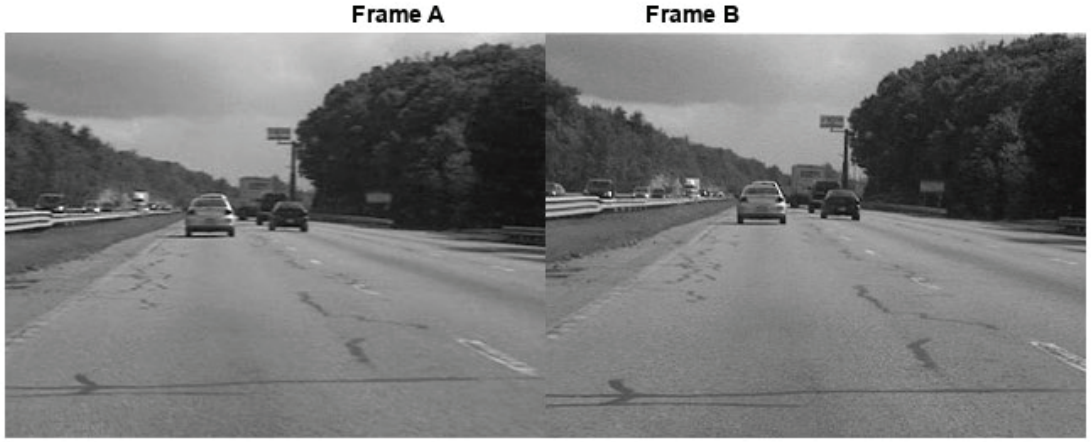
Repeticiones según número muestras



I. Gkioulekas

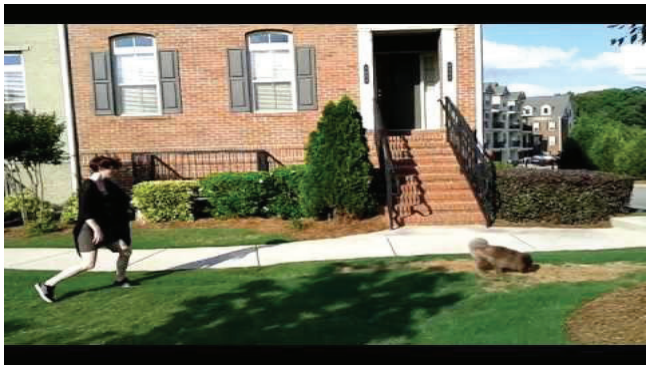
Introducción

Alineamiento en vídeo
Movimiento de píxeles



Introducción

Aplicado a estabilización de cámara
Con flujo óptico y/o seguimiento de patrón (*template*)



Introducción

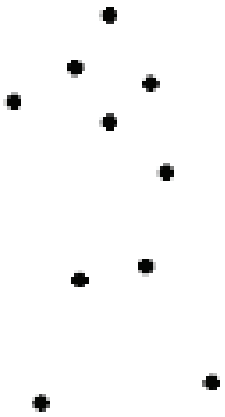
Movimiento como pista perceptiva
A veces única pista disponible

Introducción

Movimiento como pista perceptiva
A veces única pista disponible

Introducción

Potente incluso empobrecido

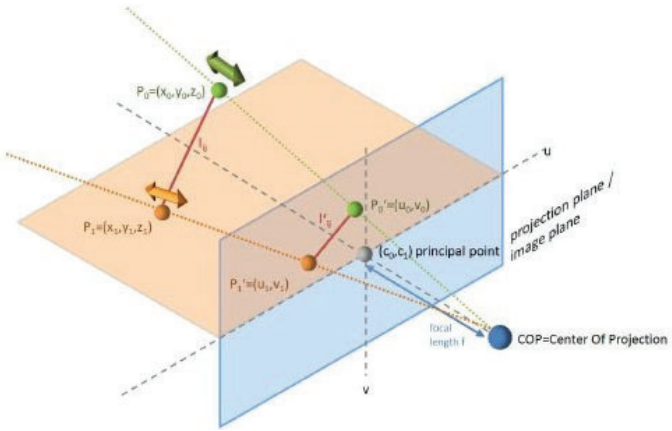


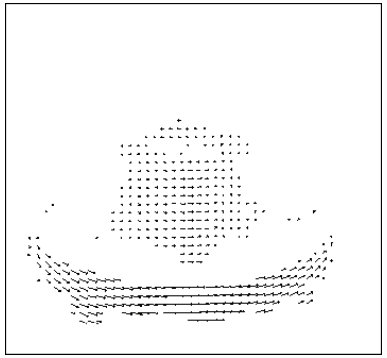
G. Johansson, “Visual Perception of Biological Motion and a Model For Its Analysis", *Perception and Psychophysics, 1973*.

Introducción

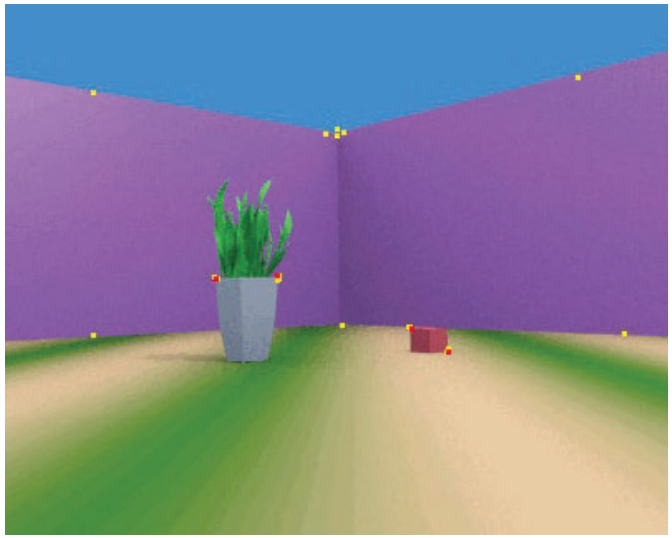
Estimación:

- campo de movimiento de secuencia
- movimiento 3D de puntos escena
- egomovimiento de la cámara





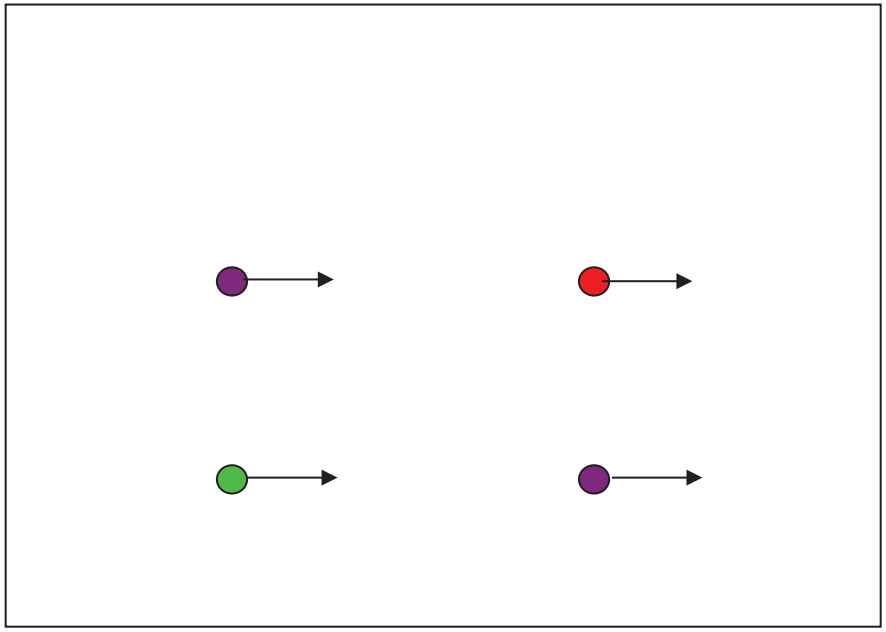
S. Seitz, R. Szeliski, M. Pollefeys



Alineamiento traslacional

Encaje más simple de dos imágenes

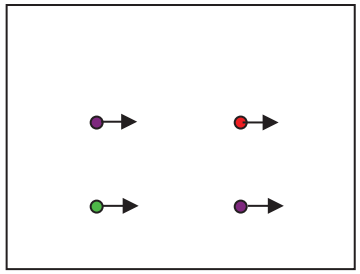
Imagen $f_0(x)$ con conjunto discreto de píxeles $p_i=(x_i,y_i)$



Alineamiento traslacional

Encaje más simple de dos imágenes

Imagen $f_0(x)$ con conjunto discreto de píxeles $p_i=(x_i,y_i)$



Minimizar suma de diferencias de cuadrados (SSD) o residuos

$$E_{SSD}(u) = \sum_i [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$

$\boldsymbol{u}$ desplazamiento (dos componentes), e_i error residual

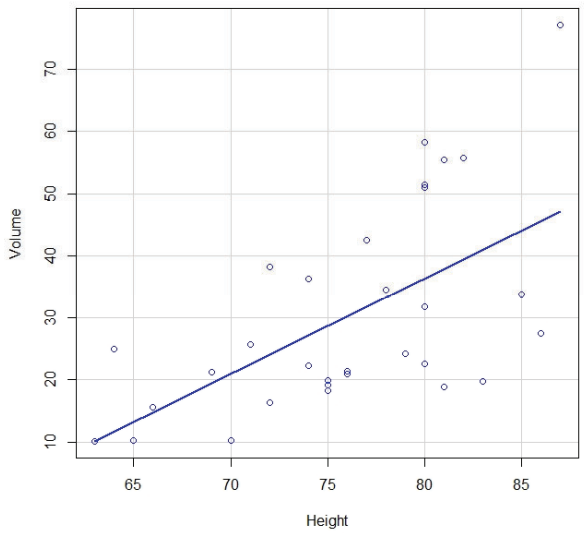
Solución de **mínimos cuadrados**

Alineamiento traslacional

Mínimos cuadrados

Regresión lineal que minimiza residuos

$$E_{SSD}(u) = \sum_i [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$



Alineamiento traslacional

E_{SSD} no contempla píxeles que *salen* de la imagen

$$E_{SSD}(u) = \sum_i [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$

Variante ponderada/ventana

$$E_{WSSD}(u) = \sum_i w_0(p_i) w_1(p_i + u) [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$

Permite eliminar píxeles/partes de la imagen

Alineamiento traslacional

Alternativa a E_{SSD} o diferencia de intensidades:
Correlación, máximo del producto

$$E_{CC}(u) = \sum_i f_0(p_i) f_1(p_i + u)$$

Alineamiento traslacional

Alternativa a E_{SSD} o diferencia de intensidades:
Correlación, máximo del producto

$$E_{CC}(u) = \sum_i f_0(p_i) f_1(p_i + u)$$

Correlación cruzada normalizada (rango [-1,1])
Obvia efecto de brillos altos:

$$E_{NCC}(u) = \frac{\sum_i \left[f_0(p_i) - \overline{f_0}\right]\left[f_1(p_i + u) - \overline{f_1}\right]}{\sqrt{\sum_i \left[f_0(p_i) - \overline{f_0}\right]^2}\sqrt{\sum_i \left[f_1(p_i + u) - \overline{f_1}\right]^2}}$$

$$\overline{f_0} = \frac{1}{N}\sum_i f_0(p_i) \qquad \overline{f_1} = \frac{1}{N}\sum_i f_1(p_i + u)$$

Alineamiento traslacional

¿Cómo buscar el mínimo (o máximo con correlación)?

Búsqueda exhaustiva por todo el rango de u

Alineamiento traslacional

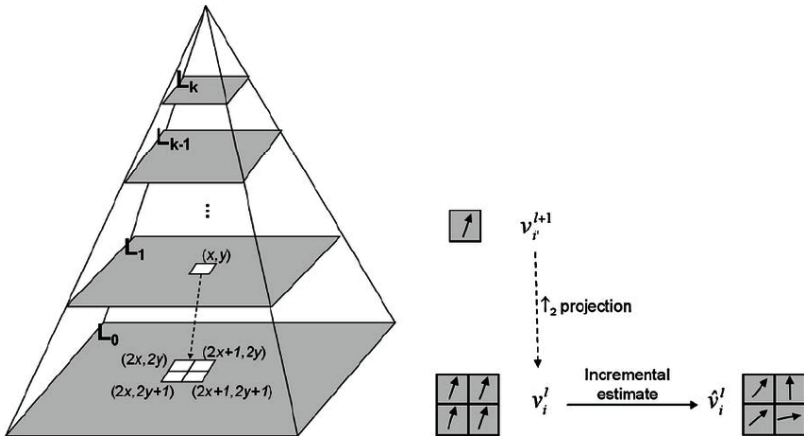
¿Cómo buscar el mínimo (o máximo con correlación)?

Búsqueda exhaustiva por todo el rango de *u*

Estimación de movimiento jerárquica

Basada en pirámides:

- Busca a menor resolución
- La estimación en un nivel, inicializa en niveles más finos
- Variantes con transformada de Fourier



Refinamiento

Lo previo alinea al píxel más cercano
Estabilización y *stitching* necesitan más exactitud
Buscar valores discretos *(u,v)* alrededor del mejor

Partiendo de E_{SSD}

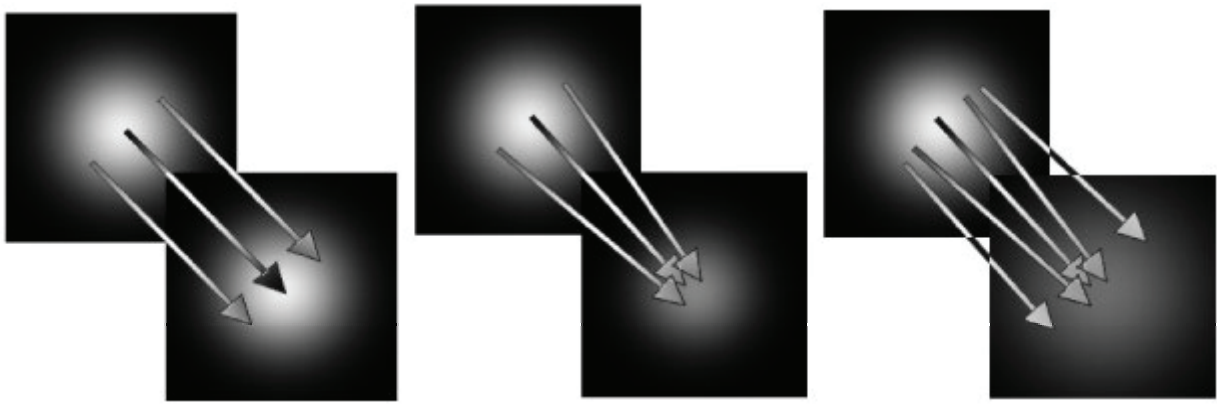
$$E_{SSD}(u) = \sum_i [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$

Refinamiento

Fotogramas consecutivos, saltos pequeños espacio-tiempo
Asumible brillo idéntico de los píxeles

$$E_{SSD}(u) = \sum_i [f_1(p_i + u) - f_0(p_i)]^2 = \sum_i {e_i}^2$$

Restricción de constancia de brillo



(a) (b) (c)

Refinamiento

Constancia de brillo en pequeños lapsos de tiempo

$$f_0\,(p_i) = f_1\,(p_i + u)$$

Linealizada con expansión de Taylor

$$f_0\,(p_i) \approx f_1\,(pi) + f_x u + f_y v$$

Términos espacial y temporal

$$f_x u + f_y v + f_t \approx 0$$
$$\nabla f u + f_t \approx 0$$

Uso frecuente para el **cálculo del flujo óptico** en una región

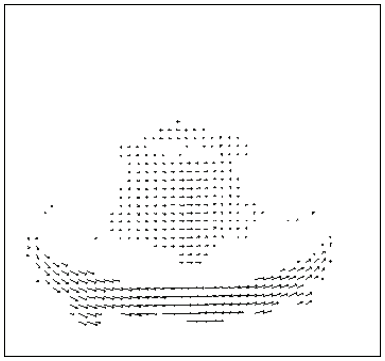
Flujo óptico

Cálculo de movimiento por píxel, más general

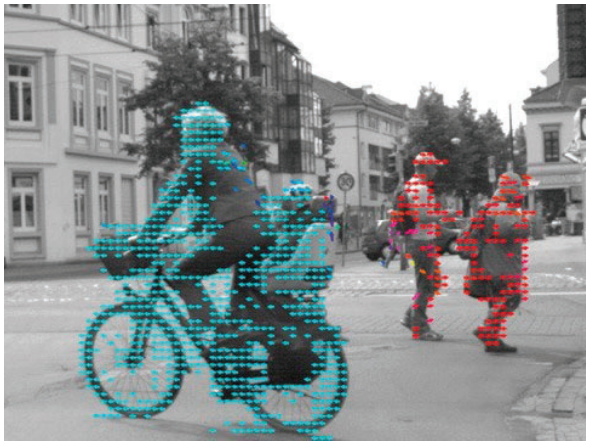
Campo de movimiento proyecta movimiento 3D en la imagen

Flujo óptico: aproxima movimiento real

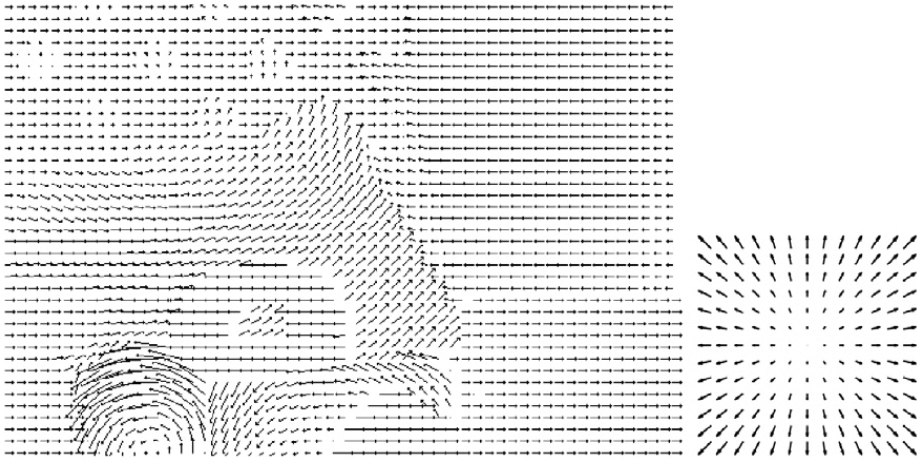
Movimiento aparente de los patrones de brillo



S. Seitz, R. Szeliski, M. Pollefeys



Flujo óptico

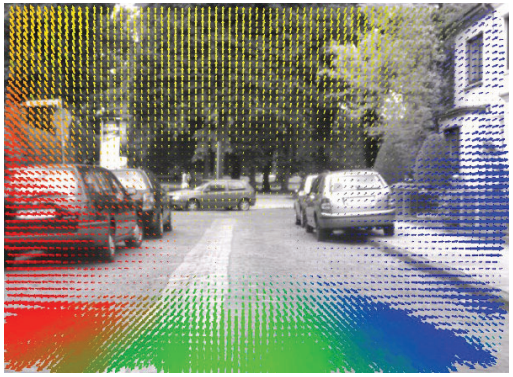


Flujo óptico

Aproximación del movimiento real (en muchos casos)

Producido por:

- movimiento de cámara
- del objeto
- o iluminación (no movimiento real)

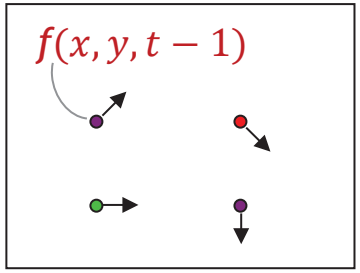


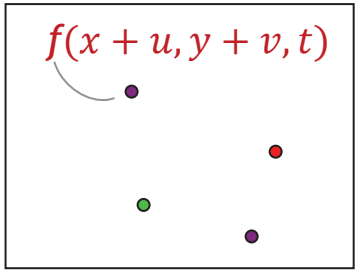
Minimizar diferencias de grises entre píxeles

$$E_{SSD-OF}(\{u_i\}) = \sum_i [f_1(p_i + u_i) - f_0(p_i)]^2$$

Estimación

Dados fotogramas en *t* y *t-1*
Campo de movimiento aparente *u(x,y)* y *v(x,y)* entre ellos





$$f(x,y,t-1)=f(x+u(x,y),y+v(x,y),t)$$

Asume:

- Constancia de brillo: Puntos similares o idénticos
- Movimiento reducido
- Coherencia espacial en vecindad

Estimación

$$f(x,y,t-1)= f(x+u(x,y),y+v(x,y),t)$$

Expansión de Taylor parte derecha

$$f(x,y,t-1) \approx f(x,y,t) + f_x\, u(x,y) + f_y\, v(x,y)$$

$$f_x\, u(x,y) + f_y\, v(x,y) + f(x,y,t) - f(x,y,t-1)=0$$

Resultando

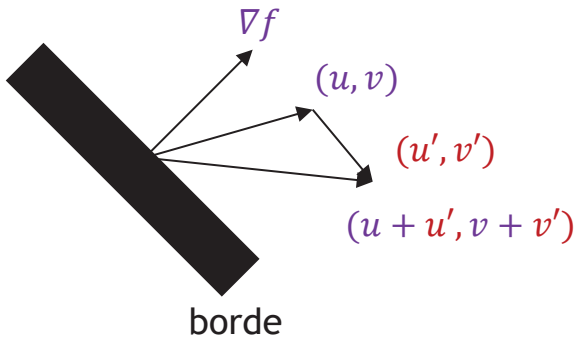
$$f_x\, u(x,y)+ f_y\, v(x,y)+ f_t=0$$

Estimación

$$f_x\, u(x,y) + f_y\, v(x,y) + f_t = 0$$

Gradientes f_x, f_y e f_t ¿determinación unívoca?

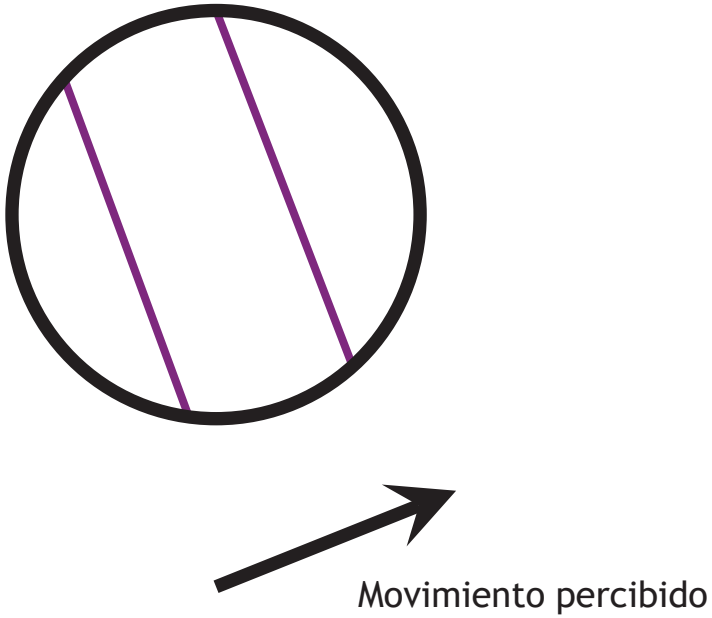
Asumiendo $\nabla f(u,v) + f_t = 0$
$\nabla f(u+u',v+v') + f_t = 0$ para todo (u',v') t.q. $\nabla f(u',v') = 0$
No puede recuperar componente perpendicular al gradiente



Problema de apertura

Mala textura afecta a que el sistema sea resoluble

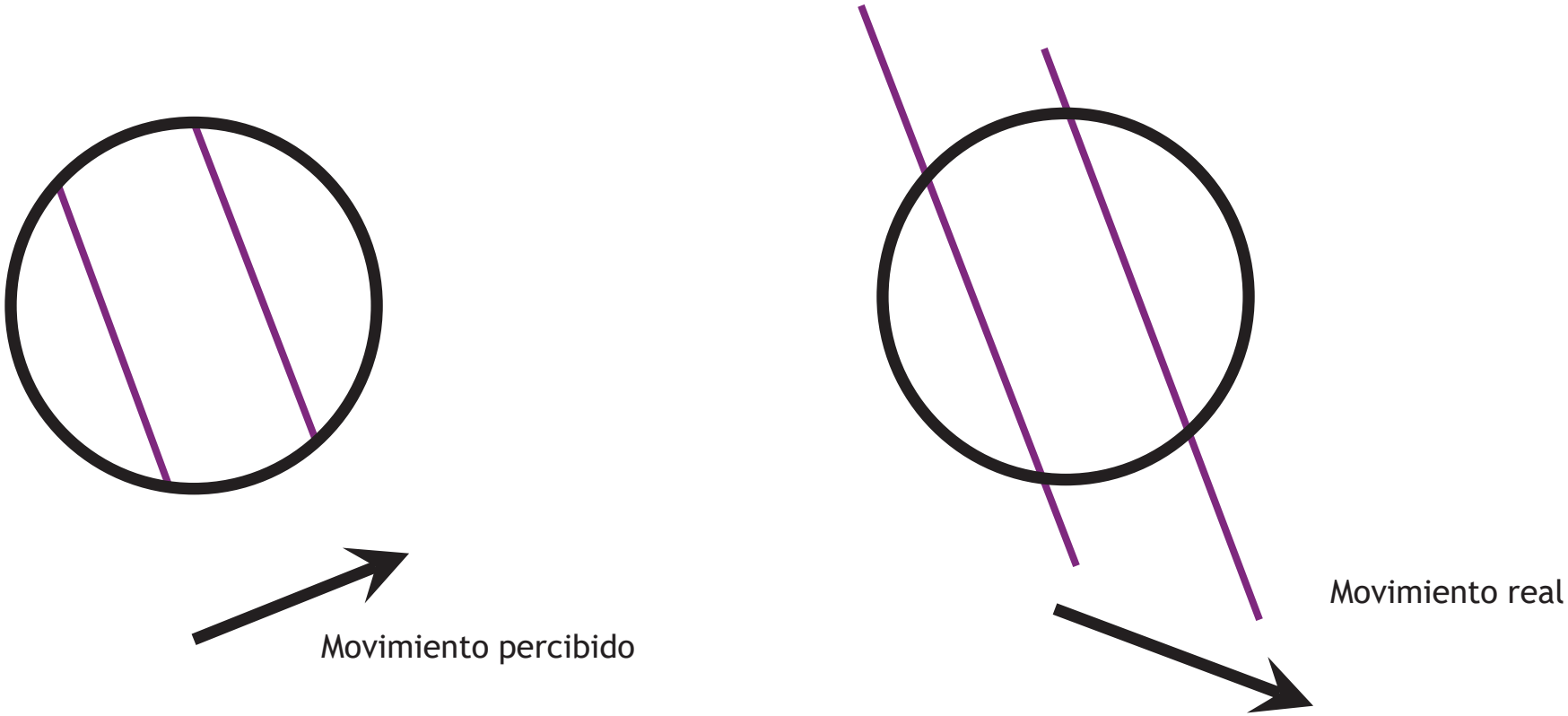
Denominado **problema de apertura**



Problema de apertura

Mala textura afecta a que el sistema sea resoluble

Denominado **problema de apertura**



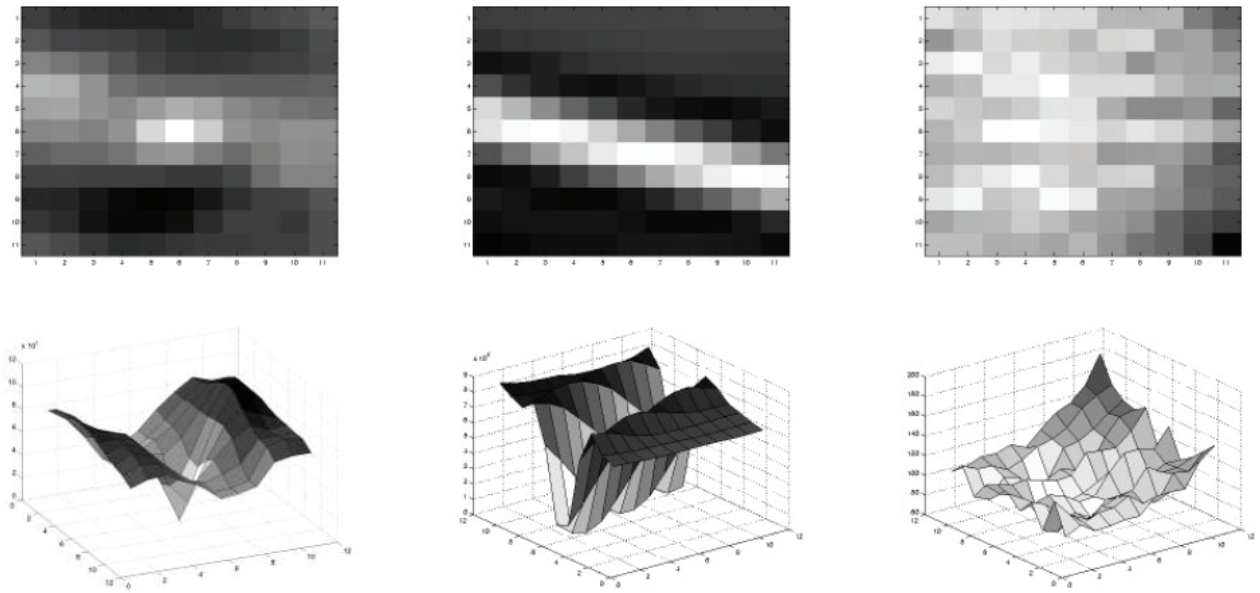
Problema de apertura

Ilusión palo de barbero

Problema de apertura

Problema de apertura

Recordatorio T4: SSD y buenos puntos de interés



(b) (c) (d)

Szeliski

Flujo óptico

Lucas-Kanade

Una ecuación por píxel, con dos incógnitas

Coherencia espacial: idéntico desplazamiento (u,v)

Más ecuaciones integrando vecindad (5x5 -> 25 ec.)

Evita problema de apertura usando esquinas

Para *n* vecinos, resolución de mínimos cuadrados

$$\begin{bmatrix} f_x(x_1, y_1) & f_y(x_1, y_1) \\ \vdots & \vdots \\ f_x(x_n, y_n) & f_y(x_n, y_n) \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{bmatrix} f_t(x_1, y_1) \\ \vdots \\ f_t(x_n, y_n) \end{bmatrix}$$

$$Ad=b$$

B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. IJCAI, 1981

Flujo óptico

Lucas-Kanade

Una ecuación por píxel, con dos incógnitas

Coherencia espacial: idéntico desplazamiento (u,v)

Más ecuaciones integrando vecindad (5x5 -> 25 ec.)

Evita problema de apertura usando esquinas

Para *n* vecinos, resolución de mínimos cuadrados

$$\begin{bmatrix} f_x(x_1, y_1) & f_y(x_1, y_1) \\ \vdots & \vdots \\ f_x(x_n, y_n) & f_y(x_n, y_n) \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{bmatrix} f_t(x_1, y_1) \\ \vdots \\ f_t(x_n, y_n) \end{bmatrix}$$

$$Ad=b$$

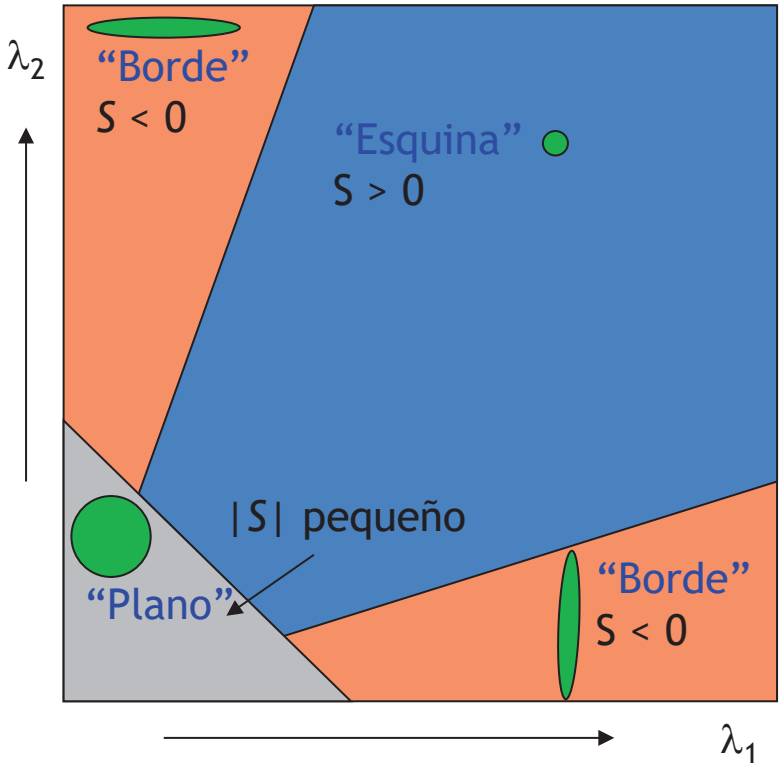
$$A^T Ad=A^t b$$

¿Mínimos cuadrados resoluble?

Flujo óptico

Lucas-Kanade

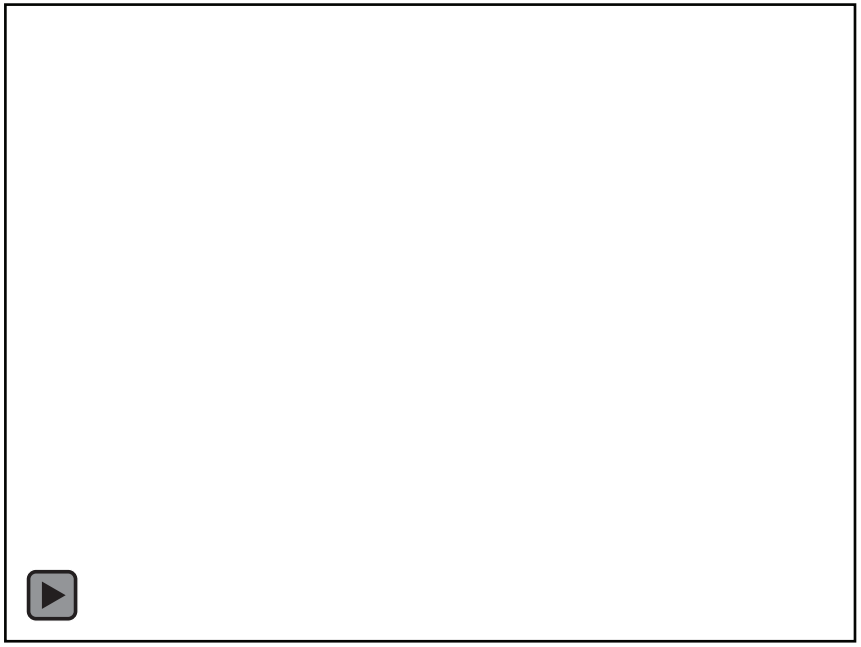
Problema bien condicionado con esquinas



Adaptado de Szeliski

Flujo óptico

Lucas-Kanade



S. Seitz, R. Szeliski, M. Pollefeys

Flujo óptico

LK tiene problemas si:

- Movimientos amplios
- Puntos que no se mueven como vecinos
- Sin constancia de brillo

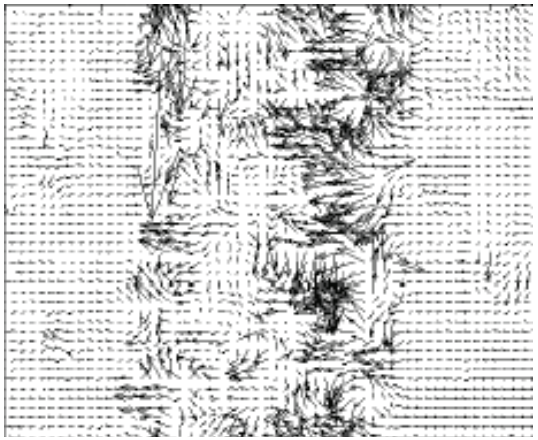
Refinamiento iterativo, multiresolución

Szeliski

Flujo óptico

Refinamiento iterativo:

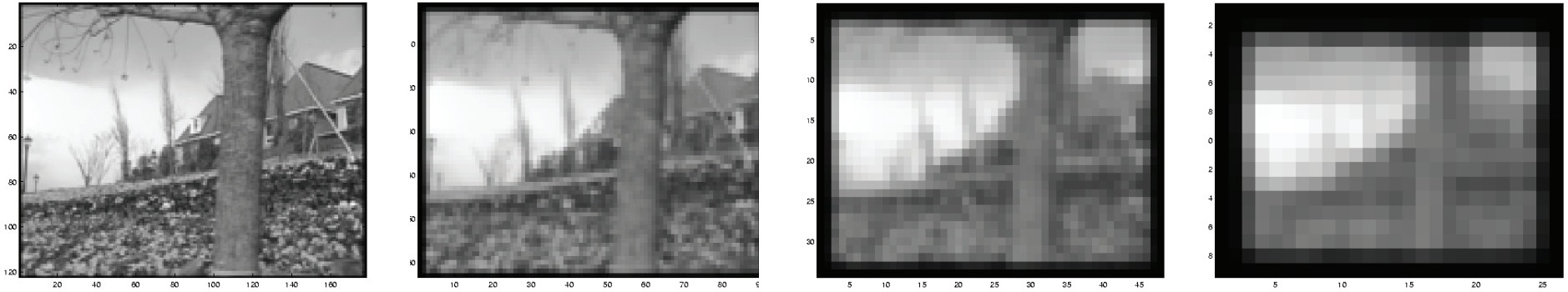
- Estima velocidad del píxel con una iteración LK
- Transforma una imagen en la otra con la estimación
- Refina repitiendo el proceso

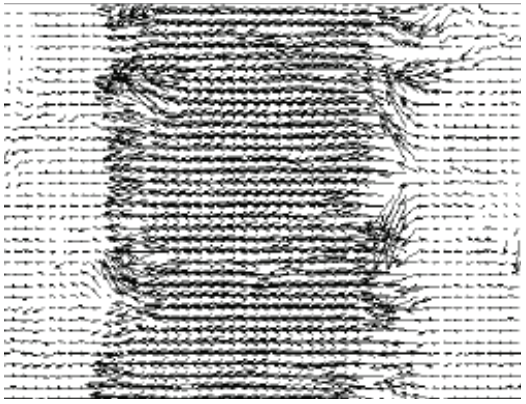


Szeliski

Flujo óptico

Estimación multiresolución y refinamiento





Szeliski

Flujo óptico

Segmentación de movimiento

Superar movimiento dispar entre vecinos

Movimiento por capas, útil para seguimiento

(a) (b) (c)

Figure 11: (a) The optic flow from multi-scale gradient method. (b) Segmentation obtained by clustering optic flow into affine motion regions. (c) Segmentation from consistency checking by image warping. Representing moving images with layers.

Szeliski

J. Wang and E. Adelson, Representing Moving Images with Layers, IEEE Transactions on Image Processing, 1994

Flujo óptico

Si no hay constancia de brillo

Encaje puntos de interés (p.e. SIFT)

Seguimiento (*tracking*) -> siguiente apartado

Figure 1: Three frame details from Woody Allen's *Manhattan*. The details are from the 1st, 11th, and 21st frames of a subsequence from the movie.

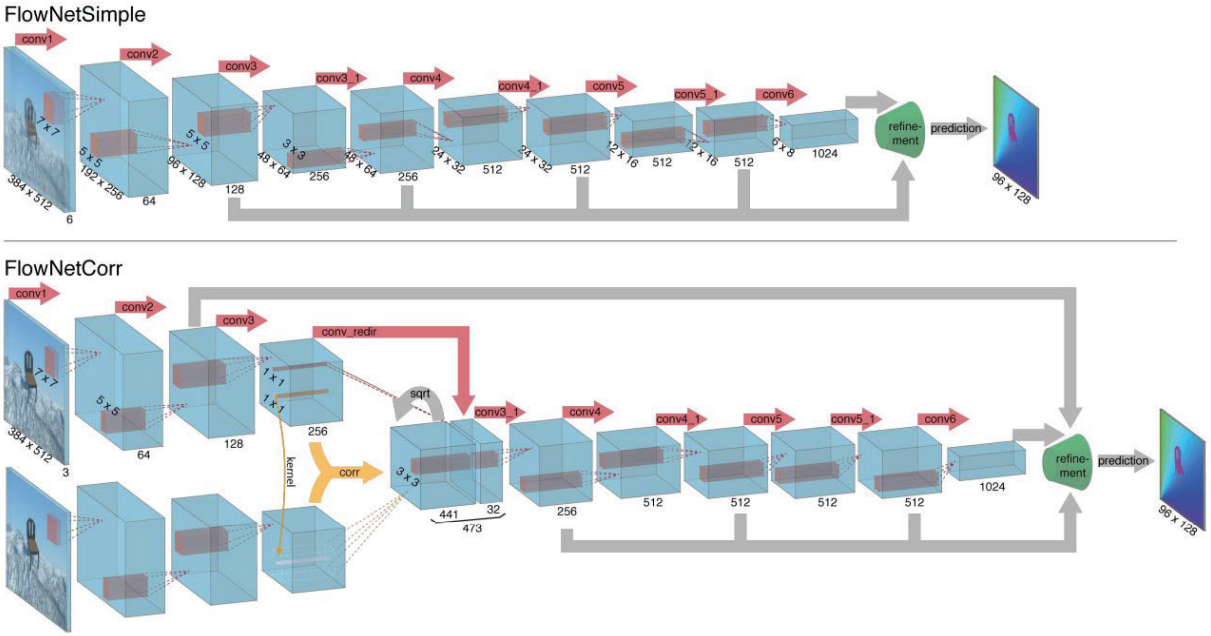
Figure 2: The traffic sign windows from frames 1,6,11,16,21 as tracked (top), and warped by the computed deformation matrices (bottom).

J. Shi and C. Tomasi. Good Features to Track. CVPR 1994

Aprendizaje

FlowNet
Datos (incluye sintéticos)

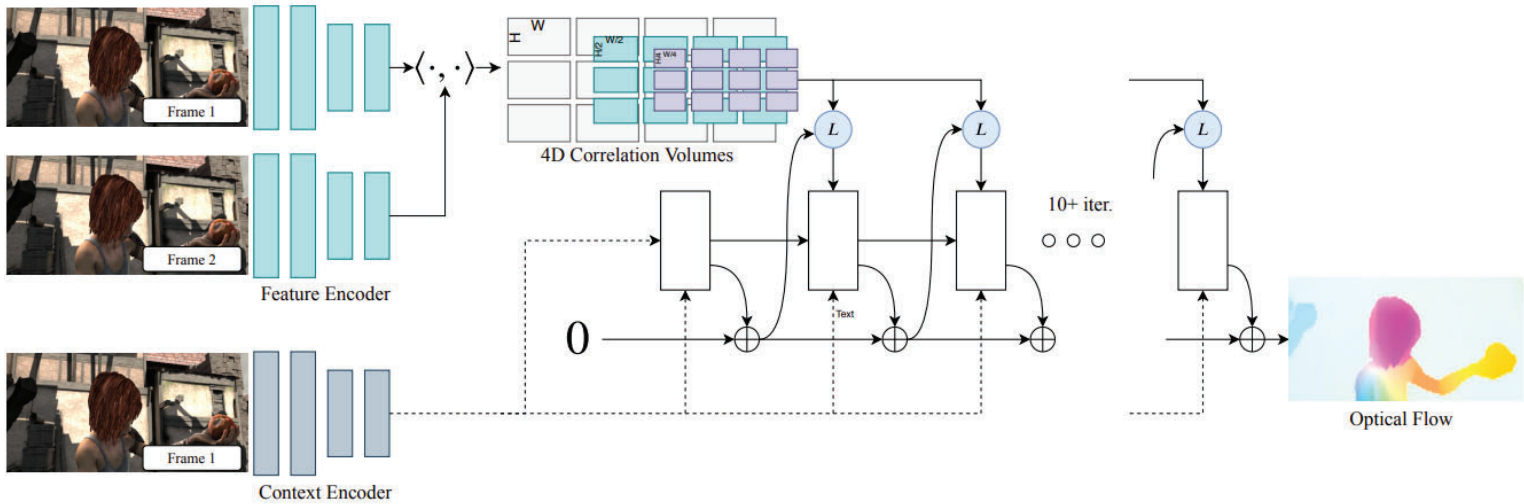
Flying Chairs

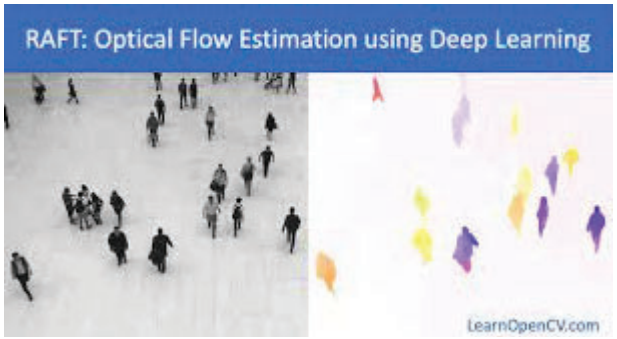


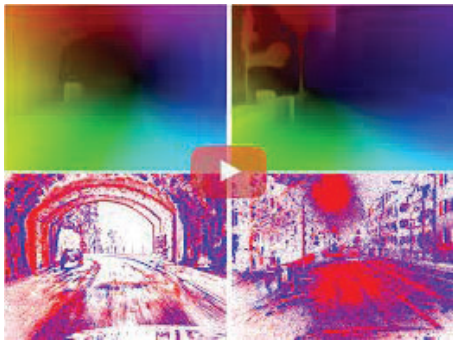
A. Dosovitskiy et al., FlowNet: Learning Optical Flow with Convolutional Networks, ICCV 2015

Aprendizaje

RAFT y E-RAFT



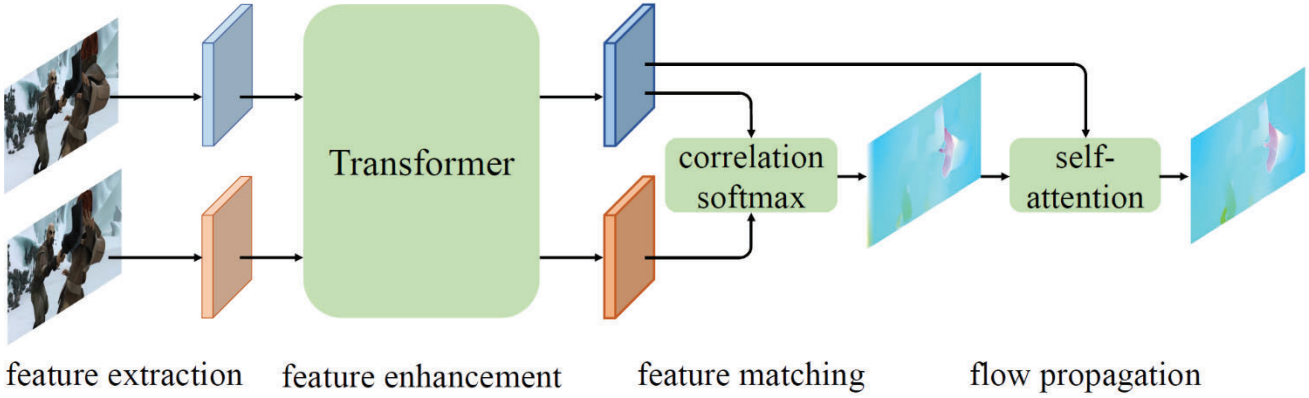




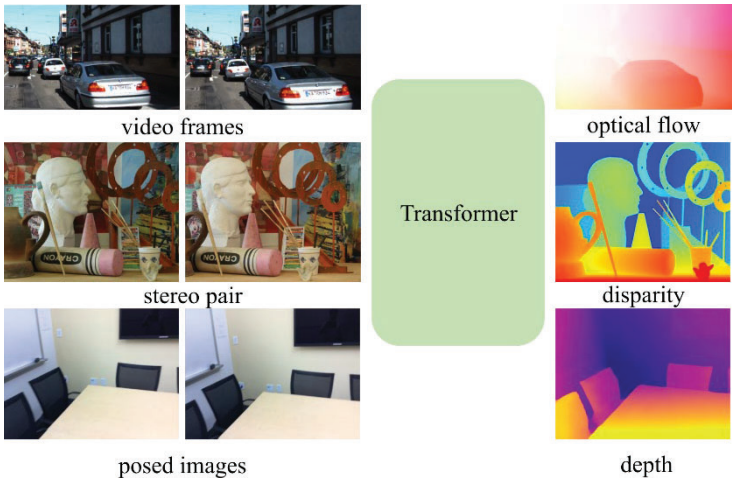
Z. Teed and J. Deng. [RAFT: Recurrent All-Pairs Field Transforms for Optical Flow](). ECCV 2020 (Best Paper Award)

Aprendizaje

GMFlow



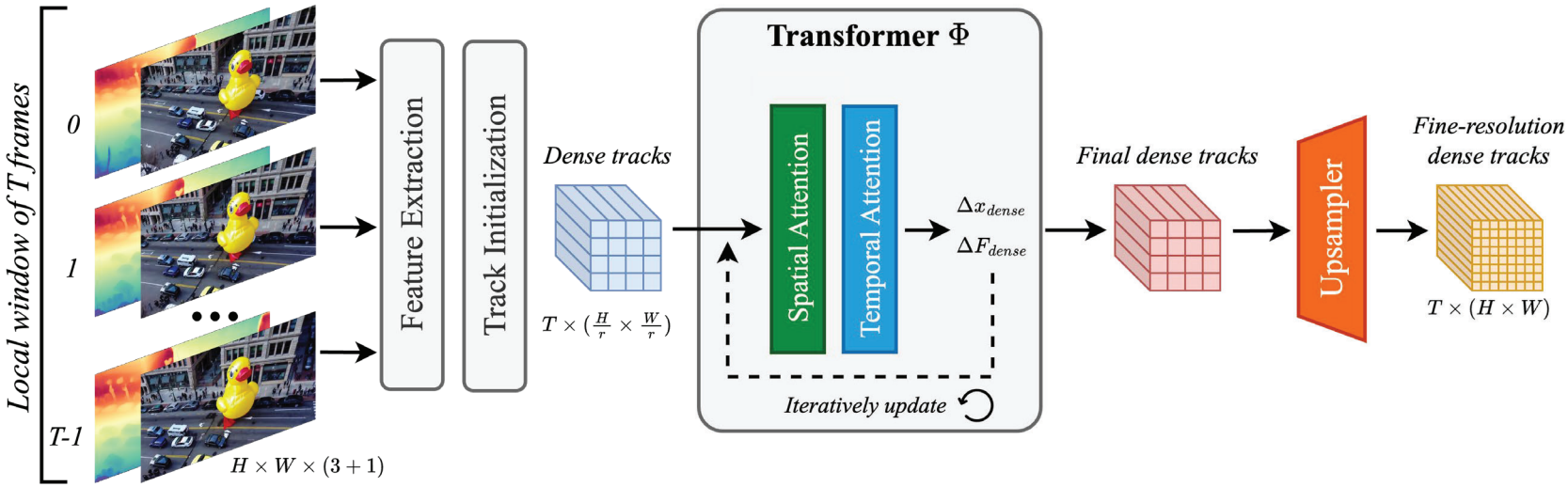
Unimatch



GMFlow: Learning Optical Flow via Global Matching, CVPR 2022
Unifying flow, stereo nd depth estimation, PAMI 2023

Aprendizaje

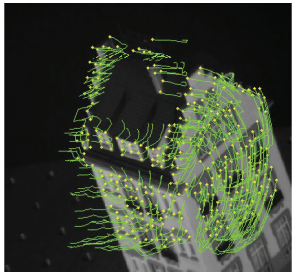
DELTA



Flujo óptico

Aplicaciones:

- Análisis de vídeo
  - Estabilización
  - Detección cambio de plano
  - Amplificación del movimiento
- Seguimiento y segmentación de objetos
- Estructura desde movimiento
- Encaje estéreo
- Reconocimiento de eventos y actividad
- Aprendizaje autosupervisado y predictivo



Reconocimiento de acciones

Flujo óptico como entrada para reconocimiento de acciones CNNs

Joao Carreira and Andrew Zisserman. "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset

Detección de planos

Magnificando

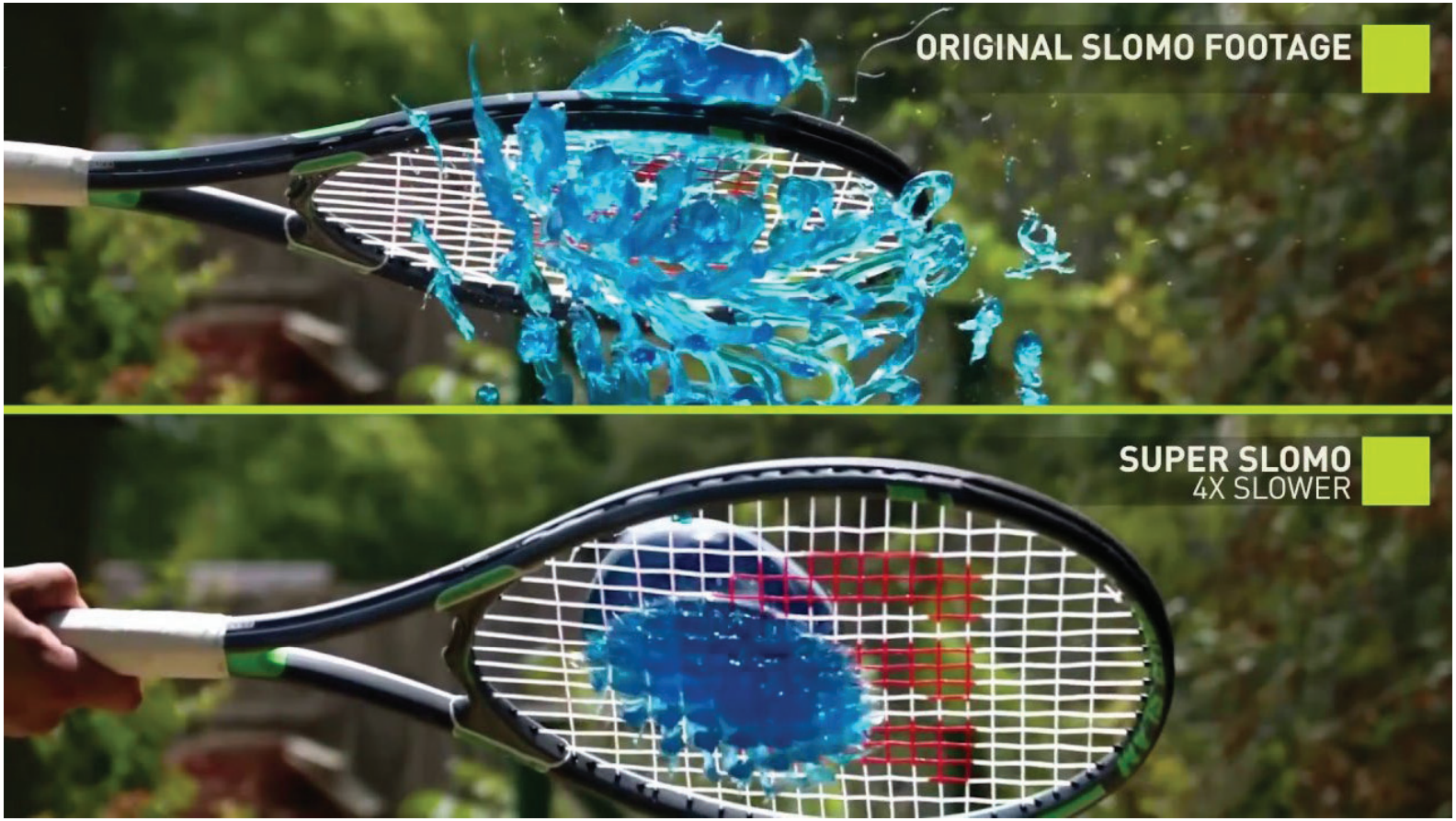
Obtener el flujo y magnificar

(a) Input

(b) Magnified

H. Wu et al. Eulerian Video Magnification for Revealing Subtle Changes in the World. SIGGRAPH 2012

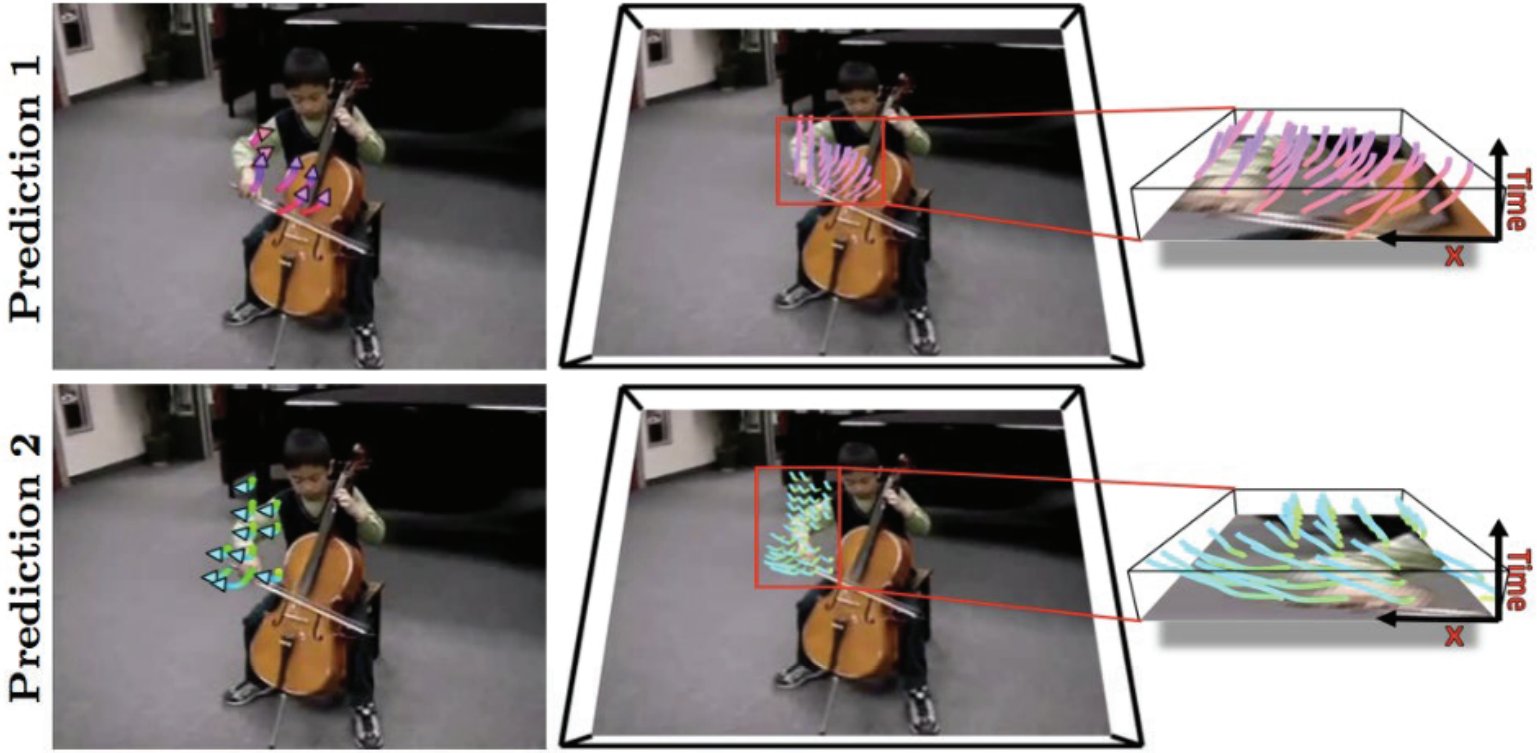
Interpolación de fotogramas



Super SloMo: High Quality Estimation of Multiple Intermediate Frames for Video Interpolation, 2018

Imagen estática

Flujo óptico desde una imagen



Walker et al. Dense Optical Flow Prediction from a Static Image, ICCV 2015

Seguimiento

Detectar uno o varios objetos (*targets*) en vídeo

Modelar cambio
Asociación de objetos en diversos fotogramas (identidades)
Útil asignar identidades y observar comportamiento temporal

Seguimiento

Detección inicial del objeto automática o manual

Seguimiento

Detección inicial del objeto automática o manual

Seguimiento sin detección *(detection-free tracking*)

- No requiere un detector
- Inicialización manual de objetivos
- Válido para objetos sin detector

Seguimiento con detección (*detection based tracking*)

- Seguimiento para fallos del detector
- Gestiona nuevos objetos
- Gestiona aquellos que desaparecen
- Detectar es costoso, seguir menos costoso

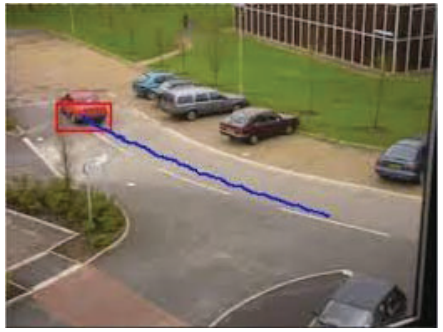
Tracking by detection

- En ocasiones sin seguimiento, sólo cajas detectadas

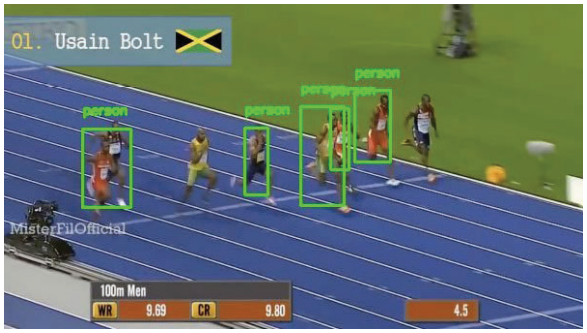
Seguimiento

Tipos:

- Un objeto (*Single Object Tracking* o *SOT*):
  Habitualmente sin detector, detección manual inicial



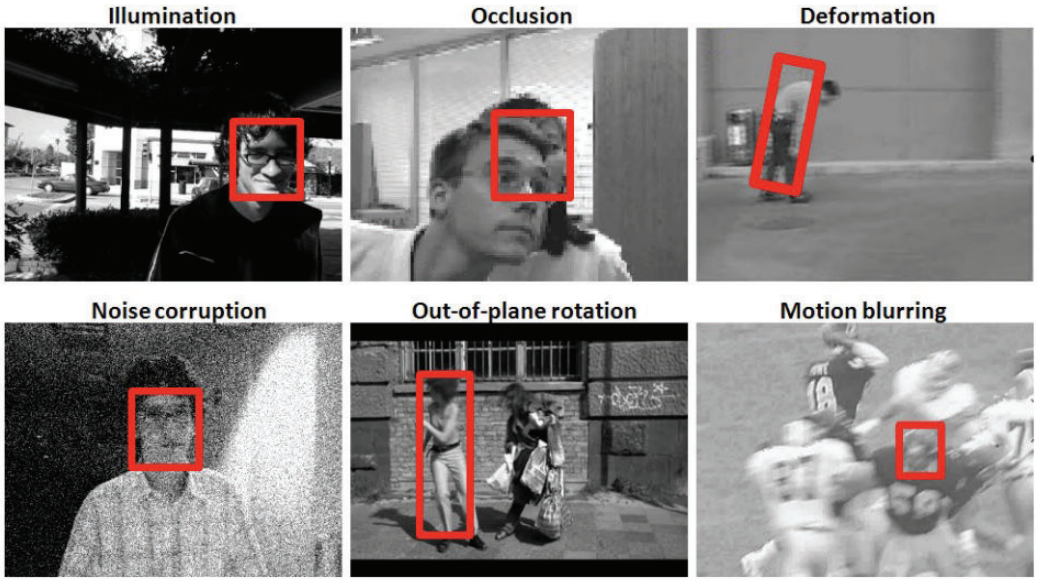
- Múltiples objetos (*Multiple Object Tracking* o *MOT*):
  Detecta nuevos objetos, y mantiene presentes



Seguimiento

Dificultades:

- Correspondencia entre fotogramas
- Oclusiones y cruces
- Deformaciones: escala, rotaciones
- Desenfoque por movimiento
- Iluminación
- Fondos complejos *(Background clutter)*
- Resolución



Seguimiento

Detectar objeto(s) en una imagen
Buscar en sucesivas (vídeo)
Representación del objeto →

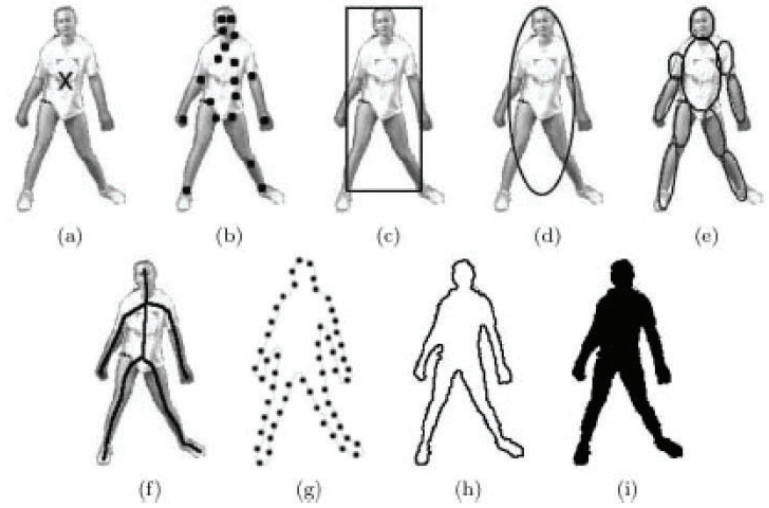


Fig. 1. Object representations. (a) Centroid, (b) multiple points, (c) rectangular patch, (d) elliptical patch, (e) part-based multiple patches, (f) object skeleton, (g) control points on object contour, (h) complete object contour, (i) object silhouette.

Jonathan Gagné

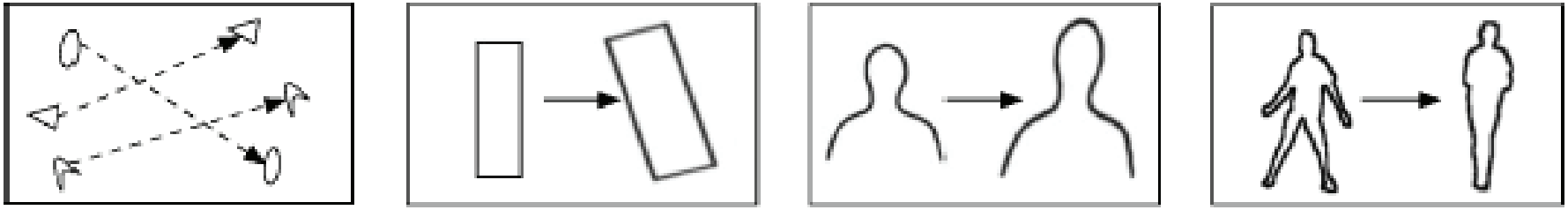
Seguimiento en base a:

- Puntos de interés
- Patrón
- Color
- Silueta
- Flujo óptico
- …

Seguimiento

Puntos, patrón o ventana, contornos, …

Encaje minimizar SSD, mínimos cuadrados, similitud histograma …



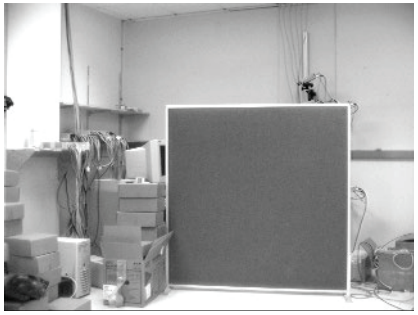
(a) (b) (c) (d)

Jonathan Gagné

Detección del objetivo

Diferencia de fotogramas

Cámara estática



1 fotograma

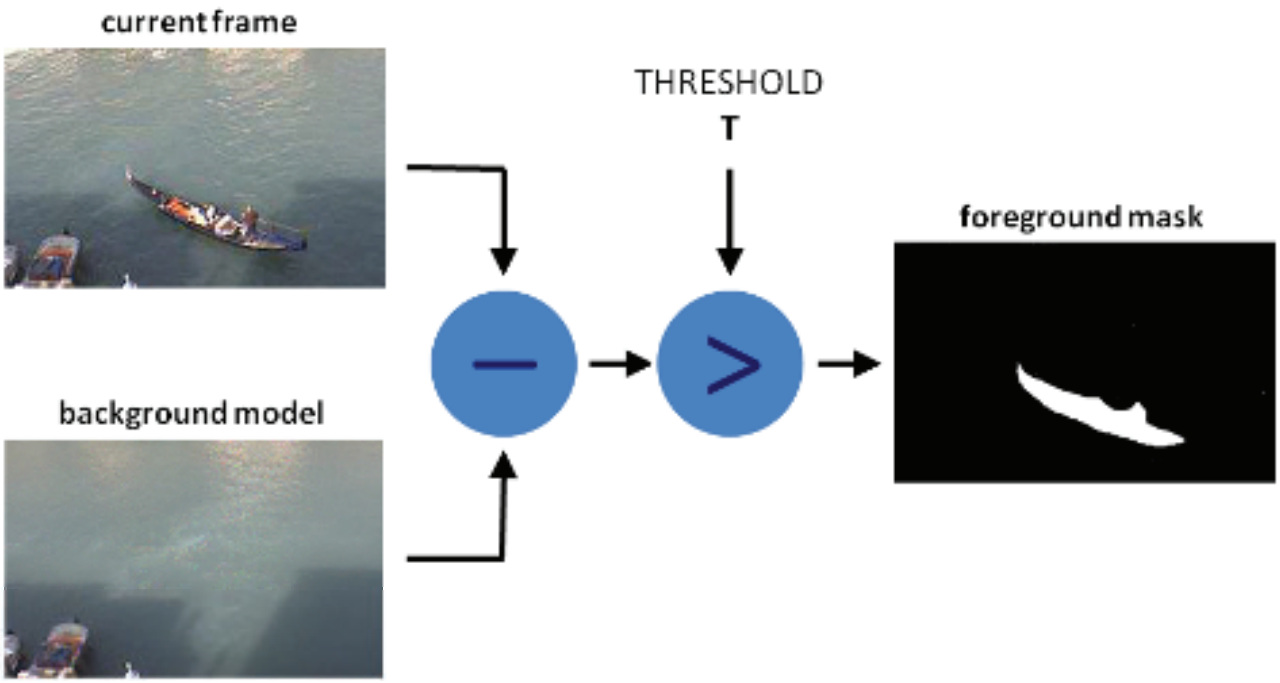
5 fotogramas

Adaptado de J. Kosecka

Detección del objetivo

Sustracción de fondo

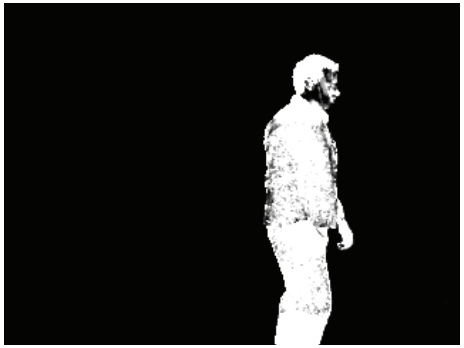
Cámara estática



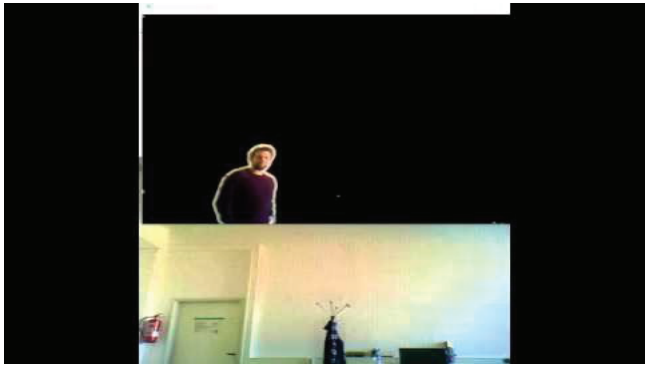
Detección del objetivo

Sustracción de fondo

Cámara estática



J. Kosecka



Detección del objetivo

Sustracción de fondo

Modelo del color del fondo (grises o RGB)

Calcula media (μ) y varianza (σ) en cada fotograma (z medida actual)

$$\mu_{t+1} = \alpha\mu_t + (1-\alpha)z_{t+1}$$
$$\sigma_{t+1}^2 = \alpha(\sigma_t^2 + (\mu_{t+1} - \mu_t)^2) + (1-\alpha)(z_{t+1} - \mu_{t+1})^2$$

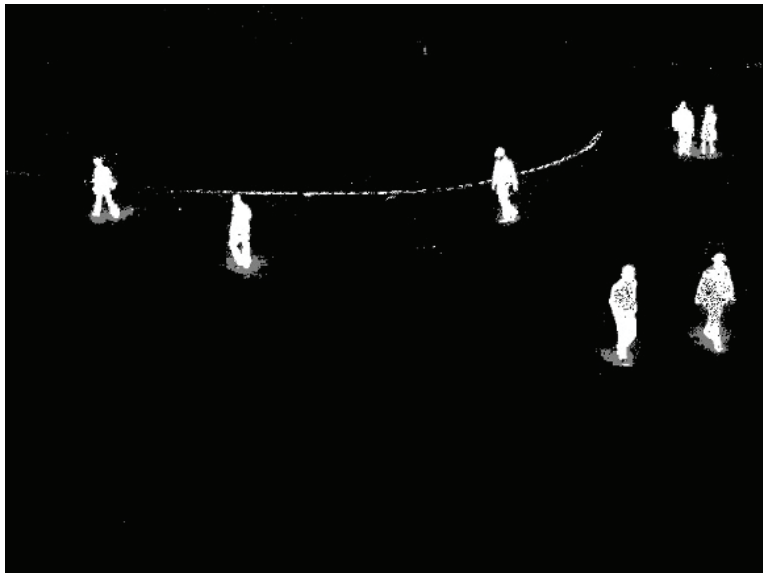
Marcar si para alguna componente

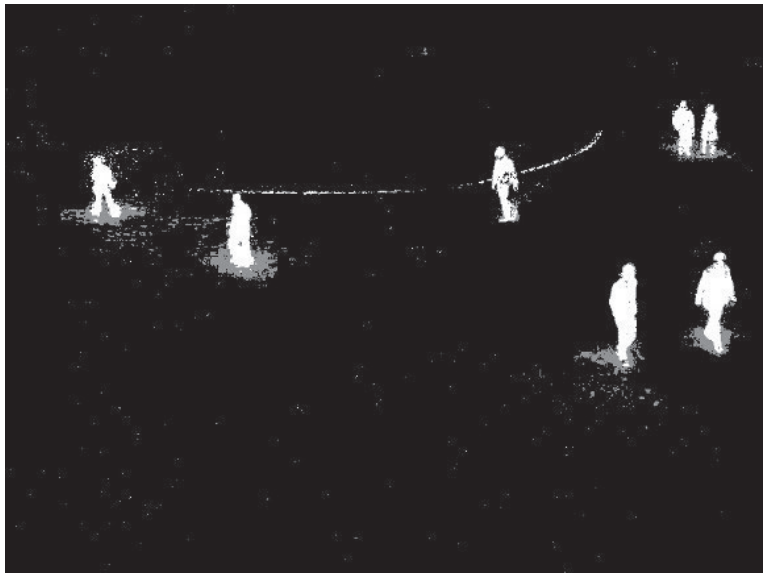
$$|r - \mu_t| > 3\max(\sigma_r, \sigma_{rcam})$$

Detección del objetivo

Sustracción de fondo

OpenCV KNN y MOG

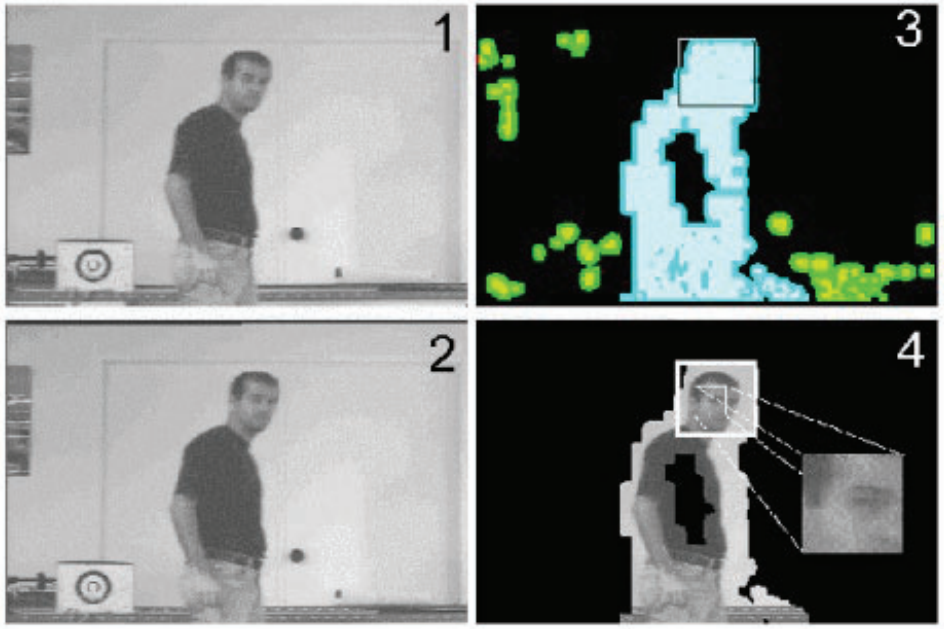




Detección del objetivo

Flujo óptico

DESEO: Detección y Seguimiento de Objetos



F. Hernández at al. DESEO: An Active Vision System for Detection, Tracking and Recognition, ICVS 1999

Detección del objetivo

Flujo óptico

OSVOS, FusionSeg, MaskTrack o SegFlow

K.K. Maninis et al. Video Object Segmentation Without Temporal Information, *(T-PAMI)*, 2018

Descripción y encaje

Modelo del objetivo (*target*):

- Zona o *patch:* Seguimiento con SSD o correlación

$$SSD = \sum_{i=1}^{n}\sum_{j=1}^{n}[M(i,j) - N(i,j)]^2$$

$$C = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} M(i,j)N(i,j)}{[\sum_{i=1}^{n}\sum_{j=1}^{n} M(i,j)^2 \sum_{i=1}^{n}\sum_{j=1}^{n} N(i,j)^2]^{1/2}}$$

Descripción y encaje

Modelo del objetivo (*target*):

- Zona o *patch:* Seguimiento con SSD o correlación

$$SSD = \sum_{i=1}^{n}\sum_{j=1}^{n}[M(i,j) - N(i,j)]^2$$

$$C = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} M(i,j)N(i,j)}{[\sum_{i=1}^{n}\sum_{j=1}^{n} M(i,j)^2 \sum_{i=1}^{n}\sum_{j=1}^{n} N(i,j)^2]^{1/2}}$$

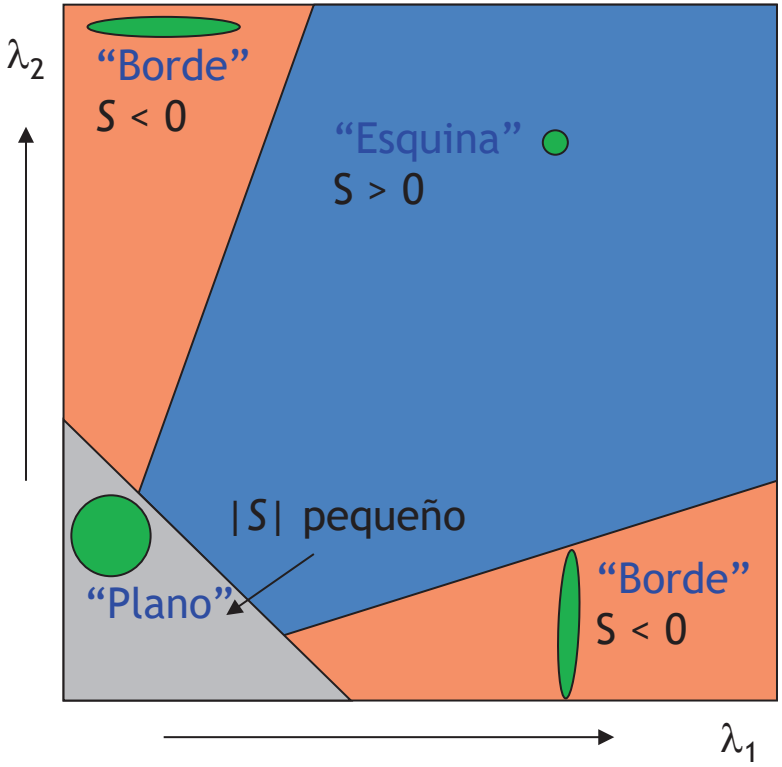
- Color: Histogramas, intersección o color

$$\chi^2(h_c, h_b) = \sum_i 2\frac{(h_c(i) - h_b(i))^2}{h_c(i) + h_b(i)} \qquad I(h_c, h_b) = \frac{\sum_i \min\{h_c(i), h_b(i)\}}{\sum_i \max\{h_c(i), h_b(i)\}}$$

Descripción

Seguimiento con puntos de interés

LK

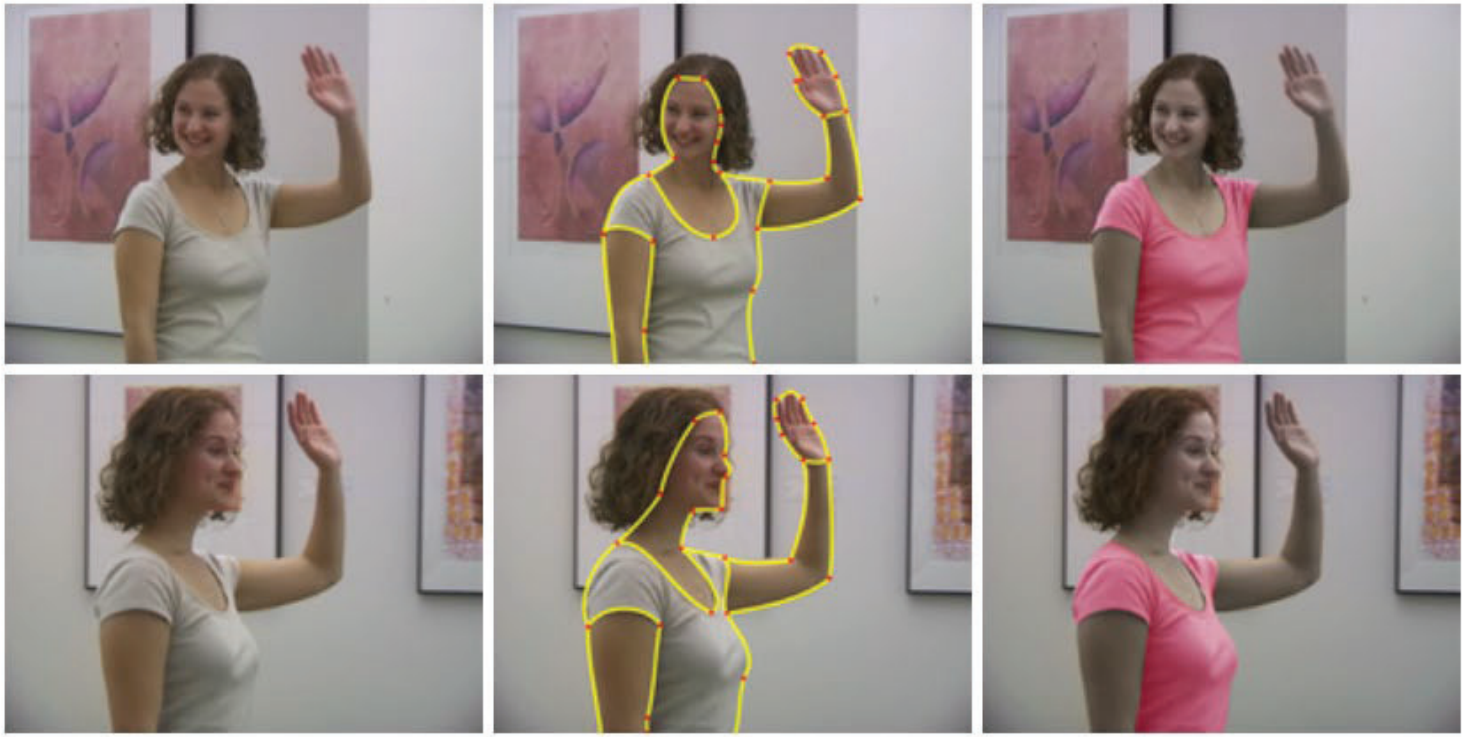


Adaptado de Szeliski

Descripción

Contornos

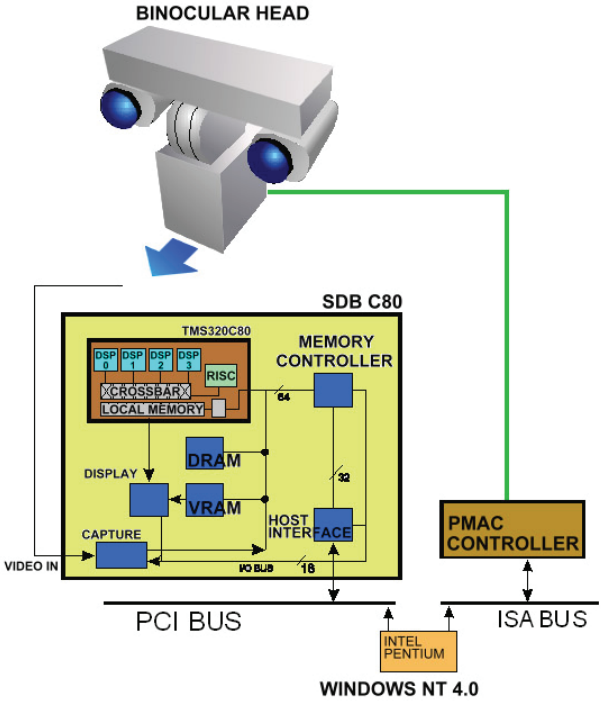
Contornos activos, *level sets*



A. Agarwala at al. **Keyframe-Based Tracking for Rotoscoping and Animation**

Detección del objetivo

[Batallita]
Patrón
DESEO: DEtección y SEguimiento de Objetos



F. Hernández at al. DESEO: An Active Vision System for Detection, Tracking and Recognition, ICVS 1999

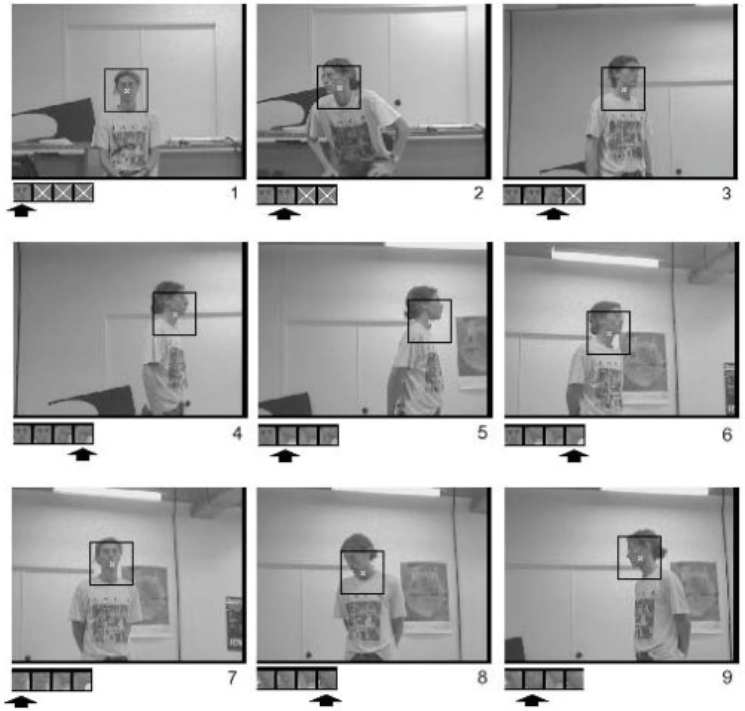
Detección del objetivo

[Batallita]
Patrón
DESEO: DEtección y SEguimiento de Objetos

Seguimiento basado en patrón (*template*)

Actualización

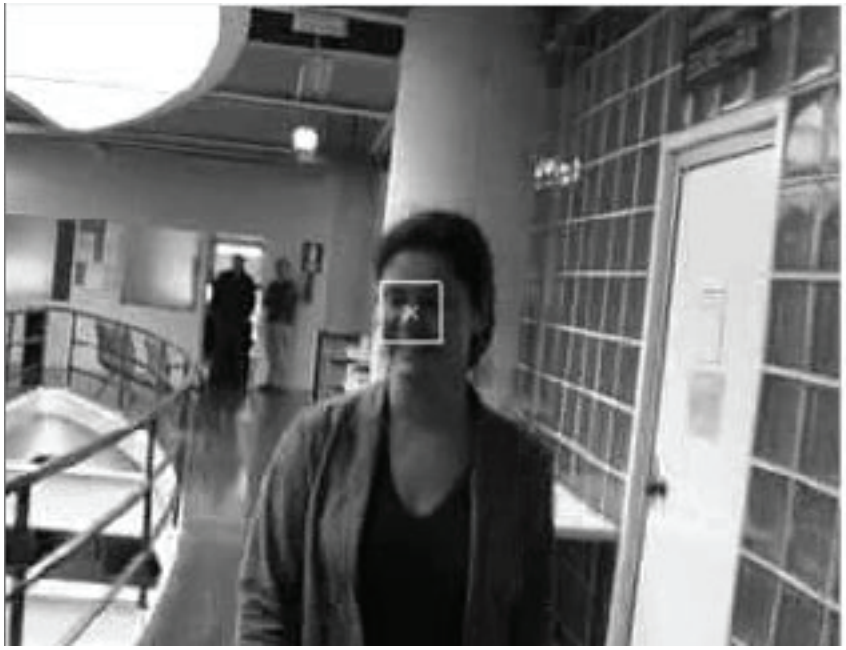
Múltiples

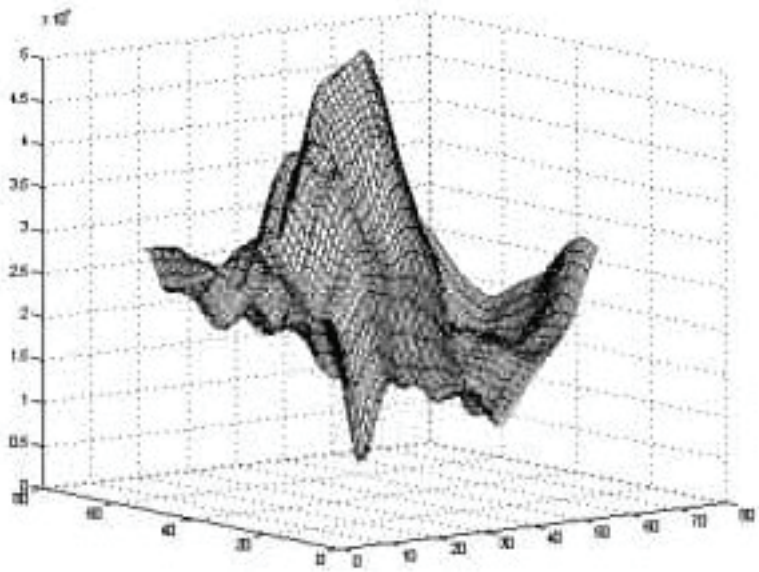


Detección del objetivo

[Batallita]
Patrón
DESEO: DEtección y SEguimiento de Objetos

Actualización basada en el patrón



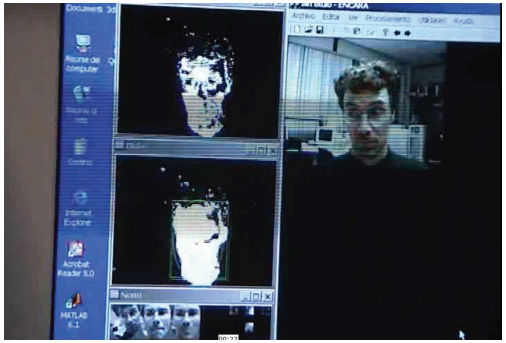


Detección del objetivo

[Batallita]
Patrón
DESEO: DEtección y SEguimiento de Objetos

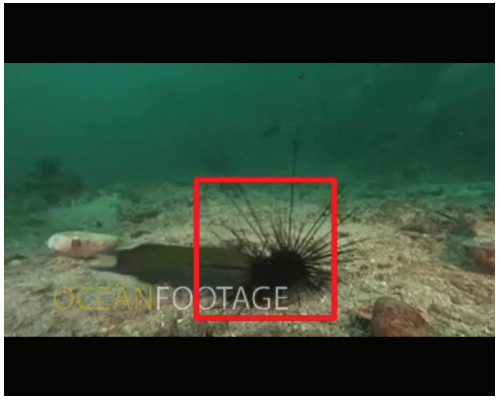
Detección del objetivo

[Batallita]
Caras



Profundidad

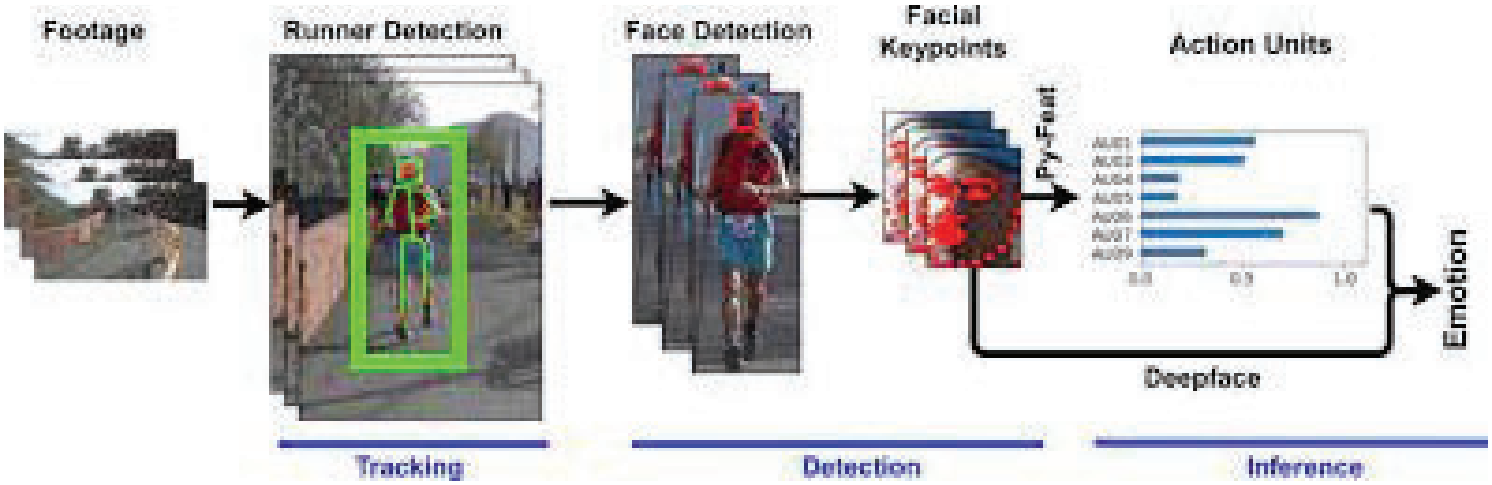
PANLAB



Tracking by detection

Detección objetivo

Usando algún detector, p.e. YOLO



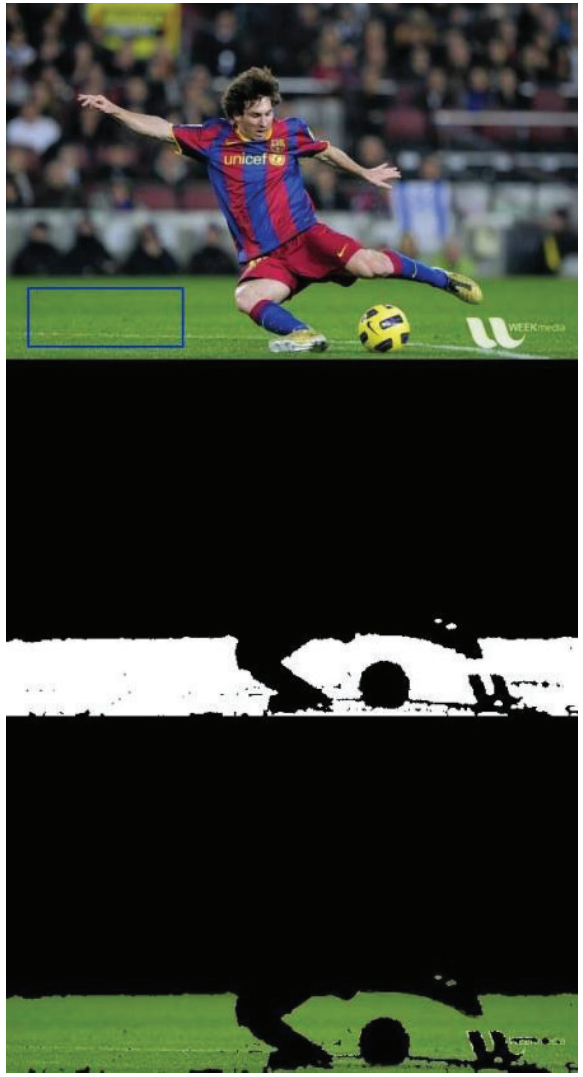
Encaje

Color: histogram backprojection

Histograma del color de objeto de interés

Probabilidad del píxel en histograma

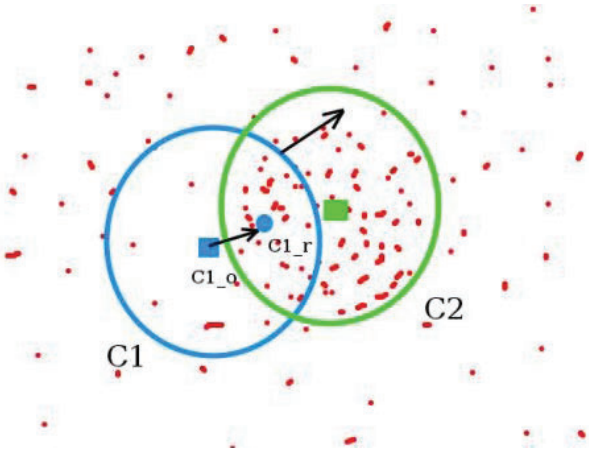
Integrado en mean-shift



Encaje

Mean shift

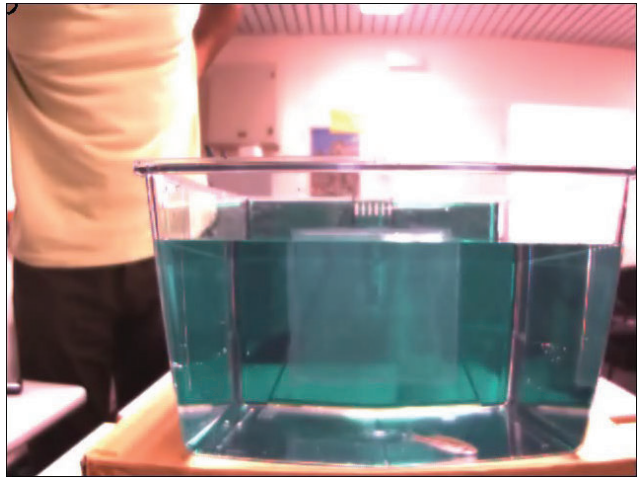
Traslada centroide solidario con densidad de puntos localizados

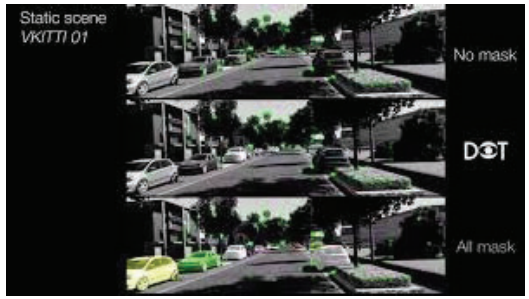


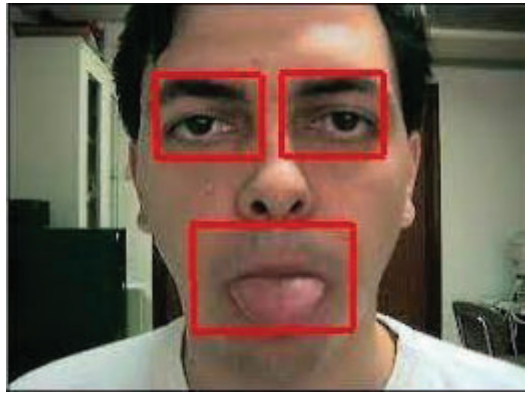
Seguimiento

Aplicaciones:

- Videovigilancia
- Monitorizacion animales
- Deportes
- Seguridad vehículos
- Visual SLAM







OpenCV (comparativa)

No DL:

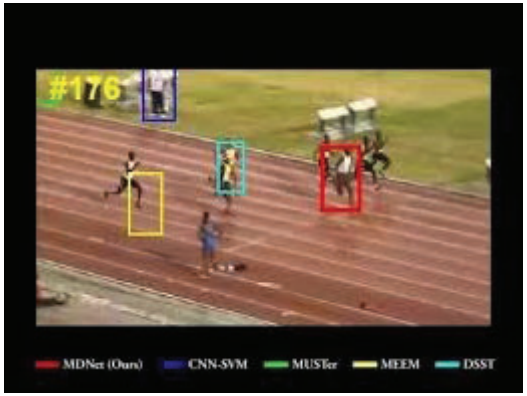
- Boosting
- MIL (Multiple instance learning)
- KCF (Kernelized Correlation Filters)
- TLD (Tracking Learning Detection)
- MedianFlow
- MOSSE (Minimum Output Sum of Squared Error)
- CSRT (Discriminative Correlation Filter with Channel and Spatial Reliability)

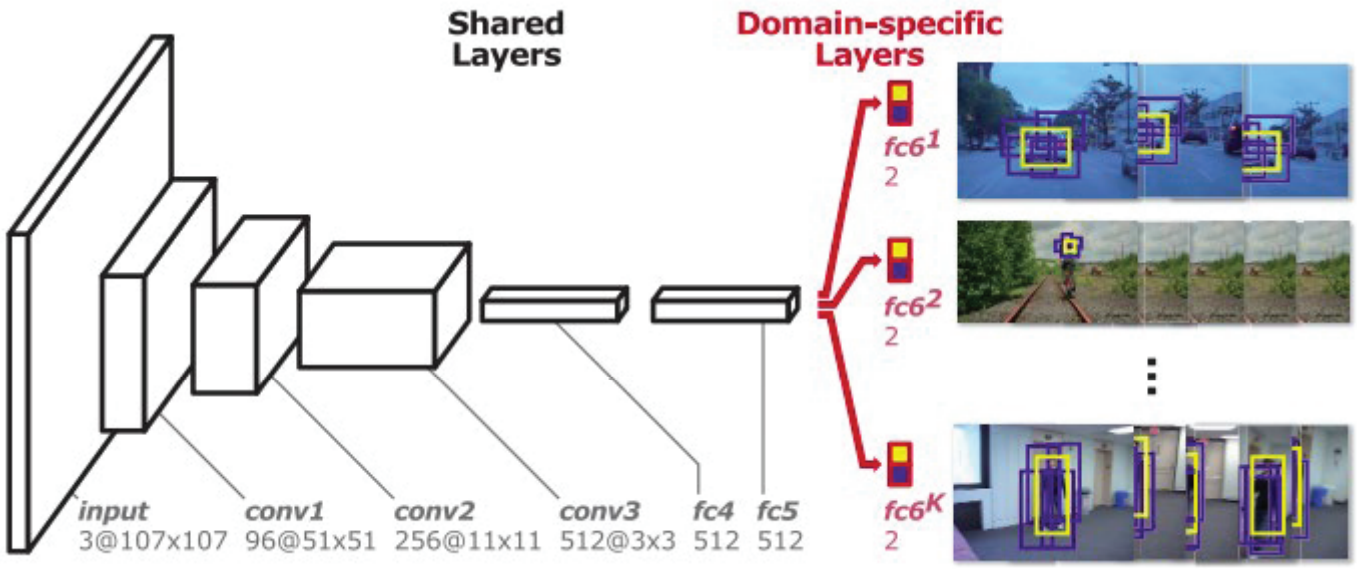
DL:

- GOTURN

MOT

MDNet
Ganador VOT2015
Persistencia
Redes poco profundas
Capas que aprenden en línea

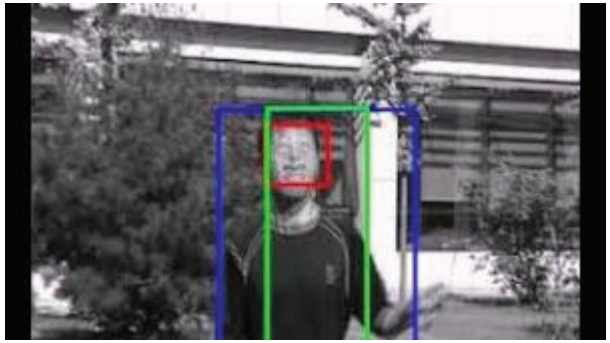


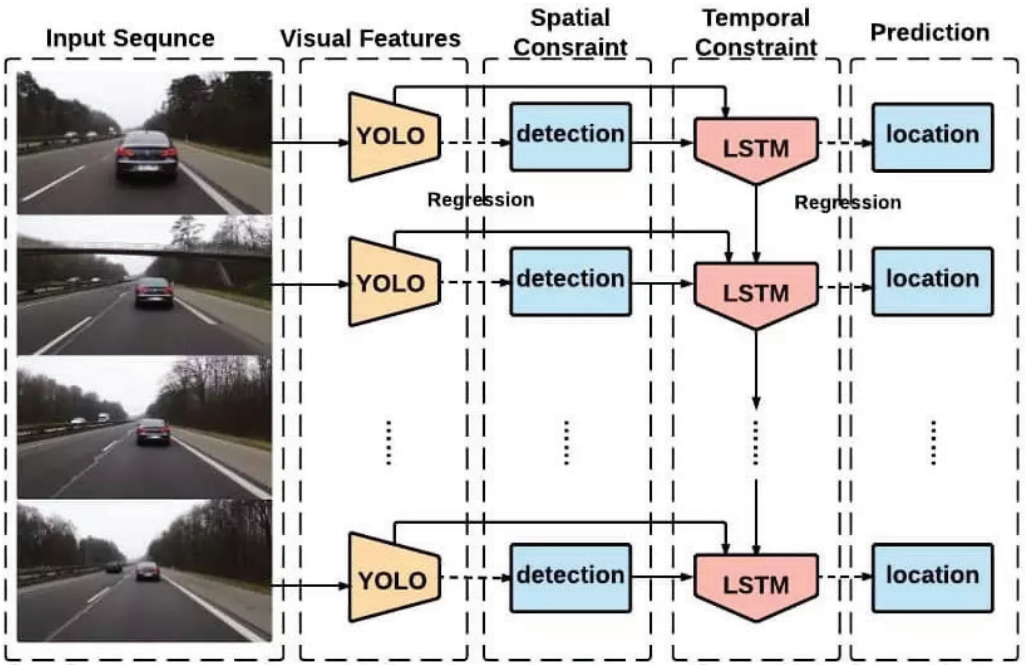


H. Nam y B. Han, Learning Multi-Domain Convolutional Neural Networks for Visual Tracking, 2016

MOT

ROLO (Recurrent YOLO)
Detección con YOLO
Long Short Term Memory (LSTM)





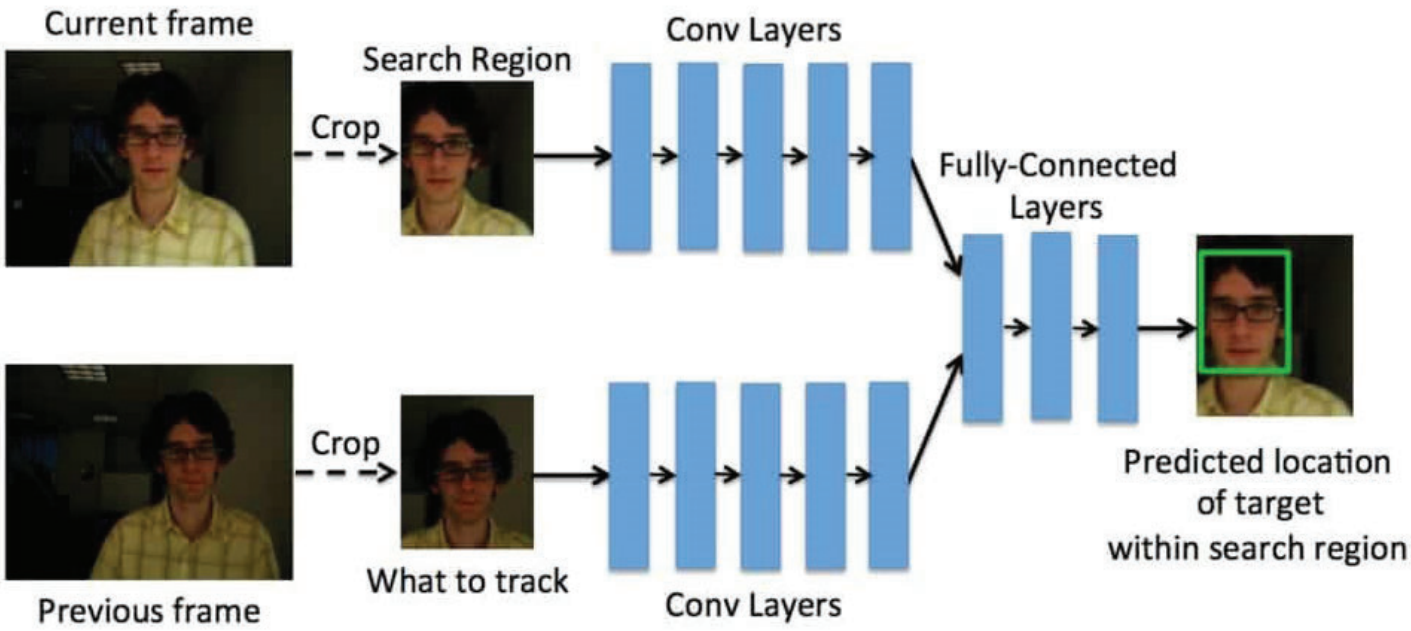
G. Ning et al., Spatially Supervised Recurrent Convolutional Neural Networks for Visual Object Tracking, 2016

MOT

GOTURN

DL y rápido

Sin adaptación en vivo



D. Held et al. Learning to Track at 100 FPS with Deep Regression Networks, 2018

MOT

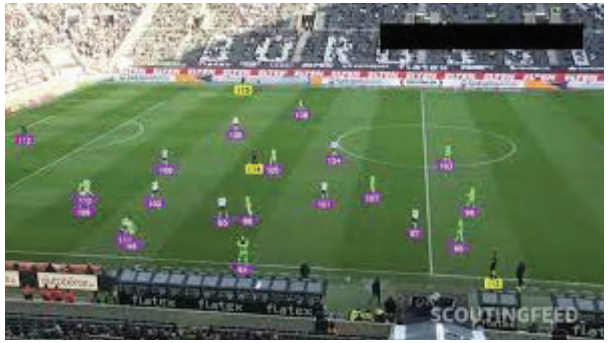
ByteTrack

Seguimiento, asocia contenedores entre fotogramas

Habitualmente con buena confianza

ByteTrack incluye de baja confianza

Filtro de Kalman



Y. Zhang. ByteTrack: Multi-Object Tracking by Associating Every Detection Box, 2022

MOT

ByteTrack
Muestra con YOLOv7

MOT

ByteTrack
Integrado por ultralytics desde YOLOv8

MOT

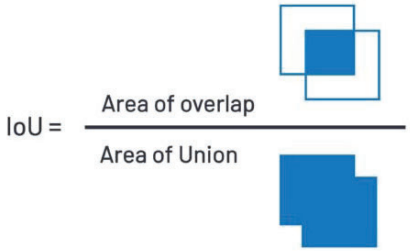
DeepSORT
Extiende SORT
No solo posición y velocidad
Integra apariencia obtenida con DL

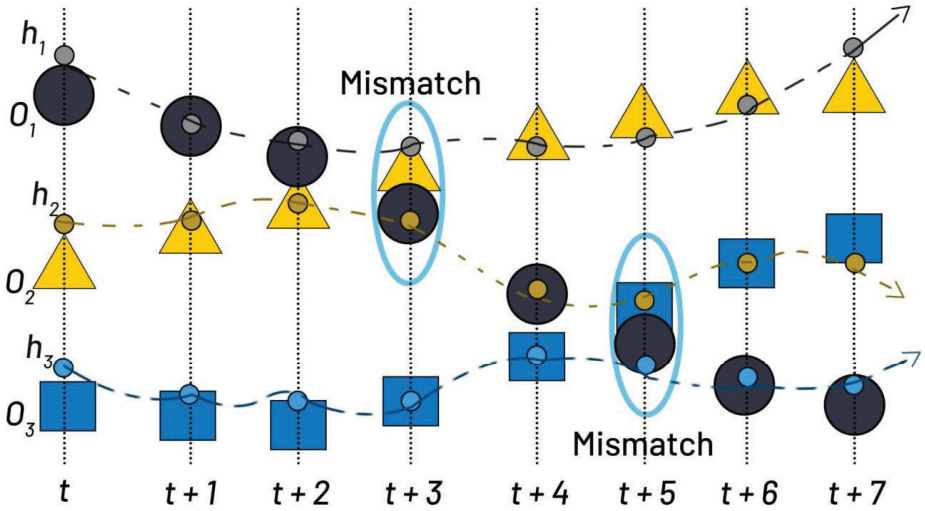
Simple Online and Realtime Tracking with a Deep Association Metric (DeepSORT), 2017

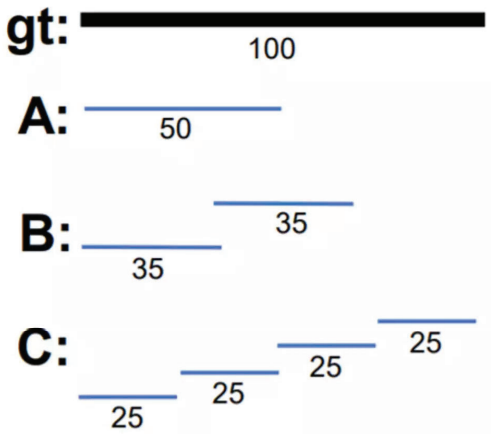
MOT

Métricas MOT

Encaje contenedores, recordar IoU







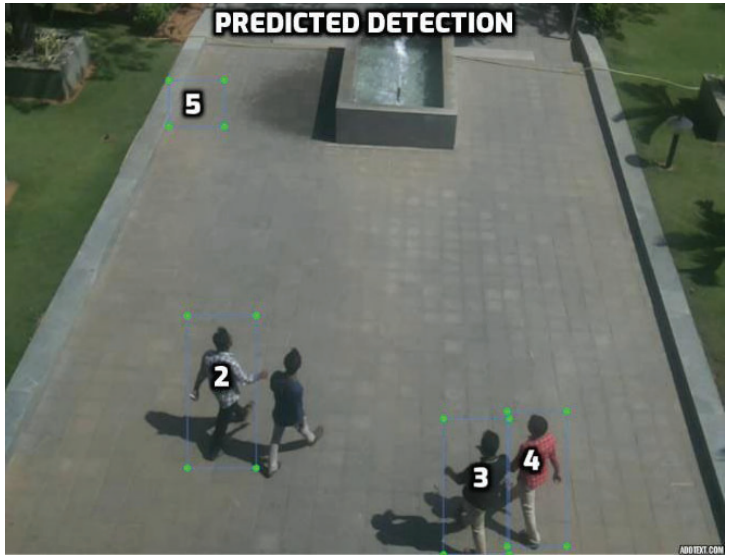
MOT

Métricas MOT

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDSW_t)}{\sum_t GT_t}$$

Se calcula para todos los fotogramas, IoU fijo

Ejemplo un fotograma $MOTA = 1 - \frac{(1+1+1)}{4} = 0.25$



MOT

Métricas MOT

$$IDF1 = \frac{2 * IDTP}{2 * IDTP + IDFP + IDFN}$$

Mejor predicción para maximizar la métrica

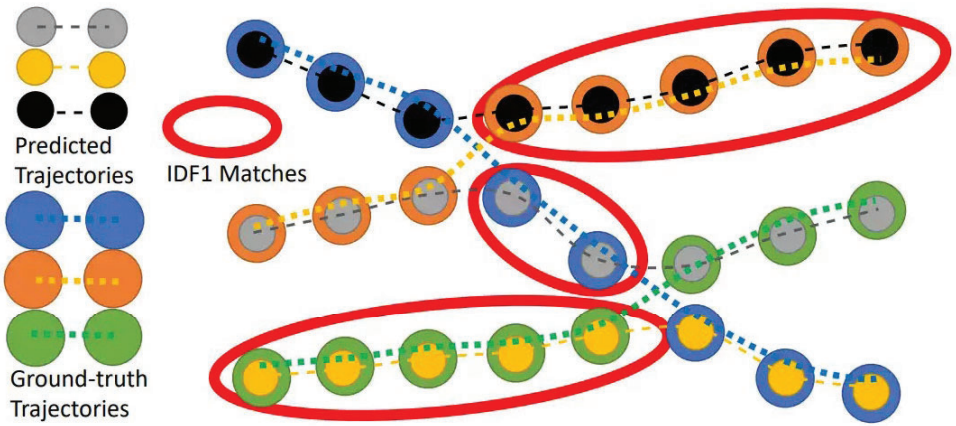


Fig. 10 A tracking example which shows how the single best trajectory matching, as performed by IDF1, can result in unintuitive matches between trajectories.

MOT

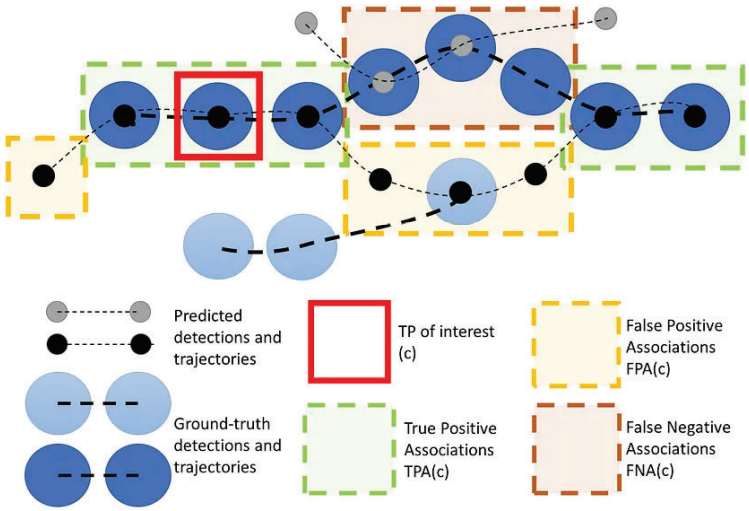
Métricas MOT

$$HOTA_{\propto} = \sqrt{DetA_{\propto} * AssA_{\propto}}$$

$$DetA = \frac{TP}{TP+FP+FN}$$

$$AssA = \frac{1}{TP}\sum_{c \in TP} Ass_{IoU(c)} \quad donde \quad Ass_{IoU} = \frac{TPA}{TPA + FPA + FNA}$$

AssA para la consistencia temporal
α se refiere a valor de IoU



MOT

Métricas MOT

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDSW_t)}{\sum_t GT_t}$$

$$IDF1 = \frac{2 * IDTP}{2 * IDTP + IDFP + IDFN}$$

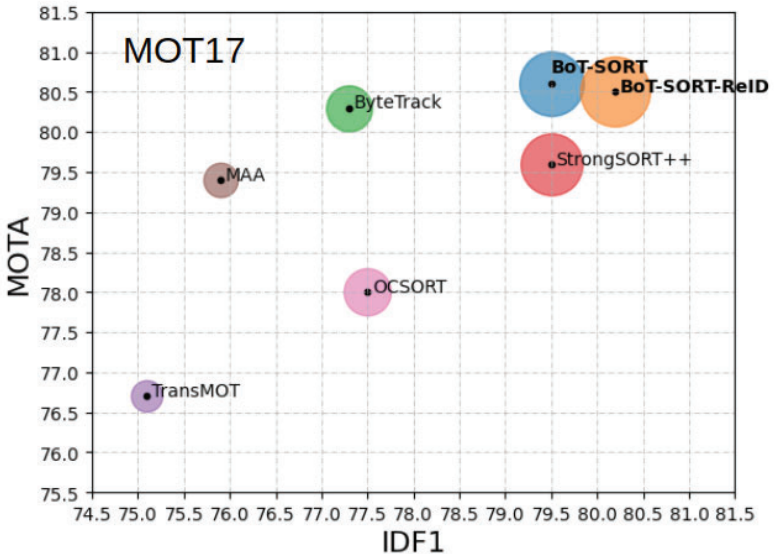
$$HOTA_\propto = \sqrt{DetA_\propto * AssA_\propto}$$

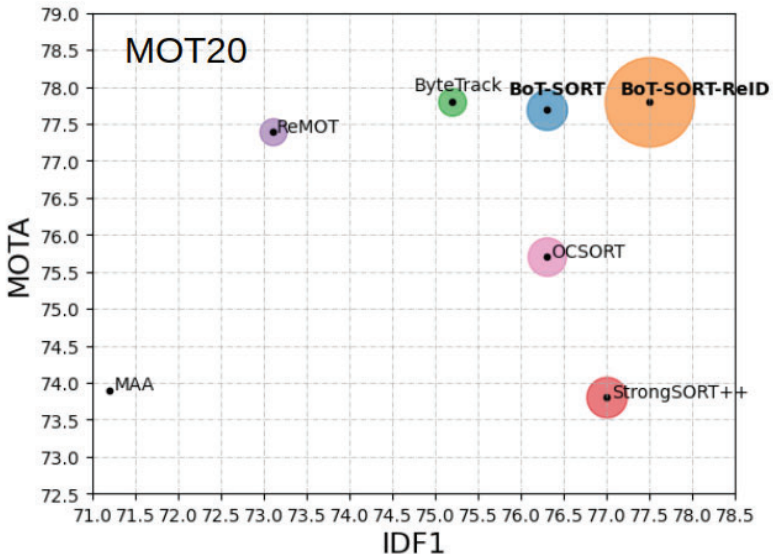
MOTA simple, demasiado en escenarios complejos
IDF1 centrada en consistencia no aprecia mejoras detección
HOTA, balance entre detección y asociación, estándar actual

MOT

BoT-SORT

Basado en ByteTrack y FastReID (integra apariencia)
Integrado por ultralytics desde YOLOv8





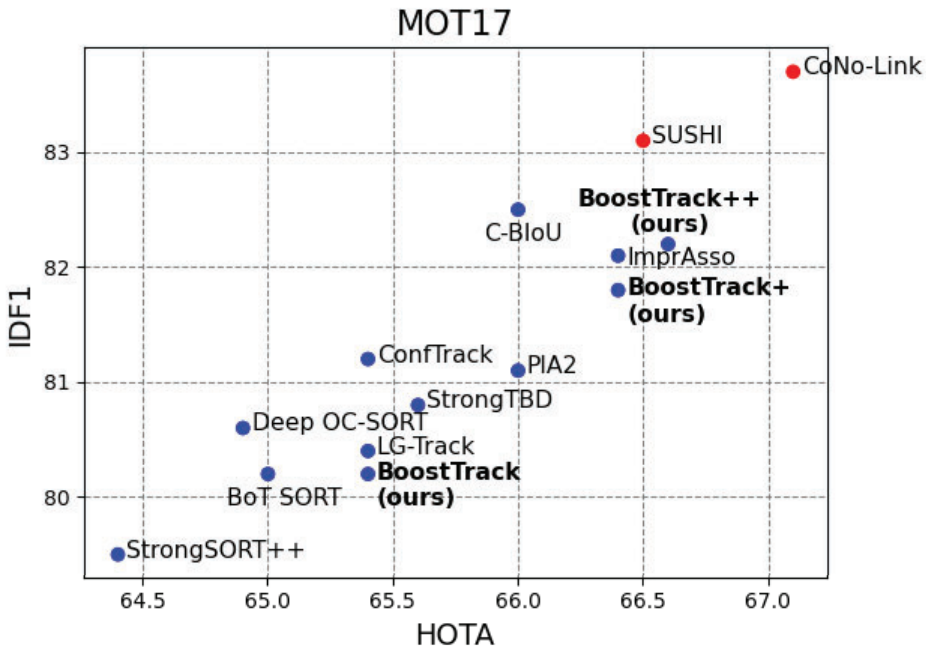
BoT-SORT: Robust Associations Multi-Pedestrian Tracking , 2022

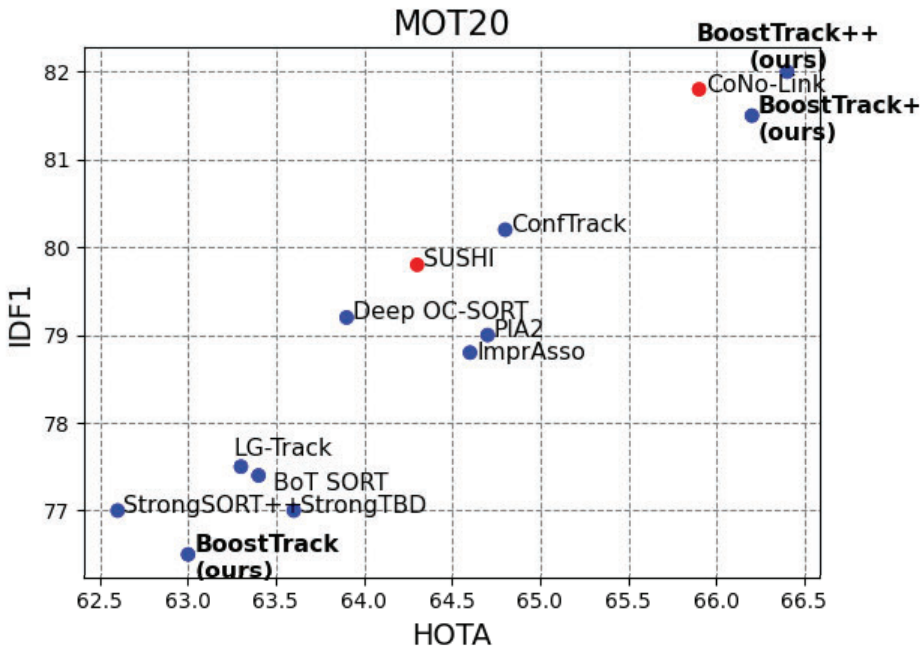
MOT

BoostTrack++

Sustituye IoU por distancia Mahalanobis y similitud forma

SOTA en MOT20

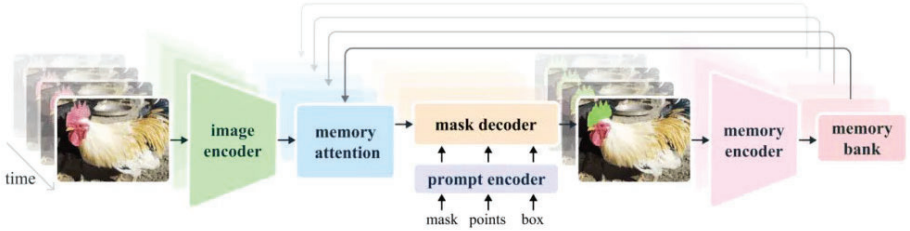


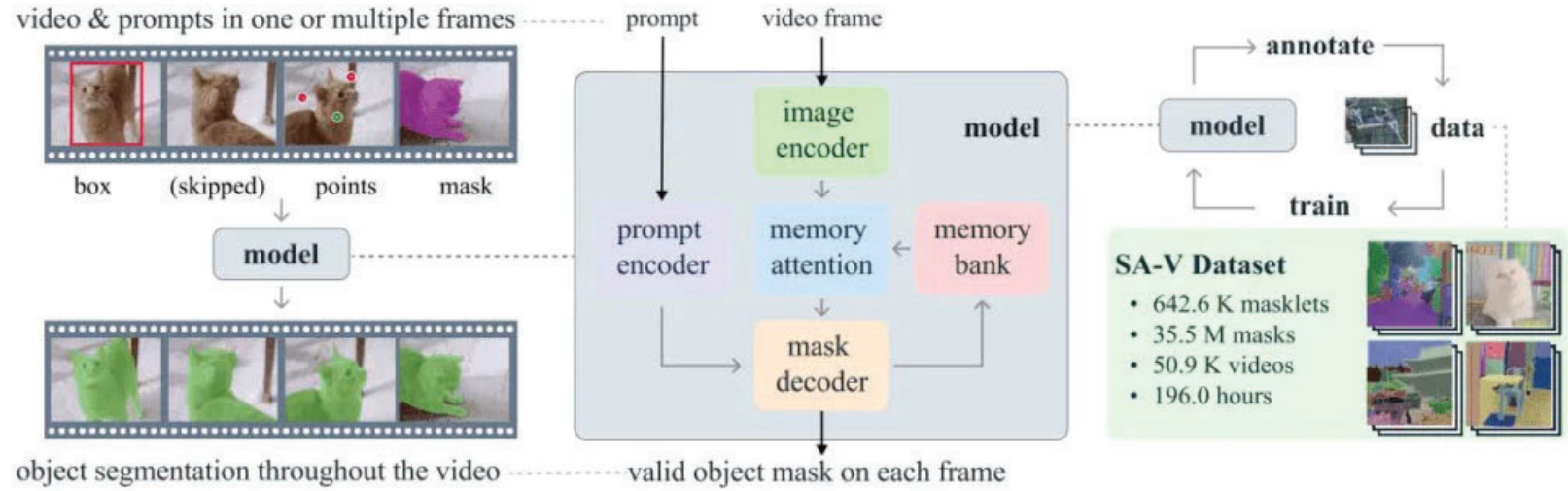


BoostTrack++: using tracklet information to detect more objects in multiple object tracking, 2024

MOT

SAM 2
Modelo fundacional
Añade módulos de memoria
Interactivo, entrenable
Tiempo real complicado

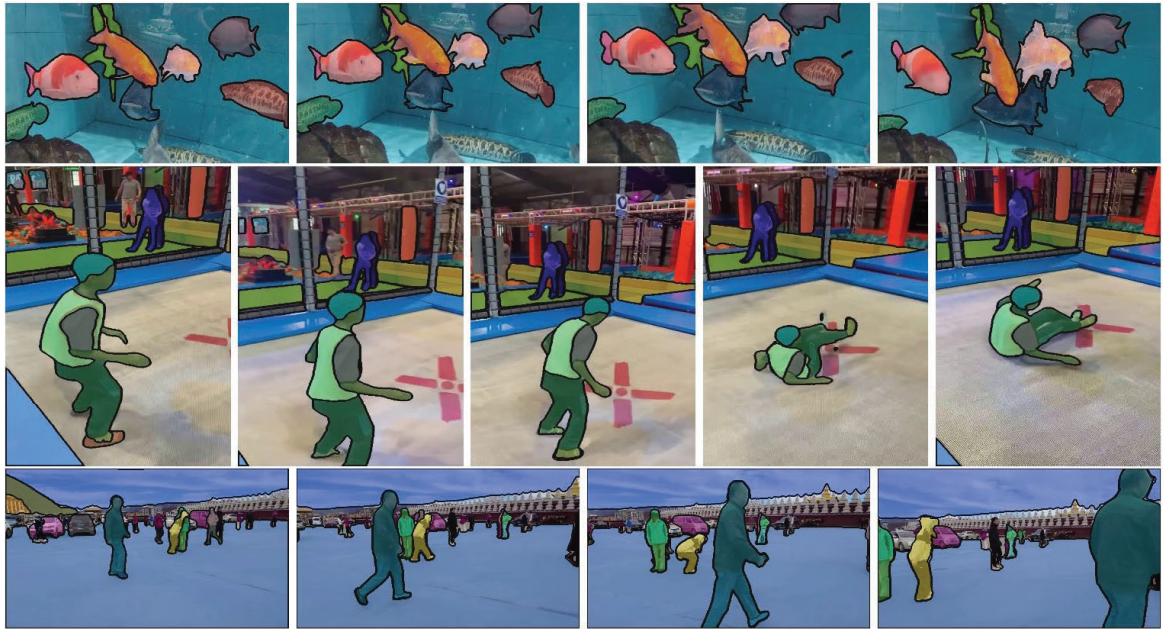




(a) Task: promptable visual segmentation **(b) Model:** Segment Anything Model 2 **(c) Data:** data engine and dataset

MOT

SAM 2
Sobre conjunto de datos SA-V
Cuadernos disponibles para vídeo
Disponible en ultralytics (recalcan que es más lento)
Nota:12/24 para imagen me va, para vídeo parece haber un *bug*



SA-V dataset

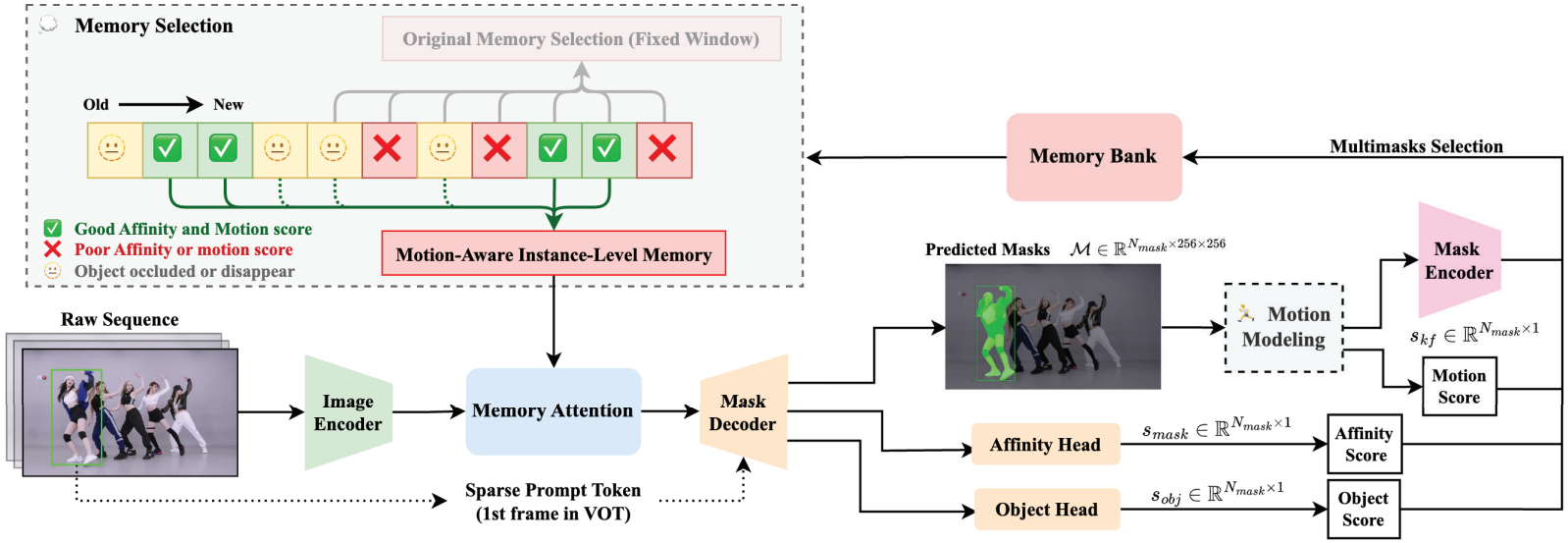
MOT

SAMURAI

Modelo de movimiento, filtro de Kalman

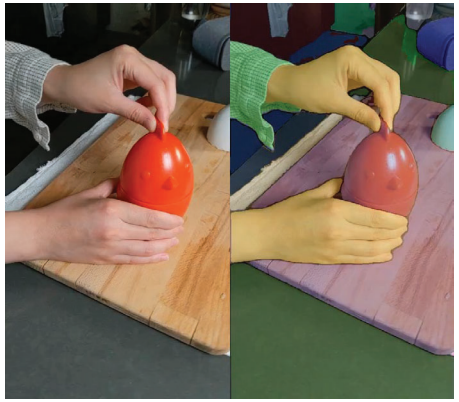
Demo SOT:

- TGC
- MOT

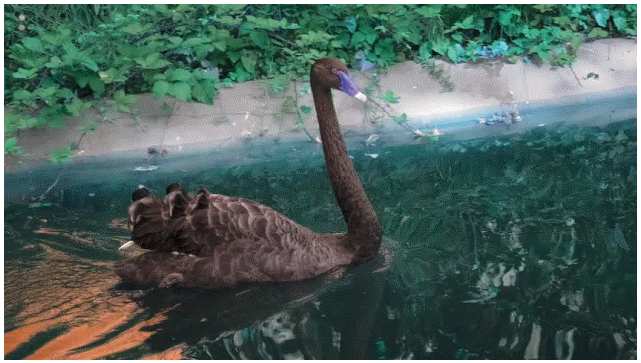


MOT

¿MOT?
AutoSeg-SAM2
Adaptado para MOT

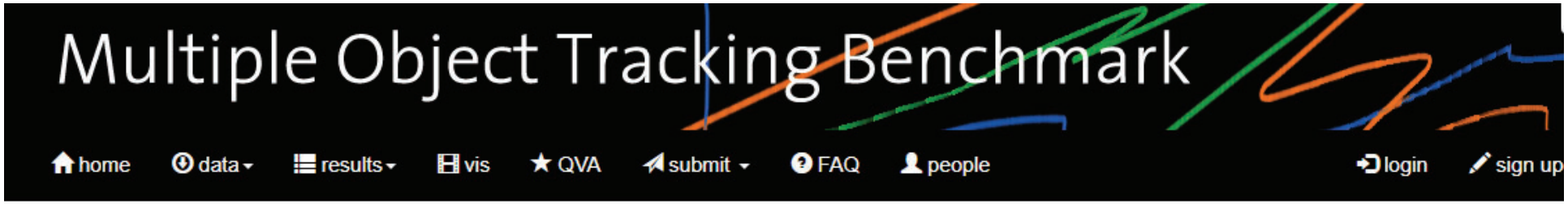


¿Tiempo real?



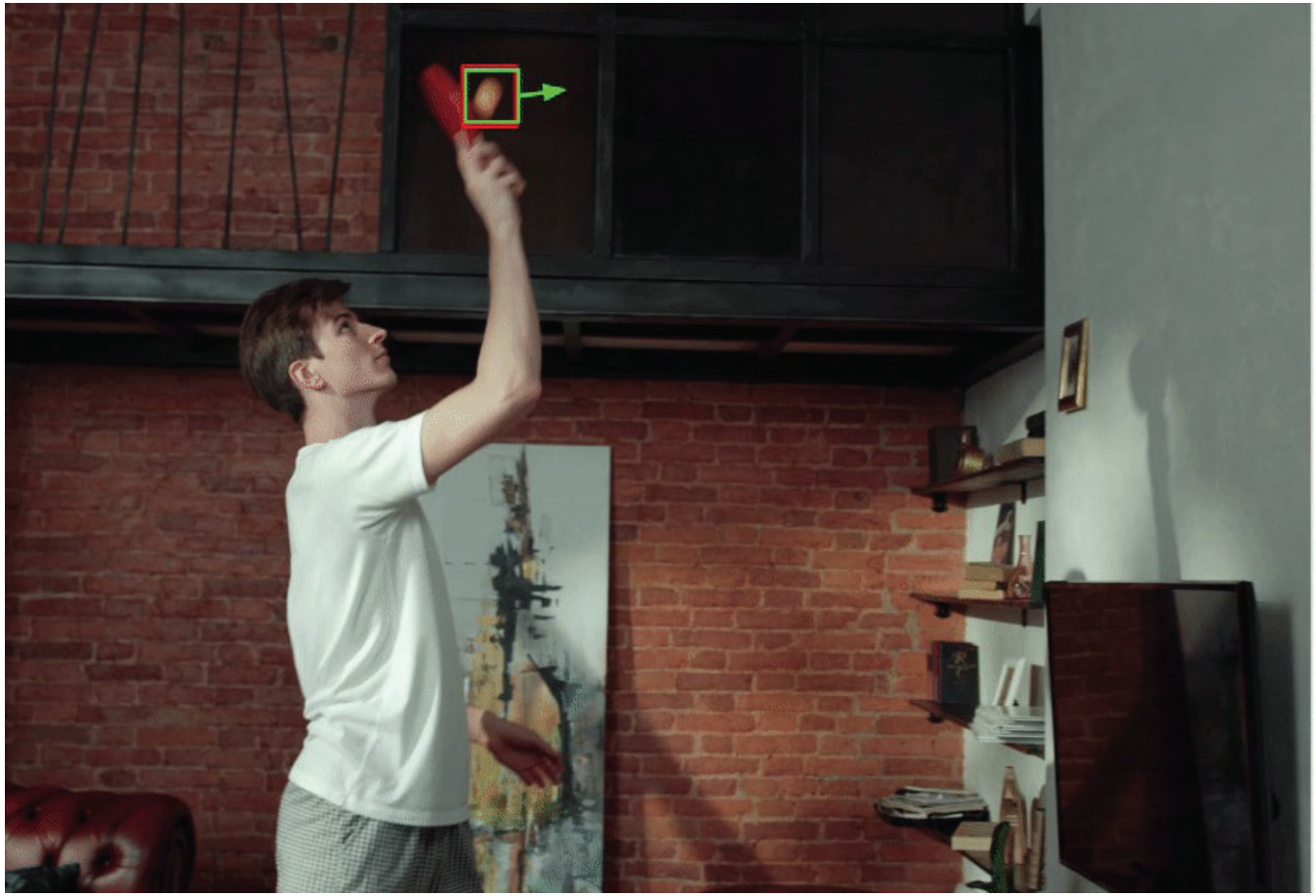
Benchmarks

MOT16
MOT17
MOT20
BDD100K
...



Filtro de Kalman

Predicción, detección

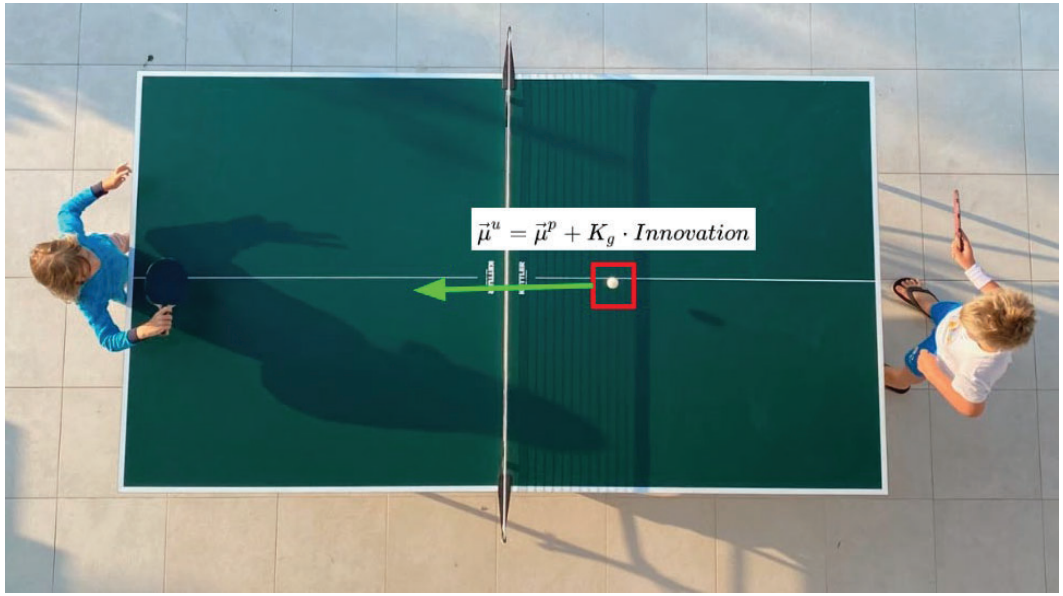


Filtro de Kalman

Durante el seguimiento, estimar posición en futuros fotogramas

Frecuente uso del filtro de Kalman (FK o KF)

Predicción estado del sistema



Filtro de Kalman

Dadas estimación y medición, estima estado futuro del objetivo

En seguimiento, medición proporcionada por la detección

Medición

Estimación previa

$$\hat{x}_n = K_n z_n + (1 - K_n)\ \hat{x}_{n-1}$$

Estimación actual

Ganancia de Kalman

Filtro de Kalman

Dadas estimación y medición, estima estado futuro del objetivo

En seguimiento, medición proporcionada por la detección

Medición

Estimación previa

$$\hat{x}_n = K_n z_n + (1 - K_n)\ \hat{x}_{n-1}$$

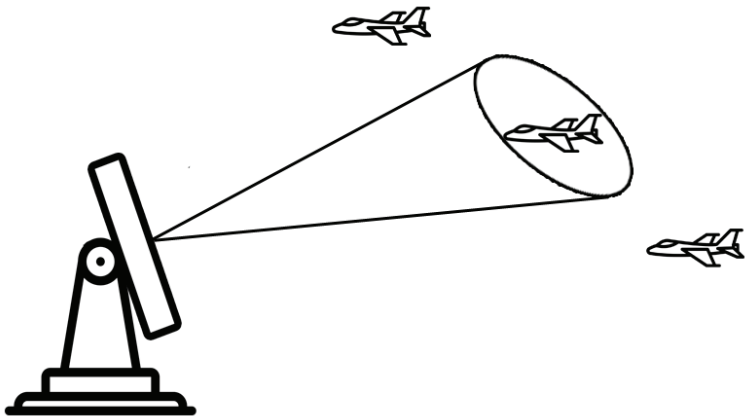
Estimación actual

Ganancia de Kalman

$$\hat{x}_n = K_n z_n + (1 - K_n)\ \hat{x}_{n-1} = \hat{x}_{n-1} + K_n (z_n - \hat{x}_{n-1})$$

Filtro de Kalman

Escenario radar y avión



Estado del sistema:

- posición
- velocidad
- aceleración

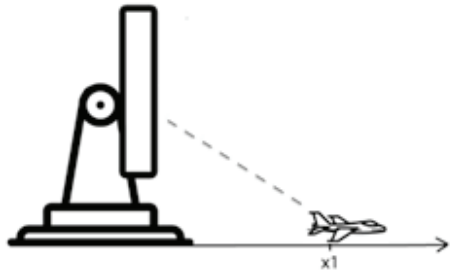
$$x = x_0 + v_0 \Delta t + \frac{1}{2} a \Delta t^2$$

Predecir estado parece trivial, pero:

- Ruido medición
- Ruido del proceso: viento, turbulencias, etc.

Filtro de Kalman

Sistema estático



Sistema el avión, estado su posición

Sin cambios de estado en t

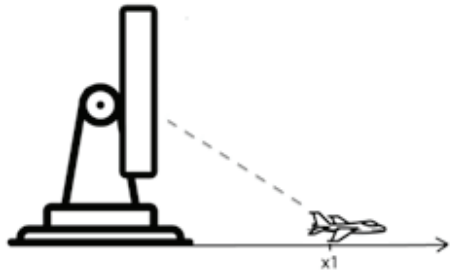
Modelo dinámico del sistema constante: su posición no cambia

Radar observa avión parado en tierra

Errores de medición

Filtro de Kalman

Sistema estático



Sistema el avión, estado su posición

Sin cambios de estado en t

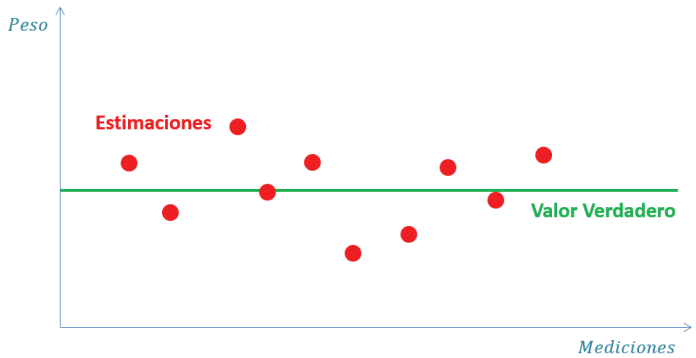
Modelo dinámico del sistema constante: su posición no cambia

Radar observa avión parado en tierra

Errores de medición

Estimación estado: múltiples mediciones y promedio

$$\hat{x}_{n,n} = \frac{1}{n}(z_1 + z_1 + \cdots + z_n) = \frac{1}{n}\sum_{i=1}^{n} z_i$$



Filtro de Kalman

Sistema estático

Estimación estado: múltiples mediciones y promedio

$$\hat{x}_{n,n} = \frac{1}{n}(z_1 + z_1 + \cdots + z_n) = \frac{1}{n}\sum_{i=1}^{n} z_i$$

$\hat{x}_{n,n}$ estimación de *x* en el tiempo *n*, con *n* medidas
$\hat{x}_{n,n-1}$ estimación de *x* en el tiempo *n*, con *n-1* medidas

$$\hat{x}_{n,n} = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n}\sum_{i=1}^{n-1} z_i + \frac{1}{n} z_n = \frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n-1} z_i + \frac{1}{n} z_n =$$

$$= \hat{x}_{n-1,n-1} + \frac{1}{n}\left(z_n - \hat{x}_{n-1,n-1}\right)$$

Filtro de Kalman

Sistema estático

Resumiendo

$$\hat{x}_{n,n} = \hat{x}_{n-1,n-1} + \frac{1}{n}\left(z_n - \hat{x}_{n-1,n-1}\right)$$

Al ser un sistema estático, estado predicho igual al estimado

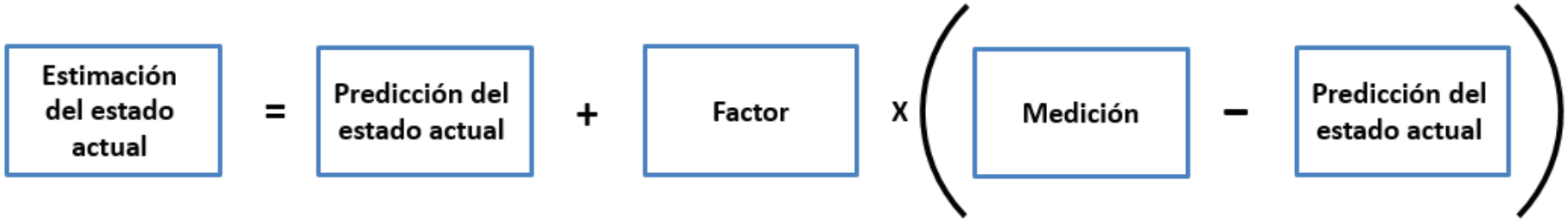
$$\hat{x}_{n,n-1} = \hat{x}_{n-1,n-1}$$

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + \frac{1}{n}\left(z_n - \hat{x}_{n,n-1}\right)$$

Filtro de Kalman

Sistema estático

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + \frac{1}{n}\left(z_n - \hat{x}_{n,n-1}\right)$$



Ecuación de actualización de estado, generaliza *Factor*

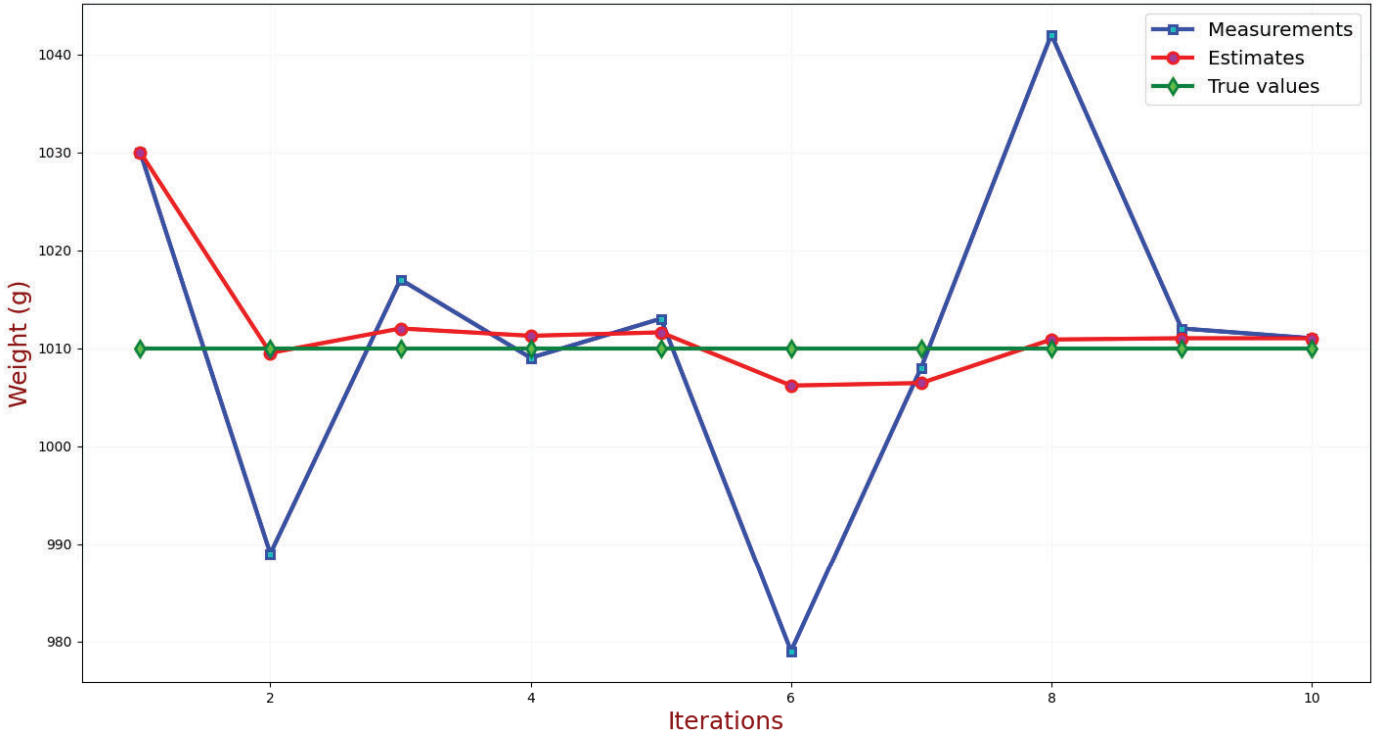
$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + \alpha_n\left(z_n - \hat{x}_{n,n-1}\right)$$

$z_n - \hat{x}_{n,n-1}$ medida residual, conocida como **innovación**

Filtro de Kalman

Sistema estático

Resultado

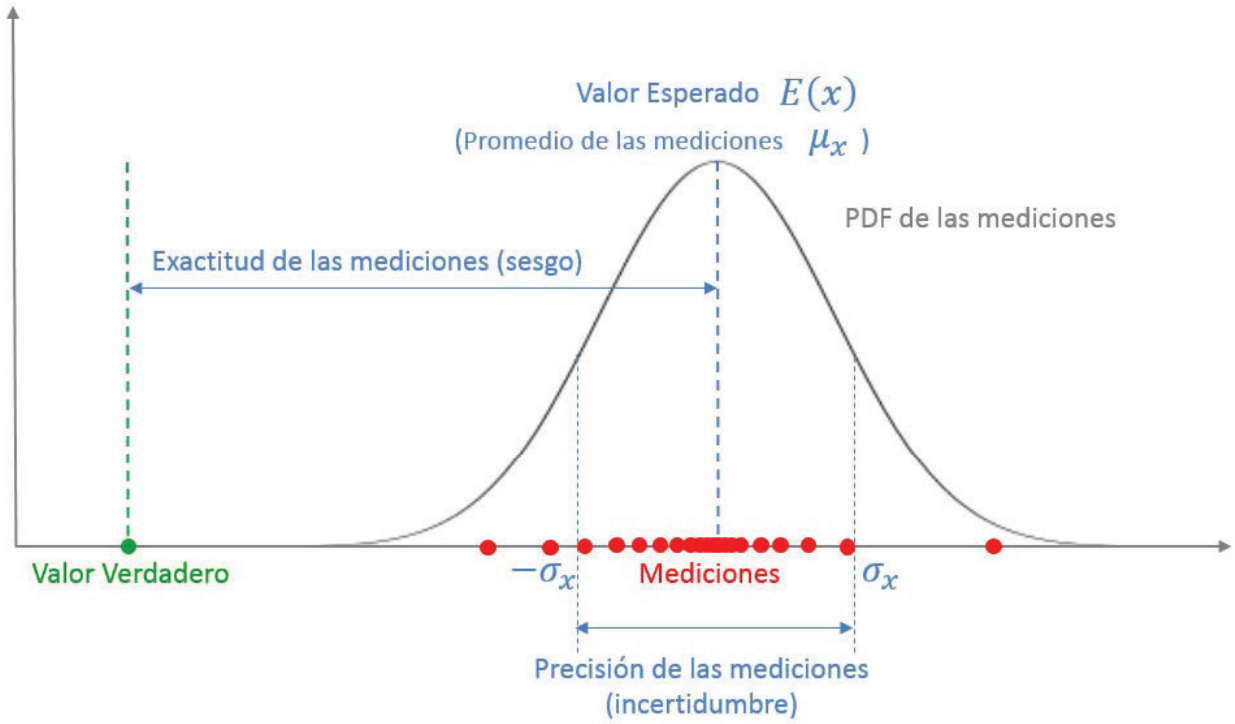


Filtro de Kalman

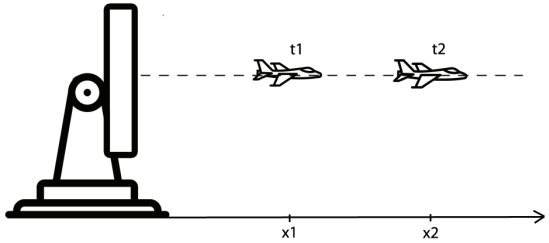
Radar presenta errores de medición

Medida caracterizada por su exactitud y precisión

Asume distribución normal, gaussiana



Filtro de Kalman



Sistema dinámico

Avión a velocidad constante
x_n posición de la nave en momento n
velocidad su derivada

$$\dot{x} = v = \frac{dx}{dt}$$

Δt intervalo entre pulsos del radar

Modelo dinámico (con velocidad constante):

$$x_{n+1} = x_n + \Delta t \dot{x}_n$$
$$\dot{x}_{n+1} = \dot{x}_n$$

Conforman la **ecuación de extrapolación de estado**

Filtro de Kalman

Sistema dinámico

Si *Δt* = 5s, la posición en *n-1* 30000 m y velocidad 40 m/s

Estimar en *n* la posición con la ec. de extrapolación de estado

$$\hat{x}_{n,n-1} = \hat{x}_{n-1,n-1} + \Delta t \hat{\dot{x}}_{n-1,n-1} = 30000 + 5 \times 40 = 30200$$

Filtro de Kalman

Sistema dinámico

Si Δt = 5s, la posición en *n-1* 30000 m y velocidad 40 m/s

Estimar en *n* la posición con la ec. de extrapolación de estado

$$\hat{x}_{n,n-1} = \hat{x}_{n-1,n-1} + \Delta t \hat{\dot{x}}_{n-1,n-1} = 30000 + 5 \times 40 = 30200$$

Sin embargo, la medida z_n ha sido 30110

Diferencia entre medida y predicción se puede deber a:

- Cambio de velocidad
- Medición del radar imprecisa

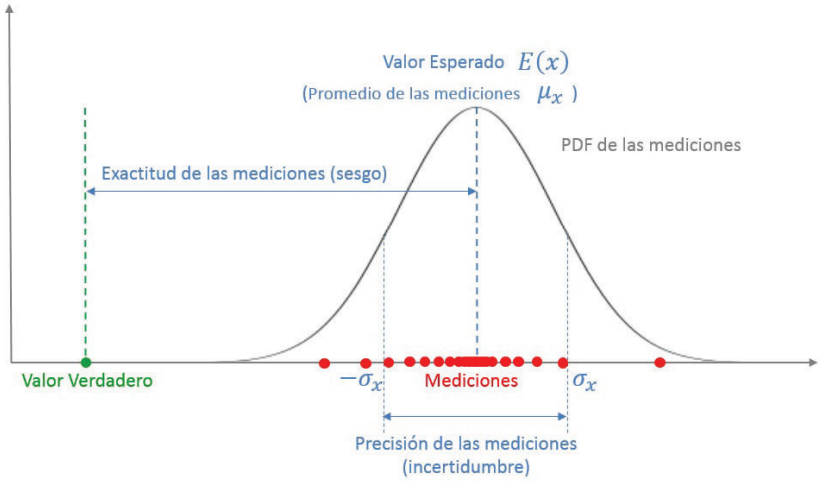
Filtro de Kalman

Sistema dinámico

Ec. de actualización de estado de la velocidad

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta\left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t}\right)$$

β dependiente de la precisión del radar (σ en figura)



Filtro de Kalman

Sistema dinámico

Diferencia entre estimación (30200) y medida (30110) es de 90 m

Filtro de Kalman

Sistema dinámico

Diferencia entre estimación (30200) y medida (30110) es de 90 m

Si σ=20 m, probable diferencia por cambio de velocidad, con β=0.9

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta\left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t}\right) = 40 + 0.9\left(\frac{30110 - 30200}{5}\right) = 23.8\, m/s$$

Filtro de Kalman

Sistema dinámico

Diferencia entre estimación (30200) y medida (30110) es de 90 m

Si σ=20 m, probable diferencia por cambio de velocidad, con β=0.9

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta\left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t}\right) = 40 + 0.9\left(\frac{30110 - 30200}{5}\right) = 23.8\, m/s$$

Si σ=150 m, probable diferencia por error de medida, con β=0.1

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta\left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t}\right) = 40 + 0.1\left(\frac{30110 - 30200}{5}\right) = 38.2\, m/s$$

Filtro de Kalman

Sistema dinámico

Ec. de actualización de estado de posición y velocidad

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + \alpha\left(z_n - \hat{x}_{n,n-1}\right)$$

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta\left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t}\right)$$

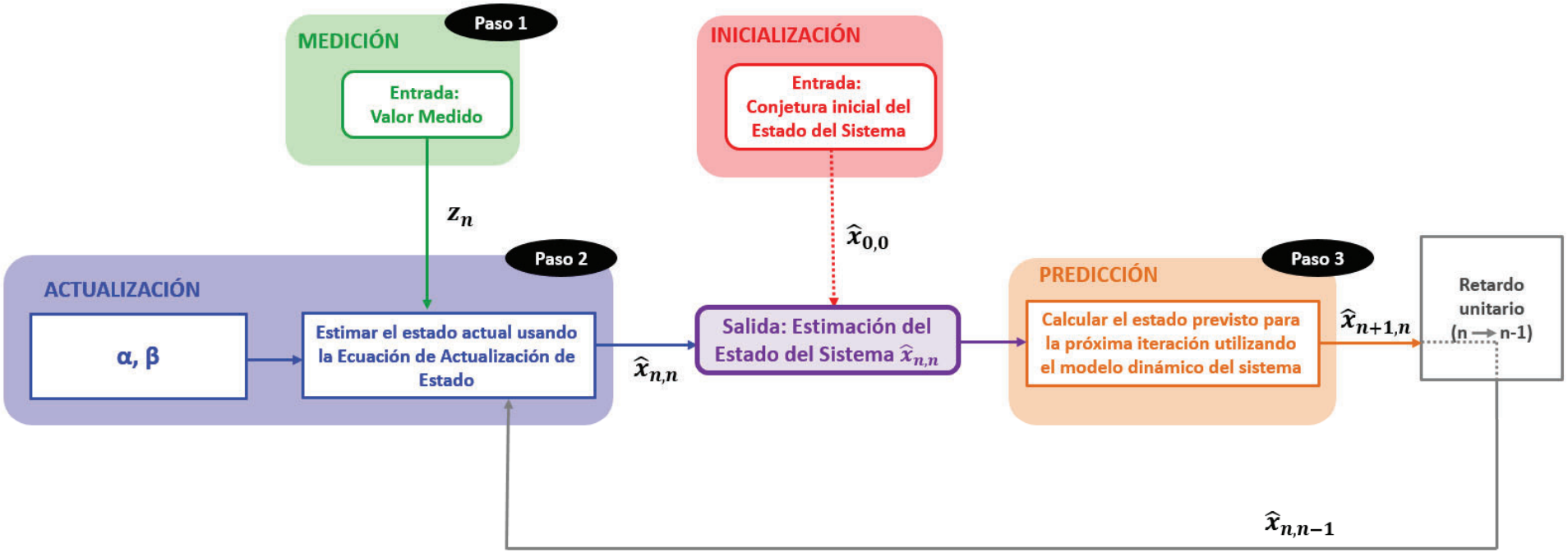
Alta precisión de la medida, α y β próximos a 1

Ecuaciones de actualización de seguimiento α–β

Filtro de Kalman

Sistema dinámico

Filtro α–β



Filtro de Kalman

Sistema dinámico

Filtro α–β-γ incluye la aceleración

Ecuaciones de extrapolación de estado o modelo dinámico

$$\hat{x}_{n+1,n} = \hat{x}_{n,n} + \Delta t \hat{\dot{x}}_{n,n} + \frac{\Delta t^2}{2} \hat{\ddot{x}}_{n,n}$$

$$\hat{\dot{x}}_{n+1,n} = \hat{\dot{x}}_{n,n} + \Delta t \hat{\ddot{x}}_{n,n}$$

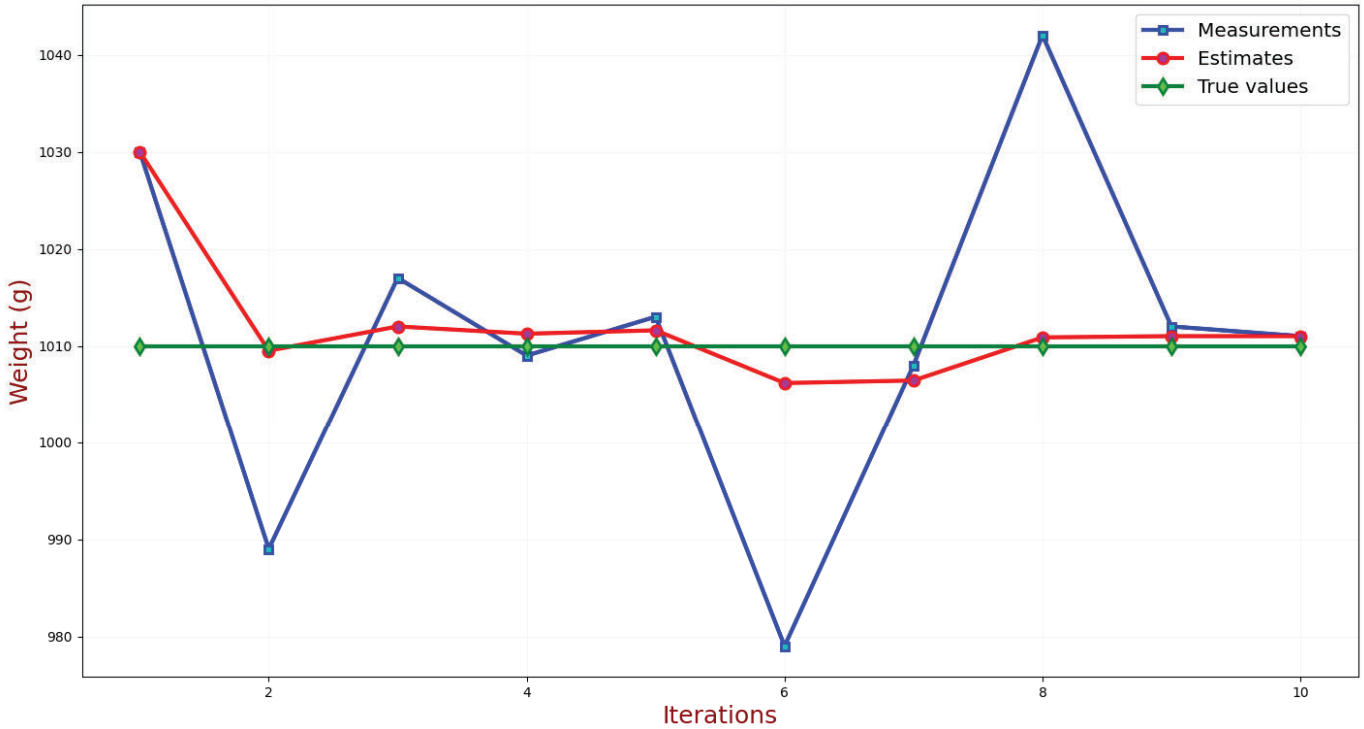
$$\hat{\ddot{x}}_{n+1,n} = \hat{\ddot{x}}_{n,n}$$

Filtro de Kalman

FK sin ruido

Calcula valores de α–β-γ con ganancia de Kalman en cada iteración

Basada en incertidumbre de medición, ***r***, y de estimación, ***p***



Filtro de Kalman

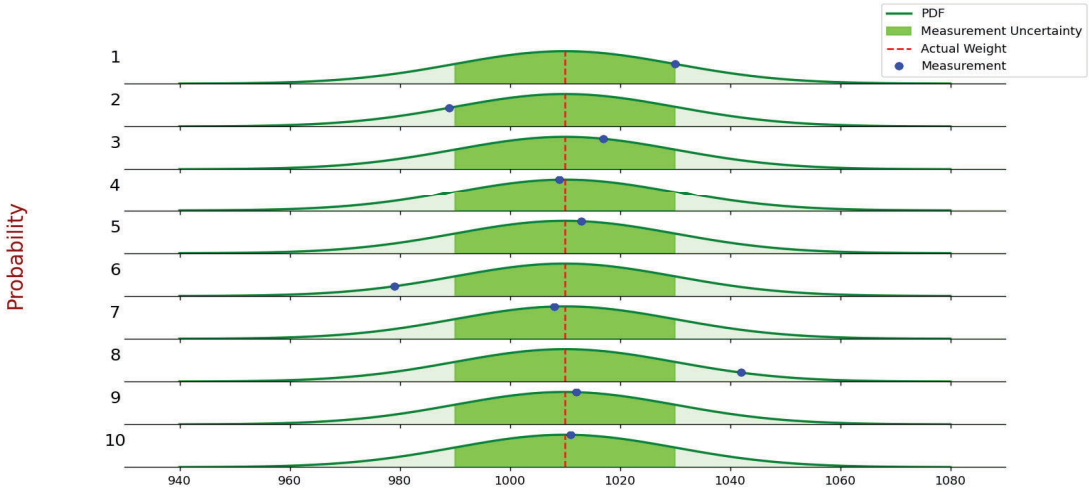
FK sin ruido

Diferencia entre medición y valor verdadero → error de medición

Incertidumbre de medición o ***r***, descrito con varianza (gaussiana)

Función de densidad de probabilidad (PDF siglas en inglés)

Proporcionada por proveedor o tras calibración



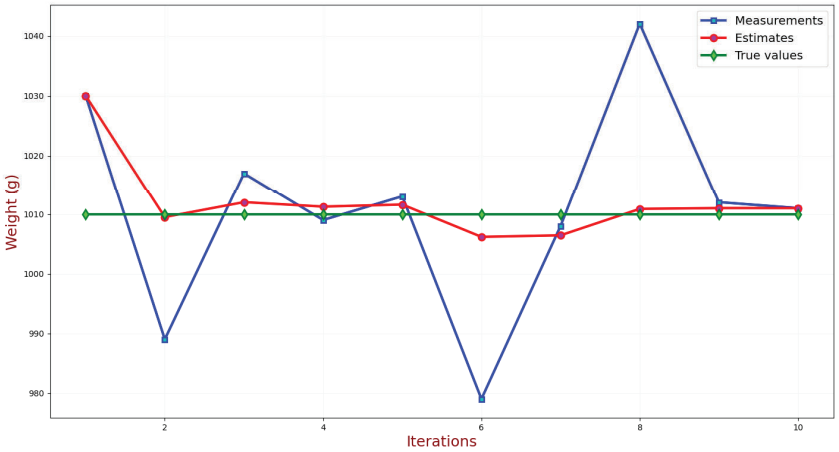
Filtro de Kalman

FK sin ruido

Diferencia estimación y valor verdadero → error de estimación

No se conoce el error, hay incertidumbre de la estimación, ***p***

Converge hacia cero



Filtro de Kalman

FK sin ruido

Kalman calcula los valores a, B (y γ) dinámicamente

Ganancia de Kalman, cambia en cada iteración

$$K_n = \frac{p_{n,n-1}}{p_{n,n-1} + r_n}$$

Filtro de Kalman

FK sin ruido

Kalman calcula los valores a, B (y γ) dinámicamente

Ganancia de Kalman, cambia en cada iteración

$$K_n = \frac{p_{n,n-1}}{p_{n,n-1} + r_n}$$

Ec. actualización de estado

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + K_n\left(z_n - \hat{x}_{n,n-1}\right) = (1 - K_n)\hat{x}_{n,n-1} + z_n K_n$$

Ponderación entre medición y estimación

Filtro de Kalman

FK sin ruido

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + K_n\left(z_n - \hat{x}_{n,n-1}\right) = (1 - K_n)\hat{x}_{n,n-1} + z_n K_n$$

Ponderación medición y estimación

Según si *p* o *r* es mayor, K_n es próximo a 0 (*r* alta) o a 1 (*p* alta)

Actualización incertidumbre de estimación

$$p_{n,n} = (1 - K_n)p_{n,n-1}$$

Menguante, porque $K_n < 1$

Con ruido, la ec. de actualización de covarianza lo integra

Filtro de Kalman

FK sin ruido

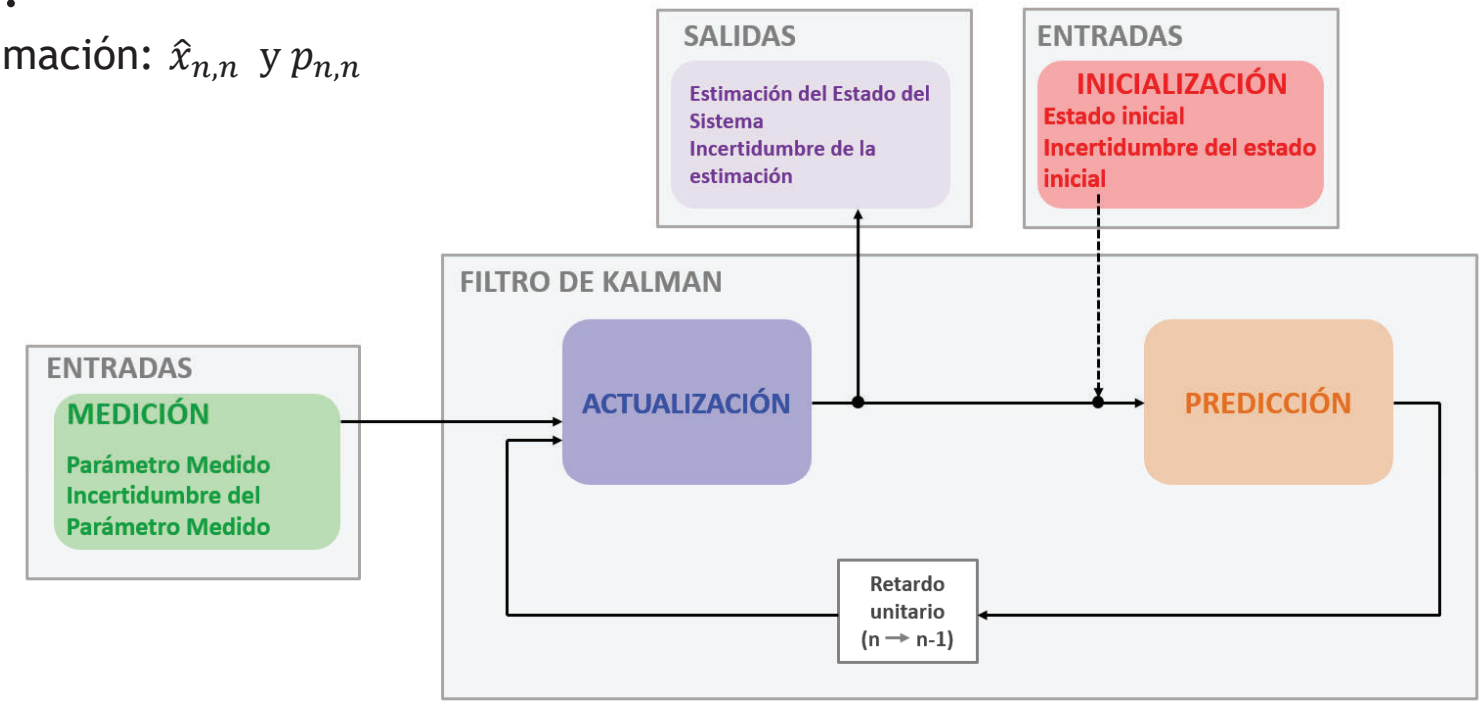
Entradas:

Inicialización: $\hat{x}_{1,0}$ y $p_{1,0}$

Medición: z_n y r_n

Salidas:

Estimación: $\hat{x}_{n,n}$ y $p_{n,n}$



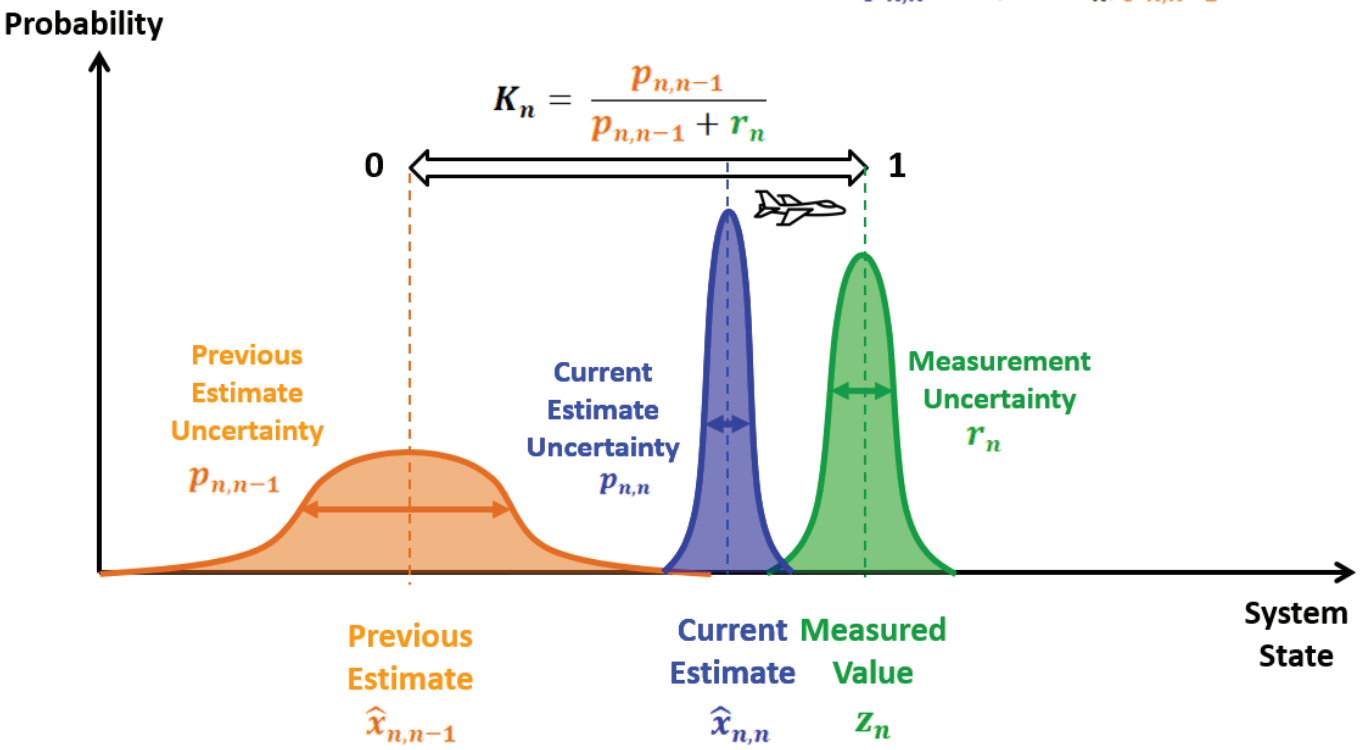
Filtro de Kalman

FK sin ruido

Alta K_n, nueva estimación cerca medición

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + K_n\left(z_n - \hat{x}_{n,n-1}\right)$$

$$p_{n,n} = (1 - K_n)p_{n,n-1}$$



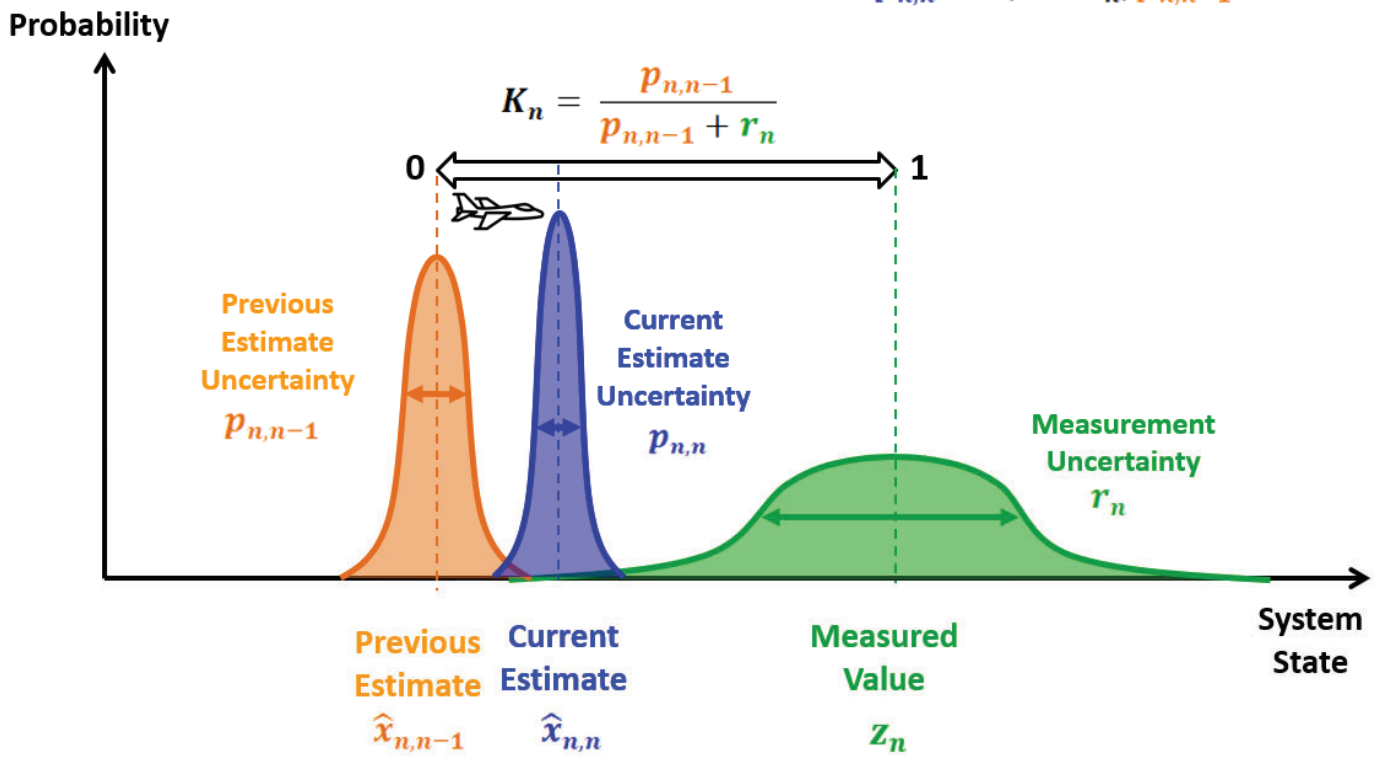
Filtro de Kalman

FK sin ruido

Baja K_n, nueva estimación cerca estimación previa

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + K_n\left(z_n - \hat{x}_{n,n-1}\right)$$

$$p_{n,n} = (1 - K_n)p_{n,n-1}$$



Filtro de Kalman

Resumen

Algoritmo que estima estado del sistema

Estima variables ocultas a partir de otras observables

Maximiza probabilidad a posteriori

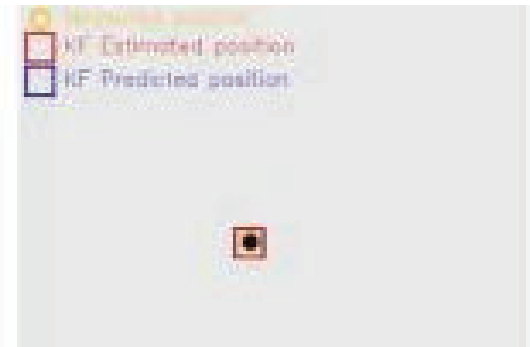
Asume:
- Dinámica lineal
- Ruido blanco, aditivo y gaussiano de las medidas

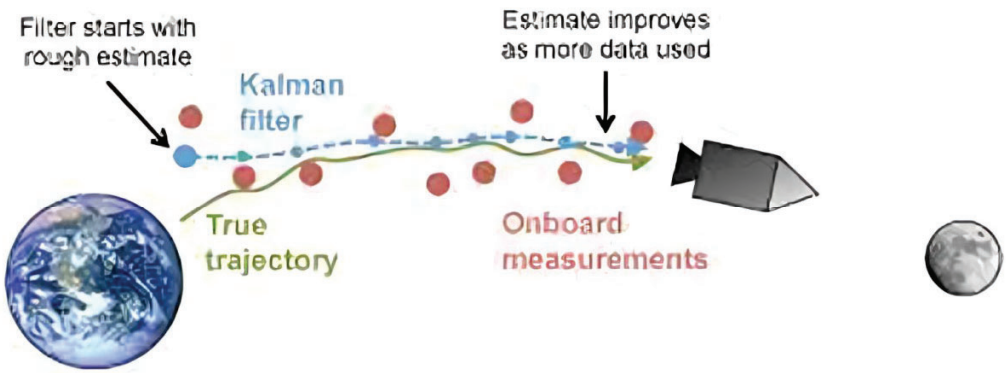
Existen variantes, alternativas como el filtro de partículas

Filtro de Kalman

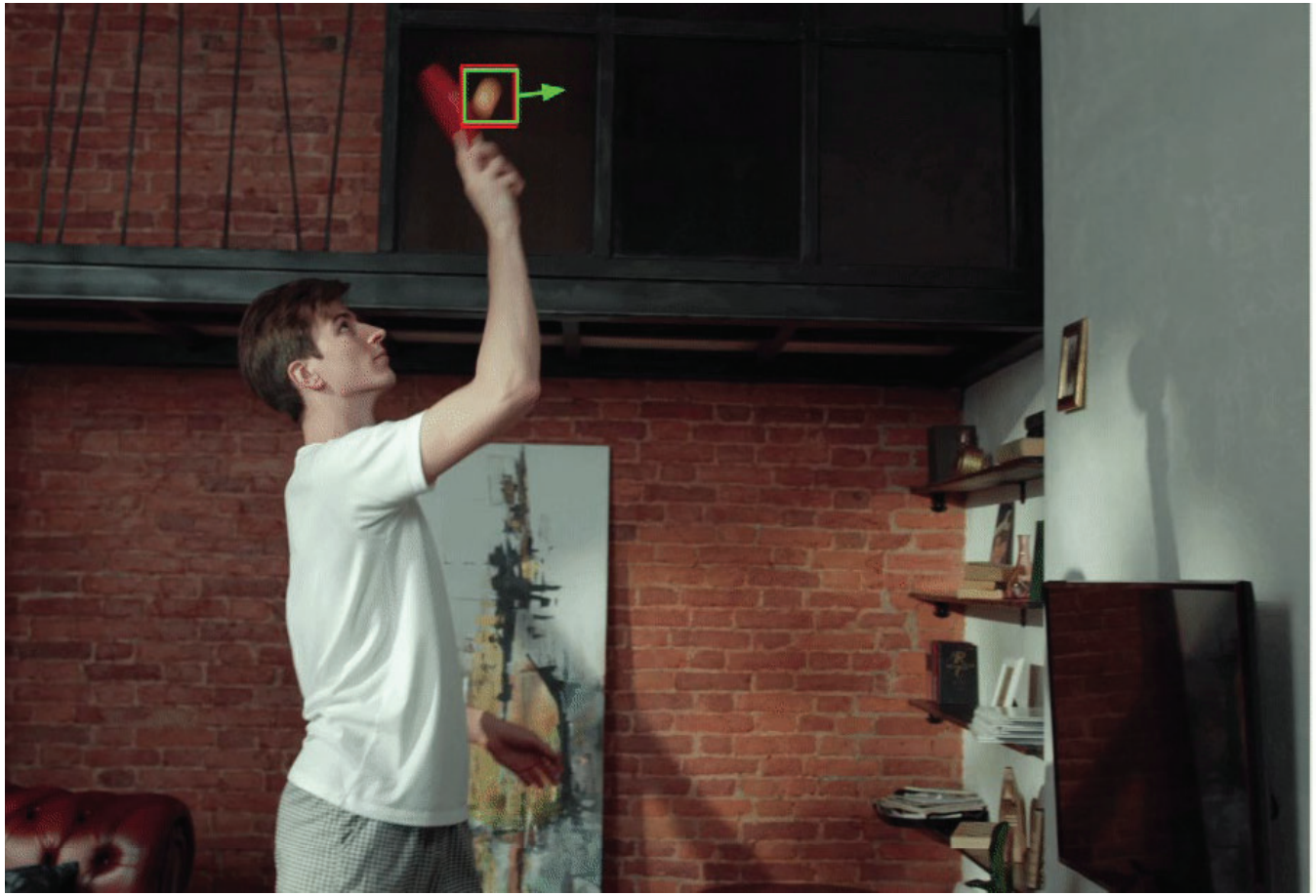
Aplicaciones:

- Seguimiento de proyectiles
- Seguimiento de cabezas/manos, etc.
- Movimiento de labios
- Ajuste de Bezier a nubes de puntos
- Economía
- Navegación





Kalman



Bibliografía

- Filtros de Kalman, Alex Becker
- R. Szeliski. Computer Vision: Algorithms and Applications, 2nd ed. 2022
- A. Torralba et al. Foundations of Computer Vision, 2024
- Awesome Optical Flow Papers