

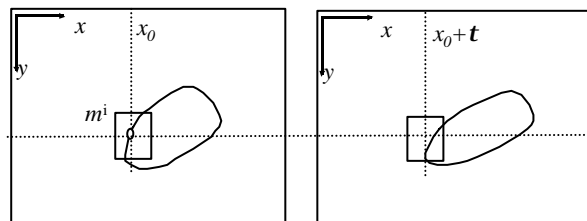
- ⇒ **El problema de Correspondencia**
- ⇒ **Restricciones aplicadas a la correspondencia**
- ⇒ **Técnicas de correspondencia**
- ⇒ **Conclusiones**

- ⇒ **Técnicas basadas en Area (“area-based”)**
  - ⇒ **Correlación de área**
  - ⇒ **Estrategias tosco-a-fino**
- ⇒ **Técnicas basadas en Características (“feature-based”)**
  - ⇒ **Teoría computacional de Marr-Poggio**
  - ⇒ **Teoría computacional de Mayhew-Frisby**
  - ⇒ **Técnicas basadas en segmentos de borde**
  - ⇒ **Correspondencia mediante características múltiples**
  - ⇒ **Proceso de relajación**
- ⇒ **Técnicas Jerárquicas**
  - ⇒ **Algoritmo de Hoff y Ahuja**
  - ⇒ **Algoritmo de Marapane y Trivedi**
- ⇒ **Programación Dinámica**
  - ⇒ **Algoritmo de Otha y Kanade**
  - ⇒ **Algoritmo de Lee y Leou**
- ⇒ **Algoritmos basados en estéreo trinocular**

**x Técnicas basadas en Area (“area-based”)**

- ⇒ Correlación de área
- ⇒ Estrategia tosco-a-fino
- ⇒ Técnicas basadas en Características (“feature-based”)
- ⇒ Técnicas Jerárquicas
- ⇒ Programación Dinámica
- ⇒ Algoritmos basados en estéreo trinocular

- ⇒ Estos métodos consideran las dos imágenes captadas como una señal bidimensional trasladada.
- ⇒ Tratan de obtener, para cada punto de la imagen, dicha traslación minimizando un cierto criterio (correlación)
- ⇒ Para cada pixel de una imagen se calcula la correlación entre la distribución de intensidades de una ventana centrada en dicho pixel y una ventana del mismo tamaño centrada en el pixel a analizar de la otra imagen.



⇒ **Restricciones aplicadas:**⇒ **Restricción lambertiana:**

⇒ Las superficies han de ser lambertianas. La intensidad proyectada de un punto 3D no depende del punto de vista

⇒ **Restricción fronto-paralela:**

⇒ Asumen que la disparidad es constante localmente (ventana de correlación). Las superficies deben ser paralelas a los planos de imagen de la cámara (pendiente pequeña)

⇒ **Restricción de continuidad:**

⇒ Las superficies son, al menos localmente, continuas.

⇒ **Criterios de correlación (minimiza el funcional):**⇒ **Suma de la diferencia de cuadrados:**

$$SSD(m^l(x, y), m^r(x', y')) = -\frac{1}{N} \sum_{\forall i, j \perp m(x+i, y+j) \in W(m)} (I^l(m^l(x+i, y+j)) - I^r(m^r(x'+i, y'+j)))^2$$

$$\Downarrow$$

$$SSD(m^l, m^r) = -\frac{1}{N} \sum_{\forall m_i \in W(m)} (I^l(m^l_i) - I^r(m^r_i))^2$$

Es muy sensible a diferencias de iluminación

⇒ **Suma de las diferencias de cuadrados centradas respecto a la media**

$$ZSSD(m^l, m^r) = -\frac{1}{N} \sum_{\forall m_i \in W(m)} \left( (I^l(m^l_i) - \overline{I^l(m^l)}) - (I^r(m^r_i) - \overline{I^r(m^r)}) \right)^2$$

⇒ **Suma normalizada de las diferencias de cuadrados centradas respecto a la media**

$$ZNSSD(m^l, m^r) = -\frac{\sum_{\forall m_i \in W(m)} \left( (I^l(m^l_i) - \overline{I^l(m^l)}) - (I^r(m^r_i) - \overline{I^r(m^r)}) \right)^2}{N s^l(m^l) s^r(m^r)}$$

Tiene en cuenta variaciones en la distribución de intensidades dentro de la ventana

⇒ **Correlación cruzada:**

$$CC(m^l, m^r) = \frac{1}{N} \sum_{\forall m_i \in W(m)} I^l(m^l_i) \cdot I^r(m^r_i)$$

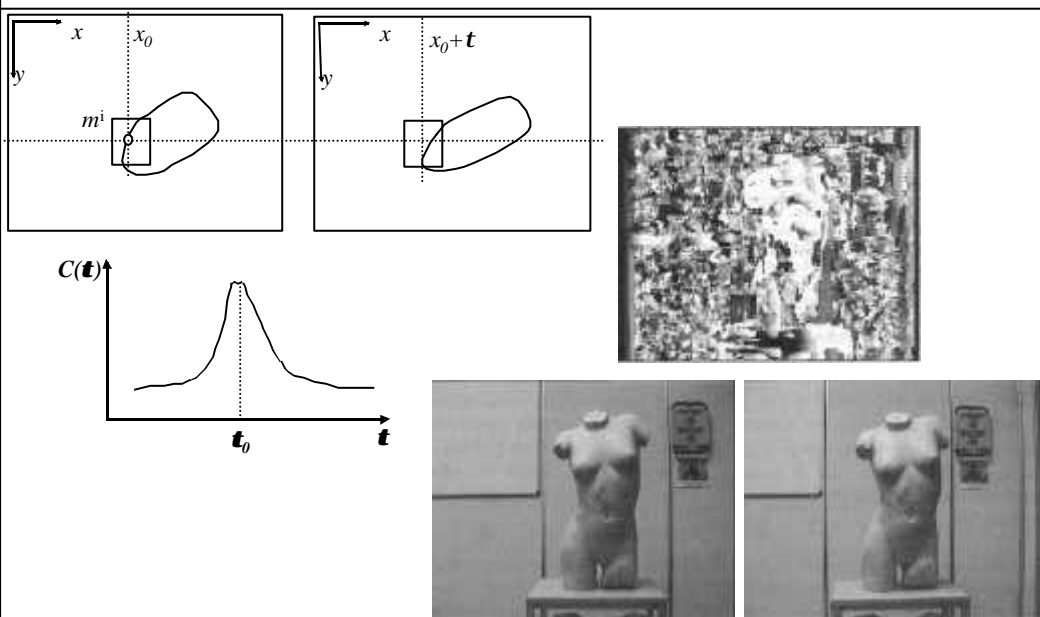
⇒ Favorece zonas de intensidad elevada

⇒ **Correlación cruzada normalizada centrada en la media**

$$CC(m^l, m^r) = \frac{\sum_{\forall m_i \in W(m)} (I^l(m^l_i) - \overline{I^l(m^l)}) \cdot (I^r(m^r_i) - \overline{I^r(m^r)})}{N \sigma^l(m^l) \sigma^r(m^r)}$$

**x Mejoras:**

⇒ **Uso de ventanas adaptativas**



⇒ **Ventajas**

- ⇒ **Obtienen buenos resultados sobre imágenes con textura importante**
- ⇒ **Permiten crear mapas densos de disparidad**
- ⇒ **Son fáciles de paralelizar**

⇒ **Inconvenientes**

- ⇒ **Presentan problemas con imágenes con elevadas discontinuidades de superficie**
- ⇒ **Son muy sensibles a variaciones fotométricas debidas a sombras**
- ⇒ **Requieren un proceso posterior de eliminación de falsas correspondencias**
- ⇒ **Problemas con las oclusiones**

- ⇒ **Se trata de una estrategia utilizada con diferentes técnicas de correspondencia**
- ⇒ **La información obtenida en una escala grosera se utiliza para guiar y limitar la búsqueda de correspondencias a una escala más fina.**
- ⇒ **Suele emplearse con las técnicas de correlación vistas anteriormente junto con el uso de ventanas adaptativas.**
- ⇒ **También se emplea con técnicas basadas en primitivas dispersas**

- ✓ **Técnicas basadas en Area (“area-based”)**
- ✗ **Técnicas basadas en Características (“feature-based”)**
  - ⇒ **Teoría computacional de Marr-Poggio**
  - ⇒ **Teoría computacional de Mayhew-Frisby**
  - ⇒ **Algoritmo de relajación de Kim y Aggarwal**
  - ⇒ **Técnicas basadas en segmentos de borde**
    - ⇒ **Algoritmo de Disparidad Diferencial Mínima (Medioni y Nevatia)**
    - ⇒ **Algoritmo de Ayache y Faverjon**
- ⇒ **Técnicas Jerárquicas**
- ⇒ **Programación Dinámica**
- ⇒ **Algoritmos basados en estéreo trinocular**

- ⇒ **Obtienen primitivas de alto nivel (puntos de borde, segmentos, curvas, regiones) que atesoran un conjunto de características invariantes a la proyección en mayor o menor medida**
- ⇒ **Ventajas**
  - ⇒ **Las primitivas de alto nivel atesoran información más rica que los niveles de intensidad**
  - ⇒ **Permite utilizar restricciones geométricas entre las primitivas.**
  - ⇒ **Son más robustas.**
- ⇒ **Inconvenientes**
  - ⇒ **Proporcionan información dispersa**

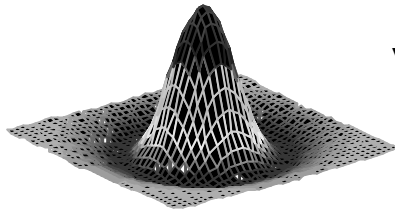
- ⇒ **Marr [Vision, 82] presenta uno de los estudios más importantes sobre el proceso visual humano estableciendo una teoría computacional completa.**
- ⇒ **Dentro de esta teoría se incluye el proceso de correspondencia estereoscópica**
- ⇒ **Postula la visión como un proceso que debe ser comprendido a tres niveles:**

Teoría computacional	Representación y algoritmo	Implementación física
¿Cuál es la finalidad? ¿por qué es apropiada? ¿cuál es la lógica y la estrategia para llevarla a cabo?	¿Cómo puede implementarse? ¿Cuál es la representación de la entrada y la salida? ¿Cuál es el algoritmo para esta transformación?	¿Cómo pueden realizarse físicamente la representación y el algoritmo?

- ⇒ **Postula que la correspondencia estereoscópica se produce en las etapas iniciales de formación de imagen sin información perceptual de alto nivel.**
- ⇒ **Etapas:**
  - 1) **Las imágenes del par estéreo son filtradas mediante filtros Laplaciano-Gaussianos ajustados a diferentes escalas espaciales, evaluando los pasos por cero “zero-crossings”**
  - 2) **Se realiza, para cada máscara, la correspondencia entre los segmentos de paso por cero extraídos de cada una de las imágenes que tienen el mismo signo y aproximadamente la misma orientación. Las ambigüedades en la correspondencia local se resuelven considerando la disparidad en el signo de correspondencias próximas no ambiguas.**
  - 3) **Las correspondencias obtenidas de las máscaras de mayor tamaño controlan la convergencia del algoritmo eliminando ambigüedades en la correspondencia de las máscaras más pequeñas.**
  - 4) **Un proceso de relajación de etiquetas impone las restricciones de continuidad de superficie y unicidad**
  - 5) **Los resultados de la correspondencia se almacenan en un buffer que denomina esquema 2.5D, representando los pasos por cero junto a la disparidad entre ambas imágenes.**

⇒ **Filtrado Laplaciano-Gaussiano / zero-crossings:**

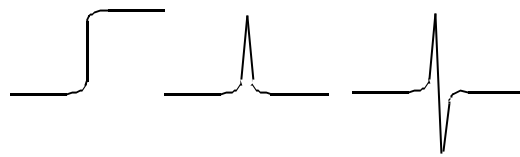
⇒ **Detecta discontinuidades de intensidad**



$$\nabla^2 G(r, \vartheta) = -\frac{1}{\pi\sigma^4} \left[ 1 - \frac{r^2}{\sigma^2} \right] e^{-\left(\frac{r^2}{2\sigma^2}\right)}$$

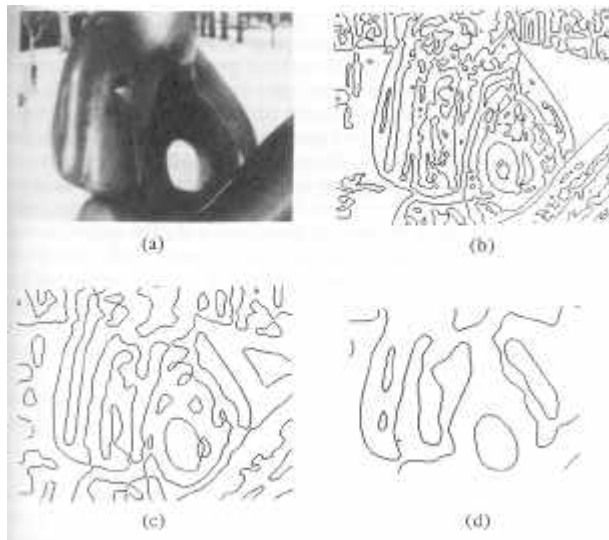
$$\omega_{2D} = 2\sqrt{2} \sigma \quad \text{Anchura del filtro}$$

Los diferentes filtros detectan discontinuidades a diferentes escalas espaciales



⇒ **La presencia de discontinuidades de superficie esta ligada a la presencia de pasos por cero a diferentes escalas espaciales**

⇒ **Detección de pasos por cero a diferentes escalas espaciales**

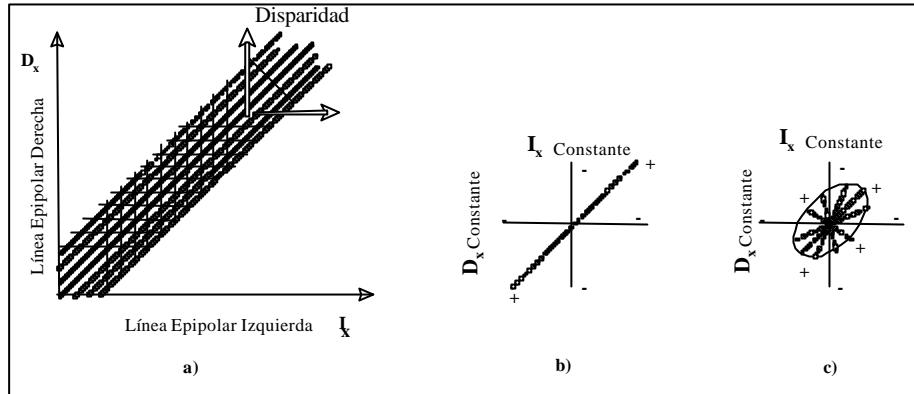




- ⇒ **Correspondencia entre los segmentos de paso por cero:**
  - ⇒ Para cada escala se evalúa la correspondencia a lo largo de la epipolar de los pasos por cero que tengan el mismo signo y aproximadamente la misma orientación.
  - ⇒ **Límite fusional de Panum:**
    - ⇒ la probabilidad de encontrar una correspondencia correcta con una disparidad menor que  $\pm(\omega_{2D}/2)$  es del 95%
  - ⇒ Las ambigüedades en la correspondencia local se resuelven considerando la disparidad en el signo de correspondencias próximas no ambiguas
  - ⇒ Las correspondencias obtenidas de las máscaras de mayor tamaño controlan la convergencia del algoritmo eliminando ambigüedades en la correspondencia de las máscaras más pequeñas.

- ⇒ **Proceso de Relajación: Algoritmo Cooperativo**
  - ⇒ **Correspondencia mediante la relajación de etiquetas:**
    - ⇒ Identificar un conjunto de características en cada imagen (denominadas nodos), y posteriormente asignar una única etiqueta (correspondencia) a cada nodo dentro de un espacio de búsqueda discreto.
    - ⇒ Para cada pareja candidata se establece un probabilidad de correspondencia que es adaptada iterativamente dependiendo de las probabilidades de correspondencia de los nodos vecinos, de forma que correspondencias fuertes de nodos vecinos aumenten la probabilidad de las correspondencias débiles
    - ⇒ Establece por tanto una consistencia global en el proceso de correspondencia, frente al proceso local de las etapas primarias comentadas anteriormente

## ⇒ Algoritmo Cooperativo de Marr y Poggio:

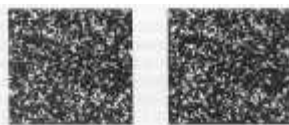


## ⇒ Algoritmo Cooperativo de Marr y Poggio:

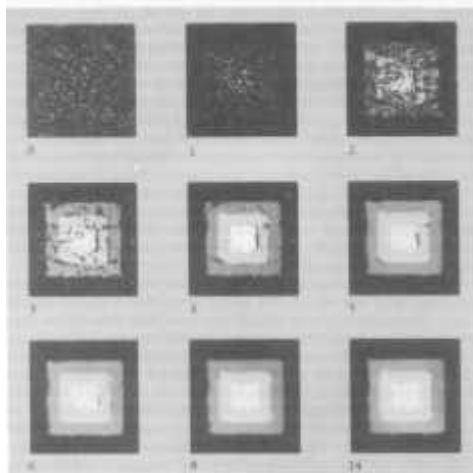
- ⇒ Para cada línea epipolar de un par estéreo se asocia una red bidimensional de nodos o células interconectados (figura a)
- ⇒ Las conexiones horizontales y verticales establecen un efecto inhibitorio, de forma que los nodos situados a lo largo de una línea horizontal o vertical se inhiben mutuamente quedando al final una única correspondencia en cada línea (Restricción de Unicidad) (figura b)
- ⇒ Las conexiones diagonales (conexiones excitatorias) favorecen la correspondencia de elementos adyacentes en la diagonal con la misma disparidad (Continuidad de Disparidad) (figura b), considerando por tanto en este modelo que las superficies son continuas y suaves.
- ⇒ La restricción anterior de continuidad de la disparidad se establece únicamente en la dirección diagonal, para considerar una restricción bidimensional considera un entorno de vecindad excitatorio en forma de disco  $S(x,y,d)$  (figura c).

- En el proceso cooperativo de generación de probabilidades de correspondencia  $C_{x,y;d}^t$  representa el estado de un nodo en el instante  $t$ , para la correspondencia de las coordenadas  $(x,y)$  en la imagen izquierda con las coordenadas  $(x+d, y)$  en la imagen derecha
- El estado inicial de la red se establece con un 1 en los nodos que representan correspondencias candidatas, y un 0 en el resto
- En cada iteración el proceso excitatorio añade al coeficiente  $C_{x,y;d}^t$  de cada nodo los estados previos de las correspondencias potenciales dentro del entorno de vecindad excitatorio  $S(x,y,d)$
- El proceso inhibitorio resta una suma ponderada de los estados de las correspondencias potenciales en el entorno de vecindad inhibitorio  $O(x,y,d)$

$$C_{x,y;d}^{t+1} = \sigma \left\{ \underbrace{\sum_{x',y',d' \in S(x,y,d)} C_{x',y',d'}^t}_{\text{Componente Excitatoria}} - \varepsilon \underbrace{\sum_{x',y',d' \in O(x,y,d)} C_{x',y',d'}^t}_{\text{Componente Inhibitoria}} + \underbrace{C_{x,y;d}^0}_{\text{Valor Inicial}} \right\}$$



Estereogramas de puntos aleatorios



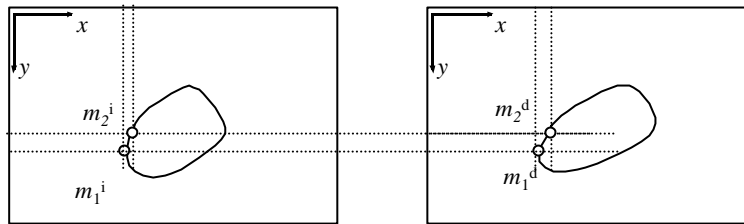
Mapas de profundidad en diferentes iteraciones del proceso de relajación

- ✓ **Técnicas basadas en Area (“area-based”)**
- ✗ **Técnicas basadas en Características (“feature-based”)**
  - ⇒ **Teoría computacional de Marr-Poggio**
  - ⇒ **Teoría computacional de Mayhew-Frisby**
  - ⇒ **Algoritmo de relajación de Kim y Aggarwal**
  - ⇒ **Técnicas basadas en segmentos de borde**
- ⇒ **Técnicas Jerárquicas**
- ⇒ **Programación Dinámica**
- ⇒ **Algoritmos basados en estéreo trinocular**

- ⇒ **Conjetura BRPS :**
  - ⇒ **“el proceso de combinación binocular humana relaciona de forma integral la extracción de información de disparidad con la construcción de los elementos del esquema primitivo en bruto”**
  - ⇒ **Elementos fundamentales:**
    - ⇒ **Utiliza los zero-crossings y los picos del filtro LoG**
    - ⇒ **Restricción de continuidad figural**
    - ⇒ **Límite del gradiente de disparidad**
    - ⇒ **Actividad de cruzada entre canales**

⇒ **Continuidad Figural:**

⇒ Establece la continuidad de las superficies formulada como un agrupamiento local binocular de correspondencias de pasos por cero o picos adyacentes con el mismo signo del contraste a lo largo de los contornos de las figuras (continuidad de pasos por cero).



$$d_1 = x_1^1 - x_1^2$$

$$d_2 = x_2^1 - x_2^2$$

$$m_1, m_2 \in C_{ob} \quad |\bar{m}_1 - \bar{m}_2| < \Delta_a \Rightarrow |d_1 - d_2| < \Delta_b$$

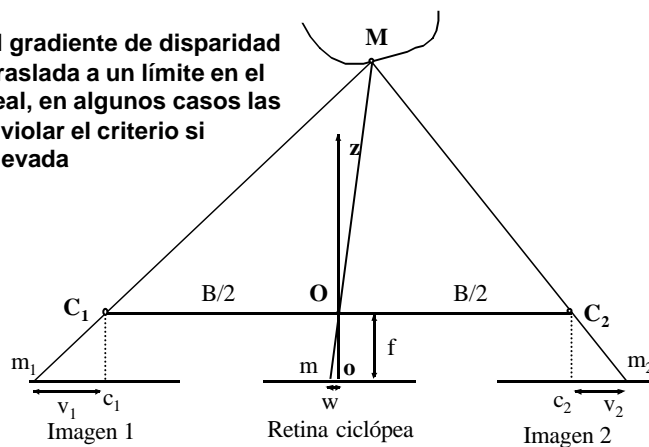
⇒ **Límite del Gradiente de Disparidad**

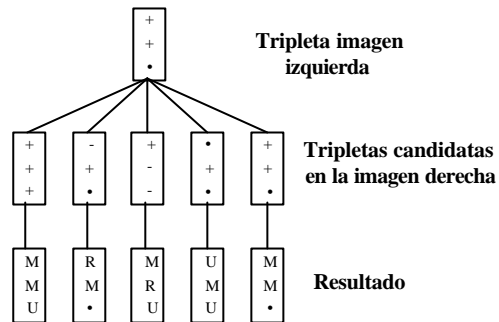
⇒ La fuerza de una correspondencia candidata es evaluada como la suma del soporte que recibe de todas las correspondencias potenciales del entorno de vecindad que satisfacen este criterio

⇒ A pesar de que el límite del gradiente de disparidad en el espacio ciclópeo se traslada a un límite en el gradiente del espacio 3D real, en algunos casos las superficies planas pueden violar el criterio si presentan una pendiente elevada

$$GD = \left| \frac{d^1 - d^2}{w^1 - w^2} \right|$$

$$GD < K$$



⇒ **Actividad de cruzada entre canales:**⇒ **Actividad de cruzada entre canales:**

- ⇒ Establece una correspondencia cruzada entre canales que requiere que un atributo de una primitiva con una disparidad dada, deba estar soportada por un atributo similar en otro canal de frecuencia espacial, dentro de un cierto rango de disparidad.
- ⇒ Como atributos utilizan el signo del contraste de los pasos por cero y los picos de la convolución de la imagen con el filtro LoG, codificando una tripleta para cada posición y para tres frecuencias espaciales diferentes
- ⇒ Cada tripleta representa, para una posición de la imagen, la medida de una primitiva a tres escalas diferentes, marcando +, - y • (nulo) para significar coincidencia, desavenencia o que no existe paso por cero
- ⇒ Para una tripleta de la imagen izquierda (superior), existe un conjunto de posiciones candidatas de correspondencia en la imagen derecha cada una de las cuales tiene asignada su propia tripleta (fila media)
- ⇒ La fila inferior representa el resultado de la correspondencia binocular entre canales. Una correspondencia correcta se marca como M (match), y las incorrectas como R (rivalous). Si solo una de las imágenes tiene primitiva para un cierto canal se marca como desconocida U (unknown); y si no existe en ninguna de las dos imágenes es ignorada (•)

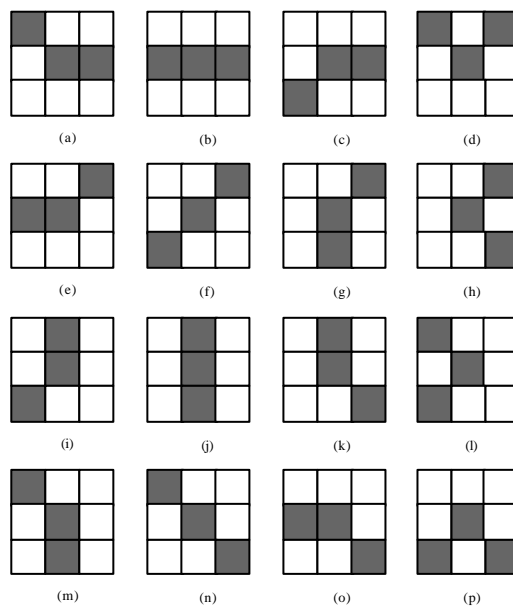
**⇒ Relajación:**

- ⇒ La restricción de unicidad en la correspondencia se propaga utilizando un esquema de relajación de forma que si dos primitivas  $a_i$  y  $b_j$  tienen el mayor soporte de correspondencia entre las listas de candidatos, entonces es considerada una correspondencia correcta
- ⇒ El soporte de las correspondencias vecinas se evalúa como la suma de aquellas que cumplen el límite del gradiente de disparidad

- ✓ Técnicas basadas en Area (“area-based”)
- ✗ Técnicas basadas en Características (“feature-based”)
  - ⇒ Teoría computacional de Marr-Poggio
  - ⇒ Teoría computacional de Mayhew-Frisby
  - ⇒ Algoritmo de relajación de Kim y Aggarwal
  - ⇒ Técnicas basadas en segmentos de borde
- ⇒ Técnicas Jerárquicas
- ⇒ Programación Dinámica
- ⇒ Algoritmos basados en estéreo trinocular

- ⇒ **Proponen un esquema de relajación que combina tres restricciones para eliminar las ambigüedades:**
  - Restricción de continuidad de la disparidad, utilizada en el modelo de Marr y Poggio
  - Restricción de continuidad figural, introducida en el modelo computacional de Mayhew y Frisby
  - Restricción de suavidad de las probabilidades de correspondencia, introducida este modelo.
- ⇒ **Considera un sistema binocular de ejes alineados, utilizando puntos de borde como primitivas características, extraídas mediante la localización de los pasos por cero del operador Laplaciano Gaussiano LoG**
- ⇒ **Cada primitiva se clasifica de acuerdo a 16 patrones de paso por cero. Se establece una medida de similitud entre los dos puntos de paso por cero candidatos, dependiendo del patrón que rodea a cada uno de ellos**

16 Patrones de paso por cero





- ⇒ Utiliza como nodos  $\{a_i\}$  aquellos pasos por cero que no tienen patrones con componente horizontal. A cada nodo se le asigna un conjunto de etiquetas  $L_i=\{l_j\}$  y una probabilidad  $p_i(l_j)$  de correspondencia entre el nodo  $\{a_i\}$  en el punto  $Z_l(x_i, y_i)$  en la imagen izquierda y el punto  $Z_d(x_j, y_j)$  en la imagen derecha
- ⇒ Para cada nodo  $\{a_i\}$  con disparidad  $l_j$  se define una función de ponderación  $w_i(l_j)$  basada en la similaridad de los patrones de paso por cero así como de la diferencia en los gradientes de intensidad. La probabilidad inicial  $p_i^0(l_j)$  se calcula mediante las funciones de ponderación  $w_i(l_j)$
- ⇒ Sobre el mapa de pasos por cero se construye una matriz tridimensional con las probabilidades de correspondencia para cada nodo en el tercer eje. Los índices de la matriz están formados el punto de la imagen izquierda  $Z_l(x_i, y_i)$  ,y la disparidad  $l_j$  correspondiente al punto candidato de la imagen derecha  $Z_d(x_j, y_j)$

- ⇒ El procedimiento iterativo implementa:
  - Continuidad de la disparidad
  - Continuidad Figural
  - Suavidad de la probabilidad de correspondencia

$$p_i^{n+1}(l_j) = p_i^n(l_j) + \underset{\substack{\uparrow \\ \text{Componente} \\ \text{Excitatoria}}}{c \cdot F(p_i^n(l_j)) \cdot (p_s^n)} - \underset{\substack{\uparrow \\ \text{Componente} \\ \text{Inhibitoria}}}{d \cdot p_i^n(l_j) \cdot I(P_{FS})}$$

**El procedimiento iterativo** Continuidad de la probabilidad

$$p_i^{n+1}(l_j) = p_i^n(l_j) \left[ + c \cdot F(p_i^n(l_j)) \cdot (p_s^n) - d \cdot p_i^n(l_j) \cdot I(P_{FS}) \right]$$

$\Uparrow$  Continuidad de la disparidad  $\Uparrow$

$F(p_i^n(l_j)) = \begin{cases} [p_i^n(l_j)]^2 & 0 \leq p_i^n(l_j) \leq 0.5 \\ p_i^n(l_j) * (1 - p_i^n(l_j)) & 0.5 < p_i^n(l_j) < 1 \end{cases}$ 

 $\leftarrow$  Reduce la tendencia a una convergencia rápida al valor de disparidad más probable

$I(P_{FS}) = \begin{cases} 0 & p_F^n + p_S^n \neq 0 \\ 1 & p_F^n + p_S^n = 0 \end{cases}$ 

 Implementa la continuidad figural  
 Si los vecinos no tienen correspondencias de alta probabilidad con disparidad  $l_j \pm 1$ , se activa

$p_F^n = \max[p_f^n(l_j - 1), p_f^n(l_j), p_f^n(l_j + 1)]$   
 $p_S^n = \max[p_s^n(l_j - 1), p_s^n(l_j), p_s^n(l_j + 1)]$ 

 $Z_i(x_f, y_f)$  y  $Z_i(x_s, y_s)$  son el primer y segundo paso por cero en la vecindad de  $Z_i(x_i, y_i)$  (de entre los ocho posibles vecinos)

- ✓ **Técnicas basadas en Area (“area-based”)**
- ✗ **Técnicas basadas en Características (“feature-based”)**
  - ⇒ Teoría computacional de Marr-Poggio
  - ⇒ Teoría computacional de Mayhew-Frisby
  - ⇒ Algoritmo de relajación de Kim y Aggarwal
  - ✦ **Técnicas basadas en segmentos de borde**
- ⇒ **Técnicas Jerárquicas**
- ⇒ **Programación Dinámica**
- ⇒ **Algoritmos basados en estéreo trinocular**

⇒ **Ventajas:**

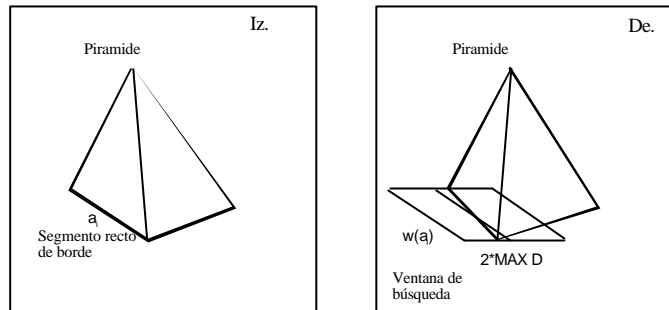
- ⇒ Cuando se agrupan los puntos de borde en segmentos lineales se reduce el efecto de errores posicionales en la detección de dichos puntos. La posición y orientación del segmento recto de bordes se ve menos afectada por errores aislados.
- ⇒ La localización de los segmentos rectos de borde puede realizarse con una resolución por debajo del nivel de pixel, mejorando la respuesta ante errores de discretización.
- ⇒ La restricción de conectividad entre los puntos de bordes en ambas imágenes, que debe ser impuesta explícitamente en los modelos basados en puntos de borde, es en este caso considerada de forma implícita en la construcción de las primitivas características.

⇒ **Inconvenientes:**

- ⇒ Plantea problemas debido a la posible fragmentación de la información correspondiente a un borde en diferentes segmentos en cada imagen
- ⇒ Esta fragmentación supone que un segmento de una imagen puede tener más de un segmento correspondiente en otra imagen, lo cual complica la restricción de unicidad

**(Medioni y Nevatia)**

- ⇒ Implementa la restricción de continuidad de la disparidad
- ⇒ Cada segmento se describe por las coordenadas de su puntos finales, orientación y contraste medio en la normal.
- ⇒ **Algoritmo:**
  - ⇒ Define los conjuntos  $A=\{a_i\}$  conjunto de segmentos de la imagen izquierda y  $B=\{b_j\}$  conjunto de segmentos de la imagen derecha
  - ⇒ Para cada segmento  $a_i$  en la imagen izquierda se define una ventana de búsqueda en la imagen derecha en forma de paralelogramo  $w(a_i)$



- Uno de los lados del paralelogramo es paralelo a  $a_i$  y el otro lado es un vector horizontal de longitud  $2 \cdot \text{MAX D}$  (límite superior de la máxima disparidad esperada)
- De forma similar para cada segmento  $b_j$  en la imagen derecha se define una ventana de búsqueda  $w(b_j)$  en la imagen izquierda
- Para asignar una correspondencia  $(a_i, b_j)$  es preciso que  $a_i$  esté dentro de la ventana  $w(b_j)$ , y que  $b_j$  caiga dentro de la ventana  $w(a_i)$

- Se considera que dos segmentos se solapan si desplazando cualquiera de ellos a lo largo de la dirección paralela a la línea epipolar llegan a intersectarse
  - Los segmentos  $a_i$  y  $b_j$  pueden corresponderse solo si se solapan, tienen un contraste de intensidad similar y tienen orientaciones similares.
  - $S(a_i) \cap w(a_i)$  representa el conjunto de todas las posibles correspondencias para el segmento  $a_i$  de la imagen izquierda
  - Un segmento  $a_i$  de una imagen puede tener dos o más segmentos correspondientes  $b_{i1}, b_{i2}, \dots, b_{im}$  en la otra imagen siempre que ninguno de los candidatos se solapen entre sí.
- ⇒ La función de evaluación de  $v^t(i, j)$  se calcula iterativamente para evaluar la bondad de cada correspondencia  $(a_i, b_j)$

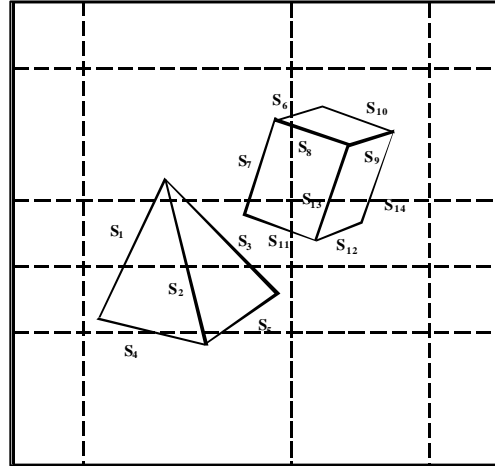
$$v^{t+1}(i, j) = \sum_{a_i \in w(b_j)} \min_{b_k \text{ verifica } C_1(a_k)} \lambda_{i \ jhk} \frac{|d_{hk} - d_{ik}|}{\text{card}(b_j)} + \sum_{b_j \in w(a_i)} \min_{a_n \text{ verifica } C_2(b_n)} \lambda_{i \ jhk} \frac{|d_{hk} - d_{ik}|}{\text{card}(b_j)}$$

$$v^{+1}(i, j) = \sum_{a_h \in w(b_j)} \min_{b_k \text{ verifica } C_1(a_h)} \lambda_{ijk} \frac{|d_{hk} - d_{ik}|}{\text{card}(b_j)} + \sum_{b_k \in w(a_i)} \min_{a_h \text{ verifica } C_2(b_k)} \lambda_{ijk} \frac{|d_{hk} - d_{ik}|}{\text{card}(b_j)}$$

- El término  $\lambda_{ijk}$  indica la longitud de solapamiento más pequeña para todas las parejas de correspondientes  $(a_i, b_j)$  y  $(a_h, b_k)$
- Las funciones  $\text{card}(a_i)$  y  $\text{card}(b_j)$  evalúan el número de segmentos en las ventanas  $w(a_i)$  y  $w(b_j)$  respectivamente
- La condición  $C_1(a_h)$  permite que  $a_i$  y  $a_h$  puedan ser correspondientes del mismo segmento  $b_j(=b_k)$  solo si  $a_i$  y  $a_h$  no se solapan, y viceversa para la condición  $C_2(a_h)$
- Esto permite la posibilidad de que si  $a_i$  y  $a_h$  son partes de un segmento fragmentado, puedan ser asignadas a un segmento no fragmentado en la otra imagen
- Esta restricción impone la condición de que los objetos 3D reconstruidos a partir de los segmentos de borde correspondientes, son suaves en su mayor parte.

- El término  $\lambda_{ijk}$  indica la longitud de solapamiento más pequeña para todas las parejas de correspondientes  $(a_i, b_j)$  y  $(a_h, b_k)$
- Las funciones  $\text{card}(a_i)$  y  $\text{card}(b_j)$  evalúan el número de segmentos en las ventanas  $w(a_i)$  y  $w(b_j)$  respectivamente
- La condición  $C_1(a_h)$  permite que  $a_i$  y  $a_h$  puedan ser correspondientes del mismo segmento  $b_j(=b_k)$  solo si  $a_i$  y  $a_h$  no se solapan, y viceversa para la condición  $C_2(a_h)$
- Esto permite la posibilidad de que si  $a_i$  y  $a_h$  son partes de un segmento fragmentado, puedan ser asignadas a un segmento no fragmentado en la otra imagen
- Para cada segmento  $a_h$  en la ventana  $w(b_j)$  (entorno de vecindad) una correspondencia preferida  $b_k$  se localiza cuando  $|d_{hk} - d_{ij}|$  alcanza un valor mínimo
- Durante la primera iteración, la selección de  $b_k$  para cada  $a_h$  se realiza a partir del conjunto completo  $S(a_h)$  ya que el conjunto de correspondencias preferidas está vacío
- Este algoritmo de correspondencia minimiza la diferencia de disparidad entre segmentos de línea correspondientes en un entorno de vecindad.
- Esta restricción impone la condición de que los objetos 3D reconstruidos a partir de los segmentos de borde correspondientes, son suaves en su mayor parte.

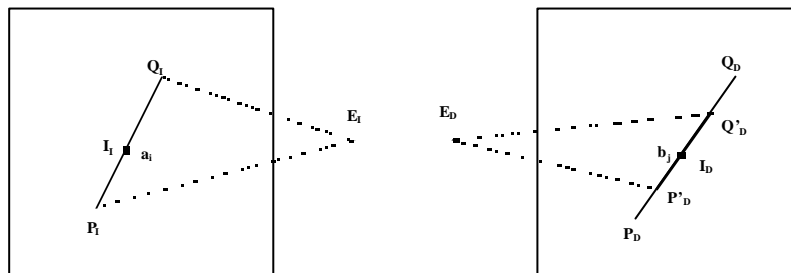
- ⇒ Basado en la restricción del límite del gradiente de disparidad
- ⇒ Grafo de vecindad para almacenar la información en función de la adyacencia de los segmentos rectos de borde en cada imagen
- ⇒ Cada segmento  $s_j$  tiene asociada una lista de vecinos que es obtenida a partir de la unión de agrupaciones de segmentos  $\{b_k\}$  enlazados a las ventanas  $\{w_k\}$  que intersectan dicho segmento  $s_j$



- ⇒ Se establece una etapa de correspondencia global utilizando una representación especializada de las correspondencias potenciales llamado **“grafo de disparidad”**
- ⇒ La idea de usar el grafo de disparidad es propagar las correspondencias dentro de los entornos de vecindad para recuperar subconjuntos de segmentos 3D que yacen sobre un trozo de superficie suave
- ⇒ Un par de segmentos  $a_i$  y  $b_j$  de las imágenes izquierda y derecha constituyen una correspondencia potencial. La correspondencia se establece por tanto como un proceso de criba de los posibles candidatos, criba que es realizada en dos etapas:
  - una primera etapa de *verificación local*,
  - una segunda de *verificación global* de las hipótesis planteadas

⇒ **Verificación local:**

- Se imponen las restricciones de correspondencia local. Los segmentos  $a_i$  y  $b_j$  constituirán una correspondencia potencial si satisfacen la restricción de *similaridad geométrica* (longitud y orientación) para segmentos rectos y sus puntos medios satisfacen la condición de epipolaridad
- Para el punto medio  $I_i$  de un segmento  $a_i$ , su punto correspondiente  $I_D$  es buscado sobre la línea epipolar cerca de un valor de disparidad esperado.



- La disparidad es definida en función de la distancia desde cada uno de los puntos correspondientes en ambas imágenes, al epipolo correspondiente.

$$d_{ij} = E_D I_D - E_I I_I$$

⇒ **Verificación global:**

- Proceso de predicción y propagación recursiva
- Se construye un grafo de disparidad cuyos *nodos* son las parejas de correspondencias potenciales  $(a_i, b_j)$  entre las imágenes izquierda y derecha, y los *enlaces* conectan parejas de nodos cuyos segmentos son adyacentes en sus respectivos grafos de vecindad
- La diferencia en la disparidad permisible entre nodos vecinos de parejas correspondientes se denomina *límite del gradiente de disparidad* y corresponde a una variación  $\bullet$  en profundidad
- Para cada nodo del grafo de disparidad  $(a_i, b_j)$ , se realiza una exploración recursiva de los grafos de vecindad de  $a_i$  y  $b_j$  para detectar parejas correspondientes que tienen disparidades dentro del intervalo de disparidad permisible
- De entre todas las correspondencias potenciales, se escoge aquella cuya disparidad está más próxima a la disparidad predicha (continuidad de las superficies)

- ✓ **Técnicas basadas en Area (“Area-Based”)**
- ✓ **Técnicas basadas en Características (“feature-based”)**
- ✗ **Técnicas Jerárquicas**
  - ⇒ **Algoritmo de Hoff y Ahuja**
  - ⇒ **Algoritmo de Marapane y Trivedi**
  
- ⇒ **Programación Dinámica**
- ⇒ **Algoritmos basados en estéreo trinocular**

- ⇒ **La utilización de una estructura jerárquica en el algoritmo de correspondencia estereoscópica supone la realización de dicho proceso a varios niveles del proceso de reconstrucción visual, intercambiando la información obtenida en cada nivel para imponer una consistencia global en el mapa de disparidad**
  - ⇒ **Algoritmo de Hoff y Ahuja**
  - ⇒ **Algoritmo de Marapane y Trivedi**

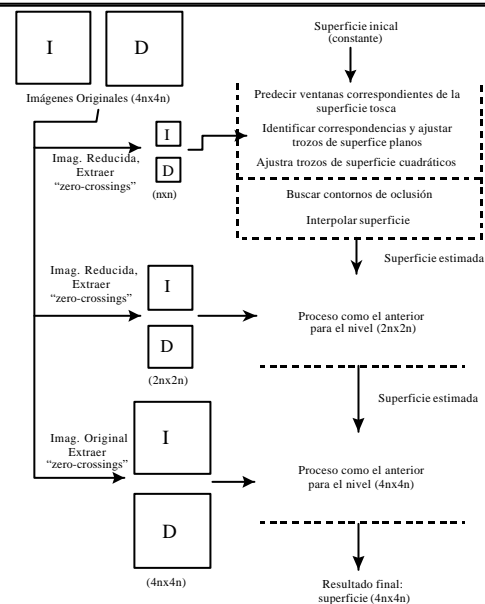


- ⇒ Integra las etapas de correspondencia estéreo y la interpolación de superficies (superficies curvadas suaves por trozos)
- ⇒ Proponen la integración de la correspondencia y el ajuste de los puntos 3D a una superficie en un único proceso, de forma de la corrección en la elección de las correspondencias pueda ser juzgada por el tipo de superficie que produce.
- ⇒ El algoritmo considera una pareja estéreo de  $4n \times 4n$  imágenes
- ⇒ Los puntos de borde son extraídos utilizando un operador Laplaciano Gaussiano ( $L^2G$ ) con tres resoluciones  $n \times n$ ,  $2n \times 2n$ , y  $4n \times 4n$ .

- ⇒ La correspondencia inicial se realiza en las dos direcciones “izquierda-derecha” y “derecha-izquierda”.
- ⇒ Para cada punto característico  $P_i$  en la imagen izquierda, se selecciona un conjunto de puntos correspondientes candidatos  $\{Q_j\}$  en función de la similitud de las propiedades locales geométricas de cada punto
- ⇒ A continuación se ajusta un conjunto de funciones paramétricas, planares y cuadráticas a regiones de la imagen circulares centradas en cada punto de una rejilla  $(x,y)$

$$z_p(x, y) = a_1x + b_1y + c_1$$

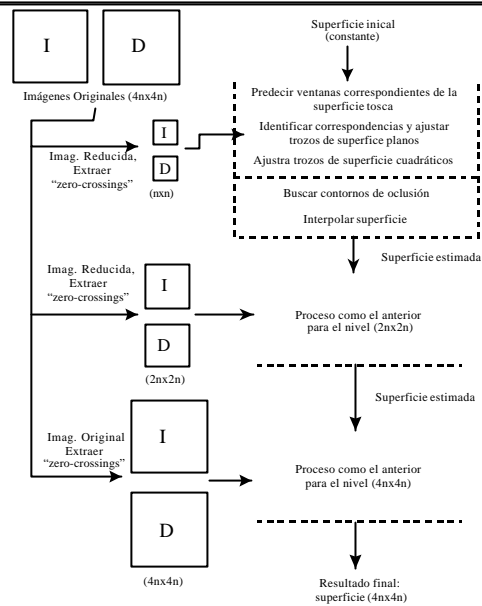
$$z_q(x, y) = a_2x^2 + b_2y^2 + c_2xy + d_2x + e_2y + f_2$$



### Algoritmo de Hoff y Ahuja (3)

88

- ⇒ En primer lugar se escogen hasta dos trozos de superficie plana en cada punto de la rejilla que ofrecen el mejor ajuste por mínimos cuadrados con la disparidad observada  $z_i$ .
- ⇒ En un segundo paso, se ajustan trozos de superficie cuadráticas para cada punto de la rejilla con las combinaciones de correspondencias anteriores.
- ⇒ La superficie cuadrática que contiene la mayor parte de los puntos es asignada como mejor ajuste para dicho punto de la rejilla.



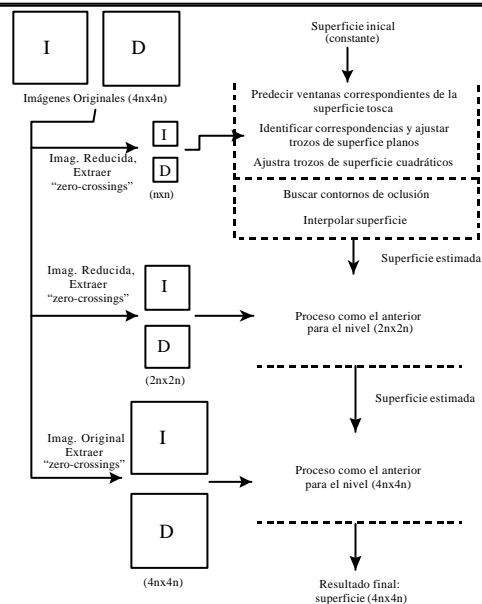
Visión Tridimensional

UPM-DISAM / UMH

### Algoritmo de Hoff y Ahuja (4)

89

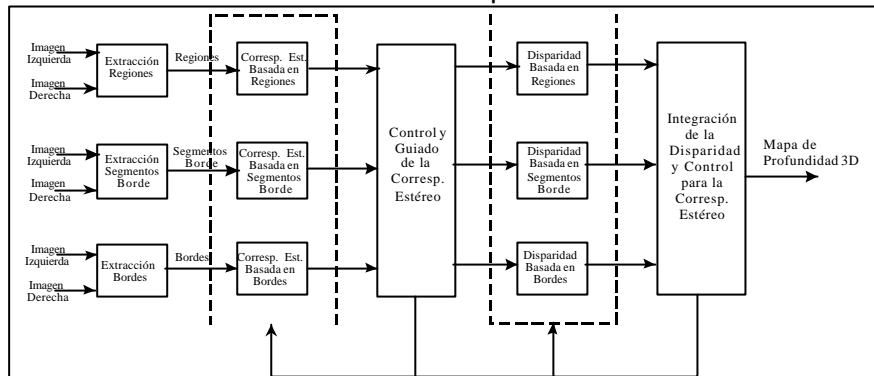
- ⇒ Una vez ajustado el trozo de superficie, se calculan la profundidad y orientación de los contornos ajustando un trozo con doble superficie plana calculando las discontinuidades entre ambas mitades.
- ⇒ Finalmente se interpola una superficie suave en zonas alejadas de los contornos para generar un mapa de superficie para cada resolución, suave a trozos.
- ⇒ La correspondencia de los bordes para resoluciones más finas es guiada por la superficie interpolada a una resolución más tosca



Visión Tridimensional

UPM-DISAM / UMH

- ⇒ Implementado en varias etapas, utilizando primitivas múltiples, y bajo una estrategia de control jerárquica (Algoritmo MPH). El algoritmo consta de tres subsistemas integrados:
  - ⇒ Un módulo de análisis basado en regiones.
  - ⇒ Un módulo de análisis basado en segmentos de borde lineales.
  - ⇒ Un módulo de análisis basado en primitivas de borde.



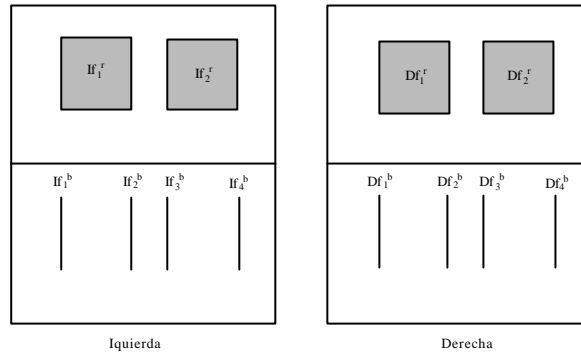
- ⇒ Los resultados del análisis en los niveles más altos de la jerarquía son utilizados para guiar el análisis en los niveles inferiores
- ⇒ El algoritmo no está ligado a un único tipo de primitiva, y por tanto funciona dentro de un rango de tipos de escenas más amplio
- ⇒ El resultado del algoritmo es un conjunto de mapas de disparidad para diferentes niveles de abstracción que son fusionados para obtener un mapa más exacto y con una mayor resolución
- ⇒ El algoritmo también provee la capacidad para analizar selectivamente regiones de la imagen con un detalle variable.

- ⇒ Utiliza primitivas de alto nivel que contienen una información mucho más rica y con mayor capacidad de discriminación.
- ⇒ La utilización de regiones presenta una mayor robustez frente al ruido, siendo su detección más fiable, y por otro lado facilitan el proceso de correspondencia al incorporar las restricciones de continuidad y suavidad
- ⇒ Para realizar el emparejamiento de regiones se utilizan las propiedades de:
  - valor medio del nivel de gris,
  - área, mínimo rectángulo conteniendo la región,
  - centroide, longitud del perímetro, eje principal de inercia ( $PA$ ),
  - ancho a lo largo de  $PA$ , altura perpendicular a  $PA$ , y razón ancho-alto.

- ⇒ *Restricciones jerárquicas* :
  - ⇒ *Restricciones jerárquicas espaciales (HSC)*: se utilizan para el guiado del análisis estéreo
    - *Restricción jerárquica de herencia (HINH)*
    - *Restricción jerárquica relacional (HREL)*
    - *Restricción jerárquica de intervalo (HINT)*.
  - ⇒ *Restricciones jerárquicas de disparidad (HDC)*: limitan la disparidad de las correspondencias potenciales generadas anteriormente.

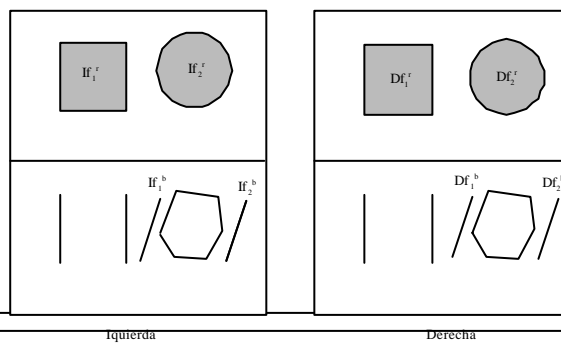
⇒ *Restricción jerárquica de herencia (HINH)*

- ⇒ Impone la restricción de que las primitivas preservan su relación herencia en ambas imágenes, entre los diferentes niveles de abstracción.
- ⇒ Si una borde en una imagen procede o es parte de una región, esta pertenencia o herencia se debe reflejar en los dos niveles de abstracción de la otra imagen, para establecer la correspondencia.



⇒ *Restricción jerárquica relacional (HREL)*

- ⇒ Extiende la restricción de orden utilizada en la correspondencia de primitivas con un solo nivel.
- ⇒ Las primitivas correspondientes en cualquier nivel de la jerarquía deben preservar cualquier relación estructural que exista entre las primitivas de las cuales proceden
- ⇒ La consistencia estructural se chequea, para cada nivel, una vez que se ha establecido la correspondencia, bien por que se ha realizado en niveles superiores, o bien por que se han establecido todas las correspondencias potenciales dentro del mismo nivel donde se realiza la correspondencia.



⇒ **Restricción jerárquica de intervalo (HINT)**

- Se impone sobre aquellas primitivas que no proceden de otra de nivel superior y por tanto no es aplicable la restricción de herencia
- En esta caso se pueden utilizar las correspondencias de las primitivas del nivel superior para limitar el intervalos de búsqueda de posibles candidatos correspondientes, limitando de este modo el proceso de búsqueda.

⇒ **Restricciones jerárquicas de disparidad (HDC)**

- Impone la similaridad entre las disparidades obtenidas en la correspondencia a distintos niveles para aquellas primitivas que proceden de correspondencias de un nivel superior
- Esta restricción no descarta posibles candidatos sino que genera una medida de la consistencia de la disparidad.

✓ **Técnicas basadas en Area (“Area-Based”)**

✓ **Técnicas basadas en características (“feature-based”)**

✓ **Técnicas Jerárquicas**

✗ **Programación Dinámica**

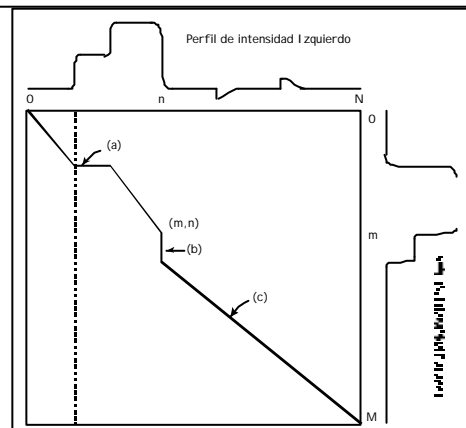
⇒ **Algoritmo de Otha y Kanade**

⇒ **Algoritmos basados en estéreo trinocular**

- ⇒ Plantean el problema de correspondencia estereoscópica como el trazado de un camino sobre un espacio de búsqueda bidimensional, que minimiza algún tipo de funcional
- ⇒ Para un sistema estereoscópico de imágenes de ejes alineados (o bien un sistema estereoscópico rectificado), los puntos correspondientes deben ser buscados dentro la misma línea horizontal, pudiendo definir un espacio bidimensional cuyos ejes son las líneas de rastreo de las imágenes izquierda y derecha
- ⇒ El problema subyace por tanto en la definición de un camino sobre dicho espacio que enlace cada punto de la línea de la imagen izquierda con un punto de la imagen derecha (búsqueda intra-scanline), así como las relaciones figurales entre primitivas de diferentes líneas de rastreo (búsqueda inter-scanline).
- ⇒ Los algoritmos se basan en la asunción de preservación de orden en los bordes de un par estereoscópico de imágenes.

Algoritmo de Otha y Kanade (1)

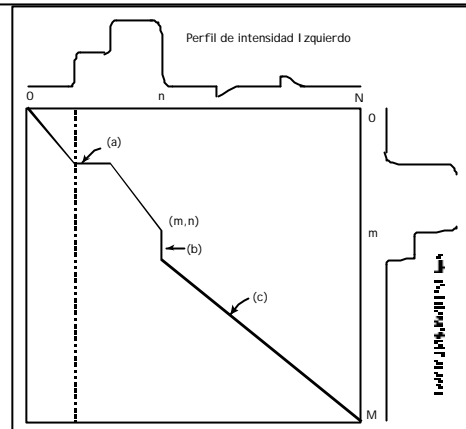
- ⇒ Utiliza los niveles de intensidad dentro de los intervalos delimitados por puntos de borde para guiar la búsqueda de correspondencia dentro de la línea ("intra-scanline")
- ⇒ La búsqueda dentro de la línea es formulada como un trazado del camino óptimo que se realiza mediante una función de coste asociada a cada tramo parcial, basado en la variación de los niveles de intensidad dentro de los intervalos de línea de rastreo que deben establecer una correspondencia
- ⇒ Los bordes se numeran de izquierda a derecha en cada línea, considerándose como nodos



## Algoritmo de Otha y Kanade (2)

100

- ⇒ Si existen  $N$  nodos en la línea izquierda y  $M$  nodos en la línea derecha, la solución para la búsqueda dentro de la línea de rastreo está formada por una secuencia de segmentos rectos desde el nodo  $(0,0)$  al nodo  $(M,N)$  con el mínimo coste
- ⇒ El coste del camino óptimo desde el nodo  $(0,0)$ , se representa mediante  $D(m,n)$  y es la suma de los costes asociados a cada segmento del camino
- ⇒ El coste del camino óptimo se obtiene recursivamente añadiendo el de cada segmento de camino nuevo al camino óptimo parcial existente.



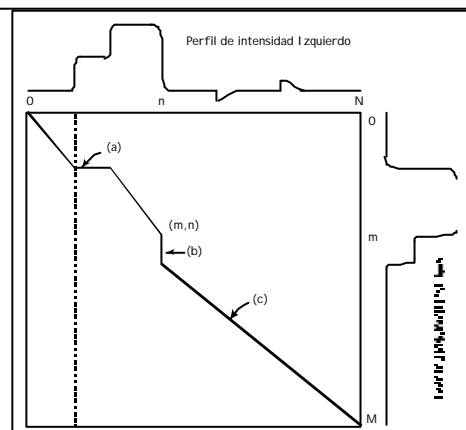
Visión Tridimensional

UPM-DISAM / UMH

## Algoritmo de Otha y Kanade (3)

101

- ⇒ El cálculo de la función de coste para cada tramo del camino (intervalo entre dos bordes sucesivos  $m$  y  $k$ ) viene dado por  $d(m,k)$
- ⇒ Este funcional se calcula en base a la similitud de los niveles de intensidad entre intervalos borde a borde de ambas líneas de rastreo
- ⇒ En caso de oclusiones los segmentos de camino serán horizontales o verticales
- ⇒ El camino vertical  $(b)$  corresponde a una parte visible en la imagen derecha pero no visible en la imagen izquierda
- ⇒ El camino horizontal  $(a)$  en cambio corresponde a una parte no visible en la imagen derecha
- ⇒ En el camino  $(c)$  se ignoran los bordes existentes hasta el nodo final (bordes que no tienen correspondientes a causa del ruido).



Visión Tridimensional

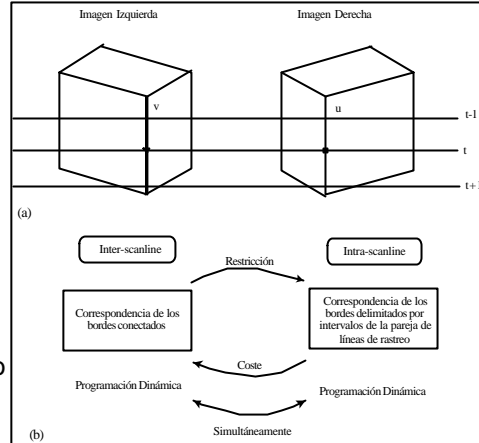
UPM-DISAM / UMH



## Algoritmo de Otha y Kanade (4)

102

- ⇒ Los resultados de esta búsqueda dentro de la línea de rastreo se utilizan para establecer una consistencia global entre correspondencias entre líneas vecinas ("inter-scanline")
- ⇒ La consistencia global utiliza la conectividad vertical entre los bordes pertenecientes a diferentes líneas de rastreo para eliminar las ambigüedades en la correspondencia
- ⇒ Si un punto en el borde conectado  $u$  de la imagen izquierda corresponde con un punto del borde conectado  $v$  de la imagen izquierda en la línea  $t$ , entonces otros puntos de estos bordes para otras líneas de rastreo deben también corresponder



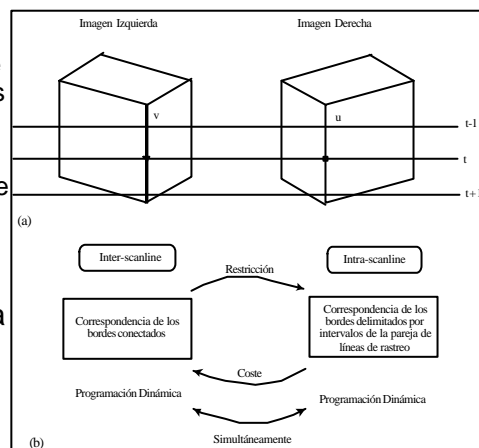
Visión Tridimensional

UPM-DISAM / UMH

## Algoritmo de Otha y Kanade (5)

103

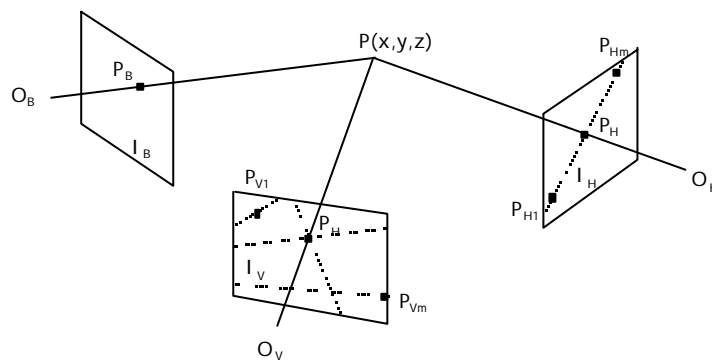
- ⇒ La búsqueda se realiza dentro de un espacio tridimensional de nodos. Cada nodo 3D está formado por una colección de nodos 2D conectados a través de las líneas de rastreo
- ⇒ El camino óptimo en el espacio tridimensional se obtiene recursivamente de forma similar a la búsqueda bidimensional.
- ⇒ Se establecen por tanto dos procesos de búsqueda
  - ⇒ El otro establece la correspondencia de bordes (más concretamente, de intervalos delimitados por bordes) en las líneas de rastreo izquierda y derecha
  - ⇒ Uno de ellos establece la correspondencia de todos los bordes conectados en las imágenes izquierda y derecha



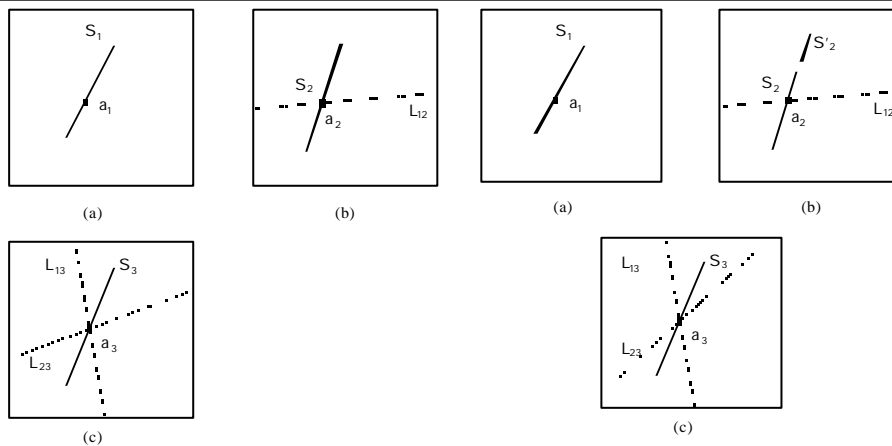
Visión Tridimensional

UPM-DISAM / UMH

- ✓ Técnicas basadas en Area (“Area-Based”)
- ✓ Técnicas basadas en características (“feature-based”)
- ✓ Técnicas Jerárquicas
- ✓ Programación Dinámica
- ✗ Algoritmos basados en estéreo trinocular
  - ⇒ Algoritmos sobre puntos de borde
  - ⇒ Algoritmos sobre segmentos de borde



- ⇒ Considera tres cámaras con centro de proyección  $O_B$ ,  $O_H$  y  $O_V$  en posiciones conocidas y con ejes ópticos de orientación conocida
- ⇒ La restricción de epipolaridad establece para cualquier punto  $P_B$  en el plano de imagen  $I_B$  sendas rectas epipolares  $R_{BH}$  y  $R_{BV}$  sobre los planos de imagen  $I_H$  e  $I_V$  respectivamente
- ⇒ Suponiendo que existen múltiples puntos candidatos  $\{P_{H1}, P_{H2}, \dots, P_{Hm}\}$  sobre la recta epipolar  $R_{BH}$ , se construye un conjunto de rectas epipolares  $\{R_{HVj}\}$  sobre  $I_V$ , a partir de cada punto candidato del plano de imagen  $I_H$
- ⇒ El punto correspondiente es comprobado en la intersección de  $R_{BV}$  con  $\{R_{HVj}\}$  ( $P_V$ ) eliminado la ambigüedad en la correspondencia
- ⇒ Cada tripleta  $(P_{Bj}, P_{Hj}, P_{Vj})$  es comprobada teniendo en cuenta la similaridad de los atributos de la primitiva, considerando la mejor correspondencia
- ⇒ En algunos casos la ambigüedad de la correspondencia puede persistir, seleccionándose en este caso el punto candidato que produce una disparidad más próxima a la de los puntos en un su entorno de vecindad (restricción de continuidad de la disparidad).



- ⇒ Ayache y Lustman presentan un algoritmo que extiende la técnica de correspondencia binocular sobre segmentos rectos de borde mediante la utilización de una geometría trinocular
- ⇒ El algoritmo emplea un esquema de predicción y verificación, usando grafos de vecindad de los segmentos rectos de borde en tres imágenes como una extensión del algoritmo binocular.
- ⇒ **Para un segmento**  $S_1$  en la imagen 1, si la tripleta  $(S_1, S_2, S_3)$  satisface la restricción epipolar trinocular (líneas epipolares  $L_{12}, L_{23}, L_{13}$ ) y presenta suficiente similitud en las propiedades geométricas locales, entonces es retenida como una correspondencia potencial.
- ⇒ Debido al ruido en el preprocesamiento se puede producir una fragmentación de los segmentos rectos de borde entre las diferentes imágenes
- ⇒ Este problema es manejado flexibilizando el orden de las imágenes para generar las hipótesis de correspondencia

- ⇒ **El problema de Correspondencia**
- ⇒ **Restricciones aplicadas a la correspondencia**
- ⇒ **Técnicas de correspondencia**
- ⇒ **Conclusiones**

- ⇒ **Existen diferentes estrategias de correspondencia**
- ⇒ **Según el tipo de información utilizada:**
  - ⇒ **Niveles de intensidad en entornos de vecindad (Métodos basados en área)**
  - ⇒ **Primitivas de alto nivel (Métodos basados en características)**
  - ⇒ **Jerarquía en la organizativa de la escena (Métodos jerárquicos)**
- ⇒ **Diferentes técnicas que tienen ciertos elementos comunes:**
  - ⇒ **Proceso de optimización (Relajación)**
  - ⇒ **Uso de información estructural de los objetos (continuidad de superficie,..)**
- ⇒ **Existen cierto porcentaje de falsas correspondencias**