Visión por Computador

Luis Baumela

http://www.dia.fi.upm.es/~lbaumela/Alcala

Departamento de Inteligencia Artificial

Universidad Politécnica de Madrid

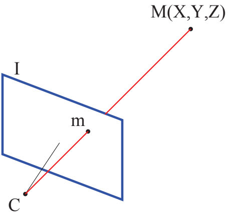
Visión Estéreo

Visión Estéreo

Índice

Planteamiento del problema

Obtener información sobre relaciones tridimensionales en una escena (p.ej. reconstrucciones) a partir de imágenes.



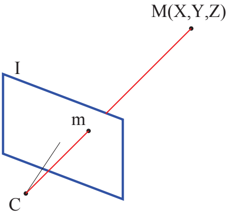
Modelo de formación imagen:

$$\bar{m} = \mathbf{K}\,[\mathbf{R}\,|\,\bar{t}]\begin{bmatrix}\tilde{M}\\ 1\end{bmatrix}$$

$$\begin{bmatrix}\lambda j\\ \lambda i\\ \lambda\end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14}\\ p_{21} & p_{22} & p_{23} & p_{24}\\ p_{31} & p_{32} & p_{33} & p_{34}\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}$$

Planteamiento del problema

Obtener información sobre relaciones tridimensionales en una escena (p.ej. reconstrucciones) a partir de imágenes.



Reconstrucción:

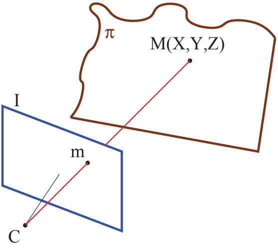
$$\tilde{M} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda[\mathbf{KR}]^{-1}\bar{m} - \mathbf{R}^{\top}\bar{t}$$

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0,$$

donde $a_{rs} = p_{rs} - p_{3s} j^{(r \bmod 2)} i^{(r+1 \bmod 2)}$

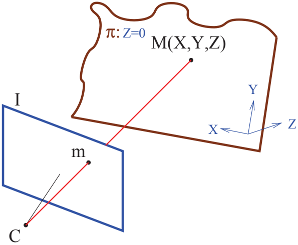
Planteamiento del problema

- Escenas planas: Homografía del plano.



Planteamiento del problema

- Escenas planas: Homografía del plano.

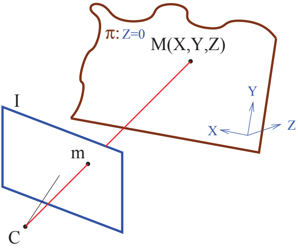


Escena plana:

$$\begin{bmatrix} \lambda j \\ \lambda i \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

Planteamiento del problema

- Escenas planas: Homografía del plano.



Escena plana:

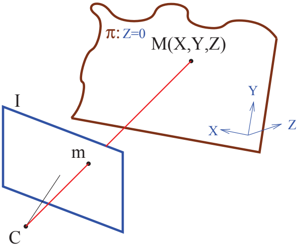
$$\begin{bmatrix} \lambda j \\ \lambda i \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

Eliminando $\bar{p}_3$ y Z

$$\bar{m} = \mathbf{H}_{3\times 3} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Planteamiento del problema

- Escenas planas: Homografía del plano.



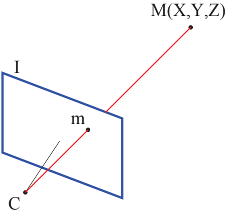
Escena plana:

$$\begin{bmatrix} \lambda j \\ \lambda i \\ \lambda \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \mathbf{H}^{-1}\bar{m}$$

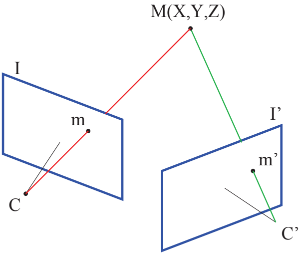
Planteamiento del problema

- Visión estéreo. Estructura a partir de X.



Planteamiento del problema

- Visión estéreo. Estructura a partir de X.



Reconstruyendo de I:

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$

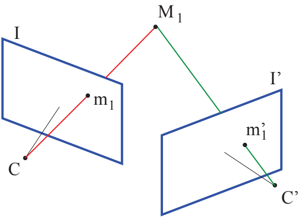
Reconstruyendo de I':

$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0,$$

donde $a_{rs} = p_{rs} - p_{3s} j^{(r \bmod 2)} i^{(r+1 \bmod 2)}$

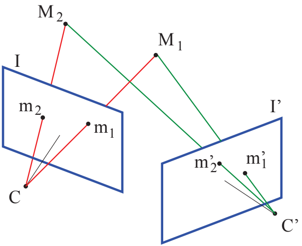
Planteamiento del problema

- Visión estéreo. La correspondencia



Planteamiento del problema

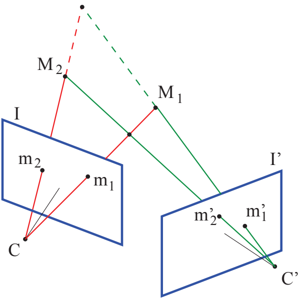
- Visión estéreo. La correspondencia



Parejas de puntos de ambas imágenes que se corresponden con un mismo punto de la escena.

Planteamiento del problema

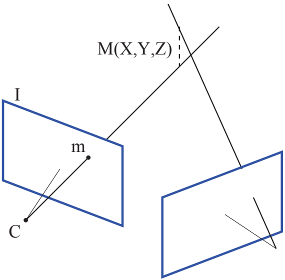
- Visión estéreo. La correspondencia



Parejas de puntos de ambas imágenes que se corresponden con un mismo punto de la escena.

La reconstrucción

En general, las visuales se cruzan.

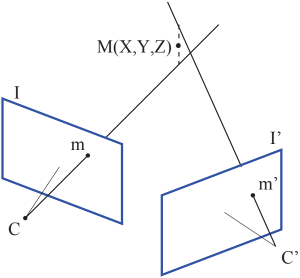


- Minimización algebráica.

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

La reconstrucción

En general, las visuales se cruzan.



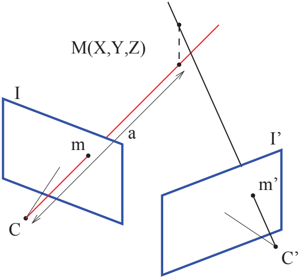
- Minimización algebráica.

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

- Solución geométrica.

La reconstrucción

En general, las visuales se cruzan.



- Minimización algebráica.

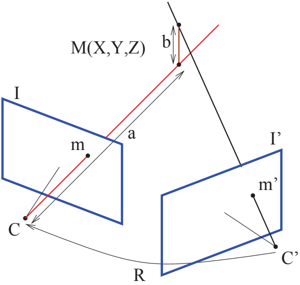
$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

- Solución geométrica.

$$a(\mathbf{K}^{-1}\bar{m})$$

La reconstrucción

En general, las visuales se cruzan.



- Minimización algebráica.

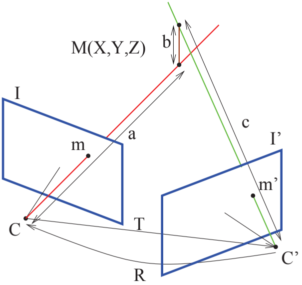
$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

- Solución geométrica.

$$a(\mathbf{K}^{-1}\bar{m}) + b\bar{p}$$

La reconstrucción

En general, las visuales se cruzan.



- Minimización algebráica.

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

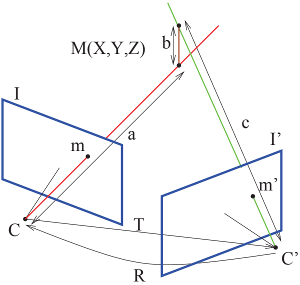
- Solución geométrica.

$$a(\mathbf{K}^{-1}\bar{m}) + b\bar{p} = c\mathbf{R}\mathbf{K}'^{-1}\bar{m}' + \bar{T},$$

donde $\bar{p} = (\mathbf{K}^{-1}\bar{m} \times \mathbf{R}\mathbf{K}'^{-1}\bar{m}')$

La reconstrucción

En general, las visuales se cruzan.



- Minimización algebráica.

$$a_{11}X + a_{12}Y + a_{13}Z + a_{14} = 0$$
$$a_{21}X + a_{22}Y + a_{23}Z + a_{24} = 0$$
$$a'_{11}X + a'_{12}Y + a'_{13}Z + a'_{14} = 0$$
$$a'_{21}X + a'_{22}Y + a'_{23}Z + a'_{24} = 0$$

- Solución geométrica.

$$\bar{M} = a(\mathbf{K}^{-1}\bar{m}) + \frac{1}{2}b(\mathbf{K}^{-1}\bar{m} \times \mathbf{RK}^{-1}\bar{m}')$$

válida sólo para dos cámaras.

La reconstrucción

Algoritmo (c-cámaras):

1. **Solución lineal.** Reconstruye linealmente cada punto de la escena, mediante alguno de los algoritmos anteriores.
2. **Minimiza el error de reproyección**. Sea $d(\cdot,\cdot)$ una distancia, $\tilde{x}_i$ la posición, en cartesianas, de un punto en la imagen i y $\tilde{\mathbf{P}}_i\bar{M}$ su reproyección, en cartesianas,

$$\min_{\tilde{M}} \sum_{i=1}^{c} d(\tilde{x}_i, \tilde{\mathbf{P}}_i \begin{bmatrix} \tilde{M} \\ 1 \end{bmatrix}),$$

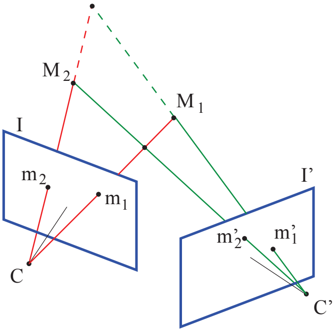
mediante un algoritmo iterativo (ej. Levenberg-Marquardt).

Nota:

1. Para mejorar el condicionamiento numérico de la solución lineal, es conveniente normalizar los datos, al igual que en el algoritmo de calibración.

La búsqueda de correspondencias

La puesta en correspondencia de imágenes es un proceso ambiguo: dado un punto en una imagen, existen multitud de puntos en las otras que pueden estar en correspondencia con él.



La búsqueda de correspondencias

La puesta en correspondencia de imágenes es un proceso ambiguo: dado un punto en una imagen, existen multitud de puntos en las otras que pueden estar en correspondencia con él.

Para su resolución hemos de abordar las cuestiones:

- ¿Qué elementos pondré en correspondencia?
  - Píxeles.
  - Bordes.
  - Píxeles + bordes.

La búsqueda de correspondencias

La puesta en correspondencia de imágenes es un proceso ambiguo: dado un punto en una imagen, existen multitud de puntos en las otras que pueden estar en correspondencia con él.

Para su resolución hemos de abordar las cuestiones:

- ¿Qué elementos pondré en correspondencia?
- ¿Qué restricciones reducen la ambigüedad del proceso?
  - Derivadas de la geometría de las cámaras.
  - Derivadas de la geometría de la escema.

La búsqueda de correspondencias

La puesta en correspondencia de imágenes es un proceso ambiguo: dado un punto en una imagen, existen multitud de puntos en las otras que pueden estar en correspondencia con él.

Para su resolución hemos de abordar las cuestiones:

- ¿Qué elementos pondré en correspondencia?
- ¿Qué restricciones reducen la ambigüedad del proceso?
- Algoritmos de puesta en correspondencia.
  - Cooperativos.
  - Programación dinámica.
  - Multivista.

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

Los más comúnmente empleados:

- *Píxel.*

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

Los más comúnmente empleados:

- *Píxel.* Características:
  - Nivel de gris.
  - Niveles de gris de una vecindad (ventana).
  - Conjunto descriptores invariantes.
- *Borde.*

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

Los más comúnmente empleados:

- *Píxel.*
- *Borde.* Características:
  - Magnitud.
  - Orientación.

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

Los más comúnmente empleados:

- *Píxel.*
- *Borde.* Segmentos rectilíneos. Características:
  - Longitud.
  - Orientación.
  - Nivel gris medio.

lementos para la puesta en correspondenc

Los elementos a poner en correspondencia han de ser:

- Fácilmente extraíbles.
- Informativos
- Invariantes a cambios de iluminación y de punto de vista.

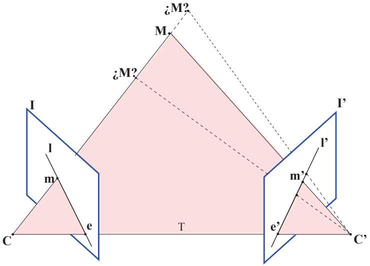
Los más comúnmente empleados:

- *Píxel.*
- *Borde.*
- *Píxel + borde.*

Restricción basada en geometría cámaras

- La restricción epipolar (I). Planteamiento.

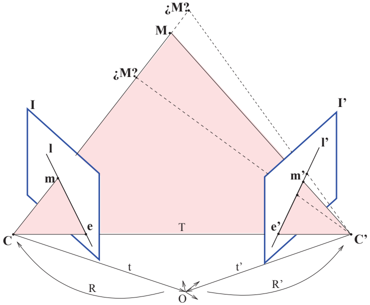
Para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo.



Restricción basada en geometría cámaras

- La restricción epipolar (I). Planteamiento.

Para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo.



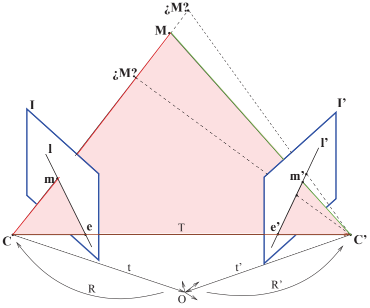
Dadas las matrices de proyección:

$$\bar{m} = [\mathbf{A} \,|\, \bar{b}\,]\bar{M}; \; \bar{m}' = [\mathbf{B} \,|\, \bar{d}\,]\bar{M},$$

Restricción basada en geometría cámaras

- La restricción epipolar (I). Planteamiento.

Para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo.



Dadas las matrices de proyección:

$$\bar{m} = [\mathbf{A} \,|\, \bar{b}]\bar{M}; \;\; \bar{m}' = [\mathbf{B} \,|\, \bar{d}]\bar{M},$$

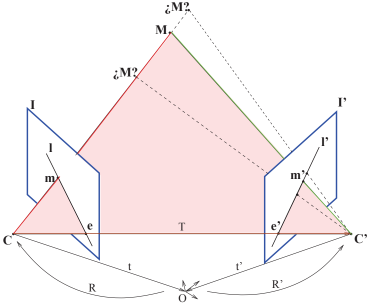
$T = \mathbf{B}^{-1}\bar{d} - \mathbf{A}^{-1}\bar{b}$ y las visuales:

$$\tilde{M} = \lambda\, \mathbf{A}^{-1}\bar{m} \; - \mathbf{A}^{-1}\bar{b}$$
$$\tilde{M} = \lambda' \mathbf{B}^{-1}\bar{m}' - \mathbf{B}^{-1}\bar{d}.$$

Restricción basada en geometría cámaras

- La restricción epipolar (I). Planteamiento.

Para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo.



Dadas las matrices de proyección:

$$\bar{m} = [\mathbf{A} \,|\, \bar{b}]\bar{M}; \ \bar{m}' = [\mathbf{B} \,|\, \bar{d}]\bar{M},$$

$T = \mathbf{B}^{-1}\bar{d} - \mathbf{A}^{-1}\bar{b}$ y las visuales:

$$\tilde{M} = \lambda\, \mathbf{A}^{-1}\bar{m} - \mathbf{A}^{-1}\bar{b}$$

$$\tilde{M} = \lambda' \mathbf{B}^{-1}\bar{m}' - \mathbf{B}^{-1}\bar{d}.$$

El producto triple:

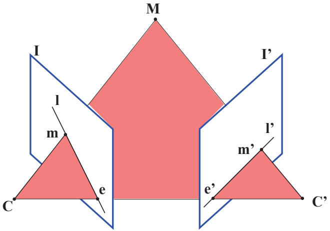
$$\bar{m}'^{\top} \underbrace{\mathbf{B}^{-\top}[\mathbf{B}^{-1}\bar{d} - \mathbf{A}^{-1}\bar{b}]_{\times}\mathbf{A}^{-1}}_{\mathbf{F}_{3\times 3}} \bar{m} = 0$$

Matriz Fundamental

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

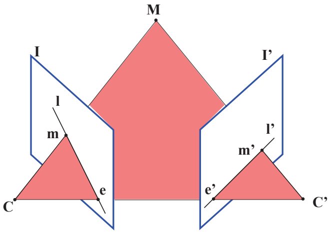
Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



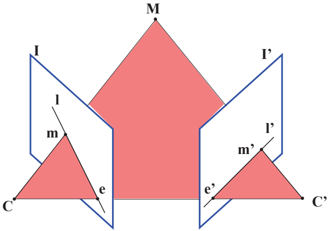
Propiedades:

- **F** es homogénea con rango 2.
  $\Rightarrow$ tiene 7 grados libertad.

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



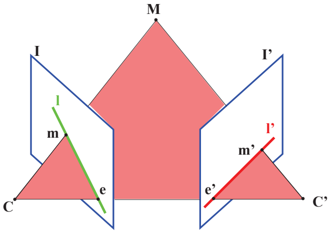
Propiedades:

- $\mathbf{F}$ es homogénea con rango 2.
- Si $\bar{m}$, $\bar{m}'$ son homólogos $\bar{m}'^{\top}\mathbf{F}\bar{m} = 0$.

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



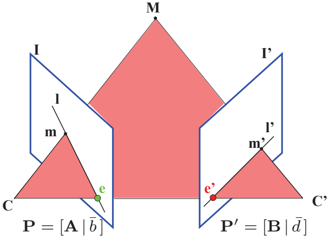
Propiedades:

- $\mathbf{F}$ es homogénea con rango 2.
- Si $\bar{m}$, $\bar{m}'$ son homólogos $\bar{m}'^{\top}\mathbf{F}\bar{m} = 0$.
- Epipolares: $\bar{l}' = \mathbf{F}\bar{m}$ y $\bar{l} = \mathbf{F}^{\top}\bar{m}'$.
  - $\bar{l}' = \mathbf{F}\bar{m}$ es la línea epipolar asociada a $\bar{m}$.
  - $\bar{l} = \mathbf{F}^{\top}\bar{m}'$ es la línea epipolar asociada a $\bar{m}'$.

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



Propiedades:

- $\mathbf{F}$ es homogénea con rango 2.
- Si $\bar{m}$, $\bar{m}'$ son homólogos $\bar{m}'^\top \mathbf{F} \bar{m} = 0$.
- Epipolares: $\bar{l}' = \mathbf{F}\bar{m}$ y $\bar{l} = \mathbf{F}^\top \bar{m}'$.
- Epipolos: $\mathbf{F}\bar{e} = \bar{0}$ y $\mathbf{F}^\top \bar{e}' = \bar{0}$.
  - $\mathbf{F}\bar{e} = \bar{0}$ imagen de C' en I.

$$\bar{e} = [\mathbf{A} \,|\, \bar{b}] \begin{bmatrix} \mathbf{B}^{-1}\bar{d} \\ 1 \end{bmatrix} = \mathbf{A}\mathbf{B}^{-1}\bar{d} + \bar{b}$$

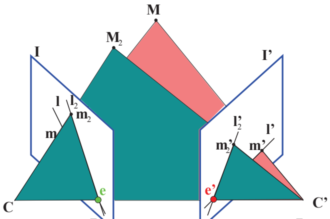
  - $\mathbf{F}^\top \bar{e}' = \bar{0}$ imagen de C en I'.

$$\bar{e}' = [\mathbf{B} \,|\, \bar{d}] \begin{bmatrix} \mathbf{A}^{-1}\bar{b} \\ 1 \end{bmatrix} = \mathbf{B}\mathbf{A}^{-1}\bar{b} + \bar{d}$$

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



$\mathbf{P} = [\mathbf{A} \,|\, \bar{b}]$ $\quad \mathbf{P}' = [\mathbf{B} \,|\, \bar{d}]$

Propiedades:

- $\mathbf{F}$ es homogénea con rango 2.
- Si $\bar{m}$, $\bar{m}'$ son homólogos $\bar{m}'^{\top}\mathbf{F}\bar{m} = 0$.
- Epipolares: $\bar{l}' = \mathbf{F}\bar{m}$ y $\bar{l} = \mathbf{F}^{\top}\bar{m}'$.
- Epipolos: $\mathbf{F}\bar{e} = \bar{0}$ y $\mathbf{F}^{\top}\bar{e}' = \bar{0}$.
  - $\mathbf{F}\bar{e} = \bar{0}$ imagen de C' en I.

$$\bar{e} = [\mathbf{A} \,|\, \bar{b}] \begin{bmatrix} \mathbf{B}^{-1}\bar{d} \\ 1 \end{bmatrix} = \mathbf{A}\mathbf{B}^{-1}\bar{d} + \bar{b}$$

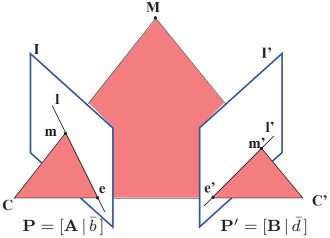
  - $\mathbf{F}^{\top}\bar{e}' = \bar{0}$ imagen de C en I'.

$$\bar{e}' = [\mathbf{B} \,|\, \bar{d}] \begin{bmatrix} \mathbf{A}^{-1}\bar{b} \\ 1 \end{bmatrix} = \mathbf{B}\mathbf{A}^{-1}\bar{b} + \bar{d}$$

Restricción basada en geometría cámaras

- La restricción epipolar (II). La matriz Fundamental.

Es la representación algebráica de la geometría epipolar. Sólo depende de los parámetros de las cámaras.



Propiedades:

- $\mathbf{F}$ es homogénea con rango 2.
- Si $\bar{m}$, $\bar{m}'$ son homólogos $\bar{m}'^{\top}\mathbf{F}\bar{m} = 0$.
- Epipolares: $\bar{l}' = \mathbf{F}\bar{m}$ y $\bar{l} = \mathbf{F}^{\top}\bar{m}'$.
- Epipolos: $\mathbf{F}\bar{e} = \bar{0}$ y $\mathbf{F}^{\top}\bar{e}' = \bar{0}$.
- Cálculo de $\mathbf{F}$ a partir de $\mathbf{P}$ y $\mathbf{P}'$.
  - $\mathbf{F} = \mathbf{B}^{-\top}[\mathbf{B}^{-1}\bar{d} - \mathbf{A}^{-1}\bar{b}]_{\times}\mathbf{A}^{-1}$.
  - $\mathbf{F} = [\bar{e}']_{\times}\mathbf{P}'\mathbf{P}^{+}$.
  - Si$C \equiv O$,$\mathbf{P} = \mathbf{K}[\mathbf{I}|\bar{0}]$,$\mathbf{P}' = \mathbf{K}'[\mathbf{R}|\bar{t}]$

$$\mathbf{F} = \mathbf{K}'^{\top}[\bar{t}]_{\times}\mathbf{R}\mathbf{K}^{-1} = [\mathbf{K}'\bar{t}]_{\times}\mathbf{K}'\mathbf{R}\mathbf{K}^{-1}$$

Restricción basada en geometría cámaras

- La restricción epipolar (III). Implementación.

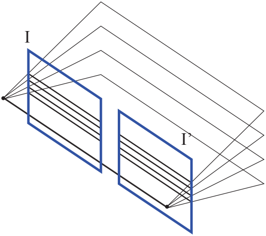
Un par estéreo está en posición estándar si los planos imagen son paralelos y sus sistemas de referencia están alineados.

Restricción basada en geometría cámaras

- La restricción epipolar (III). Implementación.

Un par estéreo está en posición estándar si los planos imagen son paralelos y sus sistemas de referencia están alineados.

Las imágenes se transforman a una posición estándar para facilitar la puesta en correspondencia. Este proceso se denomina **rectificación**.

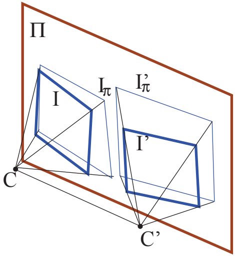


(Pollefeys, 00)

Rectificación

- Técnicas lineales

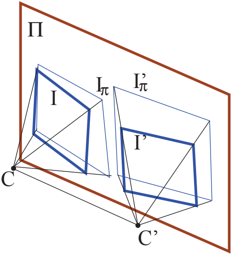
Se proyectan las imágenes I e I' sobre un plano Π paralelo a la línea base.



Rectificación

- Técnicas lineales

Se proyectan las imágenes I e I' sobre un plano Π paralelo a la línea base.



- Ventajas.
  - Modelo lineal

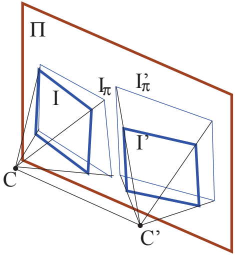
$$\bar{m} = \mathbf{HK}[\mathbf{R} \mid \bar{t}]\bar{M}$$

Se puede hacer la reconstrucción desde I_π e I'_π.

Rectificación

- Técnicas lineales

Se proyectan las imágenes I e I' sobre un plano Π paralelo a la línea base.



- Ventajas.
  - Modelo lineal

$$\bar{m} = \mathbf{HK}[\mathbf{R} \,|\, \bar{t}]\bar{M}$$

Se puede hacer la reconstrucción desde I_π e I'_π.

- Inconvenientes
  - No se puede rectificar si $\bar{e} \in I$.
  - El tamaño de I_π es muy grande si $\bar{e}$ es cercano a I.

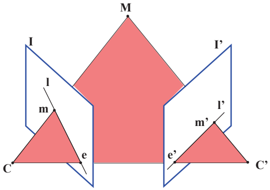
Rectificación

- Técnicas no lineales

Rectificación

- Técnicas no lineales (Chen, 2003)

No pueden expresarse como una transformación lineal de la imagen de partida.



Procedimiento:

1. Calcular $\mathbf{F}$.
   - Si $\mathbf{P} = [\mathbf{A} \,|\, \bar{b}]$ y $\mathbf{P}' = [\mathbf{B} \,|\, \bar{d}]$ son conocidas

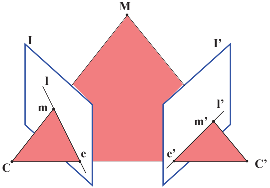
     $$\mathbf{F} = \mathbf{B}^{-\top}[\mathbf{B}^{-1}\bar{d} - \mathbf{A}^{-1}\bar{b}]_{\times}\mathbf{A}^{-1}$$

   - Si son desconocidas, a partir de al menos 7 correspondencias.

Rectificación

- Técnicas no lineales (Chen, 2003)

No pueden expresarse como una transformación lineal de la imagen de partida.



Procedimiento:

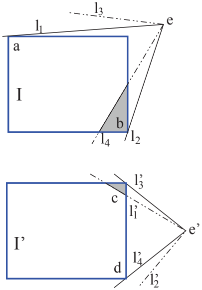
1. Calcular $\mathbf{F}$.
2. Calcular $\bar{e}$ y $\bar{e}'$.

$$\mathbf{F}\bar{e} = \bar{0}; \quad \mathbf{F}^\top \bar{e}' = \bar{0}$$

Rectificación

- Técnicas no lineales (Chen, 2003)

No pueden expresarse como una transformación lineal de la imagen de partida.



Procedimiento:

1. Calcular $\mathbf{F}$.
2. Calcular $\bar{e}$ y $\bar{e}'$.
3. Determinar las regiones comunes a I e I'.

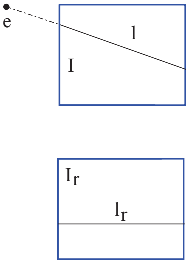
$$l_1' = \mathbf{F}\bar{a}; \quad l_2' = \mathbf{F}\bar{b};$$

$$l_3 = \mathbf{F}^\top \bar{c}; \quad l_4 = \mathbf{F}^\top \bar{d}$$

Rectificación

- Técnicas no lineales (Chen, 2003)

No pueden expresarse como una transformación lineal de la imagen de partida.



Procedimiento:

1. Calcular $\mathbf{F}$.
2. Calcular $\bar{e}$ y $\bar{e}'$.
3. Determinar las regiones comunes a I e I'.
4. Generar la imagen rectificada, recorriendo las epipolares de la de partida.

Rectificación

- Técnicas no lineales

No pueden expresarse como una transformación lineal de la imagen de partida.

- Ventajas.
  - Son aplicables inluso cuando $\bar{e} \in I$ o $\bar{e}' \in I'$.

Rectificación

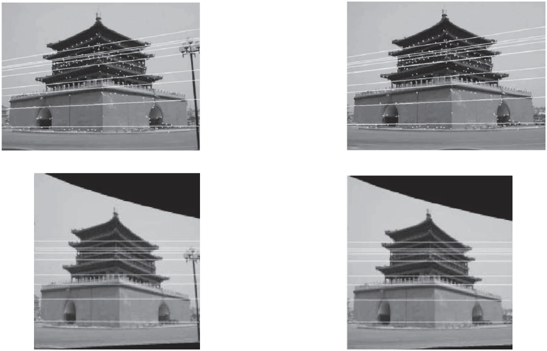
- Técnicas no lineales

No pueden expresarse como una transformación lineal de la imagen de partida.

- Ventajas.
  - Son aplicables inluso cuando $\bar{e} \in I$ o $\bar{e}' \in I'$.
- Inconvenientes.
  - No puede reconstruirse la escena desde la imagen rectificada.

Rectificación

- Resultados.



Rectificación

- Resultados.

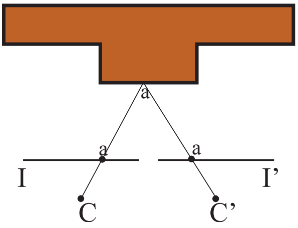
Rectificación

- Resultados.

Restricción basada en geometría escena

- Restricción de unicidad.

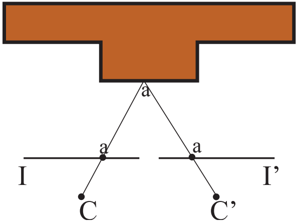
En una escena opaca, un punto de una imagen tiene como máximo una única correspondencia en la otra imagen del par.



Restricción basada en geometría escena

- Restricción de unicidad.

En una escena opaca, un punto de una imagen tiene como máximo una única correspondencia en la otra imagen del par.



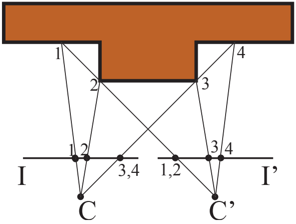
Es decir:

- Cuando **no** hay oclusión la relación de correspondencia es bidireccional.

Restricción basada en geometría escena

- Restricción de unicidad.

En una escena opaca, un punto de una imagen tiene como máximo una única correspondencia en la otra imagen del par.



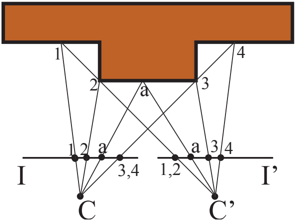
Es decir:

- Cuando **no** hay oclusión la relación de correspondencia es bidireccional.
- Cuando **sí** existe oclusión, hay puntos que no tienen correspondencias.

Restricción basada en geometría escena

- Restricción de unicidad.

En una escena opaca, un punto de una imagen tiene como máximo una única correspondencia en la otra imagen del par.



Es decir:

- Cuando **no** hay oclusión la relación de correspondencia es bidireccional.
- Cuando **sí** existe oclusión, hay puntos que no tienen correspondencias.
- No existen dos o más correspondencias.

Restricción basada en geometría escena

- Restricción de ordenación.

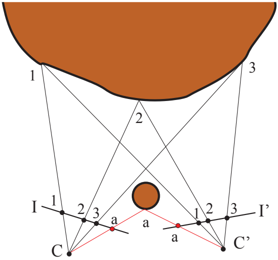
Los homólogos de un objeto opaco aparecen en el mismo orden sobre las epipolares correspondientes.

Restricción basada en geometría escena

- Restricción de ordenación.

Los homólogos de un objeto opaco aparecen en el mismo orden sobre las epipolares correspondientes.

Puede no cumplirse para varios objetos.

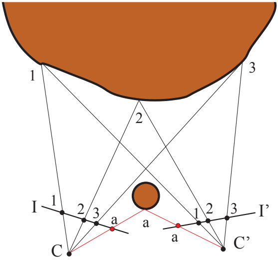


Restricción basada en geometría escena

- Restricción de ordenación.

Los homólogos de un objeto opaco aparecen en el mismo orden sobre las epipolares correspondientes.

Puede no cumplirse para varios objetos.

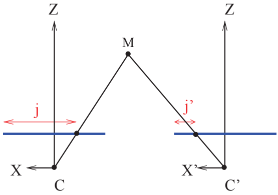


Esta restricción será el fundamento de los algoritmos de *programación dinámica*.

Restricción basada en geometría escena

- Restricción de disparidad

Dadas dos imágenes rectificadas, y dos homólogos $\bar{m} = (i, j)$ y $\bar{m}' = (i, j')$, la *disparidad*, d, viene dada por $d = j - j'$.

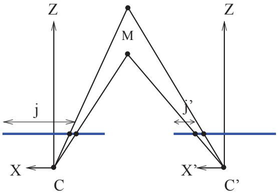


- *Disparidad:* $d = j - j'$

Restricción basada en geometría escena

- Restricción de disparidad

Dadas dos imágenes rectificadas, y dos homólogos $\bar{m} = (i, j)$ y $\bar{m}' = (i, j')$, la *disparidad*, d, viene dada por $d = j - j'$.

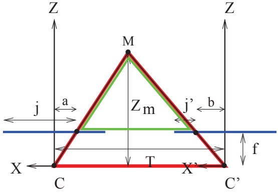


- *Disparidad:* $d = j - j'$

Restricción basada en geometría escena

- Restricción de disparidad

Dadas dos imágenes rectificadas, y dos homólogos $\bar{m} = (i, j)$ y $\bar{m}' = (i, j')$, la *disparidad*, d, viene dada por $d = j - j'$.



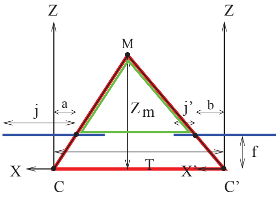
- *Disparidad:* $d = j - j'$
- Por semejanza de triángulos

$$\frac{Z_m}{T} = \frac{Z_m - f}{T - (a + b)},$$

Restricción basada en geometría escena

- Restricción de disparidad

Dadas dos imágenes rectificadas, y dos homólogos $\bar{m} = (i, j)$ y $\bar{m}' = (i, j')$, la *disparidad*, d, viene dada por $d = j - j'$.



- *Disparidad:* $d = j - j'$
- Por semejanza de triángulos

$$\frac{Z_m}{T} = \frac{Z_m - f}{T - (a + b)},$$

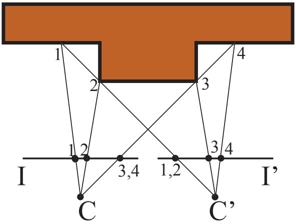
como $d = j - j' = a + b$,

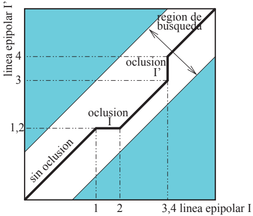
$$Z_m = f\frac{T}{d}$$

Restricción basada en geometría escena

- Restricción de disparidad

Estableciendo un límite sobre la disparidad máxima admisible, permite definir una banda de búsqueda en torno a la epipolar.





Algoritmos de puesta en correspondencia

- **Planteamiento general**
  Obtener un mapa de disparidades denso, que sea:
  - *Contino.* La disparidad de cada objeto debe variar suavemente según la estructura del objeto.
  - *Detallado.* Los contornos han de estar bien definidos y los pequeños detalles perfectamente reflejados.

  En general, estos dos requisitos son contrapuestos.

Algoritmos de puesta en correspondencia

- **Planteamiento general**
  Obtener un mapa de disparidades denso, que sea:
  - *Contino.* La disparidad de cada objeto debe variar suavemente según la estructura del objeto.
  - *Detallado.* Los contornos han de estar bien definidos y los pequeños detalles perfectamente reflejados.

  En general, estos dos requisitos son contrapuestos.

- **Algoritmos** (Sharstein, 2002)
  - Cooperativos.
  - Programación dinámica.
  - Basados en cortes de grafos.

Algoritmo Cooperativo

- **Hipótesis**:

Algoritmo Cooperativo

- **Hipótesis**: R. Epipolar,

Algoritmo Cooperativo

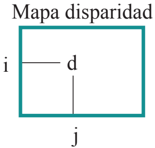
- **Hipótesis**: R. Epipolar,R. Unicidad,

Algoritmo Cooperativo

- **Hipótesis**: R. Epipolar,R. Unicidad, R. Disparidad.
  - Las imágenes están rectificadas.
  - Cada píxel tiene como máximo un homólogo.
  - La disparidad es generalmente continua.

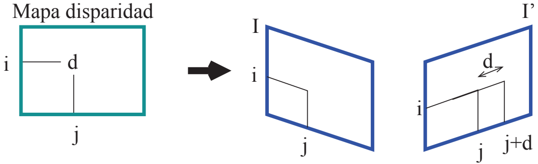
Algoritmo Cooperativo

- **Hipótesis**: R. Epipolar,R. Unicidad, R. Disparidad.
- **Objetivo:** Construir un mapa de disparidad asociado a una imagen del par estéreo.



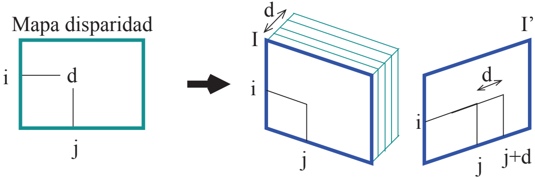
Algoritmo Cooperativo

- **Hipótesis**: R. Epipolar,R. Unicidad, R. Disparidad.
- **Objetivo:** Construir un mapa de disparidad asociado a una imagen del par estéreo.



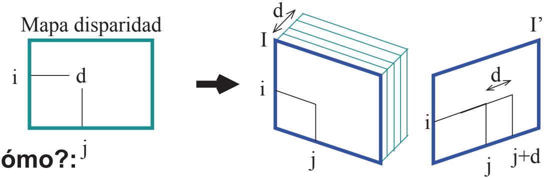
Algoritmo Cooperativo

- **Hipótesis**: R. Epipolar,R. Unicidad, R. Disparidad.
- **Objetivo:** Construir un mapa de disparidad asociado a una imagen del par estéreo.



Algoritmo Cooperativo

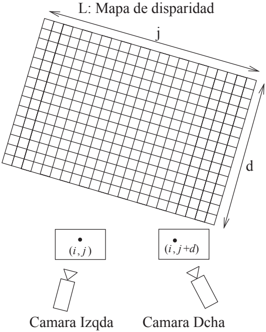
- **Hipótesis**: R. Epipolar,R. Unicidad, R. Disparidad.
- **Objetivo:** Construir un mapa de disparidad asociado a una imagen del par estéreo.



- **¿Cómo?:**
  1. Asigno a cada pixel un valor de emparejamiento inicial para cada disparidad.
  2. Iterativamente actualizo los valores de emparejamiento, al tiempo que impongo las restricciones.

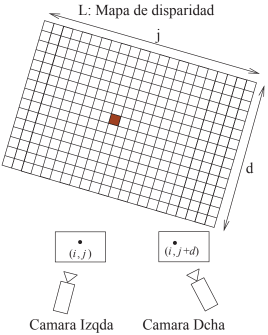
Algoritmo Cooperativo

- Procedimiento iterativo



Algoritmo Cooperativo

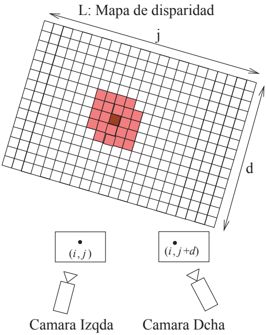
- Procedimiento iterativo



- $L_0(i,j,d) = \delta[I(i,j), I'(i,j+d)]$
  siendo $\delta(a,b)$ una medida:
  a) Valores altos para emparejamientos correctos, aunque
  b) también dé valores altos para emparejamientos incorrectos.

Algoritmo Cooperativo

- Procedimiento iterativo



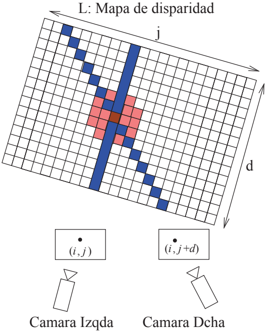
- $L_0(i,j,d) = \delta[I(i,j), I'(i,j+d)]$
- $S_n(i,j,d) = \sum_{(i',j',d') \in \Phi} L_n(i+i', j+j', d+d')$

modela la **restricción de disparidad**.

Plantea una hipótesis de continuidad que suma valores en una vecindad Φ para incrementar la consistencia local.

Algoritmo Cooperativo

- Procedimiento iterativo

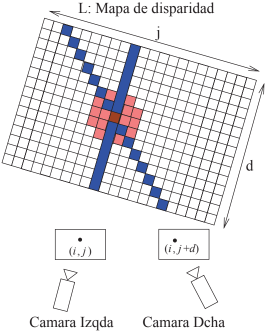


- $L_0(i,j,d) = \delta[I(i,j), I'(i,j+d)]$
- $S_n(i,j,d) = \sum_{(i',j',d') \in \Phi} L_n(i+i', j+j', d+d')$
- $R_n(i,j,d) = \left[\frac{S_n(i,j,d)}{\sum_{(i',j',d') \in \Psi} S_n(i',j',d')} \right]^{\alpha}$

  modela la **restricción de unicidad**. Siendo la constante de inhibición $\alpha > 1$.

Algoritmo Cooperativo

- Procedimiento iterativo



- $L_0(i,j,d) = \delta[I(i,j), I'(i,j+d)]$
- $S_n(i,j,d) = \sum_{(i',j',d') \in \Phi} L_n(i+i', j+j', d+d')$
- $R_n(i,j,d) = \left[\frac{S_n(i,j,d)}{\sum_{(i',j',d') \in \Psi} S_n(i',j',d')} \right]^{\alpha}$
- $L_n(i,j,d) = L_0(i,j,d) * R_n(i,j,d).$

Algoritmo Cooperativo

- Detección de oclusiones

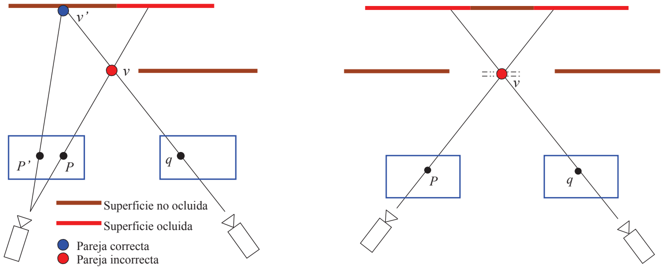
Algoritmo Cooperativo

- Detección de oclusiones

Algoritmo Cooperativo

- Detección de oclusiones

Un píxel se etiqueta como **oculto** cuando no supera un umbral, y es el máximo valor de emparejamiento en sus visuales.



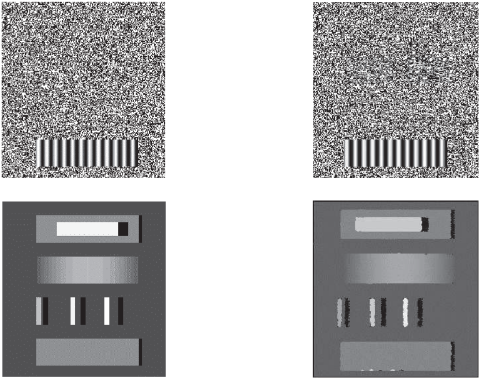
Algoritmo Cooperativo

- Resumen del algoritmo

1. Rectificamos las imágenes.
2. Determinamos una imagen de referencia $I(i,j)$ y le asociamos una matriz de disparidades $L_0(i,j,d)$, empleando una medida de similaridad (p.ej. correlación).
3. Actualizamos iterativamente L_n.
4. Para todo píxel $I(i,j)$ busco el valor el máximo $L(i,j,d)$.
5. Si el máximo supera un umbral, le asigno disparidad d, en caso contrario lo etiqueto como oculto.

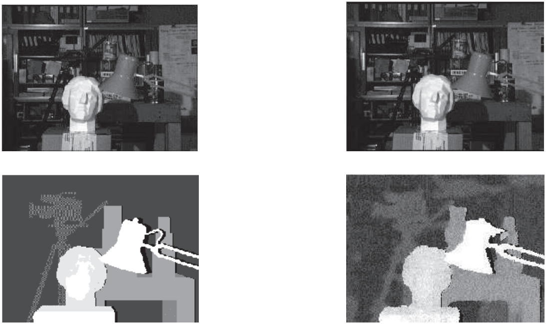
Algoritmo Cooperativo

- Resultados (I)



Algoritmo Cooperativo

- Resultados (II)



lgoritmo basado en programación dinámic

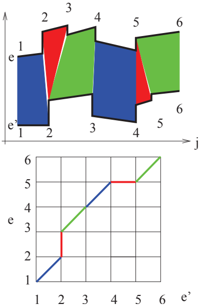
La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.

lgoritmo basado en programación dinámi

La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.



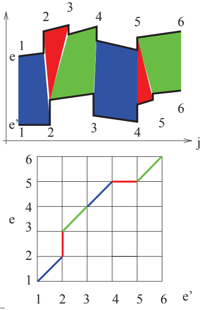
Planteamiento:

- Cada nodo representa un elemento en correspondencia.
- Cada arco modela las discrepancias entre elementos.
- La búsqueda de correspondencias consiste en encontrar un camino de coste mínimo sobre el grafo.

lgoritmo basado en programación dinámi

La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.



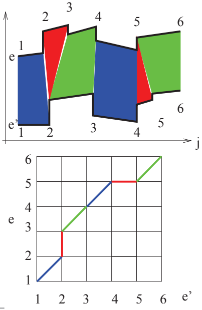
Restricciones:

- Unicidad.
- Epipolar.
- Ordenación.
- Disparidad.

lgoritmo basado en programación dinámi

La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.



Algoritmo

$\mathcal{C}(i,j)$: Coste emparejar el elemento i de I con el j de I'.

$\mathbf{C}(i,j)$: Coste de emparejar los i primeros elementos de I con los j primeros de I'.

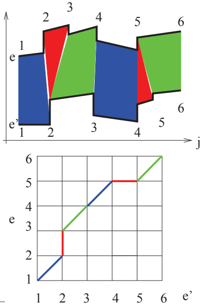
$\mathbf{B}(i,j)$: Enlace al precedesor de mínimo coste.

$$\mathbf{C} = \begin{pmatrix} \square & \square \\ \square & \square \end{pmatrix}; \mathbf{B} = \begin{pmatrix} \square & \square \\ \square & \square \end{pmatrix}$$

lgoritmo basado en programación dinámi

La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.



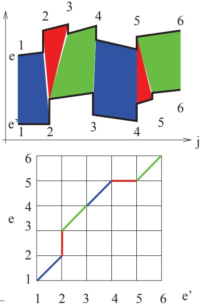
Algoritmo (para generar $\mathbf{C}$ y $\mathbf{B}$)

```
for j = 1 to m
   for i = 1 to n
      min1=C(i-1,j-1) + 𝒞(i,j)
      min2=C(i-1,j) + 𝒞oclus;
      min3=C(i,j-1) + 𝒞oclus;
      C(i,j) =min(min1,min2,min3);
      case (C(i,j)) of
          min1: B(i,j) = 1;
          min2: B(i,j) = 2;
          min3: B(i,j) = 3;
      endcase;
   endfor;
endfor;
```

lgoritmo basado en programación dinámi

La búsqueda densa de correspondencias se plantea como un problema de optimización, resoluble mediante algoritmos de Investigación Operativa.

- *Hipótesis.* Imágenes rectificadas.



Algoritmo (para camino mínimo)

ListaHomolog=nil;
$i = M$
$j = N$
while $(i \neq 0$ && $j \neq 0)$
    case $(\mathbf{B}(i,j))$ of
        1: ListaHomolog $\uplus$ (i,j);
           i-;
           j-;
        2: ListaHomolog $\uplus$ $(i,0)$;
           i-;
        3: ListaHomolog $\uplus$ $(0,j)$;
           j-;
    endcase;
endfor;

lgoritmo basado en programación dinámi

- Función de coste (Cox, 96)

La verosimilitud de que un punto $\bar{M}$ de la escena se proyecte en $\bar{z}_i$ y $\bar{z}'_j$ en cada imagen respectivamente

$$\Lambda(\bar{z}_i, \bar{z}'_j | \bar{M}) = \left[\frac{1 - P_v}{\Phi}\right]^{\delta_{ij}} \left[P_v p(\bar{z}_i|\bar{M}) P_v p(\bar{z}_j|\bar{M})\right]^{1-\delta_{ij}},$$

donde,

- P_v : Prob ser visible.
- $\delta_{ij} : \begin{cases} 1. \text{ oclusión} \\ 0. \text{ visibles} \end{cases}$
- $p(\bar{z}_k|\bar{M}) = \frac{e^{-\frac{1}{2}(\bar{z}-\bar{z}_k)^\top \Sigma_k^{-1}(\bar{z}-\bar{z}_k)}}{|(2\pi)^d \Sigma_k|^{\frac{1}{2}}}.$
- $\bar{z} \approx \frac{\Sigma_i^{-1}\bar{z}_i + \Sigma_j^{-1}\bar{z}_j}{\Sigma_i^{-1} + \Sigma_j^{-1}}.$

lgoritmo basado en programación dinámi

- Función de coste (Cox, 96)

La verosimilitud de que un punto $\bar{M}$ de la escena se proyecte en $\bar{z}_i$ y $\bar{z}'_j$ en cada imagen respectivamente

$$\Lambda(\bar{z}_i, \bar{z}'_j | \bar{M}) = \left[\frac{1 - P_v}{\Phi}\right]^{\delta_{ij}} \left[P_v p(\bar{z}_i | \bar{M}) P_v p(\bar{z}_j | \bar{M})\right]^{1-\delta_{ij}} ,$$

tomando logaritmos y asumiendo $\Sigma_i = \Sigma_j = \Sigma$,

$$\mathcal{C}_{\text{oclus}} = \ln \left[\frac{P_v^2 \Phi}{(1 - P_v)|(2\pi)^d \Sigma_k|^{\frac{1}{2}}}\right];$$

$$\mathcal{C}(i,j) = \frac{1}{4}(\bar{z}_i - \bar{z}_j)^\top \Sigma^{-1} (\bar{z}_i - \bar{z}_j)$$

lgoritmo basado en programación dinámi

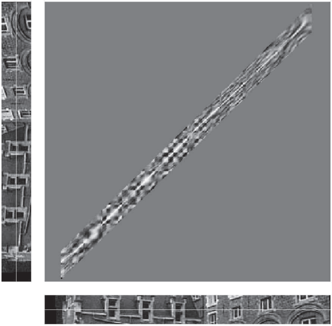
Tal como se ha presentado, esta técnica tienen el inconveniente de no considerar la coherencia entre líneas.

- Extensiones
  - Minimizando el número de discontinuidades horizontales y verticales (Cox, 96).
  - Minimizando una función de coste 3D, intra- e inter-epipolares (Ohta, 85).
  - Multivista (Pollefeys, 00).

lgoritmo basado en programación dinámi

- Resultados

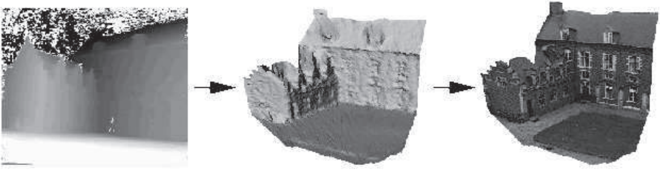
Evaluación de la función de coste para dos epipolares (Pollefeys, 00).



lgoritmo basado en programación dinámi

- Resultados

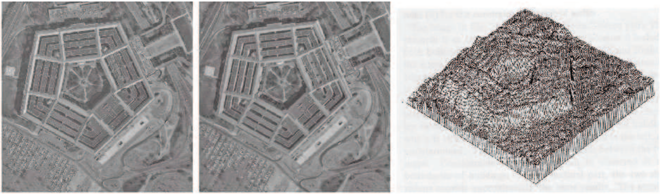
Reconstrucción (Pollefeys, 00)



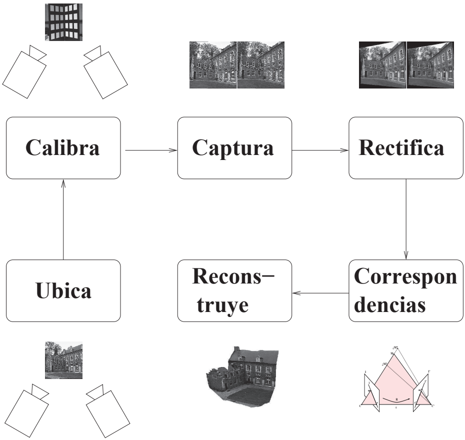
lgoritmo basado en programación dinámi

- Resultados

Reconstrucción (Ohta, 85)



Esquema de trabajo



Bibliografía

- Libros de texto:
  - O. Faugeras. “Three dimensional computer vision: a geometric viewpoint.” MIT Press. 1993.
  - D. Forsyth, J. Ponce. “Computer Vision. A modern approach.” Prentice Hall. 2003.
  - R. Hartley, A. Zisserman. “Multiple view geometry in computer vision.” Cambridge University Press, 2004.
  - M. Pollefeys. “3D modeling from images.” ECCV 2000 tutorial.
- Rectificación:
  - Z. Chen et al “A new image rectification algorith.” *Pattern Recognition Letters*. 24, pp. 251-260. 2003.
- Puesta en correspondencia densa:
  - D. Sharstein, R. Szeliski. “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms.” *Int. J. of Computer Vision*, 47(1/2/3), pp. 7-42. 2002.
  - C. Zitnick, T. Kanade. “A cooperative algorithm for stereo matching and occlusion detection.” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(7), pp. 675-684. 2000.
  - I. Cox et al. “A maximum likelihood stereo algorithm.” *Computer Vision and Image Understanding* 63(3), pp. 542-567. 1996.
  - Y. Ohta, T. Kanade. “Stereo by intra and inter-scanline search using dynamic programming.” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 7(2), pp.139-154. 1985.
  - R.T. Collins. “A space sweep approach to true multi-image matching.” *Proc. Computer Vision and Pattern Recognition*, pp. 358-363. 1996.