



Div. Ingeniería de Sistemas y Automática
Grupo de Tecnologías Industriales
Universidad Miguel Hernández

Reconstrucción Tridimensional a partir del Movimiento



UPM DISAM
Grupo de Visión



Dpto. Automática, Ingeniería
Electrónica e Informática Industrial
Universidad Politécnica de Madrid

- ☰ **Introducción**
- ↳ **El campo de movimiento (“Motion-Field”)**
- ↳ **Flujo óptico**
- ↳ **Estimación del campo de movimiento**
- ↳ **Reconstrucción 3D**

↳ Ejemplo de objetos de la escena fijo y cámara en movimiento



- ↳ **Datos de entrada: Secuencias de Imágenes en diferentes instantes de tiempo**
- ↳ **¿Qué información se puede obtener de los cambios que hay en una secuencia de imágenes tomadas en diferentes instantes de tiempo?**
- ↳ **Suponiendo iluminación constante, los cambios en la imágenes pueden ser producidos por el movimiento relativo entre cámara y escena :**
 - ↳ **Movimiento de la cámara frente a una escena estática**
 - ↳ **Movimiento de objetos de la escena frente a una cámara estacionaria**
 - ↳ **Movimiento de cámara y objetos de la escena respecto a una referencia externa**

Importancia del Movimiento visual

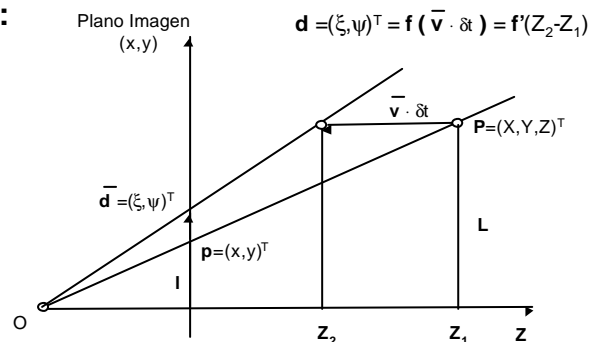
5

⇨ **2 Razones para introducir variable tiempo al procesamiento de imágenes:**

↳ **Movimiento Aparente de los objetos en la imagen → entender la estructura de los objetos y su movimiento en el espacio 3D**

↳ **Sistemas visuales biológicos utilizan el movimiento visual para inferir propiedades del espacio 3D con un pequeño conocimiento a priori del mismo**

⇨ **Ejemplo :**



Visión Tridimensional

UPM-DISAM / UMH

Información relacionada con el movimiento

6

⇨ **Detección del movimiento mediante una cámara**

↳ **Seguridad**

↳ **Cámara estática**

⇨ **Detección y localización de objetos en movimiento**

↳ **Cámara fija y objeto móvil**

↳ **Objeto móvil y cámara en movimiento**

↳ **Objetivo: detectar el movimiento y predecir su trayectoria futura**

⇨ **Obtención de características 3D de los objetos a partir de las proyecciones bidimensionales tomadas en diferentes instantes de tiempo**

Visión Tridimensional

UPM-DISAM / UMH

Problemas del análisis del Movimiento

7

- ⇐ **Se puede dividir en dos subproblemas la reconstrucción 3D:**
 - ↪ **Correspondencia** : ¿qué elementos entre diferentes imágenes de la secuencia corresponden a la misma primitiva 3D?

 - ↪ **Reconstrucción**: Dado un número de elementos correspondientes, y el conocimiento de los parámetros intrínsecos de la cámara, ¿qué se puede decir del movimiento en el espacio 3D y de la estructura del mundo observado?

Problemas del análisis del Movimiento

8

- ⇐ **Problema de correspondencia** → estimar el campo de movimiento de los objetos proyectados en la imagen.
 - ↪ **Métodos de resolución del problema de correspondencia** :
 - **Métodos Diferenciales** : medidas densas, calculadas para cada punto de la imagen. Secuencia de muestreo de imágenes alta. *Flujo Óptico*
 - **Métodos de correspondencia de características** : medidas dispersas, calculadas sólo para un conjunto de puntos de la imagen. *Filtro de Kalman*. Utilizan el flujo óptico como estimación inicial.

- ⇐ **Reconstrucción** → supone obtener los mapas de velocidad de los objetos en el escena 3D.

↳ **Restricción** : solo hay un movimiento rígido, relativo entre la cámara y la escena observada, y las condiciones de iluminación no cambian.

↳ **Ejemplo:**

→ SI: vistas de un edificio por un observador en movimiento

→ NO: secuencias de fútbol, tráfico,..

→ NO: objetos deformables.

↳ **Varios objetos se mueven. Aparece el problema de la Segmentación.**

↳ ¿Cuáles son las regiones de la imagen que corresponden a objetos diferentes moviéndose?

↳ **Planteamientos:**

Segmentar ® Correspondencia

o

Correspondencia ® Segmentar.

↳ **Introducción**

☒ **El campo de movimiento (“Motion-Field”)**

↳ **Flujo óptico**

↳ **Estimación del campo de movimiento**

↳ **Reconstrucción 3D**

El campo de movimiento (1)

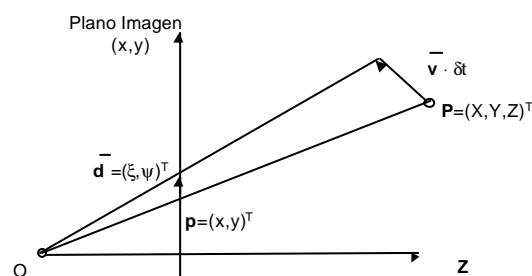
11

↳ **Campo de movimiento "Motion Field" :**

↳ Una representación bidimensional de un movimiento tridimensional se denomina campo de movimiento

↳ **Definición** : Es el campo de vectores de velocidades 2D de los puntos de la imagen, inducido por el movimiento relativo entre la cámara y la escena observada.

↳ Se puede entender como la proyección del campo de velocidades 3D en el plano de la imagen.



Visión Tridimensional

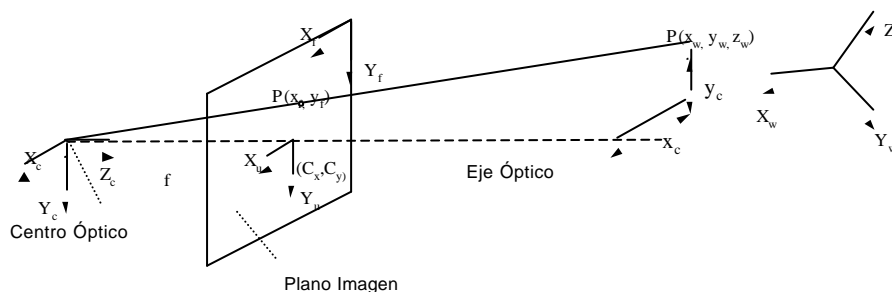
UPM-DISAM / UMH

El campo de movimiento (2)

12

↳ Las coordenadas de imagen las consideraremos respecto al S.R. centrado en el punto principal.

↳ Consideraremos los parámetros intrínsecos conocidos (correspondencia entre las coordenadas 3D en coordenadas de cámara [sistema fijo] y las coordenadas 2D de imagen centradas)



Visión Tridimensional

UPM-DISAM / UMH

El campo de movimiento (3)

13

⇐ **Ecuaciones de proyección sin considerar distorsión:**

$$\begin{bmatrix} nx_u \\ ny_u \\ n \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad \begin{aligned} x_u &= f \frac{X_c}{Z_c} \\ y_u &= f \frac{Y_c}{Z_c} \end{aligned}$$



$$\begin{aligned} \bar{p}(t) &= f \frac{\bar{P}(t)}{Z_c} & \bar{p}(t) &= f \frac{\bar{P}(t)}{\bar{P}(t) \cdot \bar{k}} \\ \bar{p} &= [x_u, y_u, f]^T & \bar{k} &= [0, 0, 1]^T \\ \bar{P} &= [X_c, Y_c, Z_c]^T \end{aligned}$$

En adelante $p=[x,y,f]^T$ $P=[X,Y,Z]^T$

El campo de movimiento (4)

14

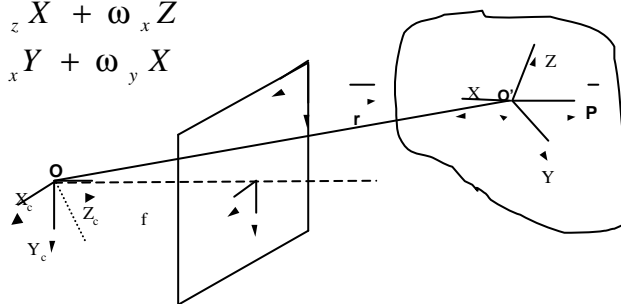
⇐ **El movimiento relativo entre el punto P y la cámara lo descomponemos en una traslación y una rotación:**

$$\bar{V} = - \frac{d \overline{OP}(t)}{dt} = -\bar{T} - \bar{\omega} \times \overline{O'P}$$

$$V_x = -T_x - \omega_y Z + \omega_z Y$$

$$V_y = -T_y - \omega_z X + \omega_x Z$$

$$V_z = -T_z - \omega_x Y + \omega_y X$$



El campo de movimiento (5)

15

↖ **Derivando respecto al tiempo:**

$$\bar{v} = \frac{d\bar{p}}{dt} = f \frac{d}{dt} \left(\frac{\bar{P}}{Z} \right) = f \frac{Z\bar{V} - V_z\bar{P}}{Z^2}$$



de forma más compacta

$$\bar{v} = \frac{d\bar{p}}{dt} = f \frac{(\bar{P} \cdot \bar{k})\bar{V} - (\bar{V} \cdot \bar{k})\bar{P}}{(\bar{P} \cdot \bar{k})^2}$$

$$\bar{v}(t) = \frac{d\bar{p}}{dt} = \frac{f}{Z(t)^2} (\bar{P}(t) \times \bar{V}(t)) \times \bar{k}$$

$$(\bar{a} \times \bar{b}) \times \bar{c} = (\bar{a} \cdot \bar{c})\bar{b} - (\bar{b} \cdot \bar{c})\bar{a}$$

Visión Tridimensional

UPM-DISAM / UMH

El campo de movimiento (6)

16

↖ **Tomando componentes :**

$$v_x = \frac{T_z x - T_x f}{Z} - \omega_y f + \omega_z y + \frac{\omega_x xy}{f} - \frac{\omega_y x^2}{f}$$

$$v_y = \frac{T_z y - T_y f}{Z} + \omega_x f - \omega_z x - \frac{\omega_y xy}{f} + \frac{\omega_x y^2}{f}$$

↖ **Dos componentes: que dependen de la traslación y la rotación respectivamente**

$$v_x^T = \frac{T_z x - T_x f}{Z}$$

$$v_y^T = \frac{T_z y - T_y f}{Z}$$

$$v_x^\omega = -\omega_y f + \omega_z y + \frac{\omega_x xy}{f} - \frac{\omega_y x^2}{f} \quad v_y^\omega = \omega_x f - \omega_z x - \frac{\omega_y xy}{f} + \frac{\omega_x y^2}{f}$$

Visión Tridimensional

UPM-DISAM / UMH

El campo de movimiento (7)

17

⇐ La componente del campo de movimiento en dirección del eje óptico es cero

$$\bar{v} = [v_x, v_y]^T \Rightarrow \bar{v} = [v_x, v_y, 0]^T$$

⇐ De las ecuaciones anteriores obtenemos que w (vel. angular) queda desacoplada de Z (profundidad)

↗ La parte del campo de movimiento que depende de la velocidad angular no tiene implícita información sobre la profundidad

$$v_x^T = \frac{T_z x - T_x f}{Z} \quad v_y^T = \frac{T_z y - T_y f}{Z}$$

Visión Tridimensional

UPM-DISAM / UMH

El campo de movimiento (8)

18

CASOS :

⇐ **Traslación Pura ($w = 0$)**

↗ T_z distinto de cero. Llamamos

$$v_x^T = \frac{T_z x - T_x f}{Z} \quad v_y^T = \frac{T_z y - T_y f}{Z}$$

$$\left\{ \begin{array}{l} \bar{p}_0 = [x_0, y_0]^T \text{ Punto de fuga} \\ x_0 = f \frac{T_x}{T_z} \quad y_0 = f \frac{T_y}{T_z} \end{array} \right.$$

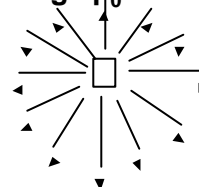
$$v_x^T = (x - x_0) \frac{T_z}{Z} \quad v_y^T = (y - y_0) \frac{T_z}{Z}$$

↗ El campo de movimiento de una traslación pura es **RADIAL**.
Vectores en forma radial desde un punto de fuga p_0

☑ Si $T_z < 0 \Rightarrow p_0$ es el foco de expansión

☑ Si $T_z > 0 \Rightarrow p_0$ es el foco de contracción

↗ p_0 es la intersección del rayo paralelo al vector de traslación con el plano de la imagen



Visión Tridimensional

UPM-DISAM / UMH

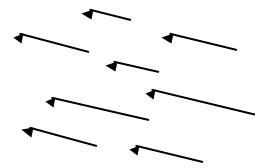
El campo de movimiento (9)

19

↪ La longitud de $\bar{v} = \bar{v}(p)$ es proporcional a la distancia entre \bar{p} y p_0 , e inversamente proporcional a la profundidad del punto P en el espacio 3D

↪ Si $T_z = 0$, todos los vectores del campo de movimiento son paralelos y sus longitudes son inversamente proporcionales a la profundidad.

$$v_x^T = -\frac{T_x f}{Z} \quad v_y^T = -\frac{T_y f}{Z}$$



El campo de movimiento (10)

20

↪ **Movimiento de un plano**

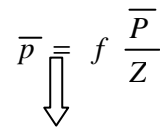
↪ Observamos una superficie plana

$$\bar{n}^T \cdot \bar{P} = d$$

$$\bar{n}^T = [n_x, n_y, n_z]$$



$$\frac{n_x x + n_y y + n_z f}{f} Z = d$$



↪ n ® Normal al plano

↪ d ® distancia del plano al centro de proyección.

↪ La superficie plana tiene un movimiento (T, w)

↪ Despejando Z y sustituyendo en la ecuación general del campo de movimiento :

$$v_x = \frac{1}{fd} (a_1 x^2 + a_2 xy + a_3 fx + a_4 fy + a_5 f^2) \quad \begin{array}{ll} a_1 = -d \omega_y + T_z n_x & a_2 = d \omega_x + T_z n_y \\ a_3 = T_z n_z - T_x n_x & a_4 = d \omega_z - T_x n_y \end{array}$$

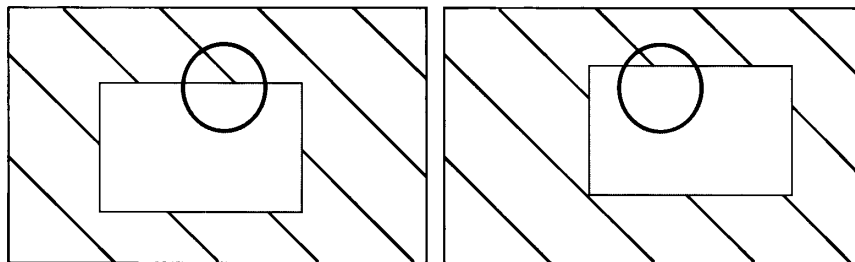
$$v_y = \frac{1}{fd} (a_1 xy + a_2 y^2 + a_6 fy + a_7 fx + a_8 f^2) \quad \begin{array}{ll} a_5 = -d \omega_y - T_x n_z & a_6 = T_z n_z - T_y n_y \\ a_7 = -d \omega_z - T_y n_x & a_8 = d \omega_x - T_y n_z \end{array}$$

⇐ Conclusiones :

- ↗ El campo de movimiento de una superficie plana, es un polinomio cuadrático en coordenadas de la imagen
- ↗ Debido a la simetría espacial de los coeficientes del polinomio, el mismo campo de velocidades puede ser producido por diferentes superficies planares bajo diferentes movimientos en el espacio
- ↗ Para recuperar el movimiento en el espacio, no pueden estar basado en puntos coplanares.

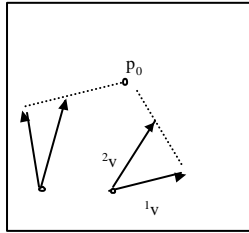
⇐ Problema de Apertura:

- ↗ La imagen es una proyección limitada de la escena. No siempre proporciona suficiente información acerca del movimiento global de un objeto



⇐ **Paradoja del movimiento :**

- ↗ El campo de movimiento relativo de dos puntos 3D solidarios que se proyectan sobre el mismo punto en la imagen (${}^1p = {}^2p$), no depende de la componente rotacional del movimiento en el espacio 3D.
- ↗ Apunta hacia p_0 , el punto de fuga de la dirección de traslación. Cuando la recta proyección coincide con T (p_0) el campo de velocidades relativo será nulo



$${}^1v_x - {}^2v_x = (T_z x - T_x f) \left(\frac{1}{{}^1Z} - \frac{1}{{}^2Z} \right)$$

$${}^1v_y - {}^2v_y = (T_z y - T_y f) \left(\frac{1}{{}^1Z} - \frac{1}{{}^2Z} \right)$$

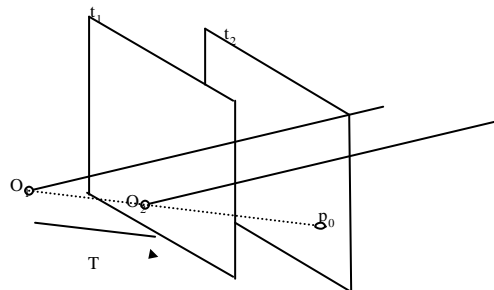
$${}^1v_x - {}^2v_x = (x - x_0) T_z \left(\frac{1}{{}^1Z} - \frac{1}{{}^2Z} \right)$$

$${}^1v_y - {}^2v_y = (y - y_0) T_z \left(\frac{1}{{}^1Z} - \frac{1}{{}^2Z} \right)$$

$$\begin{cases} \bar{p}_0 = [x_0, y_0]^T \\ x_0 = f \frac{T_x}{T_z} & y_0 = f \frac{T_y}{T_z} \end{cases}$$

⇐ **Epipolo instantáneo**

- ↗ Es posible localizar p_0 sin un conocimiento a priori de los parámetros intrínsecos de la cámara.
- Si movemos la cámara de O_1 a O_2 con una traslación pura, p_0 será aquel punto de la imagen cuyo campo de movimiento sea nulo.



- ↳ **Introducción**
- ↳ **El campo de movimiento (“Motion-Field”)**
- ☒ **Flujo óptico**
- ↳ **Estimación del campo de movimiento**
- ↳ **Reconstrucción 3D**

- ↳ **El análisis del movimiento a partir de una secuencia de imágenes se encamina hacia la estimación del movimiento relativo entre los objetos en la escena y las imágenes.**
- ↳ **Flujo óptico:**
 - ↳ **Cambios en los niveles de intensidad en la imagen debidos al movimiento en el intervalo dt**
- ↳ **Campo de flujo óptico:**
 - ↳ **Campo de velocidad que representa el movimiento tridimensional de puntos de los objetos a través del movimiento tridimensional de la imagen**

⇨ **Se quiere estimar el campo de movimiento en secuencias de imágenes, a partir de las variaciones espaciales y temporales de Intensidad de la imagen**

⇨ **Ecuación de continuidad de Intensidad de la imagen**

↖ **E (Intensidad aparente de la imagen) permanece constante**
 $dE/dt = 0$

$$E(x(t), y(t), t) \Rightarrow \frac{dE(x(t), y(t), t)}{dt} = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0$$

$$(\nabla E)^T \bar{v} + E_t = 0 \quad \text{donde } E_t = \frac{dE}{dt}$$

↖ **Problema de apertura** : “El componente del campo de movimiento en la dirección ortogonal al gradiente espacial de la imagen no está restringido por la ecuación de continuidad de intensidad de la imagen”

⇨ **Flujo Óptico** es un vector sujeto a la restricción anterior (ecuación continuidad de intensidad), y se define como el movimiento aparente del patrón de intensidad de la imagen.

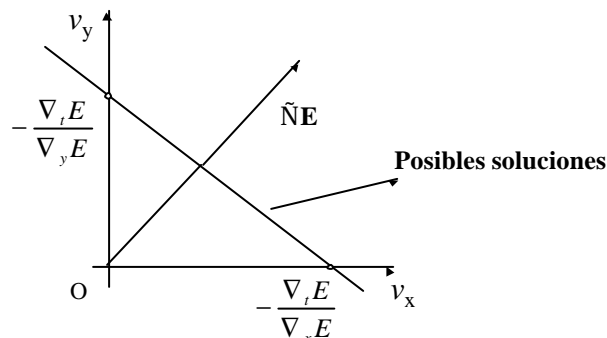
⇨ **Formalmente expresa que un punto de un objeto en movimiento se proyectará en puntos diferentes de la imagen en instantes diferentes, pero la intensidad irradiada es la misma en ambos puntos proyectados.**

$$\nabla_x E \cdot v_x + \nabla_y E \cdot v_y + \nabla_t E = 0$$

⇐ Solución de la ecuación de Continuidad de Intensidad:

$$\nabla_x E \cdot v_x + \nabla_y E \cdot v_y + \nabla_t E = 0$$

- ↗ Recta en el espacio de velocidades
- ↗ Solución ambigua debido al problema de apertura
- ↗ Es necesario imponer restricciones

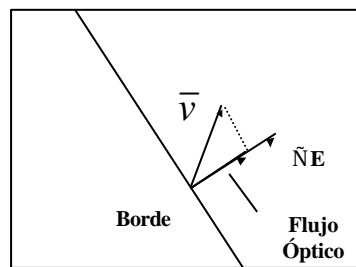


⇐ Flujo Óptico y Campo de movimiento

- ↗ El flujo óptico es una buena aproximación al campo de movimiento
- ↗ Se puede calcular a partir de las secuencias de imágenes.
- ↗ La aproximación es buena con las siguientes suposiciones :
 - Superficies Lambertianas (tiene el mismo brillo en todas las direcciones)
 - Fuente de luz puntual en el infinito
 - Sin distorsión fotométrica
- ↗ El error de esta aproximación es :
 - pequeño en los puntos con un gradiente espacial alto
 - cero solo para el movimiento traslacional o para cualquier movimiento rígido tal que la dirección de iluminación sea paralela a la velocidad angular

⇐ El flujo óptico corresponde (aproximadamente) a la proyección del campo de movimiento sobre el vector gradiente de la imagen.

$$\nabla_x E \cdot v_x + \nabla_y E \cdot v_y + \nabla_t E = 0$$



- ⇐ Introducción
- ⇐ El campo de movimiento (“Motion-Field”)
- ⇐ Flujo óptico
- ☒ Estimación del campo de movimiento
 - ↗ Técnicas diferenciales
 - ↗ Técnicas de correspondencia
 - Método de dos imágenes consecutivas
 - Método de múltiples imágenes consecutivas (Filtro de Kalman)
- ⇐ Reconstrucción 3D

↳ 2 Técnicas :

↳ **Técnicas diferenciales** : basadas en las variaciones temporales y espaciales de la Intensidad de la imagen en todos sus puntos ↳
Flujo Óptico

↳ **Técnicas de correspondencia** : determinan la disparidad entre características de secuencias de imágenes. Pueden utilizar dos o más imágenes (filtro de Kalman). Utilizan el flujo óptico como estimación inicial de la disparidad.

↳ Técnicas diferenciales (1ª opción):

↳ Estimar el flujo óptico por mínimos cuadrados (restricción de uniformidad local)

↳ Razones :

- No iterativos, menos propenso que los iterativos a posibles discontinuidades
- No utilizan derivadas de orden superior a uno, menos sensibles al ruido

↳ Suposiciones

- La ecuación de continuidad de intensidad de la imagen es una buena aproximación de la componente normal del campo de movimiento.
- El campo de movimiento se aproximará por un campo de vectores constantes dentro de una pequeña región Q del campo de la imagen.

↖ **Algoritmo para el calculo del flujo óptico (localmente constante)**

→ Región típica 5x5(Q), minimiza la función:

$$\min_{\bar{v}} (\psi(\bar{v})); \quad \psi(\bar{v}) = \sum_{p_i \in Q} [(\nabla E)^T \bar{v} + E_i]^2$$

$$A_{N^2 \times 2} = \begin{bmatrix} \nabla E(\bar{p}_1) \\ \nabla E(\bar{p}_2) \\ \cdot \\ \cdot \\ \nabla E(\bar{p}_{N \times N}) \end{bmatrix} \quad A \bar{v} = b$$

$$b_{N^2} = -[E_i(\bar{p}_1) \quad \cdot \quad \cdot \quad \cdot \quad E_i(\bar{p}_{N \times N})]^T$$

N: tamaño de la ventana de integración

→ Por mínimos cuadrados

$$\text{Sol: } \bar{v} = (A^T A)^{-1} A^T b$$

$$\bar{v} = \text{flujo optico en el centro de la región } Q$$

→ Se evalúa para todos los puntos de la imagen

↖ **Mejora del algoritmo:**

→ Se comete un error al considerar que el flujo óptico en la región es constante e igual al del punto central, el error se incrementa con la distancia de p al punto central. Introducimos una matriz de pesos $W(N^2 \times N^2)$ simétrica

$$\bar{v}_w = (A^T W^2 A)^{-1} A^T W^2 b$$

↖ **Puntos de la imagen donde falla el algoritmo**

→ Cuando la matriz (2x2) $A^T A$ es singular en la región Q

→ Gradientes espaciales en Q nulos o paralelos.

→ En este caso el problema de apertura no tiene solución.

$$A^T A = \begin{pmatrix} \sum \nabla_x E^2 & \sum \nabla_x E \nabla_y E \\ \sum \nabla_x E \nabla_y E & \sum \nabla_y E^2 \end{pmatrix}$$

↳ Técnicas diferenciales (2ª opción):

↳ **Método Iterativo:** en lugar de suponer que el flujo óptico es constante en cada región, impone una restricción de suavidad que expresa la continuidad de las superficies de un objeto rígido.

↳ **Características :**

→ Iterativo: problemas de convergencia, problemas de discontinuidades

↳ **Suposiciones**

→ La ecuación de continuidad de intensidad de la imagen es una buena aproximación de la componente normal del campo de movimiento.

→ El flujo óptico varía suavemente en las superficies de un sólido rígido.

↳ Cálculo del flujo óptico: restricción de suavidad del FO.

↳ La restricción de suavidad del flujo óptico expresa la continuidad de las superficies de un objeto rígido.

$$\min_{\bar{v}} (\Gamma_c(\bar{v})); \quad \Gamma_c(\bar{v}) = \Gamma_S(\bar{v}) + \lambda \Gamma_H(\bar{v})$$

$$\bar{v} = [v_x(x, y, t), v_y(x, y, t)]^T$$

Restricción Suavidad

Restricción Continuidad de Intensidad

$$\Gamma_S(\bar{v}) = \iint_{\Omega} \left[\|\nabla_x \bar{v}\|^2 + \|\nabla_y \bar{v}\|^2 \right] dx dy$$

$$\nabla_x \bar{v} = \left[\frac{\partial v_x}{\partial x}, \frac{\partial v_y}{\partial x} \right]^T$$

$$\Gamma_H(\bar{v}) = \iint_{\Omega} \left[(\nabla E)^T \bar{v} + E_t \right]^2 dx dy$$

$$\nabla_y \bar{v} = \left[\frac{\partial v_x}{\partial y}, \frac{\partial v_y}{\partial y} \right]^T$$

Ω: entorno de vecindad

↳ **Solución para imágenes discretas:**

$$\Gamma_S(i, j) = \frac{1}{4} \left[(v_{x_{i+1,j}} - v_{x_{i,j}})^2 + (v_{x_{i,j+1}} - v_{x_{i,j}})^2 + (v_{y_{i+1,j}} - v_{y_{i,j}})^2 + (v_{y_{i,j+1}} - v_{y_{i,j}})^2 \right]$$

$$\Gamma_H(i, j) = \left[\nabla_x E(i, j, t) \cdot v_{x_{i,j}} + \nabla_y E(i, j, t) \cdot v_{y_{i,j}} + E_t(i, j, t) \right]^2$$

$$\Gamma_C(i, j) = \sum_{\Omega} (\Gamma_S(i, j) + \lambda \Gamma_H(i, j))$$

Minimizar \Downarrow

$$\frac{\partial \Gamma_C}{\partial v_{x_{i,j}}}(i, j) = 0 ; \quad \frac{\partial \Gamma_C}{\partial v_{y_{i,j}}}(i, j) = 0$$

↳ **Solución:**

$$\Gamma_C(i, j) = \sum_{\Omega} (\Gamma_S(i, j) + \lambda \Gamma_H(i, j))$$

Minimizar \Downarrow $\frac{\partial \Gamma_C}{\partial v_{x_{i,j}}}(i, j) = 0 ; \quad \frac{\partial \Gamma_C}{\partial v_{y_{i,j}}}(i, j) = 0$

$$\frac{\partial \Gamma_S}{\partial v_{x_{i,j}}}(i, j) = 2 (v_{x_{i,j}} - \hat{v}_{x_{i,j}})$$

$$\frac{\partial \Gamma_S}{\partial v_{y_{i,j}}}(i, j) = 2 (v_{y_{i,j}} - \hat{v}_{y_{i,j}})$$

\Rightarrow Minimiza las diferencias respecto al valor promedio

$$\frac{\partial \Gamma_C}{\partial v_{x_{i,j}}}(i, j) = 2(v_{x_{i,j}} - \hat{v}_{x_{i,j}}) + 2\lambda \left[\nabla_x E(i, j, t) \cdot v_{x_{i,j}} + \nabla_y E(i, j, t) \cdot v_{y_{i,j}} + E_t(i, j, t) \right] \nabla_x E(i, j, t) = 0$$

$$\frac{\partial \Gamma_C}{\partial v_{y_{i,j}}}(i, j) = 2(v_{y_{i,j}} - \hat{v}_{y_{i,j}}) + 2\lambda \left[\nabla_x E(i, j, t) \cdot v_{x_{i,j}} + \nabla_y E(i, j, t) \cdot v_{y_{i,j}} + E_t(i, j, t) \right] \nabla_y E(i, j, t) = 0$$

⇐ **Solución iterativa:**

$$v_{x_{i,j}} = \frac{1}{1 + \lambda (\nabla_x E^2 + \nabla_y E^2)} \cdot \left[(1 + \lambda \nabla_y E^2) \cdot \hat{v}_{x_{i,j}} - \lambda \nabla_x E (\nabla_t E + \nabla_y E \cdot \hat{v}_{y_{i,j}}) \right]$$

$$v_{y_{i,j}} = \frac{1}{1 + \lambda (\nabla_x E^2 + \nabla_y E^2)} \cdot \left[(1 + \lambda \nabla_x E^2) \cdot \hat{v}_{y_{i,j}} - \lambda \nabla_y E (\nabla_t E + \nabla_x E \cdot \hat{v}_{x_{i,j}}) \right]$$

↖ **Los valores del flujo óptico \bar{v} en la posición (i,j) dependen de los valores de \bar{v} en su vecindad**

↖ **Partimos de unos valores iniciales:**

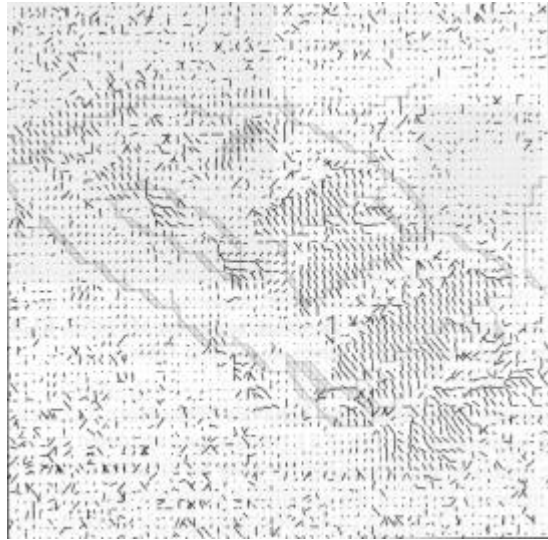
$$\begin{array}{l}
 \text{Iteración} \\
 v_{x_{i,j}}^0 = 0 \quad \Rightarrow \quad v_{x_{i,j}}^{n+1} = v_{x_{i,j}}^n - \frac{\nabla_x E \cdot \hat{v}_{x_{i,j}}^n + \nabla_y E \cdot \hat{v}_{y_{i,j}}^n + \nabla_t E}{1 + \lambda (\nabla_x E^2 + \nabla_y E^2)} \lambda \nabla_x E \\
 v_{y_{i,j}}^0 = 0 \quad \Rightarrow \quad v_{y_{i,j}}^{n+1} = v_{y_{i,j}}^n - \frac{\nabla_x E \cdot \hat{v}_{x_{i,j}}^n + \nabla_y E \cdot \hat{v}_{y_{i,j}}^n + \nabla_t E}{1 + \lambda (\nabla_x E^2 + \nabla_y E^2)} \lambda \nabla_y E
 \end{array}$$

⇐ **Cálculo de los gradientes de intensidad:**

$$\begin{aligned}
 \nabla_x E(i, j, t) &= \frac{1}{4} (E(i+1, j, t) + E(i+1, j, t+1) + E(i+1, j+1, t) + E(i+1, j+1, t+1)) \\
 &\quad - \frac{1}{4} (E(i, j, t) + E(i, j, t+1) + E(i, j+1, t) + E(i, j+1, t+1))
 \end{aligned}$$

$$\begin{aligned}
 \nabla_y E(i, j, t) &= \frac{1}{4} (E(i, j+1, t) + E(i, j+1, t+1) + E(i+1, j+1, t) + E(i+1, j+1, t+1)) \\
 &\quad - \frac{1}{4} (E(i, j, t) + E(i, j, t+1) + E(i+1, j, t) + E(i+1, j, t+1))
 \end{aligned}$$

$$\begin{aligned}
 \nabla_t E(i, j, t) &= \frac{1}{4} (E(i, j, t+1) + E(i, j+1, t+1) + E(i+1, j, t+1) + E(i+1, j+1, t+1)) \\
 &\quad - \frac{1}{4} (E(i, j, t) + E(i, j+1, t) + E(i+1, j+1, t) + E(i+1, j+1, t))
 \end{aligned}$$



- ⇐ **Introducción**
- ⇐ **El campo de movimiento (“Motion-Field”)**
- ⇐ **Flujo óptico**
- ⇐ **Estimación del campo de movimiento**
 - ▣ **Técnicas diferenciales**
 - ☒ **Técnicas de correspondencia**
 - Método de dos imágenes consecutivas
 - Método de múltiples imágenes consecutivas (Filtro de Kalman)
- ⇐ **Reconstrucción 3D**

↳ Técnicas de Correspondencia (Basadas en características)

- ↳ Estiman el campo de movimiento solamente en puntos característicos.
- ↳ Se realiza una correspondencia de estos puntos entre las sucesivas imágenes.
- ↳ La correspondencia es guiada mediante una estimación del flujo óptico.

- ↳ Dos métodos :
 - Método de dos imágenes consecutivas
 - Método de múltiples imágenes consecutivas

↳ Método de dos imágenes : "Feature Matching"

- ↳ Emparejar puntos característicos entre las imágenes de la secuencia
- ↳ Calculamos $A^T A$ sobre pequeñas regiones de la imagen: las características son los centros de aquellas regiones para las cuales el más pequeño de los valores propios de $A^T A$ es mayor que un umbral.
- ↳ Calcular el desplazamiento de los puntos característicos mediante la iteración del algoritmo visto anteriormente ⇒ Estimando con ello el flujo óptico

$$A^T A = \begin{pmatrix} \sum \nabla_x E^2 & \sum \nabla_x E \nabla_y E \\ \sum \nabla_x E \nabla_y E & \sum \nabla_y E^2 \end{pmatrix} \quad b_{N^2} = - \left[E_t(\bar{p}_1) \quad \dots \quad E_t(\bar{p}_{N \times N}) \right]^T$$

$$\bar{v} = (A^T A)^{-1} A^T b$$

$$\bar{d} = \bar{v} \cdot \delta t$$

Algoritmo Feature Point Matching

47

- ⇨ **Entrada esta formada :**
 - ↻ Por dos imágenes consecutivas I_1, I_2
 - ↻ Un conjunto de puntos correspondientes en las dos imágenes
- ⇨ **Sean :**
 - ↻ Q_1, Q_2, Q' tres regiones de la imagen (NxN)
 - ↻ "T" un número real positivo fijado (umbral). "d" desplazamiento desconocido entre I_1 e I_2 de un punto característico p, centrado en Q_1
- ⇨ **Para todos los puntos p :**
 - 1.- Poner $d = 0$ y centrar Q_1 en p
 - 2.- Estimar el desplazamiento $d_0 = \bar{v} \cdot dt$ de p, centrado en Q_1 , mediante el flujo óptico: $\bar{v} = (A^T A)^{-1} A^T b$ y $d = d + d_0$
 - 3.- Q' es la región obtenida de desplazar Q_1 un d_0 . Calcular S, la suma de las diferencias cuadradas entre Q' y la correspondiente región Q_2 en I_2
 - 4.- Si $S > T$, $Q_1 = Q'$ y volver al paso 1 y continuar con el resto de características; sino terminar

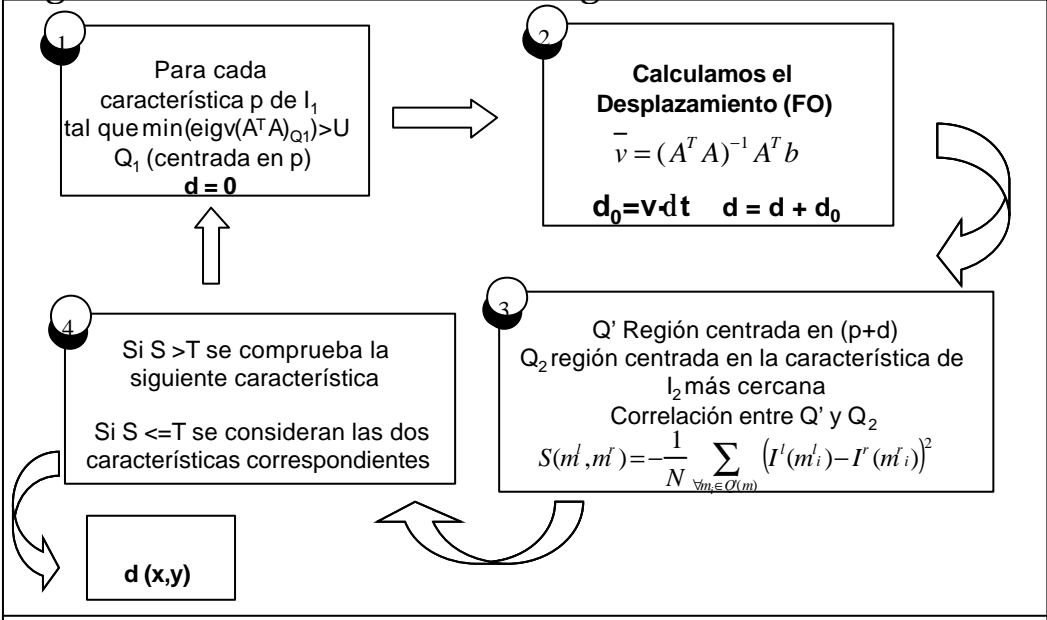
La salida es una estimación de d para todos los puntos característicos

Visión Tridimensional

UPM-DISAM / UMH

Algoritmo Feature Point Matching

48



Visión Tridimensional

UPM-DISAM / UMH

↳ **Métodos de varias imágenes (“Multiple-Frame”)**

- ↳ “Si el movimiento de la escena observada es continuo, como sucede en la mayoría de los casos, se debería poder hacer predicciones sobre el movimiento de los puntos de la imagen, en base a sus trayectorias previas”
- ↳ Se utilizan las disparidades calculadas entre I_{i-1} y I_{i-2} , I_{i-2} , I_{i-3} , ..etc para predecir las disparidades entre I_{i-1} y I_i antes de observar I_i
- ↳ Están basados en del *filtro de Kalman*.

↳ **Filtro de Kalman**

- ↳ Punto característico $p_k = [x_k, y_k]^T$ en la Imagen en t_k , con velocidad $v_k = [v_{x,k}, v_{y,k}]^T \Rightarrow$ vector de estado $x = [x_k, y_k, v_{x,k}, v_{y,k}]$
- ↳ Modelo del sistema del filtro de Kalman:

$$\left. \begin{aligned} p_k &= p_{k-1} + v_{k-1} + \xi_{k-1} \\ v_k &= v_{k-1} + \eta_{k-1} \end{aligned} \right\} \Rightarrow \text{Modelo del Sistema} \quad x_k = \Phi_{k-1} x_{k-1} + w_{k-1}$$

$$\begin{matrix} \xi_{k-1} & \eta_{k-1} & x_k = \begin{bmatrix} p_k \\ v_k \end{bmatrix} & \Phi_{k-1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & w_{k-1} = \begin{bmatrix} \xi_{k-1} \\ \eta_{k-1} \end{bmatrix} \end{matrix}$$

Ruido del sistema

$$z_k = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \mu_k \quad \text{Ecuación de Medida} \quad z_k = H_k x_k + \mu_k$$

⇨ Algoritmo Kalman_Tracking

↻ Las entradas son, en t_k , las matrices de covarianza del ruido del sistema y de la medida en t_{k-1} , Q_{k-1} y R_{k-1} respectivamente, la matriz de estado invariante en el tiempo, f , matriz de medidas invariante en el tiempo, H , y la medida de la posición en el tiempo t_k , es z_k . Las entradas de P_0 son inicializadas con valores arbitrarios.

$$P'_k = \phi_{k-1} P_{k-1} \phi_{k-1}^T + Q_{k-1}$$

$$K_k = P'_k H_k^T (H_k P'_k H_k^T + R_k)^{-1}$$

$$\hat{x}_k = \phi_{k-1} \hat{x}_{k-1} + K_k (z_k - H_k \phi_{k-1} \hat{x}_{k-1})$$

$$P_k = (I - K_k) P'_k (I - K_k)^T + K_k R_k K_k^T$$

↻ La salida es la estimación óptima de la posición y la velocidad en el tiempo t_k , \hat{x}_k y sus incertidumbres, dado por los elementos de la diagonal de P_k

⇨ Conclusiones :

- ↻ Integra las medidas del ruido con las predicciones del modelo para estimar el estado óptimo
- ↻ El filtro cuantifica la incertidumbre del estado estimado, como los elementos de la diagonal de la matriz de covarianza del estado. Esto permite al detector de características dimensionar automáticamente la región de la imagen a ser buscada para encontrar el punto característico en la próxima imagen. La región está centrada en la mejor posición estimada y tan grande como la incertidumbre.



⇐ Dos problemas en la implementación de este algoritmo :

↖ Información perdida

El filtro de Kalman está basado en el conocimiento de los siguientes datos :

- El modelo del sistema y su correspondiente matriz de covarianza Q_k
- El modelo medido y la matriz de covarianza del ruido correspondiente, R_k
- El estado del sistema inicial en t_0 , \hat{x}_0 , y la matriz de covarianza del estado, P_0

Sin embargo muchas de estas cantidades son desconocidas.

↖ Asociación de datos

- En presencia de algunas características de la imagen y múltiples medidas, ¿Qué características deberían ser asociadas las medidas observadas?

- ⇐ Introducción
- ⇐ El campo de movimiento (“Motion-Field”)
- ⇐ Flujo óptico
- ⇐ Estimación del campo de movimiento
- ☒ Reconstrucción 3D

- ↳ Suponiendo conocido el campo de movimiento (mapas de velocidades proyectados sobre la imagen) ¿qué podemos conocer acerca del movimiento 3D y localización de los objetos de la escena?
- ↳ Conocemos los desplazamientos de cada punto (o de ciertas características) de la imagen entre dos imágenes sucesivas.
- ↳ Sin ninguna restricción adicional el problema de reconstrucción no tiene solución única, por lo tanto debemos estudiar situaciones particulares:
 - ↳ Escena estática y cámara móvil rotando en torno a su centro de proyección (rotación desconocida)
 - ↳ Escena estática y cámara móvil con movimiento desconocido
 - ↳ Cámara fija calibrada y objeto móvil con movimiento conocido

- ↳ Escena estática y cámara móvil rotando en torno a su centro de proyección (rotación desconocida)
 - ↳ Dado un conjunto de puntos correspondientes entre dos imágenes I_1 e I_2 tomadas en instantes t_1 y t_2
 - ↳ Podemos calcular la rotación efectuada por la cámara si disponemos de al menos 3 puntos correspondientes
 - ↳ No podemos inferir nada acerca de la posición 3D de los puntos proyectados y por tanto de la superficie de los objetos

$$\left. \begin{aligned} x^1 &= f \frac{X^{C1}}{Z^{C1}} \\ y^1 &= f \frac{Y^{C1}}{Z^{C1}} \end{aligned} \right\} \Rightarrow \left. \begin{aligned} x^2 &= f \frac{r_{11}X^{C1} + r_{12}Y^{C1} + r_{13}Z^{C1}}{r_{31}X^{C1} + r_{32}Y^{C1} + r_{33}Z^{C1}} \\ y^2 &= f \frac{r_{21}X^{C1} + r_{22}Y^{C1} + r_{23}Z^{C1}}{r_{31}X^{C1} + r_{32}Y^{C1} + r_{33}Z^{C1}} \end{aligned} \right\}$$

$$\left. \begin{aligned} x^2 &= f \frac{r_{11}x^1 + r_{12}y^1 + r_{13}f}{r_{31}x^1 + r_{32}y^1 + r_{33}f} \\ y^2 &= f \frac{r_{21}x^1 + r_{22}y^1 + r_{23}f}{r_{31}x^1 + r_{32}y^1 + r_{33}f} \end{aligned} \right\}$$

$$[R], f \quad \leftarrow \quad \begin{aligned} \det |R| &= 1 \\ R^T &= R^{-1} \end{aligned}$$

Reconstrucción 3D (3)

57

⇐ **Escena estática y cámara móvil con movimiento desconocido**

↗ **Dado un conjunto de puntos correspondientes entre dos imágenes I_1 e I_2 tomadas en instantes t_1 y t_2**

↗ **Podemos calcular la rotación y traslación efectuada y la posición de los puntos 3D si disponemos de al menos 11 puntos correspondientes**

$$\left. \begin{aligned} x^2 &= f \frac{r_{11}X^{C1} + r_{12}Y^{C1} + r_{13}Z^{C1} + t_1}{r_{31}X^{C1} + r_{32}Y^{C1} + r_{33}Z^{C1} + t_3} \\ y^2 &= f \frac{r_{21}X^{C1} + r_{22}Y^{C1} + r_{23}Z^{C1} + t_2}{r_{31}X^{C1} + r_{32}Y^{C1} + r_{33}Z^{C1} + t_3} \end{aligned} \right\} \Rightarrow \left. \begin{aligned} x^2 &= f \frac{r_{11}x^1 + r_{12}y^1 + r_{13}f + \frac{t_1 f}{Z^{C1}}}{r_{31}x^1 + r_{32}y^1 + r_{33}f + \frac{t_3 f}{Z^{C1}}} \\ y^2 &= f \frac{r_{21}x^1 + r_{22}y^1 + r_{23}f + \frac{t_2 f}{Z^{C1}}}{r_{31}x^1 + r_{32}y^1 + r_{33}f + \frac{t_3 f}{Z^{C1}}} \end{aligned} \right\}$$

$$\left. \begin{aligned} x^1 &= f \frac{X^{C1}}{Z^{C1}} \\ y^1 &= f \frac{Y^{C1}}{Z^{C1}} \end{aligned} \right\} \Downarrow$$

$$\left. \begin{aligned} \phi_i([R], f, t, Z_i, \langle x_i^1, y_i^1, x_i^2 \rangle) &= 0 \\ \Psi_i([R], f, t, Z_i, \langle x_i^1, y_i^1, y_i^2 \rangle) &= 0 \end{aligned} \right\}$$

Visión Tridimensional

UPM-DISAM / UMH

Reconstrucción 3D (4)

58

⇐ **Cámara fija calibrada y objeto móvil con movimiento conocido**

$\overline{C}^t = [X^t, Y^t, Z^t]^T$ Coordenadas del punto 3D referidas al S.C. de cámara en el instante t

$\overline{M}^t = [X_w^t, Y_w^t, Z_w^t]^T$ Coordenadas del punto 3D referidas al S.C. del mundo en el instante t

$\overline{p}^t = [x^t, y^t, f]^T$ Coordenadas del punto proyectado sobre la imagen referidas al S.C. de cámara (3D) en el instante t

$$\overline{p}^t = f \frac{\overline{C}^t}{Z^t}$$

R_δ, T_δ Rotación y traslación entre los instantes t y t+1

$$\overline{M}^{t+1} = R_\delta \cdot \overline{M}^t + T_\delta$$

R, T Rotación y traslación entre S.C. Del mundo y el S.C. de cámara

$$\left. \begin{aligned} \overline{C}^t &= R \cdot \overline{M}^t + T \\ \overline{C}^{t+1} &= R \cdot \overline{M}^{t+1} + T \end{aligned} \right\}$$

Visión Tridimensional

UPM-DISAM / UMH

⇐ **Despejando M^{t+1} y M^t en las ecuaciones anteriores tenemos:**

$$(R \ R_\delta)^T \bar{p}^{t+1} Z^{t+1} - \frac{R^T \bar{p}^t}{f} Z^t = f [(R \ R_\delta)^T T - R^T T + R_\delta T_\delta]$$

$$\bar{a} Z^{t+1} - \bar{d} Z^t = \bar{b}$$

$$\begin{bmatrix} c_1 & -d_1 \\ c_2 & -d_2 \\ c_3 & -d_3 \end{bmatrix} \begin{bmatrix} Z^{t+1} \\ Z^t \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \Leftrightarrow AZ = b$$



Coordenadas del punto 3D referidas al S.C. del mundo en el instante t

$$\bar{Z} = \begin{bmatrix} Z^{t+1} \\ Z^t \end{bmatrix} = (A^T A)^{-1} A^T b \Leftrightarrow \bar{M}^t = R^T \left(\frac{Z^t}{f} \bar{p}^t - T \right)$$