

**Visión Estereoscópica  
Búsqueda del Conjunto Mínimo de Puntos  
Homólogos Para Rotorectificar**

Integrantes:

**Judith Donayo  
Teresa Arcomano**

Director:

**José Francisco Zelasco**

*2000*

Departamento de Computación  
Facultad de Ciencias Exactas y Naturales  
Universidad de Buenos Aires

## **ÍNDICE**

---

<b>RESUMEN .....</b>	1
<b>INTRODUCCIÓN .....</b>	2
<b>Capítulo 1 –Modelo de Sistema de Visión Estereoscópica Para AUV</b>	
1.1. Introducción .....	4
1.2. Introducción a la Estereoscopía .....	5
1.2.1. El Haz Perspectivo .....	5
1.2.2. Intersección de dos Haces Perspectivos .....	6
1.2.2.1. Elementos Internos .....	7
1.2.2.1.1. Definición de Imagen.....	8
1.2.2.2. Elementos Externos .....	9
1.2.2.3. Identificación de Rayos Homólogos.....	9
1.2.2.4. El Caso Normal .....	9
1.2.2.5. El Problema de la Paralaje Transversal .....	11
La Pseudo Intersección.....	
1.3. Búsqueda de Puntos homólogos en Imágenes no Rotorectificadas .....	13
<b>Capítulo 2 – Fundamentos de la Solución Propuesta</b>	
2.1. Introducción .....	15
2.2. Determinación del Entorno de Búsqueda.....	16
2.2.1. Recta Epipolar Conjugada.....	16
2.2.2. Coordenadas Aproximadas del Punto.....	17
2.2.3. Cooperación de la recta epipolar y el entorno de búsqueda. Estrategias de Control.....	18
2.3. Los Modelos del Fenómeno.....	18
2.3.1. Modelo Funcional.....	18
2.3.2. Modelo Estocástico.....	18
<b>Capítulo 3 - Solución Propuesta</b>	
3.1. Introducción .....	19
3.2. El Modelo Funcional de la Solución Propuesta.....	19
3.2.1. Cálculo de Diferencias.....	19
3.2.2. Matriz de Cuadrados Mínimos .....	23
3.3. Etapas en la Implementación Para la Solución Propuesta .....	27
3.4. Geometría de la Solución Propuesta .....	28
3.4.1. Rototraslación de la Secuencia .....	28
3.4.2. Determinación de la Recta Epipolar Conjugada en la Secuencia .....	29
3.4.3. Determinación del Entorno del Punto en la Secuencia .....	31
<b>Capítulo 4 - Aspectos de Implementación</b>	
4.1. Introducción .....	33
4.2. Tratamiento de las Imágenes - Cálculo de las Características.....	33
4.2.1. Detección de Contornos .....	34
4.2.2. Detección de Esquinas .....	35
4.3. Obtención de Pares Homólogos en Imágenes Simultáneas .....	36

## ÍNDICE

---

4.4. Búsqueda del Conjunto Mínimo de Puntos Homólogos .....	36
4.4.1. Reducción del Espacio de Búsqueda.....	36
4.4.2. Apareo: Algoritmo en Imágenes No Simultáneas .....	37
4.5. Características del Algoritmo.....	37
<b>Capítulo 5 – Conclusiones y Perspectivas</b>	
5.1. Resultados .....	38
5.2. Comparación de Resultados.....	42
5.3. Conclusiones .....	42
5.4. Otra Aplicación.....	43
5.5. Problemas Abiertos .....	43
<b>Anexo A: Geometría Relativa a la Automatización de la Visión Estereoscópica</b>	
A.1. Estereoscopía .....	45
A.1.1. El Caso General .....	47
A.1.2. El Problema Estereoscópico General .....	49
A.2. Rotaciones en el Espacio.....	49
A.2.1. Otras Propiedades de las Matrices de Rotación.....	51
A.3. Rotorectificación.....	57
A.3.1. Algoritmos de Búsqueda Lineal.....	58
A.3.2. Linearización del Problema de la Rotación Relativa .....	58
<b>Anexo B: Estereopsis Automática. Panorama del Estado del Arte</b>	
B.1. Introducción .....	65
B.2. Aspectos del Problema.....	66
B.2.1. Presentación de Dificultades .....	66
B.2.2. Primitivas .....	67
B.2.3. Restricciones .....	69
B.3. Estrategias y Soluciones Habituales al Problema del Apareo.....	71
B.3.1. Estrategias de Control .....	72
B.3.2. Técnicas de Búsqueda de Correspondencias .....	73
B.4. Clasificación: Algoritmos Basados en Áreas y en Características .....	75
B.4.1. Basados en Áreas .....	76
B.4.2. Basados en Características .....	76
B.4.3. Algoritmos Basados en Áreas Vs. Basados en Características .....	77
B.5. Tendencias .....	78
<b>Publicaciones</b> .....	79
<b>Referencias Bibliográficas</b> .....	83
<b>C.1. Bibliografía de Consulta</b> .....	84

## RESUMEN

Este trabajo ha sido desarrollado en el marco de un proyecto de estéreo visión para vehículos autónomos submarinos (AUV) provistos con sensores ópticos, simulando el sistema de visión binocular humana y que dio lugar a varias líneas de trabajo. Para obtener un modelo numérico de una escena subacuática de mayor precisión que el que se obtendría a partir de un par de imágenes simultáneas, provistas por los sensores ópticos, se trata de recalcularlo a partir de dos imágenes tomadas en forma sucesiva. En efecto, el aumento de la distancia entre los puntos de vista mejoraría dicha precisión. En el caso de tomas sucesivas, las líneas de imágenes no coinciden con las líneas epipolares como es el caso de las imágenes simultáneas desde los dos sensores calibrados. Esto complica la búsqueda de puntos homólogos respecto de las imágenes simultáneas, problema este último que ha sido exitosamente resuelto como resultado del trabajo de otra línea de las mencionadas [1]. Los parámetros de rotación relativa y traslación de dos imágenes sucesivas son conocidos en forma aproximada. Conocerlos en forma precisa permite realizar la roto rectificación de las imágenes y así utilizar los resultados de los trabajos mencionados. A partir del conocimiento de un cierto número de pares de puntos homólogos, entre cinco y diez, y como resultado de otra línea de trabajo [2], se ha resuelto la forma de conocer con precisión suficiente los parámetros de rotación y realizar una correcta roto rectificación. El objetivo de este trabajo es proveer esos puntos homólogos necesarios. Para ello es posible conocer la posición de las líneas epipolares en ambas imágenes sucesivas que por supuesto no coincidirán con las líneas de imagen, y tener el conocimiento aproximado de las perspectivas de ciertos puntos objeto en ambas imágenes sucesivas. Con estos elementos que limitan apropiadamente el espacio de búsqueda y basándose en un algoritmo que realiza una compensación por cuadrados mínimos se determinan los puntos homólogos buscados en éstas imágenes que, como se ha indicado, no satisfacen la condición de epipolaridad de sus líneas.

## ABSTRACT

This work has been developed in the frame of a project of stereo vision for autonomous underwater vehicles (AUV) provided with optic sensors, simulating the system of human binocular vision and that gave an origin to several work lines. To obtain a numeric model of a underwater scene of more precision than the one that would be obtained starting from a couple of simultaneous images provided by the optic sensors, it is re-calculate starting from two images taken in successive form. Really, the increase of the distance among the points of view would improve this precision. In case of successive takes, the lines of images don't coincide with the epipolar lines as in case of simultaneous images from the two gauged sensors. This complicates the search of homologous points respect the simultaneous images, this last problem has been successfully solved as a result of another work line of those mentioned [1]. The parameters of relative rotation and translation of two successive images are knowing in approximate form. To know them in precise form allows carrying out the roto-rectification of the images and using the results of the mentioned works. Starting from the knowledge of a certain number of couples of homologous points, between five and ten, and as a result of another work line [2], it has been solved the form of knowing with enough accuracy the rotation's parameters and achieve a correct roto-rectification. This work objective is providing those necessary homologous points. For these is possible to know the position of the epipolar lines in both successive images that of course won't coincide with the image lines, and achieve the approximate knowledge of the perspectives of certain object points in both successive images. With these elements that fittingly limit the search space, and based on an algorithm that carries out a compensation for minimum squares, the searches homologous points are determined in these images that, like it has been indicated, don't satisfy the epipolar condition of their lines.

# Introducción

A causa de la creciente automatización de tareas la necesidad de mejorar los sistemas de visión artificial ha produciendo un continuo desarrollo de nuevas herramientas para tratar imágenes estereoscópicas. El problema de la estereopsis automática (stereo matching), apareo o reconocimiento de puntos homólogos, puntos de dos o más imágenes que son perspectivas de un mismo punto en la escena, está dando lugar a un importante esfuerzo en investigación en estereoscopía digital. Se lo considera, generalmente, parte de los temas referentes a la puesta en correspondencia [3][4][5][6][7][8][9].

Este proceso, esencial en visión estéreo, permite, integrado con un aspecto geométrico de posicionamiento de las imágenes en el espacio, la reconstrucción tridimensional, que consiste en determinar las coordenadas de la superficie de un objeto a partir del conocimiento de puntos homólogos es decir, la conformación de un modelo numérico. Es de notar que las imágenes pueden haber sido tomadas o no simultáneamente.

En el marco del proyecto “Visión artificial en medio submarino y para AUV” y considerando un móvil submarino autónomo dotado de visión binocular que permite la obtención de dos imágenes simultáneas, se trata de mejorar la precisión en la obtención de las coordenadas de una escena calculando las mismas a partir de dos imágenes sucesivas. En efecto, se dispone de algoritmos satisfactorios para el apareo de puntos homólogos de dos imágenes simultáneas tomadas por los sensores del vehículo y cuyas posiciones relativas están calibradas [10][11][12][13]. Esta calibración implica que las líneas de ambas imágenes coinciden con las líneas epipolares (intersección de los planos imagen con el plano determinado por un punto de la escena y los dos puntos de vista). A partir de esta información y contando con ciertos parámetros es posible realizar el cálculo del modelo numérico de una escena. Sin embargo, para modelizar de alguna manera la respuesta perceptual, es decir, tener información de referencia y con el objeto de mejorar la precisión aumentando la distancia entre los puntos de vista de las dos imágenes, puede resultar interesante encontrar puntos homólogos tomando, v.g., la imagen derecha de dos tomas sucesivas. En este caso seguramente las líneas de imagen no coinciden con las líneas epipolares. Si se quiere utilizar algoritmos de probada eficiencia, para los que es necesario contar con la condición de epipolaridad de las imágenes, es necesario roto rectificar este nuevo par de imágenes .

En el caso de imágenes de líneas no epipolares, la búsqueda de puntos homólogos resulta más compleja y la literatura es bastante menos abundante. Para asegurar la calidad de resultados y limitar razonablemente el espacio de búsqueda se debe contar, además, con valores aproximados de ciertos parámetros ligados a la roto traslación relativa de las imágenes en cuestión. Si a esta información se le agrega, el conocimiento aproximado de las perspectivas de ciertos puntos del objeto a partir de las tomas simultáneas, se puede calcular, por un lado la posición de las líneas epipolares en las imágenes y por otro delimitar el conjunto de píxeles en cada imagen donde se deberá ubicar el par de puntos homólogos. El objetivo de este trabajo es, entonces, encontrar un número limitado de puntos homólogos en imágenes que no respetan la condición de epipolaridad de manera de mejorar los parámetros de rotación relativa. Esto permite la roto rectificación del par de imágenes en cuestión, con el objetivo de dejar el par de imágenes en condiciones que permitan realizar la búsqueda de puntos homólogos utilizando los algoritmos de probada eficiencia mencionados, y que como se ha dicho requieren la condición de epipolaridad.

El proyecto comporta las siguientes líneas de trabajo:

- Plataforma de prueba
- Estudio del estado del arte en identificación de imágenes estereoscópicas.
- Evaluación de resultados de algoritmos.
- Rotorectificación de imágenes.[14]
- Obtención de un conjunto mínimo de homólogos para rotorectificar [11].

Se podría completar con la modelización de la respuesta perceptual a los efectos de utilizar toda la información puesta en juego y acercarse, con un modelo más elaborado, a los mecanismos de la visión humana.

# Modelo De Sistema De Visión Estereoscópica Para Un AUV

## 1.1. Introducción

La visión estereoscópica automática o estereopsis automática, en visión por computadora, es un área de un interés particular. En efecto, su interés no se limita al reconocimiento de objetos como sucede con el tratamiento de imágenes monoculares [21] sino que se extiende a la obtención de un modelo numérico de la escena u objeto cuya perspectiva se encuentra en dos o más imágenes.

La identificación automática de puntos homólogos de dos o más imágenes de la misma escena con el objeto de obtener sus coordenadas 3D, que es objeto de desarrollo desde hace más de diez años [17], tiene un extenso campo de aplicaciones entre los que podemos mencionar:

- Conducción automática de vehículos, que realicen tareas pesadas, etc. que permiten mejorar de alguna manera la vida cotidiana de los seres humanos.
- Sistema de diagnóstico médico mediante imágenes, que permita visualizaciones tridimensionales de diferentes partes del cuerpo humano combinadas con resultados de tomografías o de resonancia magnética.
- Metroología industrial, inspección automatizada de piezas industriales, control de calidad.
- Aplicaciones militares
- Aplicaciones geodésicas, trazado de curvas de nivel a partir de imágenes estereoscópicas satelitales.

En particular para AUV:

- Prospección y explotación: plataformas submarinas para empresas petroleras.
- Problemas de cavitación: marina mercante y de guerra.
- Estudio de carenas: astilleros.
- Búsqueda de objetos hundidos: marina, prefectura
- Inspección de monoboyas: prefectura.
- Inspección de túneles de descarga.
- Recuperación de minas o su desactivación.

Dos aspectos son necesarios de resolver a los efectos de obtener las coordenadas de la superficie  $\Sigma$  de una escena u objeto grande:

1. Uno es cuantitativo o geométrico, es decir, corresponde al que se realiza contando con el conocimiento de los parámetros que definen y fijan la posición de los rayos perspectivos, y el de los parámetros que fijan los haces perspectivos correspondientes a cada imagen respecto del referencial general de la escena u objeto.
2. El otro es cualitativo o de identificación de cuales son los rayos homólogos de las distintas perspectivas.

Debido a que se dispone de una imagen es decir de un conjunto de perspectivas de los puntos de la escena, este último aspecto se puede resolver como la búsqueda de puntos homólogos en las imágenes. En efecto, si se reconocen en un par de imágenes los puntos perspectivos de un mismo punto de la escena se están identificando los rayos homólogos de los haces perspectivos.

Para esto se requiere automatizar eficazmente modelos apropiados de los mecanismos de la percepción visual.

La modelización de la visión binocular o estereoscópica parece un enfoque apropiado para resolver este problema. Se entiende por visión estereoscópica la que se produce cuando un individuo observa dos imágenes de una misma escena tomadas desde dos puntos de vista diferentes y ve un modelo estereoscópico, es decir, tiene sensación de relieve. Los procesos de estereoscopía se basan en fundamentos geométricos análogos a la visión humana. El principio es usar dos o más sistemas ópticos ubicados desde dos puntos de vista diferentes y proveer dos o más imágenes de una misma escena.

## 1.2. Introducción a la Estereoscopía

La percepción de relieve es una de las capacidades del sistema de visión humano. En él cooperan diferentes tipos de información tal como la distancia de puesta a punto, el ángulo de convergencia entre ambos ojos y la disparidad binocular.

La metroología con imágenes estereoscópicas se propone proveer con precisión conocida formas, dimensiones y orientación de objetos en el espacio utilizando medidas hechas en dos o más imágenes.

### 1.2.1. El Haz Perspectivo

Sea un objeto  $\sigma$  de superficie  $\Sigma$  y un punto  $S$  exterior a  $\Sigma$ , se llama *haz perspectivo*  $T$ , relativo a  $S$  y a  $\Sigma$  al conjunto de semirectas que unen  $S$  con cada uno de los puntos de  $\Sigma$  visibles desde  $S$ . El vértice  $S$  del haz es el *punto de vista* y cada elemento del haz es denominado *rayo perspectivo*.(fig. 1.1)



Figura 1.1 – Haz perspectivo

El conocimiento de  $S$  y  $\Sigma$  permite determinar  $T$ , pero el conocimiento de  $S$  y  $T$  no es suficiente para conocer  $\Sigma$ . La posición del punto  $M$ , intersección de  $\Sigma$  y el rayo perspectivo, está indeterminada.

### 1.2.2. Intersección de Dos Haces Perspectivos.

Sean dos haces perspectivos  $T_1$  y  $T_2$  relativo al mismo objeto  $\sigma$  de superficie  $\Sigma$  y dos puntos de vista  $S_1$  y  $S_2$  (fig. 1.2).

El espacio objeto estará referido a un sistema de coordenadas ortogonales  $OXYZ$ . A todo punto de  $\Sigma$  visible desde  $S_1$  y  $S_2$  corresponde un solo rayo perspectivo de  $T_1$  y un solo rayo perspectivo de  $T_2$ . Este par de rayos es llamado *homólogos*.

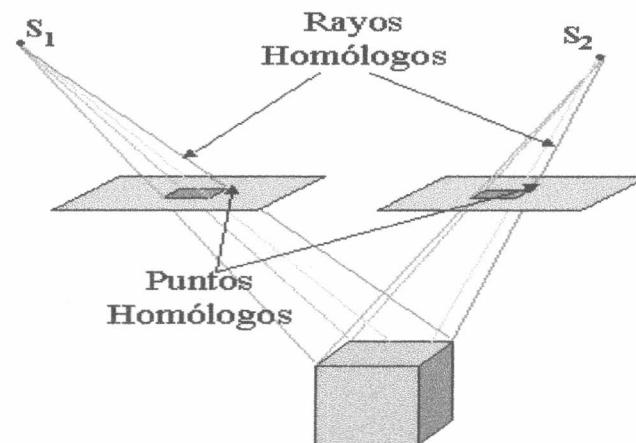


Fig. 1.2: Rayos y puntos homólogos

Recíprocamente el conocimiento de  $T_1$  y  $T_2$  permite determinar  $\Sigma$  como el lugar geométrico de los puntos  $M$  intersección de rayos homólogos.

El conocimiento de  $T_1$  y  $T_2$  implica:

1. Para cada uno se debe conocer los elementos internos del haz, los cuales no dependen de  $OXYZ$ .
2. Conocer la posición en el espacio de cada haz, los parámetros necesarios son los elementos externos, dependen del referencial adoptado.

### 3. Identificar los rayos homólogos.

Con estos tres puntos resueltos el problema de la determinación de  $\Sigma$  se reduce a encontrar la intersección de pares de rayos homólogos. Esto es conocido como *restitución*.

#### 1.2.2.1. Elementos Internos

La intersección de un plano  $\Pi$  con un haz  $T$  determina una *perspectiva de la superficie*.

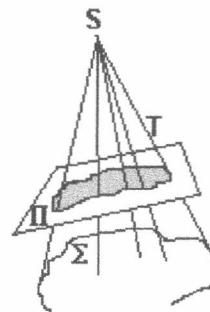


Figura 1.3 – Perspectiva de un objeto

La normal al plano perspectivo  $\Pi$  que pasa por el punto de vista se denomina *eje principal* y la intersección con el plano perspectivo es el *punto principal w* o *pie de perpendicular*. La distancia entre  $S$  y el *pie de perpendicular* es conocida como *distancia principal p*. (fig. 1.4)

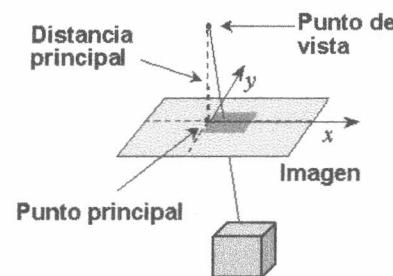


Fig. 1.4: Elementos internos

Si se conoce la perspectiva de un objeto, la *distancia principal* y la posición del punto principal, se puede reconstruir el haz perspectivo.

De acuerdo a esto, se tienen dos puntos de vista  $S_1$  y  $S_2$  que originan dos conjuntos de haces perspectivos  $T_1$  y  $T_2$  los cuales proyectan la superficie  $\Sigma$  en dos planos  $\Pi_1$  y  $\Pi_2$  no necesariamente coincidentes. Cada par de rayos que parten de  $S_1$  y  $S_2$  y coinciden en el mismo punto sobre la superficie  $\Sigma$  es denominado *rayos homólogos* y los puntos en que estos rayos intersecan los planos perspectivos son puntos

homólogos. Es decir perspectivas del mismo punto de la superficie proyectadas en los planos  $\Pi_1$  y  $\Pi_2$ . (fig. 1.5)

La construcción de la perspectiva puede no ser perfecta, la diferencia entre la posición teórica del punto  $m$  y la posición real se llama *distorsión*.

Se dice entonces que  $w$ ,  $p$  y la *distorsión* son los elementos internos del haz. El

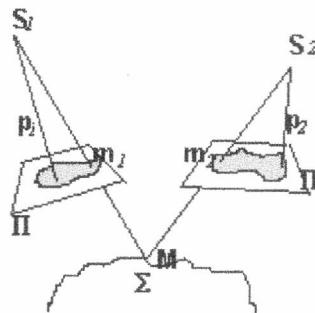


Figura 1.5 Objeto visto desde dos perspectivas

plano  $\Pi$  estará referido al referencial

#### 1.2.2.1.1. Definición de Imagen

Imagen es una aplicación de un espacio de tridimensional llamado espacio objeto en un espacio de bidimensional llamado espacio imagen.

Un elemento del espacio objeto será notado  $XYZ$  y uno del espacio imagen  $x, y$ .

Sea  $J$  una tal imagen, se puede entonces expresar las coordenadas  $x, y$  en función del  $XYZ$ :

$$J(x, y) = \begin{cases} x = F(X, Y, Z) \\ y = G(X, Y, Z) \end{cases}$$

En general se considera que  $F$  y  $G$  satisfacen condiciones de continuidad y derivabilidad deseables.

#### Par Estereoscópico. Imagen Estereoscópica

Dadas dos imágenes

$$J_1(x, y) = \begin{cases} x_1 = F_1(X, Y, Z) \\ y_1 = G_1(X, Y, Z) \end{cases}$$

$$J_2(x, y) = \begin{cases} x_2 = F_2(X, Y, Z) \\ y_2 = G_2(X, Y, Z) \end{cases}$$

y un objeto  $\sigma$  de ecuación de superficie  $Z=H(X,Y)$ , se llama imagen estereoscópica de  $\sigma$  al conjunto de cuádruplas  $(x_1,y_1,x_2,y_2)$  cuando  $XYZ$  se desplaza por todo el espacio objeto  $\sigma$  (o su superficie).

Si se identifica una cuádrupla  $(x_1,y_1,x_2,y_2)$  a partir de  $T_1$  y  $T_2$ , se definen entonces dos rayos perspectivos cuya intersección es un punto  $XYZ$  de  $\sigma$  o de su superficie

#### 1.2.2.2. Elementos Externos

Una vez elegido el referencial  $OXYZ$ , el problema se reduce a la orientación en el espacio del haz perspectivo, es decir de un sólido en el espacio lo que implica el conocimiento de sus parámetros independientes

Una translación: tres parámetros.

Una rotación: tres parámetros.

Es decir, se deben conocer la coordenadas tridimensionales en un sistema de coordenadas. Los cuales pueden ser conocidos por medidas directas o indirectas.

#### 1.2.2.3. Identificación de Rayos Homólogos.

Existen distintos aparatos que con la intervención de un operador y por medios ópticos monoculares o estereoscópicos permiten la identificación de puntos homólogos.

También esto se puede realizar con mayor o menor éxito por medio de algoritmos aplicados a pares estereoscópicos.

La restitución es entonces, una vez conocidos los puntos homólogos, la obtención de las coordenadas objeto de los puntos de una escena como resultado de la intersección de los rayos homólogos.

#### 1.2.2.4. El Caso Normal

Sean dos referenciales con la misma orientación, esto es con  $R_1$  igual a  $R_2$ . Las distancias principales también coinciden. Los planos perspectivos  $\Pi_1$  y  $\Pi_2$  son coplanares. Los centros de perspectiva  $S_1$  y  $S_2$  están separados por una distancia  $B$ . Sean  $x_1$  y  $x_2$  un par de puntos homólogos de un punto  $M$  de la superficie  $\Sigma$ . Se asume sin pérdida de generalidad que los referenciales están rotados según la rotación identidad  $I$ , es decir que no tienen giro respecto a un referencial general. También que  $S_1$  coincide con el origen de coordenadas del referencial genérico. Y que  $S_2$  está desplazado con respecto a  $S_1$  a lo largo de la dimensión  $x$ . La distancia entre los puntos de vista  $S_1$  y  $S_2$  se llama *distancia base*  $B$ . Por último y a los efectos de simplificar las ecuaciones, se considera equivalentes a  $y_1$  e  $y_2$ .

Se trata de hallar las coordenadas  $XYZ$  del los puntos  $M$  en el referencial general, conocidas sus perspectivas  $(x_i, y_i)$  y  $(x_2, y_2)$  en la imagen izquierda  $I_i$  y derecha  $I_d$ .

En la fig. 1.6 puede observarse que se cumplen las siguientes igualdades:

$$\frac{N}{x_2} = \frac{Z}{P} \quad \frac{X}{x_1} = \frac{Y}{y} = \frac{Z}{p} \quad \text{con} \quad y = y_1 = y_2 \quad p = p_1 = p_2$$

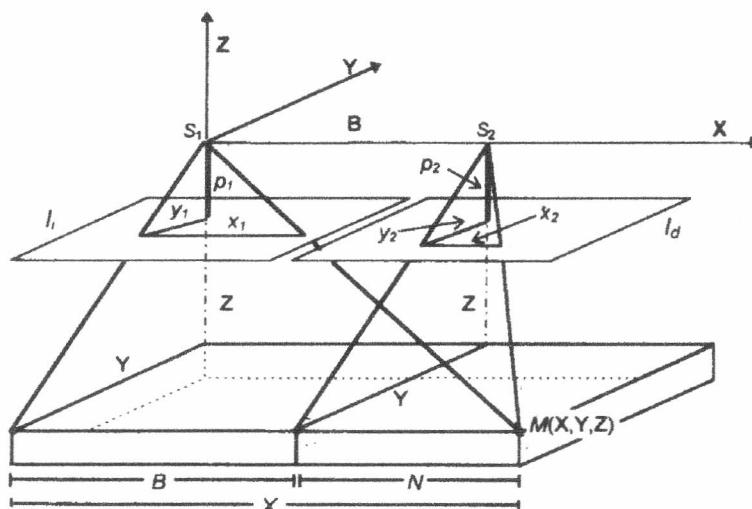


Fig. 1.6: Caso Normal  
Parámetros para la obtención de coordenadas espaciales

De la primera igualdad es posible determinar los valores de  $X$  e  $Y$  como:

$$X = \frac{Z}{p} x_1 \quad Y = \frac{Z}{p} y_1$$

Sólo resta calcular  $Z$ . Despejando  $N$  de la segunda igualdad se tiene:

$$N = \frac{Z}{p} x_2$$

Restando  $N$  a  $X$ :

$$X - N = \frac{Z}{p} (x_1 - x_2) \quad \text{con} \quad X - N = B$$

Finalmente reemplazando y despejando  $Z$  obtenemos:

$$Z = \frac{B}{(x_1 - x_2)} p$$

Haciendo, se obtiene finalmente:

$$Z = \frac{B}{\pi} p$$

Donde  $\pi$  es la diferencia entre  $x_1$  y  $x_2$ , también llamada *paralelaje longitudinal o disparidad*. La disparidad ( $x_1 - x_2$ ) está definida tomando valores desde otro referencial, frecuentemente el borde izquierdo inferior de la imagen desplazando distancias  $a$  y  $b$ . Se tiene:

$$x'_1 = x_1 - a \quad x'_2 = x_2 - b$$

restando:

$$x'_1 - x'_2 = x_1 - a - x_2 + b = \pi$$

resulta:

$$x'_1 - x'_2 = x_1 - x_2 = \pi$$

Así se puede despejar  $Z$  como producto de datos conocidos. Por lo tanto las coordenadas del punto  $M$  en la superficie  $\Sigma$  resultan:

$$M = \frac{b}{\pi} \begin{bmatrix} x_1 \\ y_2 \\ -p \end{bmatrix} \quad (1.1)$$

#### 1.2.2.5. El Problema de la Paralaje Transversal - La Pseudo- Intersección

No obstante, de ocurrir las condiciones del caso Normal, los rayos homólogos podrían resultar alabeados. Esto implica, que para esos valores de  $m_1$  y  $m_2$ , no existe un  $M$  en la superficie  $\Sigma$ .

Utilizar la ecuación (1.1) en esta oportunidad arrojará valores arbitrarios. Pero se puede minimizar el error de esta ecuación si se pseudo-intersecan los rayos.

Se consideran los infinitos planos  $xy$  (planos cuya ecuación es  $z$  equivalente a una constante  $k$ , con  $k$  real). En cada uno de esos planos se pondrá atención en el segmento determinado por los rayos, perteneciente al mismo, como se puede ver en la figura. 1.7.

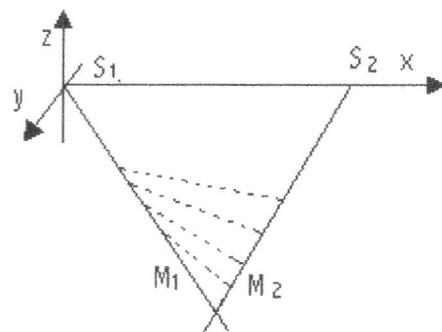


Fig 1.7 : Distancias de rectas alabeadas en diversos planos paralelos

Se entiende como paralaje transversal al segmento mínimo que une ambos rayos, perteneciente a algunos de los planos considerados. Y se llama pseudo-intersección de los rayos al punto medio de este paralaje.

$$M \approx \min(\overline{M_1 M_2})^2 \quad (1.2)$$

Se considera a  $M$  según la ecuación (1.2). Para encontrar este mínimo se deriva e iguala a 0 la longitud del segmento  $M_1 M_2$ .

Asumiendo sin pérdida de generalidad, que la distancia principal es igual a (-1), la función a minimizar es

$$\overline{M_1 M_2} = \rho m_1 - (b + \rho m_2)$$

En notación matricial la misma resulta

$$\overline{M_1 M_2} = \begin{bmatrix} \rho x_1 \\ \rho y_1 \\ -\rho \end{bmatrix} - \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \rho x_2 \\ \rho y_2 \\ -\rho \end{bmatrix}$$

Con esto la fórmula a minimizar es

$$(\overline{M_1 M_2})^2 = (\rho(x_1 - x_2) - b)^2 + \rho^2(y_1 - y_2)^2 \quad (1.3)$$

Diferenciando e igualando a 0, queda

$$\partial(\overline{M_1 M_2})^2 / \partial \rho = 2(\rho(x_1 - x_2) - b)(x_1 - x_2) + 2\rho(y_1 - y_2)^2$$

Así se puede despejar  $\rho$  según

$$\rho = \frac{b(x_1 - x_2)}{(x_1 - x_2) + (y_1 - y_2)}$$

Despreciando el término  $(y_1 - y_2)^2$  que es de segundo orden, se simplifica  $\rho$  en

$$\rho = \frac{b}{(x_1 - x_2)} = \frac{Z}{-p} \quad (1.4)$$

La expresión para  $\rho$  en la ecuación (1.4) puede usarse en la ecuación (1.3) de paralaje transversal

$$(\overline{M_1 M_2})^2 = \frac{b(y_1 - y_2)}{(x_1 - x_2)}$$

Acá se puede distribuir  $b$  para obtener

$$\left(\overline{M_1 M_2}\right)^2 = \frac{by_1}{(x_1 - x_2)} - \frac{by_2}{(x_1 - x_2)}$$

Y esto basándonos en la ecuación (1.1) no es otra cosa que

$$\left(\overline{M_1 M_2}\right) = Y_1 - Y_2$$

Con este último resultado, se puede fijar el valor de la pseudo - intersección en

$$M \approx \frac{b}{\pi} \begin{bmatrix} x_1 \\ \frac{y_1 + y_2}{2} \\ -P \end{bmatrix} \quad (1.5)$$

### 1.3. Búsqueda de Puntos Homólogos en Imágenes no Rotorectificadas

La búsqueda de puntos homólogos puede presentar dos alternativas. Una es utilizar un algoritmo de búsqueda de homólogos sin contar con la restricción de epipolaridad [16]. La otra alternativa, consiste en tratar de hallar pares homólogos contando con dicha restricción [3][4][5][6][7][8][9][10].

En general, en visión estereoscópica, las imágenes observadas cumplen con la condición de epipolaridad en el sentido de que las imágenes están en el mismo plano y que las líneas epipolares coinciden con las líneas de imagen. Esto es de particular importancia puesto que normalmente los algoritmos de apareo automático cuentan con la condición de epipolaridad de las imágenes. Es interesante notar que aun con una pequeña rotación de una imagen respecto de la otra, una persona con algún entrenamiento, consigue visualizar el modelo estereoscópico. Sin embargo no es corriente encontrar bibliografía en la que se modelice este proceso y por lo tanto contempla el caso en que la condición de epipolaridad de líneas no es respetada.

El objetivo del proyecto general es el desarrollo un sistema de visión para un robot submarino (AUV). Se ha pretendido simular la visión humana proveyéndolo de dos sensores ópticos ubicados (como los ojos en los humanos) a ambos lados. En este caso para la toma simultánea de imágenes (ojos derecho e izquierdo del móvil) los puntos de vista  $S_1$  y  $S_2$  (sensores en nuestro caso de estudio) están ubicados a una distancia  $B$  conocida y limitada por las características del vehículo, las imágenes que se obtienen con los sensores razonablemente calibrados son coplanares, cumpliendo además la restricción de epipolaridad que, como se ha dicho, establece que las líneas de imagen son paralelas a la línea base ( $S_1, S_2$ ) (fig. 1.6). A las exigencias derivadas del modelo descripto es necesario introducir la idea de respuesta perceptual y considerar problema de la precisión. En efecto, a profundidad constante del objeto respecto de los sensores, la precisión depende de la distancia entre los puntos de vista, a mayor distancia mayor precisión. Como la distancia entre los sensores es una restricción de la arquitectura del vehículo para aumentar la precisión es razonable utilizar dos imágenes no simultáneas. Estás circunstancias son las que originan que este trabajo se focalice en

la identificación automática de puntos homólogos en el primer caso mencionado ut supra, y a los efectos de proveer los elementos que permiten realizar la roto rectificación de imágenes. Su finalidad es alimentar el desarrollo complementario que ya ha sido realizado y que a partir de la información producida por éste calcula las imágenes rotorectificadas. De acá surgen las tres primeras etapas realizadas:

- Reconocimiento de puntos homólogos en imágenes roto rectificadas [DON97] (geometría epipolar) para las tomas simultáneas de los dos sensores del vehículo.
- Reconocimiento de un número limitados de puntos homólogos a los efectos de permitir la roto rectificación de imágenes sucesivas (mayor precisión y modelo de respuesta perceptual es decir comparación con un modelo numérico existente). [22]
- Roto rectificación [23]

# Fundamentos de la Solución Propuesta

## 2.1. Introducción

Si contamos con dos imágenes de una misma escena tomadas desde diferentes puntos de vista ya sea en forma simultánea o, desplazando el sensor, sucesivas, es posible, como ya se ha hecho notar, obtener las coordenadas  $XYZ$  de un punto de la escena conociendo las coordenadas  $(x_1, y_1)$ ,  $(x_2, y_2)$  en sus perspectivas de las imágenes.

La búsqueda de pares de puntos homólogos en forma automática cuando el par de imágenes, par estereoscópico, no cumple con la condición de que las líneas de imagen coincidan con las líneas epipolares, resulta extremadamente oneroso

Si consideramos como solución la comparación de cada punto de la primera imagen con cada punto de la segunda. Dado que los movimientos de rototraslación del móvil pueden dejar al punto analizado muy lejos de la posición original, o fuera de la escena capturada en la segunda imagen, esta búsqueda exhaustiva además de resultar muy costosa en tiempo y procesamiento, puede producir gran cantidad de intentos fallidos.

Como ya mencionamos, no abunda bibliografía respecto a la búsqueda de puntos homólogos en imágenes no roto - rectificadas, es decir que no cumplen con la condición mencionada en el párrafo anterior. El algoritmo descripto en [16] basa la búsqueda de puntos homólogos en el cálculo del campo de desplazamiento de los puntos de una imagen a la otra. La búsqueda no se realiza sobre las imágenes originales, sino segmentadas con las que se obtuvieron objetos definidos por características de la escena, invariantes a pequeños desplazamientos, como los bordes y las esquinas. Para cada punto de la primera imagen el algoritmo busca el homólogo en la segunda imagen, a partir de la misma posición del punto de la primera. Si bien este procedimiento no es tan costoso como una búsqueda exhaustiva, pues no compara cada punto de una imagen con cada punto de la otra, puede ocurrir, que el homólogo correspondiente se encuentre muy lejos de la posición inicial de búsqueda, o que en realidad haya quedado fuera de la segunda imagen.

El problema, entonces, es hallar un entorno adecuado donde buscar el homólogo. Es decir, reducir el espacio de búsqueda a un entorno de la nueva posición del punto en lugar de partir de la búsqueda en un entorno a la posición original. Una posibilidad interesante de la visión binocular en un vehículo es la de disponer de dos imágenes calibradas y simultáneas de la escena que respetan la restricción de epipolaridad y por lo tanto es factible obtener puntos homólogos con cualquier algoritmo de los existentes [DON/97]. El problema que se presenta, si quisieramos obtener el modelo numérico con este tipo de imágenes, es que la posición  $XYZ$  del punto en el espacio puede no tener precisión suficiente por ser pequeña la distancia entre ambos sensores. Sin embargo, puede utilizarse para obtener la posición aproximada de la perspectiva en la siguiente secuencia de imágenes (fig. 2.1). Es interesante observar que, este mecanismo resulta modelo de la respuesta perceptual.

## 2.2. Determinación del Entorno de Búsqueda

A los efectos de reducir el entorno de búsqueda se recurre a tres estrategias. La primera consiste en determinar la recta epipolar en la imagen elegida en la secuencia siguiente donde se debe encontrar el homólogo del punto seleccionado en la imagen de la secuencia en curso. La segunda recurre a la determinación de las coordenadas  $XYZ$  de un punto a partir de la secuencia en curso. Con esas coordenadas se calcula la posición aproximada de la perspectiva en la siguiente secuencia. La perspectiva es, naturalmente, el punto homólogo buscado. Finalmente, para la tercera estrategia, se recurre a la cooperación de los resultados de las dos primeras estrategias aplicadas. Se intersecan la línea epipolar hallada con el entorno determinado, para lograr así, una mayor restricción del espacio de búsqueda.

### 2.2.1. Recta Epipolar Conjugada.

Sean dos perspectivas de una misma escena y un punto  $M$  en la superficie  $\Sigma$  (fig 2.1),

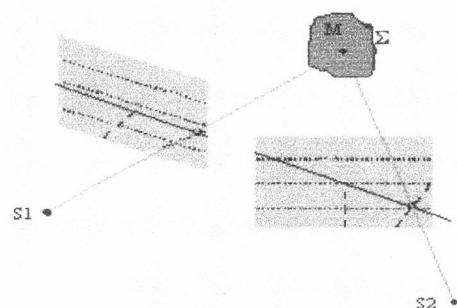


Fig. 2.1 : Rectas epipoles

Como ya hemos mencionado, la intersección del plano de cada perspectiva con el plano formado por el punto  $M$  y los centros ópticos  $S_1$  y  $S_2$  es denominado líneas epipolares conjugadas. Se conoce entonces como líneas epipolares conjugadas a aquellas que corresponden a un mismo punto del espacio.

La obtención de las líneas epipolares conjugadas, para imágenes simultáneas que satisfacen la condición de epipolaridad, es trivial dado que se corresponden a una misma línea de imagen o de pixeles. Para la imagen derecha del par de la secuencia siguiente, requiere del conocimiento aproximado de la rotación y traslación del móvil. Datos que se obtienen a partir de los sensores del móvil

Esto implica que, conocidas en forma aproximada la rotación y el desplazamiento relativo de dos imágenes, es posible, dado un punto en una de las imágenes, hallar la línea epipolar en la otra dentro de los límites de precisión de la rotación y el desplazamiento

### 2.2.2. Coordenadas Aproximadas del Punto

Se considera que el par de imágenes simultáneas satisface las condiciones del Caso Normal. Las coordenadas del punto en el objeto, para dicho par, se obtienen, como ya se mencionó, aplicando las ecuaciones descriptas en el Capítulo 1.

Considerando despreciables las rotaciones de los ejes, producidas por defectos de calibración, es de esperar una aproximación suficientemente buena a la posición real del objeto.

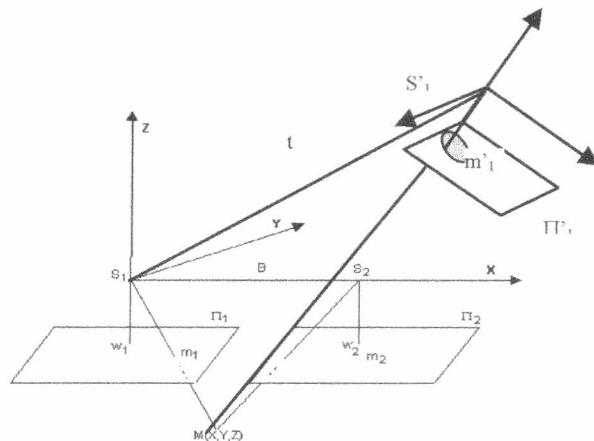


Fig. 2.2.Ubicación del entorno del homólogo en la secuencia.

Para obtener, en la secuencia siguiente, el entorno dónde buscar el punto homólogo, basta conocer la posición del punto objeto, y la rototraslación del móvil, es decir la posición de la imagen en la secuencia siguiente, respecto del referencial objeto. (fig. 2.2).

### 2.2.3. Cooperación de la recta epipolar y el entorno de búsqueda. Estrategia de Control

Si los cálculos del entorno posición del punto y la recta epipolar conjugada se realizan con suficiente precisión, es de esperar que ambos se intersequen en la segunda toma de la secuencia.

Como efecto de la rototraslación del móvil, ya mencionamos que el punto buscado puede quedar fuera de la segunda toma. Esto se evidencia cuando la recta epipolar conjugada cae fuera de la segunda imagen o bien la intersección entre la recta epipolar y el entorno hallado tampoco se encuentran en ella. En ambos casos, la búsqueda para el punto en la primera imagen es desechada. De esto resulta una estrategia de control que puede reducir notablemente los intentos fallidos.

## 2.3. Los Modelos del Fenómeno

### 2.3.1. Modelo Funcional

La Implementación del algoritmo que se describe ut infra produce un sistema sobre-dimensionado procediéndose entonces a compensar por *cuadrados mínimos*.

El objetivo de la teoría de cuadrados mínimos, es estimar lo mejor posible los valores de los parámetros (o incógnitas) ligados a un fenómeno físico.

Se realiza a partir de mediciones, en principio redundantes, de magnitudes: los observables de dicho fenómeno; y de hipótesis sobre las leyes físicas que los rigen y sobre los errores de medida de los observables.

Se trata de plantear hipótesis sobre las leyes que rigen el comportamiento del sistema. Es decir, obtener el *modelo funcional*. Dicho modelo describe la manera en que observables y parámetros están ligados. Un fenómeno puede ser estimado bajo distintas hipótesis que se traducirán en distintos modelos funcionales. [25]

### 2.3.2 Modelo Estocástico

Como se ha mencionado, a partir de las hipótesis sobre las leyes físicas que rigen el fenómeno se elabora el modelo funcional.

El modelo estocástico resulta de hipótesis sobre los errores de medida de las magnitudes observadas llamadas también “observables”

A los efectos de poder aplicar un modelo estocástico es necesario un modelo funcional canónico. Este consiste esencialmente en plantear las ecuaciones de manera que los observables queden como funciones de los parámetros incógnitas. Si estas funciones no son lineales, se lineariza como es nuestro caso. [25]

# Solución Propuesta

## 3.1. Introducción

El algoritmo propuesto por [16] se basa en la búsqueda de todos los pares homólogos en imágenes que no cumplen con la condición de epipolaridad, y sin contar con ningún tipo de conocimiento previo, realizando una búsqueda exhaustiva en diferentes niveles de resolución.

La búsqueda de un conjunto mínimo de puntos homólogos, aplicando el algoritmo arriba mencionado representa un costo muy alto.

Para nuestro problema en particular, se cuenta con el conocimiento de los parámetros de rototraslación del vehículo, permitiendo limitar razonablemente el espacio de búsqueda. Como contamos además, con las perspectivas de ciertos puntos del objeto a partir de las tomas simultáneas, se puede calcular, por un lado la posición de las líneas epipolares en las imágenes y por otro delimitar el conjunto de píxeles en cada imagen donde se deberá ubicar el par de puntos homólogos.

## 3.2. El Modelo Funcional de la Solución Propuesta.

### 3.2.1. Cálculo de Diferencias.

El vector de que se busca minimizar esta conformado por las siguientes diferencias:

a. Diferencia de Intensidad

$$r_i(u, d) = i(u + d) - i(u)$$

b. Diferencia de Borde

$$r_b(u, d) = b(u + d) - b(u)$$

c. Diferencia de Esquina Positiva

$$r_p(u, d) = p(u + d) - p(u)$$

d. Diferencia de Esquina Negativa

$$r_n(u, d) = n(u + d) - n(u)$$

e. Diferencia de Ángulo Máximo

$$r_A(u, d) = A(u + d) - A(u)$$

f. Diferencia de Ángulo Mínimo

$$r_a(u, d) = a(u + d) - a(u)$$

g. Diferencia de Radio del Ángulo Máximo

$$r_L(u, d) = L(u + d) - L(u)$$

h. Diferencia de Radio del Ángulo Mínimo

$$r_l(u, d) = l(u + d) - l(u)$$

i. Diferencia de Peso del Ángulo Máximo

$$r_W(u, d) = W(u + d) - W(u)$$

j. Diferencia de Peso del Ángulo Mínimo

$$r_w(u, d) = w(u + d) - w(u)$$

k. Diferencia de Promedio en el entorno.

$$r_{\bar{x}}(u, d) = \bar{x}(u + d) - \bar{x}(u)$$

l. Diferencia de Mediana

$$r_m(u, d) = m(u + d) - m(u)$$

### a) Intensidad:

Consideremos dos imágenes  $I_1$  e  $I_2$ . La expresión  $I_1(x, y)$  indica el valor de la intensidad de  $I_1$  en la posición  $(x, y)$ . Si ambas imágenes son tomas de una misma escena con una disparidad en el espacio, bajo ciertas condiciones se espera que

$$i_1(x, y) = i_2(x + \delta x, y + \delta y)$$

$$i_1(x, y) = i_2(x, y) + \frac{\partial i_2(x, y)}{\partial x} \delta x + \frac{\partial i_2(x, y)}{\partial y} \delta y$$

$$i_1(x, y) - i_2(x, y) = \frac{\partial i_2(x, y)}{\partial x} \delta x + \frac{\partial i_2(x, y)}{\partial y} \delta y = r_i(x, y)$$

Con:  $\frac{\partial i_2(x, y)}{\partial x} = i_2(x + 1, y) - 2i_2(x, y) + i_2(x - 1, y)$

$$\frac{\partial i_2(x, y)}{\partial y} = i_2(x, y + 1) - 2i_2(x, y) + i_2(x, y - 1)$$

Resulta entonces:

$$[i_2(x + 1, y) - 2i_2(x, y) + i_2(x - 1, y)] \delta x + [i_2(x, y + 1) - 2i_2(x, y) + i_2(x, y - 1)] \delta y = r_i(x, y)$$

### b) Bordes:

$$b_1(x, y) = b_2(x, y) + \frac{\partial b_2(x, y)}{\partial x} \delta x + \frac{\partial b_2(x, y)}{\partial y} \delta y$$

Por lo tanto:

$$(b_1(x, y) - b_2(x, y)) = \frac{\partial b_2(x, y)}{\partial x} \hat{\alpha} + \frac{\partial b_2(x, y)}{\partial y} \hat{\beta} = r_b(x, y)$$

Con :

$$\begin{aligned}\frac{\partial b_2(x, y)}{\partial x} &= b_2(x+1, y) - 2b_2(x, y) + b_2(x-1, y) \\ \frac{\partial b_2(x, y)}{\partial y} &= b_2(x, y+1) - 2b_2(x, y) + b_2(x, y-1)\end{aligned}$$

Resulta entonces:

$$[b_2(x+1, y) - 2b_2(x, y) + b_2(x-1, y)] \hat{\alpha} + [b_2(x, y+1) - 2b_2(x, y) + b_2(x, y-1)] \hat{\beta} = r_b(x, y)$$

### c) Esquinas positivas

$$p_1(x, y) = p_2(x, y) + \frac{\partial p_2(x, y)}{\partial x} \hat{\alpha} + \frac{\partial p_2(x, y)}{\partial y} \hat{\beta}$$

Por lo tanto:

$$p_1(x, y) - p_2(x, y) = \frac{\partial p_2(x, y)}{\partial x} \hat{\alpha} + \frac{\partial p_2(x, y)}{\partial y} \hat{\beta}$$

Con :

$$\begin{aligned}\frac{\partial p_2(x, y)}{\partial x} &= p_2(x+1, y) - 2p_2(x, y) + p_2(x-1, y) \\ \frac{\partial p_2(x, y)}{\partial y} &= p_2(x, y+1) - 2p_2(x, y) + p_2(x, y-1)\end{aligned}$$

Resulta entonces:

$$[p_2(x+1, y) - 2p_2(x, y) + p_2(x-1, y)] \hat{\alpha} + [p_2(x, y+1) - 2p_2(x, y) + p_2(x, y-1)] \hat{\beta} = r_p(x, y)$$

### d) Esquinas negativas:

$$n_1(x, y) = n_2(x, y) + \frac{\partial n_2(x, y)}{\partial x} \hat{\alpha} + \frac{\partial n_2(x, y)}{\partial y} \hat{\beta}$$

Por lo tanto:

$$n_2(x, y) - n_1(x, y) = \frac{\partial n_2(x, y)}{\partial x} \hat{\alpha} + \frac{\partial n_2(x, y)}{\partial y} \hat{\beta}$$

Con :

$$\begin{aligned}\frac{\partial n_2(x, y)}{\partial x} &= n_2(x+1, y) - 2n_2(x, y) + n_2(x-1, y) \\ \frac{\partial n_2(x, y)}{\partial y} &= n_2(x, y+1) - 2n_2(x, y) + n_2(x, y-1)\end{aligned}$$

Resulta entonces:

$$[n_2(x+1,y) - 2n_2(x,y) + n_2(x-1,y)]\delta x + [n_2(x+1,y) - 2n_2(x,y) + n_2(x-1,y)]\delta y = r_n$$

**e) Diferencia de Ángulo Máximo**

Para cada esquina analizada, se construye un grafo que la conecta con cada una de las otras esquinas. y se obtienen

- a) Se calcula la norma de cada uno de los arcos del grafo.
- b) Se calcula el ángulo formado por cada par de arcos del grafo.
- d) Se obtiene el ángulo máximo de todos los ángulos calculados.

En forma análoga a los puntos anteriores se obtiene:

$$[A_2(x+1,y) - 2A_2(x,y) + A_2(x-1,y)]\delta x + [A_2(x+1,y) - 2A_2(x,y) + A_2(x-1,y)]\delta y = r_A(x,y)$$

**f) Diferencia de Ángulo Mínimo**

Se procede en forma análoga pero para obtener el ángulo mínimo.

- a) Se calcula la norma de cada uno de los arcos del grafo.
- b) Se calcula el ángulo formado por cada par de arcos del grafo.
- d) Se obtiene el ángulo mínimo de todos los ángulos calculados.

En forma análoga a los puntos anteriores se obtiene:

$$[a_2(x+1,y) - 2a_2(x,y) + a_2(x-1,y)]\delta x + [a_2(x+1,y) - 2a_2(x,y) + a_2(x-1,y)]\delta y = r_a(x,y)$$

**g) Diferencia de Radio del Ángulo Máximo.**

Se definió como radio del ángulo al cociente entre las normas de sus lados. Para esto se toma el Ángulo Máximo obtenido en el punto e.

En forma análoga a los puntos anteriores se obtiene:

$$[L_2(x+1,y) - 2L_2(x,y) + L_2(x-1,y)]\delta x + [L_2(x+1,y) - 2L_2(x,y) + L_2(x-1,y)]\delta y = r_L(x,y)$$

**h) Diferencia de Radio del Ángulo Mínimo.**

En este caso se calculó el radio tomando en ángulo mínimo resultante del cálculo realizado en el punto f.

En forma análoga a los puntos anteriores se obtiene:

$$[l_2(x+1,y) - 2l_2(x,y) + l_2(x-1,y)]\delta x + [l_2(x+1,y) - 2l_2(x,y) + l_2(x-1,y)]\delta y = r_l(x,y)$$

### i) Diferencia de Peso del Ángulo Máximo

Se definió como peso del ángulo a la suma de las normas de sus lados. Para esta restricción se toma en ángulo máximo resultante del ítem e.

En forma análoga a los puntos anteriores se obtiene:

$$[W_2(x+1,y) - 2W_2(x,y) + W_2(x-1,y)]\delta x + [W_2(x+1,y) - 2W_2(x,y) + W_2(x-1,y)]\delta y = r_w(x,y)$$

### j) Diferencia de Peso del Ángulo Mínimo

Para esta restricción se toma en ángulo mínimo resultante del ítem f.

En forma análoga a los puntos anteriores se obtiene:

$$[w_2(x+1,y) - 2w_2(x,y) + w_2(x-1,y)]\delta x + [w_2(x+1,y) - 2w_2(x,y) + w_2(x-1,y)]\delta y = r_w(x,y)$$

### k) Diferencia de Promedio en el Entorno.

En cooperación con la imagen original, se obtiene la diferencia del promedio en el entorno del punto tratado tomando una ventana de 3x3. Para el cálculo del promedio no se considera el punto analizado.

En forma análoga a los puntos anteriores se obtiene:

$$[av_2(x+1,y) - 2av_2(x,y) + av_2(x-1,y)]\delta x + [av_2(x+1,y) - 2av_2(x,y) + av_2(x-1,y)]\delta y = r_{av}(x,y)$$

### l) Diferencia de Mediana

En este caso, la cooperación con la imagen original es utilizada para obtener la diferencia de la mediana en el entorno del punto tratado, tomando también una ventana de 3x3.

En forma análoga a los puntos anteriores se obtiene:

$$[m_2(x+1,y) - 2m_2(x,y) + m_2(x-1,y)]\delta x + [m_2(x+1,y) - 2m_2(x,y) + m_2(x-1,y)]\delta y = r_m(x,y)$$

### 3.2.2. Matriz De Cuadrados Mínimos

El sistema de ecuaciones a resolver de nuestro modelo funcional tiene la forma:

$$\begin{aligned} i_x \partial x + i_y \partial y &= r_i \\ b_x \partial x + b_y \partial y &= r_b \\ p_x \partial x + p_y \partial y &= r_p \\ n_x \partial x + n_y \partial y &= r_n \\ A_x \partial x + A_y \partial y &= r_A \end{aligned}$$

$$a_x \partial x + a_y \partial y = r_a$$

$$L_x \partial x + L_y \partial y = r_o$$

$$l_x \partial x + l_y \partial y = r_l$$

$$W_x \partial x + W_y \partial y = r_W$$

$$w_x \partial x + w_y \partial y = r_w$$

$$\bar{x}_x \partial x + \bar{x}_y \partial y = r_{\bar{x}}$$

$$m_x \partial x + m_y \partial y = r_m$$

Se trata de hallar los  $\partial x$  y  $\partial y$  que minimicen el error. Se debe encontrar entonces el mínimo de:

$$F(\partial x, \partial y) =$$

$$\begin{aligned} &= \left( i_x \partial x + i_y \partial y - r_i \right)^2 + \left( b_x \partial x + b_y \partial y - r_b \right)^2 + \left( p_x \partial x + p_y \partial y - r_p \right)^2 + \\ &+ \left( n_x \partial x + n_y \partial y - r_n \right)^2 + \left( A_x \partial x + A_y \partial y - r_A \right)^2 + \left( a_x \partial x + a_y \partial y - r_a \right)^2 + \\ &= \left( L_x \partial x + L_y \partial y - r_L \right)^2 + \left( l_x \partial x + l_y \partial y - r_l \right)^2 + \left( W_x \partial x + W_y \partial y - r_W \right)^2 + \\ &+ \left( w_x \partial x + w_y \partial y - r_w \right)^2 + \left( av_x \partial x + av_y \partial y - r_{av} \right)^2 + \left( m_x \partial x + m_y \partial y - r_m \right)^2 = 0 \end{aligned}$$

Minimizamos anulando las dos derivadas parciales:

$$\frac{\partial F}{\partial x} = 2 \left[ \begin{aligned} &\left( i_x \partial x + i_y \partial y - r_i \right) i_x + \left( b_x \partial x + b_y \partial y - r_b \right) b_x + \left( p_x \partial x + p_y \partial y - r_p \right) p_x + \\ &+ \left( n_x \partial x + n_y \partial y - r_n \right) n_x + \left( A_x \partial x + A_y \partial y - r_A \right) A_x + \left( a_x \partial x + a_y \partial y - r_a \right) a_x + \\ &+ \left( L_x \partial x + L_y \partial y - r_L \right) L_x + \left( l_x \partial x + l_y \partial y - r_l \right) l_x + \left( W \partial x + W \partial y - r_W \right) W_x + \\ &+ \left( w_x \partial x + w_y \partial y - r_w \right) w_x + \left( av_x \partial x + av_y \partial y - r_{av} \right) av_x + \left( m_x \partial x + m_y \partial y - r_m \right) m_x = 0 \end{aligned} \right]$$

(1.a)

$$\frac{dF}{\partial y} = 2 \left[ \begin{aligned} & \left( i_x \partial x + i_y \partial y - r_i \right) i_y + \left( b_x \partial x + b_y \partial y - r_b \right) b_y + \left( p_x \partial x + p_y \partial y - r_p \right) p_y + \\ & + \left( n_x \partial x + n_y \partial y - r_n \right) n_y + \left( A_x \partial x + A_y \partial y - r_A \right) A_y + \left( a_x \partial x + a_y \partial y - r_a \right) a_y + \\ & + \left( L_x \partial x + L_y \partial y - r_L \right) L_y + \left( I_x \partial x + I_y \partial y - r_I \right) I_y + \left( W_x \partial x + W_y \partial y - r_W \right) W_y + \\ & + \left( w_x \partial x + w_y \partial y - r_w \right) w_y + \left( \bar{x}_x \partial x + \bar{x}_y \partial y - r_{\bar{x}} \right) \bar{x}_y + \left( m_x \partial x + m_y \partial y - r_m \right) m_y = 0 \end{aligned} \right] = 0$$

(1.b)

Distribuyendo

de (1.a)

$$\begin{aligned} \frac{dF}{d\partial x} = & i_x i_x \partial x + i_y i_x \partial y - r_i i_x + \\ & + b_x b_x \partial x + b_y b_x \partial y - r_b b_x + \\ & + p_x p_x \partial x + p_y p_x \partial y - r_p p_x + \\ & + n_x n_x \partial x + n_y n_x \partial y - r_n n_x + \\ & + A_x A_x \partial x + A_y A_x \partial y - r_A A_x + \\ & + a_x a_x \partial x + a_y a_x \partial y - r_a a_x + \\ & + L_x L_x \partial x + L_y L_x \partial y - r_L L_x + \\ & + l_x l_x \partial x + l_y l_x \partial y - r_l l_x + \\ & + W_x W_x \partial x + W_y W_x \partial y - r_W W_x + \\ & + w_x w_x \partial x + w_y w_x \partial y - r_w w_x + \\ & + \bar{x}_x \bar{x}_x \partial x + \bar{x}_y \bar{x}_x \partial y - r_{\bar{x}} \bar{x}_x + \\ & + m_x m_x \partial x + m_y m_x \partial y - r_m m_x = 0 \end{aligned}$$

Reagrupando obtenemos

$$\begin{aligned} \frac{dF}{d\partial x} = & \left( i_x i_x + b_x b_x + p_x p_x + n_x n_x + A_x A_x + a_x a_x + L_x L_x + l_x l_x + W_x W_x + w_x w_x + \bar{x}_x \bar{x}_x + m_x m_x \right) \partial x + \\ & \left( i_y i_x + b_y b_x + p_y p_x + n_y n_x + A_y A_x + a_y a_x + L_y L_x + l_y l_x + W_y W_x + w_y w_x + \bar{x}_y \bar{x}_x + m_{xy} m_x \right) \partial y = \\ & = r_i i_x + b_x + r_b r_p p_x + r_x n_x + r_A A_x + r_a a_x + r_L L_x + r_l l_x + r_W W_x + r_w w_x + r_{\bar{x}} \bar{x}_x + r_m m_x \end{aligned}$$

(2.a)

Procediendo en forma análoga para la otra derivada tenemos:

de (1.b)

$$\begin{aligned}
 \frac{dF}{d\partial y} = & i_x i_y \partial x + i_y i_y \partial y - r_i i_y + \\
 & + b_x b_y \partial x + b_y b_y \partial y - r_b b_y + \\
 & + p_x p_y \partial x + p_y p_y \partial y - r_p p_y + \\
 & + n_x n_y \partial x + n_y n_y \partial y - r_n n_y + \\
 & + A_x A_y \partial x + A_y A_{xy} \partial y - r_A A_y + \\
 & + a_x a_y \partial x + a_y a_y \partial y - r_a a_y + \\
 & + L_x L_y \partial x + L_y L_y \partial y - r_L L_y + \\
 & + l_x l_y \partial x + l_y l_y \partial y - r_l l_y + \\
 & + W_x W_y \partial x + W_y W_y \partial y - r_W W_y + \\
 & + w_x w_y \partial x + w_y w_y \partial y - r_w w_y + \\
 & + \bar{x}_x \bar{x}_y \partial x + \bar{x}_y \bar{x}_y \partial y - r_{\bar{x}} \bar{x}_y + \\
 & + m_x m_y \partial x + m_y m_y \partial y - r_m m_y = 0
 \end{aligned}$$

Reagrupando obtenemos

$$\begin{aligned}
 \frac{dF}{d\partial y} = & \left( i_x i_y + b_x b_y + p_x p_y + n_x n_y + A_x A_y + a_x a_y + L_x L_y + l_x l_y + W_x W_y + w_x w_y + \bar{x}_x \bar{x}_y + m_x m_y \right) \partial x + \\
 & + \left( i_y i_y + b_y b_y + p_y p_y + n_y n_y + A_y A_y + a_y a_y + L_y L_y + l_y l_y + W_y W_y + w_y w_y + \bar{x}_y \bar{x}_y + m_y m_y \right) \partial y = \\
 & r_i i_y + r_b b_y + r_p p_y + r_n n_y + r_A A_y + r_a a_y + r_L L_y + r_l l_y + r_W W_y + r_w w_y + r_{\bar{x}} \bar{x}_y + r_m m_y
 \end{aligned} \tag{2.b}$$

Si consideramos las siguientes matrices

$$A = \begin{bmatrix} i_x & i_y \\ b_x & b_y \\ p_x & p_y \\ n_x & n_y \\ A_x & A_y \\ a_x & a_y \\ L_x & L_y \\ l_x & l_y \\ W_x & W_y \\ w_x & w_y \\ av_x & av_y \\ m_x & m_y \end{bmatrix} \quad x = \begin{bmatrix} \partial x \\ \partial y \end{bmatrix} \quad L = \begin{bmatrix} r_i \\ r_b \\ r_p \\ r_n \\ r_A \\ r_a \\ r_L \\ r_l \\ r_W \\ r_w \\ r_{av} \\ r_m \end{bmatrix}$$

Entonces el sistema formado por las ecuaciones (2.a) y (2.b) puede expresarse como

$$A^T A x = A^T L$$

Resultando un modelo funcional de cuadrados mínimos canónico cuya solución es

$$\begin{aligned} A^T Ax &= A^T L \\ x &= (A^T A)^{-1} A^T L \end{aligned}$$

Esto es un sistema lineal de dos ecuaciones con dos incógnitas cuya resolución es trivial.

### **3.3. Etapas en la Implementación Para la Solución Propuesta.**

La solución buscada es alcanzada en las siguientes etapas.

#### **I. Para cada par de puntos homólogos hallados:**

- 1) Encontrar la línea epipolar conjugada, en la segunda imagen.
- 2) Hallar la posición  $XYZ$
- 3) Determinar el entorno de la nueva posición en la segunda imagen.
- 4) Determinar la intersección de la recta con el entorno.

#### **II. Buscar el homólogo en la segunda imagen**

Para esta etapa se desarrollaron dos aproximaciones posibles:

##### **Aproximación 1:**

Calcular el campo desplazamiento de toda la grilla con un cierto número de iteraciones (Ej. 20) Aquí es donde se busca minimizar el vector de diferencias. Como el sistema resulta sobre determinado, se aplica cuadrados mínimos[20].

##### **Aproximación 2:**

En esta aproximación, se calcula la norma de la diferencia entre el vector de atributos de cada esquina en el entorno de búsqueda determinado y el vector de atributos de la esquina tratada, resultando como candidata aquella que origine la menor norma. .

En este caso la búsqueda está restringida solo a esquinas y, en particular, esquinas de la misma orientación.

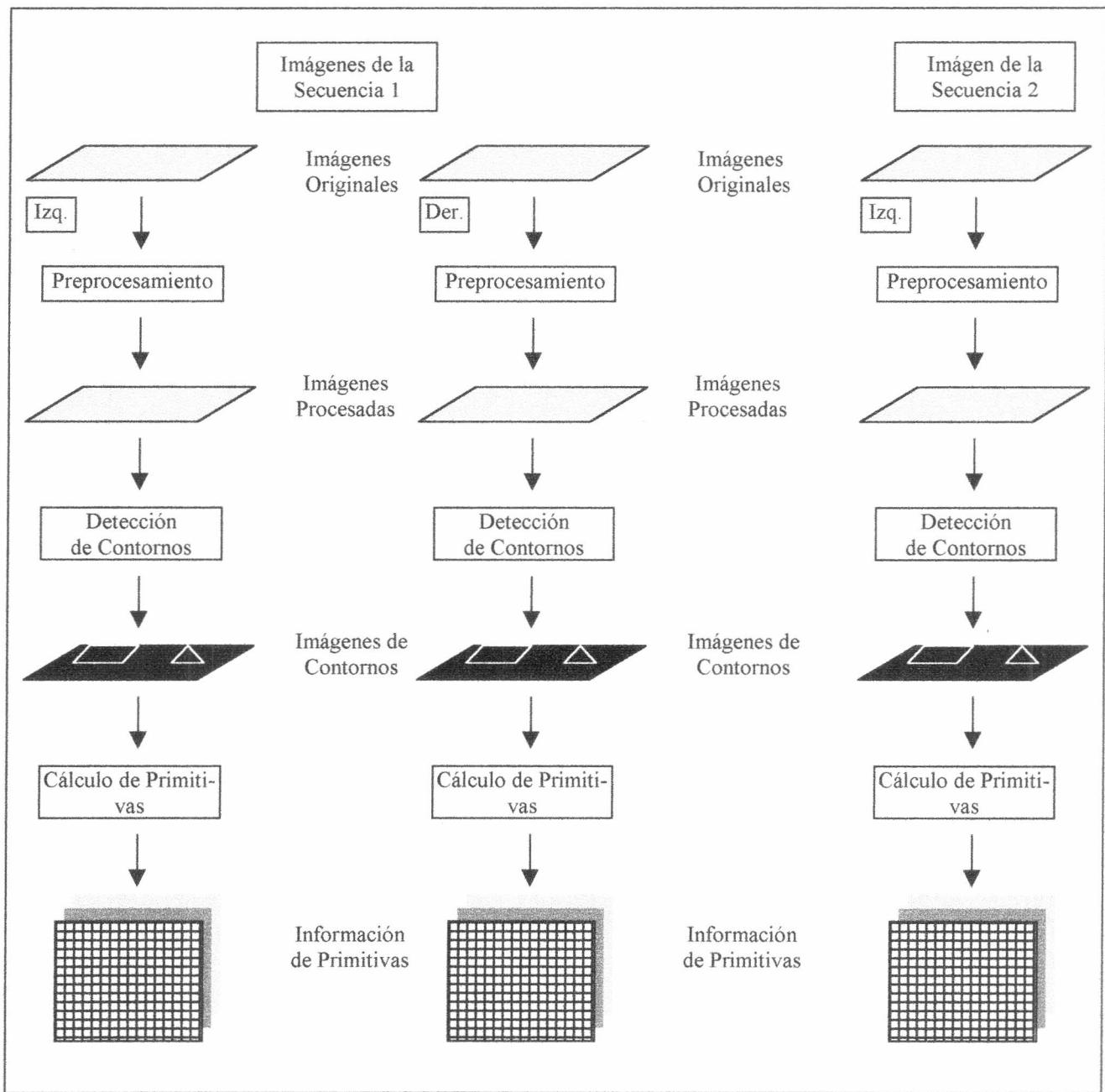


Fig. 3.1 Estructura de las etapas de tratamiento

### 3.4.-Geometría de la solución propuesta

#### 3.4.1. Rototraslación de la Secuencia

Sean dos puntos de vista  $S_1$  y  $S_3$ . Conocidos la traslación  $t = (t_1, t_2, t_3)$  y la rotación  $R$  respecto del primer referencial, es posible expresar las coordenadas del segundo punto de vista con respecta al primero.

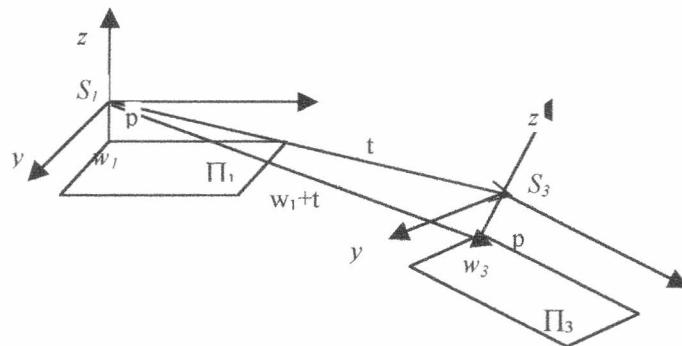


Fig. 2.3. Secuencia Rototrasladada.

En la figura tenemos los referenciales \$S\_1xyz\$ con \$w\_1\$ el pie de perpendicular del referencial al plano \$\Pi\_i\$. Sea además la rotación \$R\$ alrededor del eje \$x\$. Es decir

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

Resulta entonces que

$$\overline{S_3} = \tilde{t}$$

$$\overline{w_{3S_1}} = R[\overline{w_1} + \tilde{t}]$$

La resultante de \$\tilde{t} - w\_{3S\_1} = \eta\_3\$ es la normal al plano \$\Pi\_2\$.

### 3.4.2. Determinación de la Recta Epipolar Conjugada en la Secuencia.

Se trata de calcular la línea epipolar en la secuencia actual, conocidas las coordenadas de un punto en la secuencia anterior, y la rotación y traslación relativas entre las imágenes.

La matriz de rotación \$R\$ traduce coordenadas del referencial \$S1\$ al referencial \$S3\$, la distancia del foco: \$-p\$.

Dado un punto \$m\_1 = (m\_{1x} \ m\_{1y} \ -p)\$ en la secuencia anterior, las rectas epipoles conjugadas están ahora definidas por la intersección del plano determinado por los puntos de vista \$S\_1S\_3\$ y \$m\_1\$ y las correspondientes perspectivas.

En la secuencia actual son los  $(x, y)$  que satisfacen:

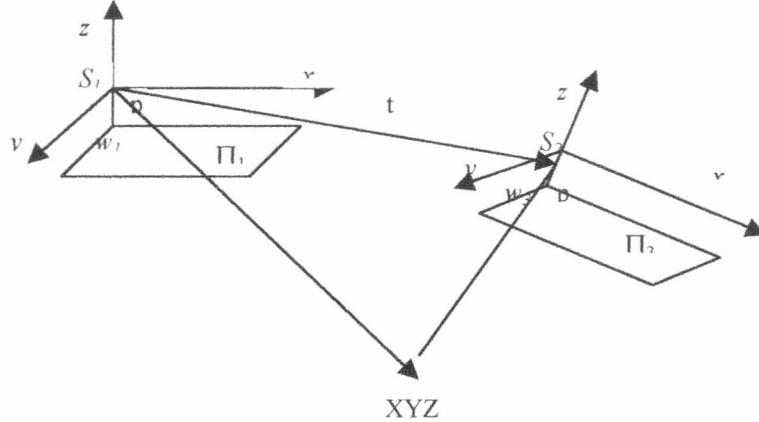


Fig 2.4. Intersección entre el plano de la secuencia y el plano formado por los puntos de vista y las coordenadas del punto en la escena

$$[R \cdot (t \times m_1)] \begin{bmatrix} x \\ y \\ -p \end{bmatrix} = 0$$

donde  $R$  es la matriz rotación

$t$  es la traslación relativa

$(x, y)$  son las coordenadas del punto en la secuencia actual.

$$\begin{aligned} t \times m_1 &= \begin{pmatrix} i & j & k \\ t_x & t_y & t_z \\ m_{1x} & m_{1y} & -p \end{pmatrix} = \begin{pmatrix} -p \cdot t_y - m_{1y} \cdot t_z \\ p \cdot t_x + m_{1x} \cdot t_z \\ m_{1y} \cdot t_x - m_{1x} \cdot t_y \end{pmatrix} \\ R \cdot \begin{pmatrix} -p \cdot t_y - m_{1y} \cdot t_z \\ p \cdot t_x + m_{1x} \cdot t_z \\ m_{1y} \cdot t_x - m_{1x} \cdot t_y \end{pmatrix} &= \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix} \cdot \begin{pmatrix} -p \cdot t_y - m_{1y} \cdot t_z \\ p \cdot t_x + m_{1x} \cdot t_z \\ m_{1y} \cdot t_x - m_{1x} \cdot t_y \end{pmatrix} = \\ &= \begin{pmatrix} -R_{xx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{xy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{xz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y) \\ -R_{yx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{yy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{yz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y) \\ -R_{zx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{zy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{zz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y) \end{pmatrix} \\ y &= \frac{-R_{xx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{xy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{xz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y)}{-\left(R_{yx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{yy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{yz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y)\right)} \cdot x + \\ &\quad \frac{-R_{zx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{zy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{zz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y)}{-\left(R_{yx} \cdot (-p \cdot t_y - m_{1y} \cdot t_z) + R_{yy} \cdot (p \cdot t_x + m_{1x} \cdot t_z) + R_{yz} \cdot (m_{1y} \cdot t_x - m_{1x} \cdot t_y)\right)} \end{aligned}$$

Resta considerar que porción de la recta pertenece a la perspectiva. Se debe satisfacer el echo que  $\forall(xyz \in L_3(\gamma)) \quad 0 \leq x \leq A1 \text{ y } 0 \leq y \leq A2$

Donde A1 y A2 representan la columnas y de filas de las perspectivas tratadas.

### 3.4.3. Determinación del Entorno del Punto en la Secuencia.

En este caso se trata de hallar la intersección entre el plano determinado por  $\eta_3$  y la recta que pasa por el punto  $t$  en dirección  $\vec{v} = \vec{M} - \vec{t}$  y donde

$$M = \frac{b}{\pi} \begin{bmatrix} x_1 \\ y_1 \\ -p \end{bmatrix} \text{ con } \pi = x_1 - x_2$$

Entonces

$$L_3(\gamma) = \gamma(\vec{M} - \vec{t}) + \vec{t}$$

$$\text{con } P_3 = \eta_{3x}x + \eta_{3y}y + \eta_{3z}z = D_3$$

sean

$$(x_0, y_0, z_0) \in L_3(\gamma) \text{ y } (x, y, z) \in P_3$$

Debemos hallar  $\gamma_0$  tq

$$\eta_3 \cdot [\gamma_0(\vec{M} - \vec{t}) + \vec{t}] = D_3$$

Distribuyendo el producto vectorial con respecto a la suma

$$\eta_3 \cdot [\gamma_0(\vec{M} - \vec{t})] + \eta_3 \cdot \vec{t} = D_3$$

$$\eta_3 \cdot [\gamma_0(\vec{M} - \vec{t})] = D_3 - \eta_3 \cdot \vec{t}$$

$$\gamma_0[\eta_3 \cdot (\vec{M} - \vec{t})] = D_3 - \eta_3 \cdot \vec{t}$$

Despejando  $\gamma_0$

$$\gamma_0 = \frac{D_3 - \eta_3 \cdot \vec{t}}{\eta_3 \cdot [\vec{M} - \vec{t}]}$$

Resulta entonces

$$(m_x, m_y, m_z) = L(\gamma_0) = \left[ \frac{D_3 - \eta_3 \cdot \vec{t}}{\eta_3 \cdot [\vec{M} - \vec{t}]} \cdot (\vec{M} - \vec{t}) \right] + \vec{t}$$

Este es el punto de intersección de la recta con el plano. Es decir el centro del entorno de búsqueda.

Todo lo calculado está descripto desde  $S_1$ . Para describir desde  $S_2$  hay que restarle las coordenadas de  $S_2$ .

- a) Recta epipolar conjugada descripta desde  $S_1$

$$L_3(\gamma) = \gamma(\eta_x, \eta_y, \eta_z) + (x_0, y_0, z_0)$$

$$L_{3s2}(\gamma) = \gamma(\eta_x, \eta_y, \eta_z) + (x_0, y_0, z_0) - (t_1, t_2, t_3)$$

- b) Posición del punto en la secuencia

desde  $S_1$ :

$$L(\gamma_0) = \gamma_0(\vec{M} - \vec{t}) + \vec{t}$$

desde  $S_2$ :

$$L(\gamma_0) = \gamma_0(\vec{M} - \vec{t}) + \vec{t} - \vec{t} = \gamma_0(\vec{M} - \vec{t})$$

Falta ver que  $m_3 = (m_{3x}, m_{3y}, m_{3z})$  pertenezca a un entorno de la recta  $L(\gamma)$ .

- 1) Calcular las paralelas a  $L_3(\gamma), L_3^{(1)}(\gamma), L_3^{(2)}(\gamma)$ . Son tres rectas con el mismo vector dirección con  $\pm \varepsilon_y$  en la ordenada al origen.
- 2) Calcular la recta ortogonal a las anteriores y que pasa por  $m_3$ . Basta buscar el vector ortogonal  $v$  y luego hacer  $L(\gamma) = \gamma v + m_3$
- 3) Calcular la intersección de  $L(\gamma)$  con sus rectas ortogonales (obtener a y b)

# Aspectos de implementación

## 4.1. Introducción

La solución propuesta se desarrolla, como ya se mencionó en capítulos anteriores, en varias etapas. Con el objeto de contar con pares de puntos homólogos en imágenes simultáneas se implementaron algoritmos que desarrollan los siguientes procesos.

- Preprocesamiento de las imágenes.
- Detección de contornos
- Cálculo de características
- Obtención de homólogos candidatos en pares de imágenes simultáneas

La fig.3.1 muestra estas etapas y su flujo de control. Las tres primeras etapas operan en forma independiente sobre cada una de las perspectivas. Para la obtención de homólogos candidatos en imágenes simultáneas, como ya se mencionó, las perspectivas derechas e izquierda deben respetar la restricción de epipolaridad, con una traslación  $t$  respecto al eje  $x$ .

Esto dio lugar a la implementación de diversos algoritmos de procesamiento de imágenes monoculares. Entre estos se encuentran los de detección de contornos, y detección de esquinas.

## 4.2. Tratamiento de las Imágenes - Cálculo de Características

La etapa de tratamiento previo de las imágenes, es decir de preprocesamiento, se realizó a los efectos de mejorar la calidad de las perspectiva de la escena. Para ello se recurrió a eliminación del ruido, y a la acentuación de las transiciones que ocurren en las imágenes con el aumento del contraste y/o normalización de la imagen.

El proceso de segmentación se basa en la información de los bordes y en información de la perspectiva original. De esta cooperación resultaron las características utilizadas en el Modelo Funcional.

#### 4.2.1. Detección de Contornos

La detección de contornos es un área ampliamente tratada en la digitalización de imágenes [17][18][19]. Para esta propuesta se desarrollaron herramientas para la detección de contornos basados en gradientes. Estos algoritmos de detección de bordes fueron ampliamente probados en escenas reales y sintéticas variando la eficacia de cada una de acuerdo a la escena.

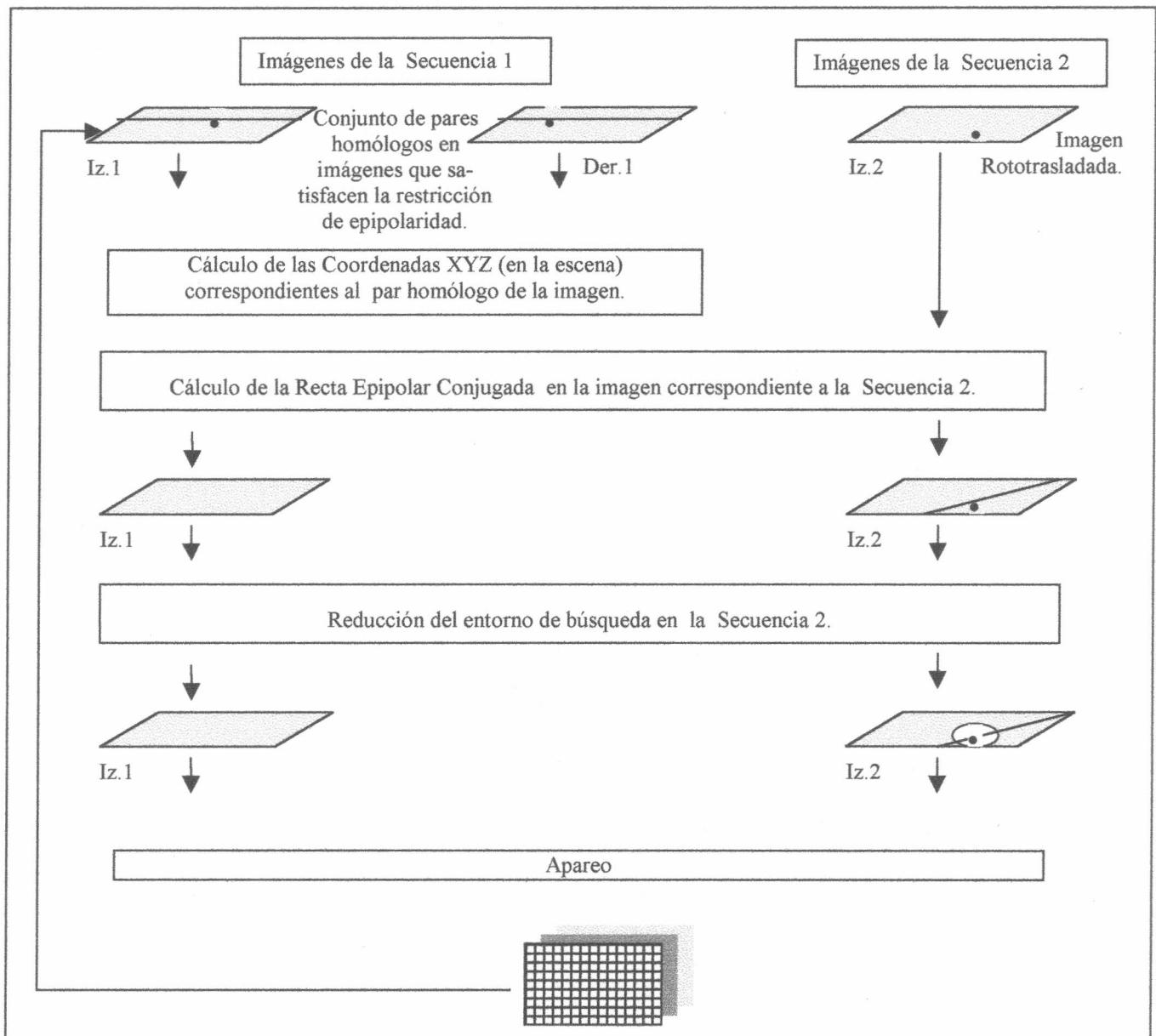
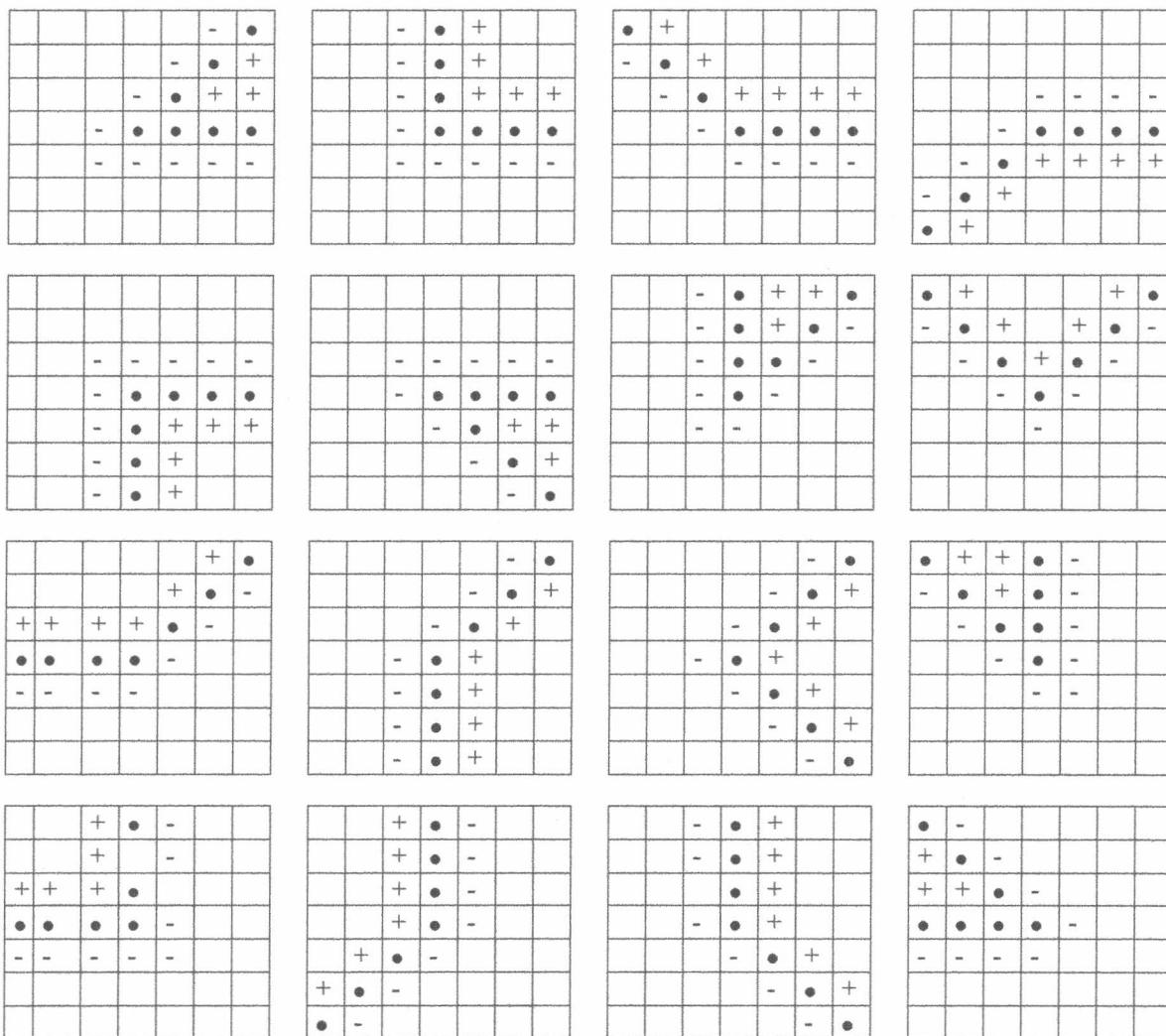


Fig 3.1: Esquema de Procesamiento

#### 4.2.2. Detección de Esquinas.

Para la determinación de las esquinas, positivas o negativas se utilizaron, en cooperación con la perspectiva original, las veinticuatro máscaras de la fig. 3.2. En primer lugar se busca el apareo de las máscaras con el mapa de bordes tomando como centro de la ventana la esquina tratada. En caso de lograr algún apareo se examina también la imagen original para relevar los atributos del entorno de la esquina. Para esto se analizan los niveles de gris de los puntos marcados con (-) y con (+) en la máscara. Si se logra un nivel uniforme de gris para los puntos con (-) la esquina se considera negativa, si la uniformidad corresponde a los (+) se considera esquina positiva. En caso de ambigüedad se las considera ambas.



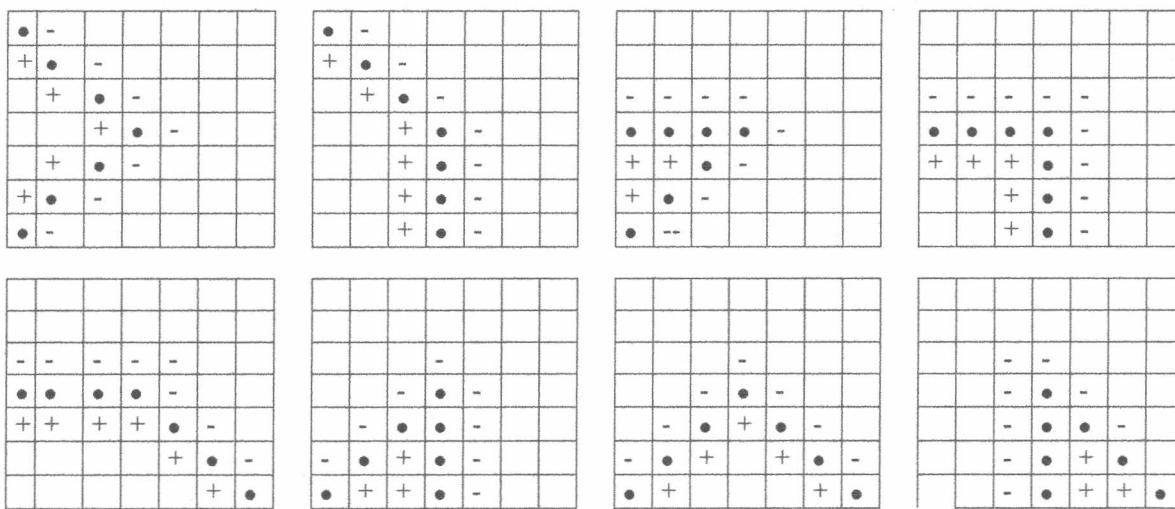


Fig.3.2. Máscaras para la detección de esquinas

### 4.3. Obtención de Pares Homólogos en Imágenes Simultáneas

La rototraslación del sensor tiene como consecuencia que una porción de la imagen en curso no aparezca en la secuencia. La selección de un punto candidato del cual su homólogo no se encuentre en la secuencia debido a la variación de la escena por la rototraslación resulta en un intento fallido. Para minimizar este tipo de casos se procedió a eliminar una porción de los bordes de la escena, la cual se determinó en forma dinámica según la rototraslación efectuada.

### 4.4. Búsqueda del Conjunto Mínimo de Puntos Homólogos.

- Reducción del espacio de búsqueda.
- Apareo
- Cálculo del campo de desplazamiento.

La reducción del espacio de búsqueda se logra con la cooperación de las tres perspectivas.

Finalmente, se logra el apareo con dos secuencias del mismo sensor, en este caso, el izquierdo.

#### 4.4.1. Reducción del Espacio de Búsqueda

Restricción del espacio de búsqueda:

Cálculo del desplazamiento inicial a partir de la cooperación entre la línea epipolar conjugada y el entorno de la posición esperada del punto en la secuencia

#### 4.4.2. Apareo: Algoritmo en Imágenes No Simultáneas

Si bien la primera aproximación propuesta es factible para cualquier punto, dadas las características particulares del problema (encontrar un número restringido de puntos homólogos) su aplicación se restringió a esquinas.

Dado que el proceso de rotorectificación se favorece si el conjunto hallado está formado por puntos cuyas ubicación no es muy próxima, se procedió a ordenar los candidatos de la primer secuencia de acuerdo al módulo de la distancia que los separa. De acuerdo a esto se pueden realizar pequeñas variantes en la primera aproximación para tratar de hallar todos los pares homólogos posibles o detenerse después de hallar el conjunto mínimo requerido. Obteniendo en este último caso aquellos con ubicación óptima para lograr la rotorectificación del par de perspectivas. El algoritmo retorna una matriz con el desplazamiento de los pares homólogos hallados.

### 4.5. Características del Algoritmo:

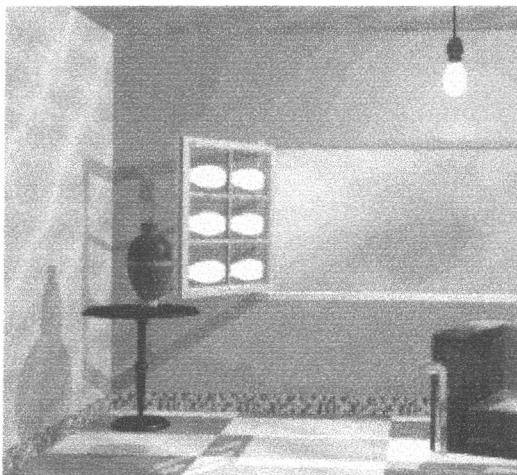
- Flexibilidad para la incorporación de nuevas características.
- Análisis local y global
- Facilidad de integración con estrategias de control adicionales

# Conclusiones y Perspectivas

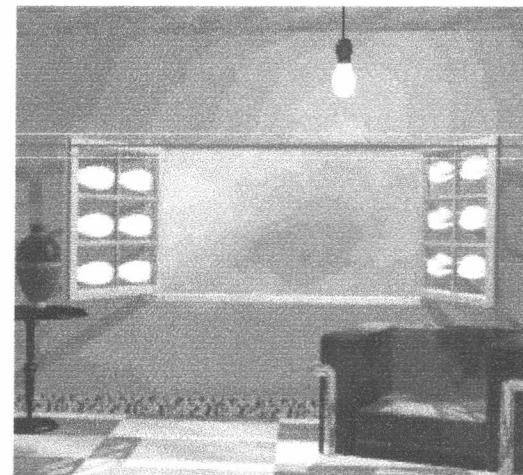
## 5.1. Resultados

La evaluación de las dos soluciones propuestas se realizó sobre imágenes sintéticas y reales.

### Caso1 : Escena Artificial Traslación

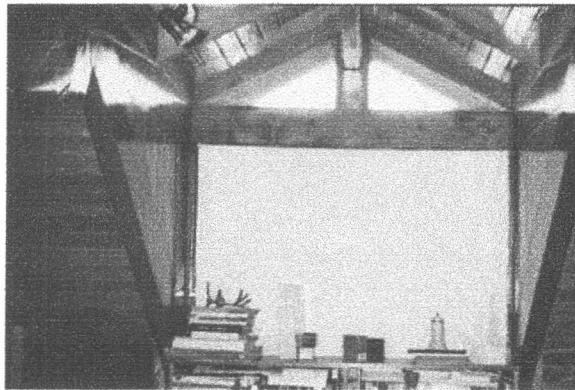


**Secuencia 1:**  
Simulación de toma del sensor izquierdo.



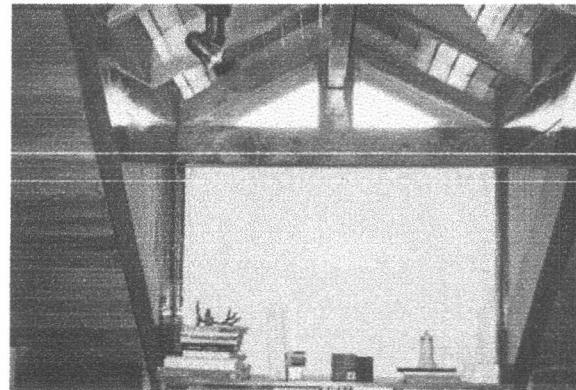
**Secuencia 2:**  
Simulación de toma del sensor izquierdo.  
Traslación:15 píxeles sobre el eje de las X.

## Caso 2 : Escena Biblioteca Traslación



**Secuencia 1:**

Escena adquirida por el sensor Izquierdo

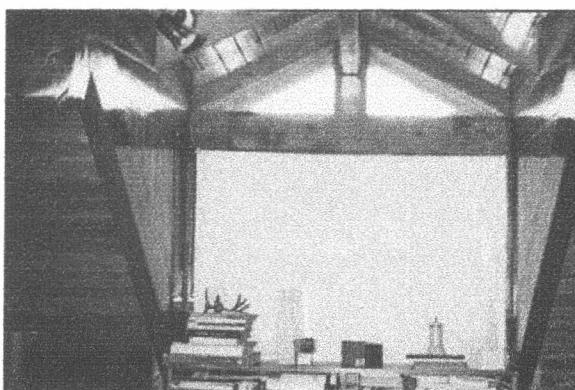


**Secuencia 2:**

Escena tomada por el Sensor Izquierdo .

**Traslación :** en dirección al eje X, 35 cm.  
respecto a la Secuencia 1.

## Caso 3 : Escena Biblioteca Rototraslación



**Secuencia 1:**

Escena adquirida por el sensor Izquierdo



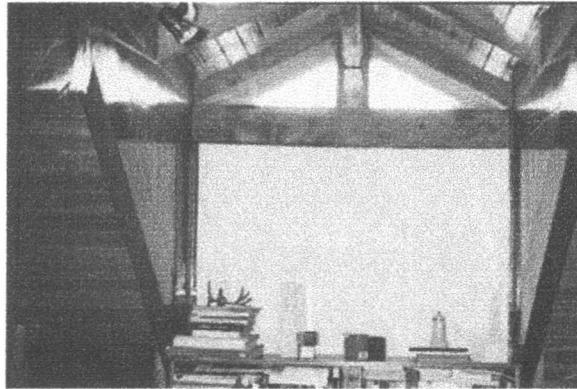
**Secuencia 3:**

Escena tomada por el Sensor Izquierdo.

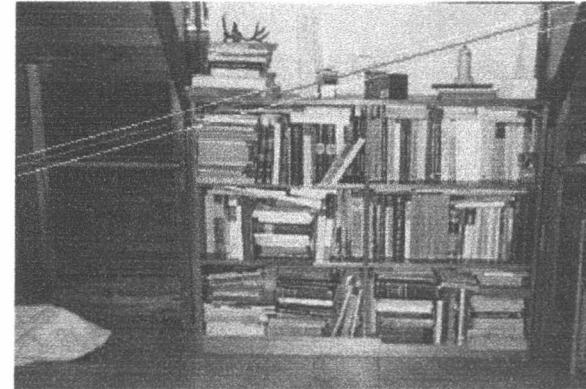
**Traslación:** en dirección al eje X, 35 cm.  
respecto a la Secuencia 1

**Rotación:** alrededor del eje X, 15° respecto a  
la Secuencia 1.

### Caso 4 : Escena Biblioteca Rototraslación

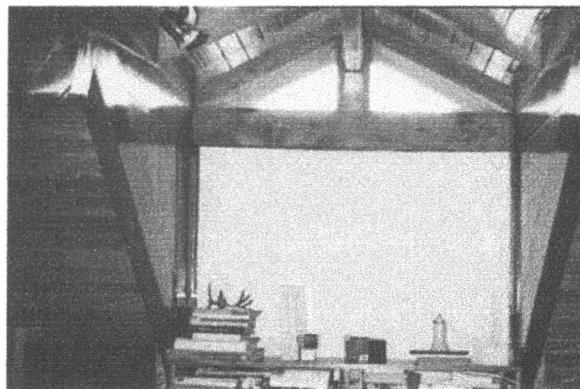


**Secuencia 1:**  
Escena adquirida por el sensor Izquierdo

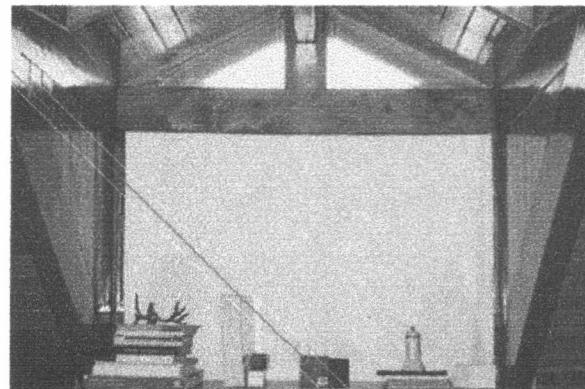


**Secuencia 4:**  
Escena adquirida por el Sensor Izquierdo.  
**Traslación:** en dirección al eje X, 35 cm.  
respecto a la Secuencia 1  
**Rotación:** alrededor del eje Z, 15° respecto a  
la Secuencia 1.

### Caso 5 : Escena Biblioteca Traslación

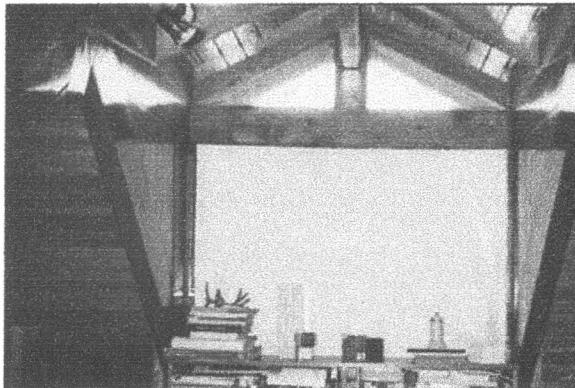


**Secuencia 1:**  
Escena adquirida por el sensor Izquierdo



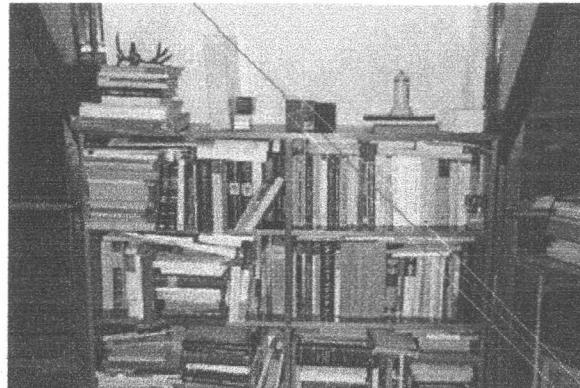
**Secuencia 5:**  
Escena adquirida por el Sensor Izquierdo.  
**Traslación:** en dirección al eje Z, 111 cm.  
respecto a la Secuencia 1

## Caso 6 : Escena Biblioteca Rototraslación



**Secuencia 1:**

Escena adquirida por el sensor Izquierdo



**Secuencia 6:**

Escena adquirida por el Sensor Izquierdo.

**Traslación:** en dirección al eje Z, 111 cm.  
respecto a la Secuencia 1

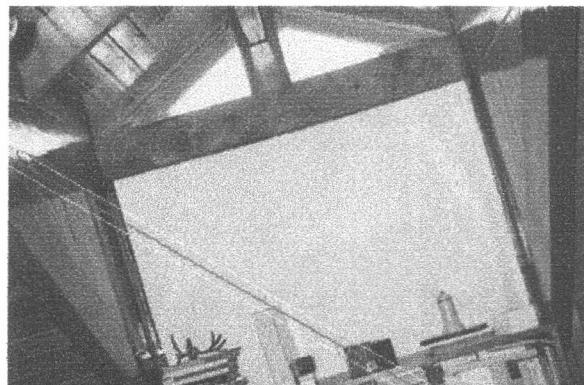
**Rotación:** alrededor del eje X 15° respecto a  
la Secuencia 1.

## Caso 7: Escena Biblioteca Rototraslación



**Secuencia 1:**

Escena adquirida por el sensor Izquierdo



**Secuencia 7:**

Escena adquirida por el Sensor Izquierdo.

**Traslación:** en dirección al eje Z, 111 cm.  
respecto a la Secuencia 1

**Rotación:** alrededor del eje Z 15° respecto a  
la Secuencia 1.

## 5.2. Comparación de Resultados

Imágenes	Dimensión Imagen	Cuadrados Mínimos		Secuencial	
		Homólogos Encontrados	Tiempo Ejecución	Homólogos Encontrados	Tiempo Ejecución
Figura 3D	259 x 240	90 %	12 min	30 %	3 min
Ventana Secuencia 2	346 x 235	85 %	14 min	25 %	6 min
Ventana Secuencia 3	346 x 235	60 %	14 min	15 %	6 min
Ventana Secuencia 4	346 x 235	60 %	14 min	10 %	6 min
Ventana Secuencia 5	346 x 235	55 %	14 min	10 %	6 min
Ventana Secuencia 6	346 x 235	60 %	14 min	15 %	6 min
Ventana Secuencia 7	346 x 235	55 %	14 min	15 %	6 min

## 5.3. Conclusiones.

En todos los casos analizados el método de cuadrados mínimos encontró una mayor cantidad de puntos homólogos. Esto es, debido a que este método no está restringido a buscar sólo esquinas dado que busca una aproximación por cuadrados mínimos de la diferencia de sus características. En el segundo método sólo se aparean entre sí los puntos caracterizados como esquinas, y luego se evalúa que las diferencias de sus características sea, menor que un epsilon suficientemente pequeño. Con este método, en los casos en que solo se produjeron traslaciones se han hallado el 100% de las esquinas que cumplen la condición de ser homólogas, porcentaje que disminuye en el caso en que la imagen se deforme por medio de rotaciones.

Si se desconoce la apariencia de la escena, es aconsejable usar el método de cuadrados mínimos dado que, aunque es más lento, es más factible encontrar los puntos sin depender de la superficie sobre la que estamos evaluando.

Los tiempos resultantes fueron obtenidos por medio de monoprocesamiento. Para cuadrados mínimos el tiempo de ejecución es mayor que los empleados por el otro método, debido a las iteraciones que realiza.

Para este problema en particular, es decir, encontrar un conjunto mínimo de puntos, es posible detener la búsqueda cuando dicho conjunto es hallado.

Los tiempos de ejecución pueden disminuirse, también, aplicando multiprocesamiento. Para ello es posible dividir las perspectivas en subimágenes y cada una de ellas puede ser evaluada en un procesador diferente, con tantos procesadores como puntos a obtener (en nuestro caso es 5).

## 5.4. Otra Aplicación.

Cuando las líneas de píxeles o líneas de imagen coinciden con las líneas epipolares del par estéreo, el apareo automático se simplifica ya que la búsqueda de puntos homólogos, se efectúa esencialmente recorriendo líneas de imagen.

Las imágenes aéreas no cumplen esta condición por lo que se requiere la roto rectificación de pares de imágenes. Con esta operación se logra que las líneas de imagen coincidan con las líneas epipolares.

Para lograr la roto rectificación del par de imágenes, se deben conocer un número mínimo de 5 o 6 pares de puntos. Su búsqueda automática no es simple ya que las líneas de imagen no coinciden con las líneas epipolares. Se limita el espacio de búsqueda, calculando, en la segunda imagen de cada par de una corrida de imágenes aéreas, la línea epipolar correspondiente al punto de la primera imagen del par.

Por otra parte, con el par de imágenes anteriores a una imagen dada, se puede conocer las coordenadas objeto de algunos puntos ubicados en el 20 % común al triple estereoscópico. Esto limita aún más el entorno de búsqueda.

Cuando no se dispone de esta información, se reduce el entorno de búsqueda sobre la línea epipolar, estimando un  $\pm \Delta h$  sobre y por debajo de la altura media del vuelo  $H$ , dato que en general es conocido al estudiar el proyecto.

Con estos elementos, el conocimiento aproximado de los parámetros de orientación externa etc., y contando con puntos de control de la corrida, se obtiene, usando algoritmos de apareo automático, el modelo numérico de terreno.

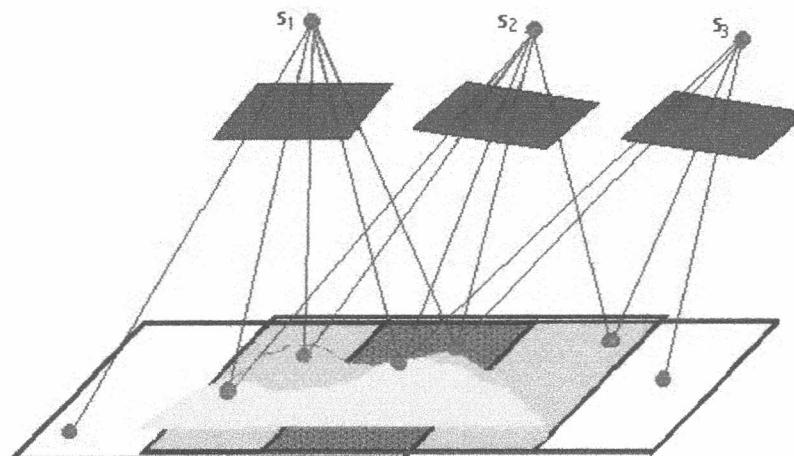


Fig. 5.1: Visión aérea

## 5.5. Problemas Abiertos

Finalmente quedaría pendiente para un trabajo ulterior la modelización de la respuesta perceptual como resultado de la comparación del modelo numérico resultado de los dos sensores con el que resulte de las tomas no sucesivas. En efecto si hubo movimientos en la escena entre una toma y otra podría suceder que algún objeto se hubiera desplazado y esto podría ser reconocido teniendo los resultados de los tres pares: primer par de imágenes simultáneas, segundo par de imágenes simultáneas, imagen izquierda (derecha) de la primera secuencia, imagen izquierda (derecha) de la segunda secuencia.

# Geometría Relativa a la Automatización de la Visión Estereoscópica.\*

## A.1. Estereoscopía

Consideremos un referencial genérico  $Q$  y sea  $S(XS, YS, ZS)$  el punto origen de un sistema referenciado en  $Q$  (fig. A1.1). Si tomamos el origen de  $S$  como punto de vista a una superficie  $M$ ,  $Ms$  un punto de  $M$  visible desde  $S$  y  $m$  la proyección de  $M$  en el plano perspectivo de  $M$  respecto de  $S$ ,  $SM$  puede obtenerse por un factor de escala de  $Sm$ . Se puede hablar de  $M_q(X, Y, Z)$  con respecto al referencial con origen en  $Q$ , o bien de  $M_s(X', Y', Z')$  con respecto al referencial cuyo origen sea  $S$ .

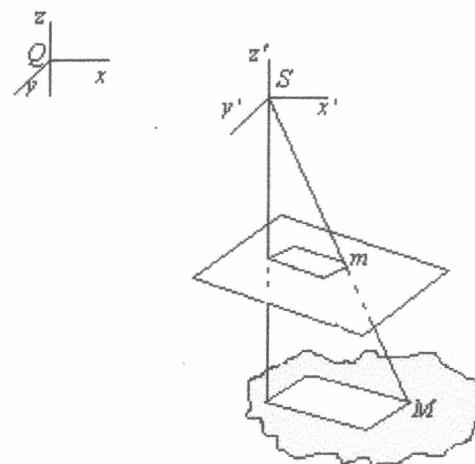


Fig. A1.1: Dos sistemas de coordenadas.

\* Diego A. Dagúm.

Tesis de Grado. Imágenes Estereoscópicas. Visión 3D. UBA, FCEyN, Dpto. de Computación, 1998

Si definimos  $R$  como la matriz de rotación con origen en  $Q$  del referencial  $S$ , entonces  $R^T$  es la matriz de rotación con origen  $S$  del referencial  $Q$  (ver Matrices de rotación).

Así, las coordenadas de  $M$  con referencia a  $S$  surgen de rotar el vector  $Ms$  referido a  $Q$  de acuerdo a

$$M_s = R(M_q - S) \quad (\text{A1.1})$$

La transformación inversa del pasaje de  $M$  referenciado desde  $S$  a  $M$  relativo a  $Q$  es

$$M_q = R^T M_s + S \quad (\text{A1.2})$$

El segmento  $Sm$  es igual a  $Ms$  reducido en un factor de escala  $\lambda$ . Por esto y por la ecuación (A1.1)  $m$  relativo a  $S$  puede obtenerse a partir de  $MQ$  por medio de

$$\begin{bmatrix} x \\ y \\ -p \end{bmatrix} = \lambda R \begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} \quad (\text{A1.3})$$

de (A1.1) y (A1.3) se pueden derivar las siguientes proporcionalidades (fig. A1.2)

$$\frac{x}{-p} = \frac{X}{Z} \quad (\text{A1.4})$$

$$\frac{y}{-p} = \frac{Y}{Z} \quad (\text{A1.5})$$

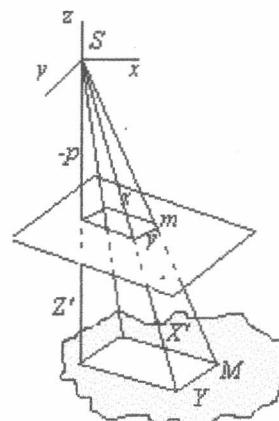


Fig. A1.2: Proporcionalidad de imagen y objeto

Basandonos en (A1.3) y (A1.5) puede expresarse la coordenada  $x$  de  $m$  como

$$x = -p \frac{r_{11}(X - X_s) + r_{12}(Y - Y_s) + r_{13}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} \quad (\text{A1.6})$$

O también considerando  $iR = (r_{11}, r_{12}, r_{13})$ ,  $jR = (r_{21}, r_{22}, r_{23})$  y  $kR = (r_{31}, r_{32}, r_{33})$

$$x = -p \frac{iR(M_Q - S)}{kR(M_Q - S)} \quad (\text{A1.7})$$

Análogamente, la coordenada  $y$  de  $m$  queda expresada en la ecuación

$$y = -p \frac{jR(M_Q - S)}{kR(M_Q - S)} \quad (\text{A1.8})$$

La imagen monocular consiste, entonces, en una transformación de coordenadas sobre un plano  $z = -p$  de una superficie  $\Sigma$ . En si misma no permite conocer con exactitud el relieve.

### A.1.1. El Caso General

La fig. (A1.3) muestra dos referenciales en el espacio de igual distancia principal. El punto  $M(X, Y, Z)$  de la superficie  $S$  puede ser denotado como  $Ms_1(X_1, Y_1, Z_1)$  o también como  $Ms_2(X_2, Y_2, Z_2)$  según se esté refiriendo a las respectivas proyecciones  $m_1$  o  $m_2$ .

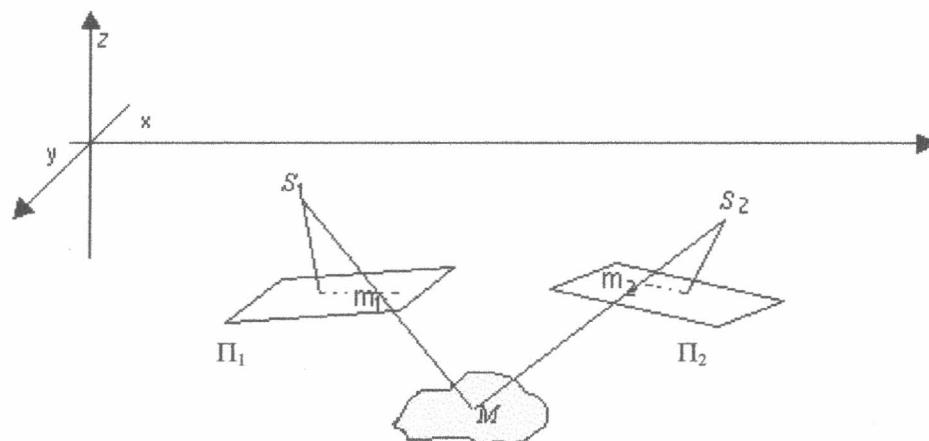


Fig. A1.3: Perspectivas no alineadas.

De la ecuación (A1.1) tenemos que

$$M_{S1} = R_l M - R_l S_1 \quad (\text{A1.9})$$

Por simplicidad se establecerá lo siguiente:

$$L^{(1)} = \begin{bmatrix} L_1^{(1)} \\ L_2^{(1)} \\ L_3^{(1)} \end{bmatrix} = R_1 S_1 \quad (\text{A1.10})$$

El supraíndice se usa para distinguirlo del análogo del segundo referencial:  $L^{(2)}$ . Hay que notar que tanto la rotación  $R_1$  como el centro de perspectiva  $S_1$  son datos conocidos, por lo tanto es conocido  $L^{(1)}$ . Así se puede remplazar en la ecuación (A1.8)

$$M_1 = R_1 M - L_1^{(1)} \quad (\text{A1.11})$$

Por la ecuación (A1.4) puede derivarse la relación entre la componente  $X_1$  de  $M_{S1}$ , el punto  $m_1$  en la perspectiva  $\Pi$ , y la componente  $Z_1$ .

$$X_1 - \frac{x_1}{-p} Z_1 = 0 \quad (\text{A1.12})$$

Cabe hacer mención al hecho de que  $p$ , la distancia principal, es común a ambos referenciales. Ahora se aplicará parte de la ecuación (A1.11) en la (A1.12)

$$iR_1 M - L_1^{(1)} - \frac{x_1}{-p} (\bar{k}R_1 M - L_3^{(1)}) = 0 \quad (\text{A1.13})$$

Se ha sustituido  $X_1$  por la parte de la ecuación (A1.11) que atañe a la primera componente de  $M_{S1}$ . Asimismo se hizo lo propio con  $Z_1$ , la tercera componente de  $M_{S1}$ .

Con un razonamiento análogo, partiendo de la ecuación (A1.5) —como recién se hizo con la ecuación (A1.4)— y por medio de la ecuación (A1.11) se puede concluir que

$$jR_1 M - L_2^{(1)} - \frac{y}{-p} (\bar{k}R_1 M - L_3^{(1)}) = 0 \quad (\text{A1.14})$$

Se pueden utilizar dos sustituciones,

$$U_1 = \frac{x_1}{-p} \quad V_1 = \frac{y}{-p}$$

y así simplificar los resultados de las ecuaciones (A1.13) y (A1.14) en la siguiente forma matricial

$$\begin{bmatrix} iR_1 - U_1 kR_1 \\ jR_1 - V_1 kR_1 \end{bmatrix} M = \begin{bmatrix} L_1^{(1)} - U_1 L_3^{(1)} \\ L_2^{(1)} - V_1 L_3^{(1)} \end{bmatrix} \quad (\text{A1.15})$$

Esto nos deja un grado de libertad ya que son sólo dos ecuaciones contra las tres incógnitas de  $M$ . Ahora bien, todo esto en cuanto al primer referencial. Para el segundo referencial se procede análogamente para concluir que

$$\begin{bmatrix} iR_2 - U_2 kR_2 \\ jR_2 - V_2 kR_2 \end{bmatrix} M = \begin{bmatrix} L_1^{(2)} - U_2 L_3^{(2)} \\ L_2^{(2)} - V_2 L_3^{(2)} \end{bmatrix} \quad (\text{A1.16})$$

Esto hace cuatro ecuaciones contra tres incógnitas, por lo que una solución que aproxime  $M$  contemplando las cuatro ecuaciones juntas bien podría ser el método lineal de

mínimos cuadrados. Este mecanismo minimiza el error cuadrático medio de una estimación. Aunque no se ahondará en la filosofía del mismo, para este problema puntual se plantean dos matrices  $A$  y  $l$  como se ve a continuación:

$$A = \begin{bmatrix} iR_1 - U_1 k R_1 \\ jR_1 - V_1 k R_1 \\ iR_2 - U_2 k R_2 \\ jR_2 - V_2 k R_2 \end{bmatrix} \quad l = \begin{bmatrix} L_1^{(1)} - U_1 L_3^{(1)} \\ L_2^{(1)} - V_1 L_3^{(1)} \\ L_1^{(2)} - U_2 L_3^{(2)} \\ L_2^{(2)} - V_2 L_3^{(2)} \end{bmatrix}$$

El punto  $M$  en la superficie del objeto surge de resolver el sistema

$$A^t A M = A^t l$$

### A.1.2. El Problema - Estereoscópico General

Este escenario queda abierto a cualquier posibilidad. Tomando en cuenta todas las ecuaciones a lo largo de este capítulo, puede plantearse que ahora son conocidos los valores que antes eran incógnita y viceversa. Puede ocurrir, por ejemplo, que no se conozcan los puntos de vista, las rotaciones o ambos, pero sí un puñado de puntos  $M$  en la superficie de  $\Sigma$ , con sus respectivas parejas homólogas  $m_1$  y  $m_2$ . Esto nos puede facilitar el armar los sistemas de ecuaciones necesarios para resolver las incógnitas de las distintas situaciones.

En general, en última instancia el problema consiste en reconocer el relieve del objeto. Esto es, todos los puntos  $M$  de su superficie  $\Sigma$ . Cuando se parte de la base que algunos puntos  $M$  son conocidos junto con sus correspondientes pares de puntos homólogos, conocidos los parámetros de rigor, podría considerarse la posibilidad de rotorectificar las imágenes, para luego hallar todos los pares de homólogos y, finalmente, proceder a la restitución.

## A.2. Rotaciones en el Espacio

Una rotación es un giro alrededor de un eje. Este giro determina una transformación de las coordenadas espaciales. En la fig. 1.1, se ve una rotación en torno al eje  $z$  (considerando este como la normal al plano determinado por los ejes  $x$  e  $y$  de la figura). El ángulo de giro será, en este caso, normalmente simbolizado por la letra griega  $\kappa$  (kapa).

Dado un punto  $(x_i, y_i, z_i)$  con respecto a los ejes cartesianos previos a la rotación, las coordenadas del mismo luego de la rotación se calculan:

$$x_i = x_i \cos(\kappa) + y_i \sin(\kappa) \quad (\text{A2.1a})$$

$$y_i = x_i (-\sin(\kappa)) + y_i \cos(\kappa) \quad (\text{A2.1b})$$

$$z_i = z_i$$

Las ecuaciones (A2.1) pueden re expresarse conjuntamente en forma matricial como

$$\begin{bmatrix} x' \\ y' \\ z \end{bmatrix} = R_z \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

con  $R_z$  matriz de rotación alrededor del eje  $z$ ; equivalente a

$$R_z = \begin{bmatrix} \cos k & \sin k & 0 \\ -\sin k & \cos k & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{A2.2 a})$$

En forma análoga una rotación de ángulo  $\omega$  alrededor del eje  $x$  puede representarse en notación matricial como

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{bmatrix} \quad (\text{A2.2 b})$$

Finalmente la matriz asociada a una rotación de ángulo  $\varphi$  (phi) alrededor del eje  $y$  es equivalente a

$$R_y = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{bmatrix} \quad (\text{A2.2 c})$$

Toda rotación alrededor de un eje arbitrario puede descomponerse en tres rotaciones primitivas alrededor de los ejes cartesianos, las cuales, aplicándose consecutivamente, producen la rotación original. Pero la descomposición de rotaciones no es conmutativa. Intuitivamente, si primero se rota alrededor de  $x$  y después alrededor de  $y$ , la segunda rotación se aplica sobre el eje  $y'$  generado por la primera rotación y no sobre el eje  $y$  previo a la rotación de  $x$ .

La matriz  $R$  asociada a una rotación alrededor a un eje cualquiera, por los anteriormente vistos, puede representarse como el producto de las matrices asociadas a rotaciones en los tres ejes de coordenadas. La no conmutatividad del producto matricial reafirma el concepto de que la composición de rotaciones no puede permutarse. En caso de hacerse, los valores de las rotaciones serán diferentes.

Aplicar una rotación de matriz  $R$  a un punto  $x$ , es multiplicar a izquierda la matriz  $R$  al punto  $x$ . Sin pérdida de generalidad, si  $R$  es la matriz asociada a una rotación, se acostumbra designar a la misma como "la rotación  $R$ ". La rotación compuesta  $r$  de dos rotaciones consecutivas,  $R_1$  y  $R_2$ , aplicadas a un punto  $x$ , es el producto a izquierda de  $R_2$  sobre  $R_1$ , como se muestra a continuación:

$$R_x = (R_2 R_1)x = R_2(R_1x) = R_2x' \quad (\text{A2.3})$$

### A.2.1. Otras propiedades de las matrices de rotación:

**A.2.1.1)**  *$R^t$ , la matriz asociada a la rotación inversa de  $R$ , es equivalente a  $R^T$ .*

Puede probarse la propiedad A2.1.1) para  $R_x$ ,  $R_y$ ,  $R_z$ , las matrices de las ecuaciones (A2.2), simplemente realizando los productos  $R_i R_i^t$ ,  $R_i^t R_i$  (con  $i = x, y, z$ ) verificando en ambos casos que resulta la identidad. Sean  $R_1, R_2, R_3$  la descomposición de  $R$  en rotaciones consecutivas alrededor de los tres ejes de coordenadas. No necesariamente  $R_1$  rotando alrededor de  $x$ ,  $R_2$  alrededor de  $y$ , etc. Luego se plantea que

$$RR^t = (R_3 R_2 R_1)(R_3 R_2 R_1)^t$$

La transposición de un producto es el producto de las transposiciones. Se aplica esta propiedad y también la propiedad asociativa del producto matricial, obteniendo

$$= (R_3 R_2 R_1)(R_1^t, R_2^t, R_3^t) = R_3 R_2 (R_1 R_1^t) R_2^t R_3^t$$

Aquí viene el paso inductivo: como  $RR^t$  es la identidad (ya que en  $R_1$  se cumple la propiedad), puede eliminarse. Del mismo modo se procede con  $R_2$  y con  $R_3$ , quedando demostrado que para una rotación cualquiera  $RR^t$  es la identidad.

La misma operatoria se repite para  $R^t R$  quedando demostrada la propiedad A2.1.1).

**A.2.1.2)** *El determinante de la matriz  $R$  es 1.*

La demostración es análoga a la propiedad A2.1.1). Se comprueba que para  $R_x$ ,  $R_y$ ,  $R_z$ , las matrices de las ecuaciones (A2.2), la propiedad se verifica. Luego se plantea  $R$  como el producto de tres rotaciones (nuevamente  $R_1$ ,  $R_2$  y  $R_3$ ) alrededor de los tres ejes cartesianos, para los cuales se cumple la propiedad. Se sigue con el hecho de que el determinante de un producto es el producto de los determinantes.

$$\text{Det}(R) = \text{Det}(R_3 R_2 R_1) = \text{Det}(R_3)\text{Det}(R_2)\text{Det}(R_1)$$

Y se aplica aquí el paso inductivo

$$\text{Det}(R) = I * I * I = I$$

Quedando demostrado que la propiedad A2.1.2) se cumple para cualquier rotación  $R$ .

**A.2.1.3)** *Fórmula de Euler para las matrices de rotación*

Sea  $u(\alpha, \beta, \gamma)$  un versor (un vector de norma unitaria) en el espacio. Sea  $R$  un giro (rotación) de un ángulo  $\theta$  con eje en  $u$ . Se define la matriz  $A$  de la ecuación (A2.4) como axiador del versor  $u$ .

$$A = \begin{bmatrix} 0 & -\lambda & \beta \\ \lambda & 0 & -\alpha \\ -\beta & \alpha & 0 \end{bmatrix} \quad (\text{A2.4})$$

El axiador cumple para todo vector  $v = (v_1, v_2, v_3)$ , la siguiente propiedad, respecto del producto vectorial

$$u \wedge v = Av \quad (\text{A2.5})$$

Esto puede comprobarse fácilmente si se desarrollan ambos miembros de la ecuación

$$\begin{bmatrix} i & j & k \\ \alpha & \beta & \gamma \\ v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} -\gamma v_2 + \beta v_3 \\ \lambda v_1 - \alpha v_3 \\ -\beta v_1 + \alpha v_2 \end{bmatrix} = \begin{bmatrix} 0 & -\lambda & \beta \\ \lambda & 0 & -\alpha \\ -\beta & \alpha & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Se va a tratar de encontrar una expresión para  $R$  a partir de  $\alpha$  y  $\theta$ . El desarrollo en series de Taylor de la función  $e^x$  es el que sigue

$$e^x = \sum_n \frac{x^n}{n!}$$

Así el desarrollo de  $e^{A\theta}$  tendrá la forma de la ecuación (A2.6)

$$e^{A\theta} = I + \frac{A\theta}{1!} + \frac{(A\theta)^2}{2!} + \frac{(A\theta)^3}{3!} + \frac{(A\theta)^4}{4!} + \dots \quad (\text{A2.6})$$

Cabe destacar que el axiador  $A$  tiene la característica de ser exponencialmente periódico. En efecto, el lector podrá verificar que

$$A^3 = -A$$

$$A^4 = -A^2$$

Y, en general

$$A^i = -A^{i-2} \quad (\text{A2.7})$$

donde corresponda se puede usar la ecuación (A2.7) para simplificar la ecuación (A2.6) con

$$e^{A\theta} = I + \frac{A\theta}{1!} + A \frac{\theta^2}{2!} - A \frac{\theta^3}{3!} - A^2 \frac{\theta^4}{4!} + \dots \quad (\text{A2.8})$$

Agrupando los términos de exponente par por un lado y de exponente impar por otro,

$$e^{A\theta} = \left( A\theta - A \frac{\theta^3}{3!} + A \frac{\theta^5}{5!} - \dots \right) + I + \left( A^2 \frac{\theta^2}{2!} - A^2 \frac{\theta^4}{4!} + A^2 \frac{\theta^6}{6!} - \dots \right) \quad (\text{A2.9})$$

La expresión del primer paréntesis, teniendo en cuenta el desarrollo de Taylor de  $\sin(x)$  en la ecuación (A2.10)

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (\text{A2.10})$$

nos permite reemplazar en la ecuación (A2.9), como se muestra a continuación

$$e^{A\theta} = Asin\theta + I + \left( A^2 \frac{\theta^2}{2!} - A^2 \frac{\theta^4}{4!} + A^2 \frac{\theta^6}{6!} - \dots \right) \quad (\text{A2.11})$$

Considerando el desarrollo de Taylor del  $\cos(x)$  de la ecuación (A2.12)

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (\text{A2.12})$$

se va a sumar y restar  $A^2$  en la expresión entre paréntesis de la ecuación (A2.11)

$$e^{A\theta} = Asin\theta + I + \left( A^2 - A^2 + A^2 \frac{\theta^2}{2!} - A^2 \frac{\theta^4}{4!} + A^2 \frac{\theta^6}{6!} - \dots \right) \quad (\text{A2.11})$$

para poder simplificar la serie con la expresión (A2.12)

$$e^{A\theta} = Asin\theta + I + (A^2 - A^2 \cos\theta)$$

y sacando  $A^2$  como factor común, se obtiene la expresión conocida como fórmula de Euler o de D'Olindes – Rodríguez.

$$e^{A\theta} = Asin\theta + I + A^2(1 - \cos\theta) \quad (\text{A2.13})$$

Esta fórmula define la rotación de u con giro  $\theta$ . Presenta las siguientes propiedades:

**A.2.1.4)** *Rotar en el espacio en  $\theta$  (grados, radianes) en un sentido y luego en el otro,  $(-\theta)$  componen la rotación identidad.*

En efecto, si se expresa la composición de rotaciones según la fórmula de Euler

$$e^{A\theta} e^{-A\theta} = (I + Asin\theta + A^2(1 - \cos\theta))(I + Asin(-\theta) + A^2(1 - \cos(-\theta))) \quad (\text{A2.14})$$

se pueden aplicar las siguientes propiedades trigonométricas

$$\sin \theta = -\sin(-\theta) \quad (\text{A2.15 a})$$

$$\cos \theta = \cos(-\theta) \quad (\text{A2.15 b})$$

con lo que el producto (A2.14) se transforma en

$$e^{A\theta} e^{A\theta} = (I + A\sin\theta + A^2(1 - \cos\theta))(I - A\sin\theta + A^2(1 - \cos\theta))$$

Aplicando propiedad distributiva se llega a

$$\begin{aligned} e^{A\theta} e^{-A\theta} &= I + A\sin\theta + A^2(1 - \cos\theta) - A\sin\theta - A^2\sin^2\theta - A^3\sin\theta(1 - \cos\theta) + \\ &+ A^2(1 - \cos\theta) + A^3\sin\theta(1 - \cos\theta) + A^4(1 - \cos\theta)^2 \end{aligned}$$

Aquí se pueden ir simplificando los términos que involucran al axiador elevando a potencia impar, ya que tienen signo contrario

$$e^{A\theta} e^{-A\theta} = I + A^2(1 - \cos\theta) - A^2\sin^2\theta + A^2(1 - \cos\theta) + A^4(1 - \cos\theta)^2$$

También se van a juntar términos idénticos, a la par que se aprovechará la periodicidad exponencial del axiador (ecuación A2.7) para sustituir donde corresponda

$$e^{A\theta} e^{-A\theta} = I + 2A^2(1 - \cos\theta) - A^2\sin^2\theta - A^2(1 + \cos^2\theta - 2\cos\theta)$$

Ahora se distribuye el factor  $-A^2$  con respecto a la suma

$$e^{A\theta} e^{-A\theta} = I + 2A^2 - 2A^2\cos\theta - A^2\sin^2\theta - A^2 - A^2\cos^2\theta + 2A^2\cos\theta$$

Seguidamente se pueden eliminar términos que involucran a  $A^2$  y que son inversos aditivos

$$e^{A\theta} - e^{-A\theta} = I + A^2 - A^2\sin^2\theta - A^2\cos^2\theta \quad (\text{A2.16})$$

Ahora se puede usar la siguiente propiedad trigonométrica

$$\cos^2\theta + \sin^2\theta = 1 \quad (\text{A2.17})$$

para reemplazar en (A2.16) quedándonos

$$e^{A\theta} e^{-A\theta} = I + A^2 - A^2(1 - \cos^2\theta) - A^2\cos^2\theta$$

Finalmente se distribuye y resta obteniendo el resultado esperado

$$e^{A\theta} e^{-A\theta} = I + A^2 - A^2 + A^2\cos^2\theta - A^2\cos^2\theta = I$$

**A.2.1. 5) Aplicar la rotación al vector  $u$  no le cambia sus coordenadas**

$$e^{A\theta}u = Iu + Asin \theta u + A^2(1 - \cos \theta)u$$

Se puede usar la ecuación (A2.5) y sustituir, obteniendo

$$e^{A\theta}u = u + sin \theta u \wedge u + A(1 - \cos \theta)u \wedge u$$

El producto vectorial de un vector por si mismo es cero, esto es

$$u \wedge u = 0$$

Así se anulan los dos últimos términos, resultando lo que se quería ver

$$e^{A\theta}u = u$$

**A.2.1.6) Siendo  $v$  ortonormal a  $u$ , la fórmula de Euler aplicada a  $v$  se simplifica como sigue**

$$e^{A\theta}\vec{v} = I\vec{v} + sin \theta A\vec{u} + A(1 - \cos \theta)A^2\vec{v}$$

Cuando  $\vec{v}$  es ortonormal a  $u$ , se cumple que

$$A^2\vec{v} = -\vec{v}$$

Con esto en vista se puede sustituir en (A2.18) resultando

$$e^{A\theta}\vec{v} = \vec{v} + sin \theta A\vec{v} + A\vec{v} - (1 - \cos \theta)\vec{v}$$

Distribuyendo, se podrán anular los sueltos, llegando a la ecuación

$$e^{A\theta}\vec{v} = sin \theta A\vec{v} + \cos \theta\vec{v}$$

Por supuesto, la fórmula de Euler también verifica las dos propiedades (A2.1) y (A2.2) que cumplen todas las matrices de rotación. Asimismo puede comprobarse que

**A.2.1.7) El coseno del ángulo de giro se relaciona con la traza de  $e^{A\theta}$  según**

$$\cos \theta = \frac{1}{2} \left( Tr(e^{A\theta}) - 1 \right)$$

Según la fórmula de Euler, en la diagonal de  $e^{A\theta}$  se va a encontrar

$$e^{A\theta} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix} \sin \theta + \begin{bmatrix} -\gamma^2 - \beta^2 & & \\ & -\gamma^2 - \alpha^2 & \\ & & -\beta^2 - \alpha^2 \end{bmatrix} (1 - \cos \theta)$$

La traza, por consiguiente es igual a

$$(\gamma^2 + \beta^2)(\cos \theta - 1) + 1 + (\gamma^2 + \alpha^2)(\cos \theta - 1) + 1 + (\alpha^2 + \beta^2)(\cos \theta - 1) + 3$$

Si se aplica la propiedad distributiva del producto respecto de la suma, se llega a que

$$(\gamma^2 + \beta^2 + \alpha^2 + \gamma^2 + \alpha^2 + \beta^2)(\cos \theta - 1) + 3$$

Como  $\mathbf{u} = (\alpha, \beta, \gamma)$  es un versor de norma unitaria, la suma de los cuadrados de sus coordenadas es la unidad.

$$(1+1)(\cos \theta - 1) + 3$$

Distribuyendo se tiene una traza equivalente a

$$\text{Tr}(e^{A\theta}) = 2 \cos \theta - 2 + 3 = 2 \cos \theta + 1$$

Lo que demuestra la proposición.

#### A.2.1.8)

$$A \sin \theta = \frac{1}{2} (e^{A\theta} - e^{-A\theta})$$

En efecto

$$e^{A\theta} - e^{-A\theta} = (I + A \sin \theta) + A^2 (1 - \cos \theta) - (I - A \sin \theta + A^2 (1 - \cos \theta))$$

Anulando términos repetidos de signo contrario, se llega a

$$e^{A\theta} - e^{-A\theta} = A \sin \theta + A \sin \theta = 2 A \sin \theta$$

Lo que prueba la proposición.

Una interesante observación es que las dos propiedades A2.6) y A2.7 sirven para conocer  $A$  y  $\theta$ , conocida la matriz  $e^{A\theta}$  que responde a la fórmula de Euler.

#### A.2.1.9) Pequeñas rotaciones

Cuando el ángulo de giro es muy pequeño, se cumple

$$\sin \theta \approx \theta$$

$$\cos \theta \approx 1$$

Con esta premisas, en la fórmula de *Euler* puede despreciarse el término cuadrático ya que

$$1 - \cos \theta \approx 0$$

en tanto que el término lineal puede estimarse con  $A\theta$ . La fórmula puede aproximarse como sigue

$$e^{A\theta} \approx I + A\theta$$

En base a esto se puede decir que las pequeñas rotaciones son aproximadamente conmutativas. La explicación resulta ser la siguiente: siendo  $e^{A\theta}$  y  $e^{A'\theta'}$  dos rotaciones, su composición resulta

$$e^{A\theta} e^{A'\theta'} \approx (I + A\theta)(I + A'\theta')$$

Distribuyendo y despreciando el término  $A\theta A'\theta'$  se obtiene

$$e^{A\theta} e^{A'\theta'} \approx I + A\theta + A'\theta' = I + A'\theta'' + A\theta \approx e^{A'\theta'} e^{A\theta}$$

### A.3. Rotorectificación

Existe un gran número de algoritmos de apareo de puntos homólogos en pares estéreo que requieren que las imágenes respeten la *restricción de epipolaridad* con el objeto de reducir la dimensión de búsqueda. Esta restricción consiste en que ambas imágenes sean coplanares y que tanto dicho plano como las líneas de las imágenes (líneas de píxel en la práctica), sean paralelas al segmento llamado *base*, que es el que une los puntos de vista. Para llevar a que dos imágenes de un objeto en posiciones arbitrarias, cumplan esta restricción busca una posición común para ambas imágenes, y con una *distancia principal* apropiada - que debe ser la misma que la de las imágenes originales- procede a proyectarlas en un plano situado a esa distancia, y en la nueva orientación. Esto se conoce como *rectificación*. Para alcanzar la epipolaridad de las líneas de píxeles, la proyección debe hacerse *gimando* esas líneas en las imágenes de modo que, como se ha dicho, las líneas de píxel sean paralelas a la línea que pasa por los dos puntos de vista. El proceso conjunto se denomina *rotorectificación*.

La restricción a cumplir, como se mencionó antes, es que las imágenes iniciales estén ambas situadas a igual distancia principal de sus respectivos puntos de vista. Primero determina la orientación relativa de una imagen respecto de otra. Para ello utiliza un algoritmo de los llamados *de búsqueda lineal*, algoritmos tales que en forma cíclica e incremental devuelven estimadores. Posteriormente, a partir de las rotaciones relativas propone un posible *plano óptimo* donde roto-rectificar las imágenes

### A.3.1. Algoritmos de búsqueda lineal

Estos algoritmos se utilizan para estimar un  $x$  perteneciente a un espacio de dimensión cualquiera, sujeto a ciertas restricciones (los llamados *problemas de optimización*). Generalmente estos algoritmos comienzan asignando un estimador inicial  $x^{(0)}$

Los diferentes métodos se basan en la información disponible en el modelo funcional. Diversos criterios de parada pueden ser usados para detener el cómputo. El criterio adecuado es cuando los errores de estimación son inferiores a un  $e$  en todos los casos. Pero nada garantiza esa convergencia en un número acotado de iteraciones, por lo que es válido combinar ese criterio ideal con la premisa de no iterar más que un máximo. De ambas condiciones, la que ocurría en primer término.

### A.3.2. Linearización del Problema de la Orientación Relativa.

Con un esquema similar al de la fig. A3.1, se tienen dos referentes de origen  $S_1$  y  $S_2$  representados por las matrices de rotación  $R_1$  y  $R_2$  respectivamente. Se puede asumir para las mismas un valor inicial equivalente a la matriz identidad, que va a ir ajustándose con las iteraciones.

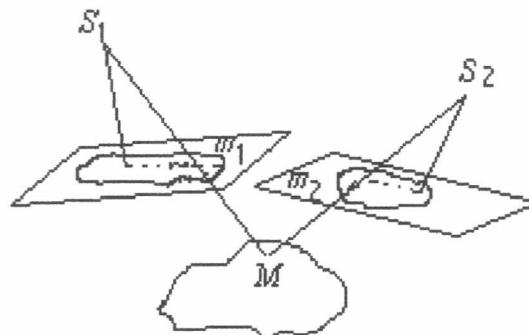


Fig A3.1 : Perspectivan que no cumplen la restricción de epipolaridad.

Para calcular la orientación absoluta se necesitan conocer  $S_1$  y  $S_2$ . Es posible asumir valores prefijados para los mismos. Para lo que sigue, se requiere elegir  $S_1$  y  $S_2$  de modo que la base sea paralela a la dirección  $x$  del referencial. Esto nos limita a conocer sólo la orientación relativa entre las imágenes, dado que ésta es independiente de la traslación.

La distancia principal  $p$  podría ser un valor medido, o bien uno arbitrario. El impacto de usar cualquier  $p$  no alcanza a la rotorectificación en su objetivo de alcanzar la epipolaridad, no obstante producir efectos de anamorfosis a consecuencias de que los rayos homólogos se cortan a mayor o menor distancia, según se elija un  $p$  más grande o más chico (respectivamente) que el verdadero. Las orientaciones relativas van a depender del  $p$  elegido. Las componentes  $\omega$  y  $\varphi$  de las matrices de rotación  $R_x$  y  $R_y$  respectivamente verán afectadas. Si el  $p$  elegido es considerablemente mayor al real, los dos valores de  $\omega$  y  $\varphi$  serán 0, es decir, prácticamente  $R_1$  y  $R_2$  presentarán giros  $k$  alrededor del eje z única-

mente. Por el contrario, tomar un  $p$  más chico que el verdadero sobre dimensiona los giros alrededor de los ejes  $x$  e  $y$ . Por último en cuanto al objeto, un  $p$  chico lo sitúa más cerca de los puntos de vista y reduce sus dimensiones a lo largo del eje  $z$ . Un  $p$  grande lo aleja de la línea base y alarga las dimensiones. Debido a esta deformación el objeto reconstruido es anamorfo respecto al observado.

Se dispone de un conjunto  $q$  de pares de puntos homólogos en sendos planos imágenes. Estos puntos de base observados son relativos a sus referenciales. Esto es, sus coordenadas toman como origen el punto de vista correspondiente  $S_1$  o  $S_2$ . Se pueden obtener, por ejemplo por pseudo-intersección de rectas alabeadas (o por métodos más generales), valores iniciales de los  $M_i$  en la superficie  $\Sigma$  del objeto correspondiente a cada par- ya que se asume la matriz identidad como valores iniciales para  $R_1$  y  $R_2$ . Se cuenta, pues, con valores iniciales  $R_1(0)$ ,  $R_2(0)$  y  $M_i(0)$  – el supraíndice indica el número de iteración -. Para cada iteración, se pretende que

$$R_1^{n+1} = r_1^n R_1^n \quad (\text{A3.1.a})$$

$$R_2^{(n+1)} = r_2^{(n)} R_2^{(n)} \quad (\text{A3.1.b})$$

$$M_q^{(n+1)} = M_q^{(n)} + \partial M_q^{(n)} (\forall q) \quad (\text{A3.2})$$

Donde  $r_1$  y  $r_2$  designan matrices de rotación diferencial (ajustes de estimación), los  $\partial M_q$  son ajustes de estimación de cada  $M_q$ . Las ecuaciones (A3.1) y (A3.20) expresan que al final de cada iteración  $n$  se ajustan los valores estimados, en base a los errores calculados en ese paso. Con las nuevas estimaciones se procede a la iteración siguiente, y así sucesivamente hasta que se cumpla el criterio de parada.

Tenemos:

$$x_{1,q}^{(n)} = -p \frac{iR_1^{(n)}(M_q^{(n)} - S_1)}{kR_1^{(n)}(M_q^{(n)} - S_1)} \quad (\text{A3.3 a})$$

$$y_{1,q}^{(n)} = -p \frac{jR_1^{(n)}(M_q^{(n)} - S_1)}{kR_1^{(n)}(M_q^{(n)} - S_1)} \quad (\text{A3.3b})$$

y análogamente

$$x_{2,q}^{(n)} = -p \frac{iR_2^{(n)}(M_q^{(n)} - S_2)}{kR_2^{(n)}(M_q^{(n)} - S_2)} \quad (\text{A3.3c})$$

$$y_{2,q}^{(n)} = -p \frac{jR_2^{(n)}(M_q^{(n)} - S_2)}{kR_2^{(n)}(M_q^{(n)} - S_2)} \quad (\text{A3.3d})$$

Donde  $i, j, k$  son los tres versores directores de los ejes  $x, y, z$  respectivamente.

Por la ecuación (A1.2), el factor de escala entre  $m1, q(n)$  y  $(Mq(n) - S_1)$  representado por  $\lambda_{1,q}(n)$ , está dado por un cociente de normas al cuadrado, según

$$\lambda_{1,q}^{(n)} = \frac{|R_{1,q}^{\pm(n)} m_{1,q}^{(n)}|}{|M_{1,q}^{(n)} - S_1|}$$

La ecuación para  $\lambda_{2,q}(n)$  es análoga.

Teniendo en cuenta el  $x_{1,q}$  observado y el  $x_{1,q}$  calculado, se puede designar a su diferencia como el incremento  $\delta x_{1,q}(n)$  según

$$x_{1,q} - x_{1,q}^{(n)} = \delta x_{1,q}^{(n)}$$

Por otra parte,  $\delta x_{1,q}(n)$  es la diferenciación de  $x_{1,q}(n)$  en la ecuación (A3.3 a), es decir

$$\delta x_{1,q}^{(n)} = -P \left( \frac{i\partial(R_1^{(n)}(M_q^{(n)} - S_1))}{k\partial(R_1^{(n)}(M_q^{(n)} - S_1))} - \frac{iR_1^{(n)}(M_q^{(n)} - S_1)}{(kR_1^{(n)}(M_q^{(n)} - S_1))^2} k\partial(R_1^{(n)}(M_q^{(n)} - S_1)) \right)$$

Luego, se saca factor común  $\partial(iR_1^{(n)}(M_q^{(n)} - S_1))$  resultando

$$\delta x_{1,q}^{(n)} = \left( \frac{-pi}{k\partial(R_1^{(n)}(M_q^{(n)} - S_1))} + \frac{piR_1^{(n)}(M_q^{(n)} - S_1)k}{(kR_1^{(n)}(M_q^{(n)} - S_1))^2} \right) \partial(R_1^{(n)}(M_q^{(n)} - S_1))$$

Efectuando la suma fraccionaria se obtiene

$$\delta x_{1,q}^{(n)} = \left( \frac{-pkR_1^{(n)}(M_q^{(n)} - S_1)i + piR_1^{(n)}(M_q^{(n)} - S_1)k}{(kR_1^{(n)}(M_q^{(n)} - S_1))^2} \right) \partial(R_1^{(n)}(M_q^{(n)} - S_1))$$

Aquí se puede usar la ecuación (A1.2) para sustituir algunos términos

$$\delta x_{1,q}^{(n)} = \left( \frac{-p \frac{-p}{\lambda_{1,q}^{(n)}} i + p \frac{x_{1,q}^{(n)}}{\lambda_{1,q}^{(n)}} k}{\left( \frac{-p}{\lambda_{1,q}^{(n)}} \right)} \right) \partial(R_1^{(n)}(M_q^{(n)} - S_1))$$

Distribuyendo

$$\partial x_{1,q}^{(n)} = \begin{pmatrix} -p \frac{-p}{\lambda_{1,q}^{(n)}} i & p \frac{x_{1,q}^{(n)}}{\lambda_{1,q}^{(n)}} k \\ \left( \frac{-p}{\lambda_{1,q}^{(n)}} \right) & \left( \frac{-p}{\lambda_{1,q}^{(n)}} \right) \end{pmatrix} \partial (R_1^{(n)} (M_q^{(n)} - S_1))$$

Entonces la ecuación se simplifica con

$$\partial x_{1,q}^{(n)} = \left( \lambda_{1,q}^{(n)} + \frac{x_{1,q}^{(n)} \lambda_{1,q}^{(n)}}{p} k \right) \partial (R_1^{(n)} (M_q^{(n)} - S_1))$$

Así, la expresión resultante para  $\delta x_{1,q}(n)$  es

$$\partial x_{1,q}^{(n)} = \lambda_{1,q}^{(n)} \left( 1,0, \frac{x_{1,q}^{(n)}}{p} \right) \partial (R_1^{(n)} (M_q^{(n)} - S_1)) \quad (\text{A3. 4 a})$$

con un razonamiento análogo, la expresión resultante para  $\delta y_{1,q}(n)$  es

$$\partial y_{1,q}^{(n)} = \lambda_{1,q}^{(n)} \left( 0,1, \frac{y_{1,q}^{(n)}}{p} \right) \partial (R_1^{(n)} (M_q^{(n)} - S_1)) \quad (\text{A3.4.b})$$

Aún queda como incógnita. Puede expresarse como el incremento

$$\partial (R_1^{(n)} (M_q^{(n)} - S_1)) = r_1^{(n)} R_1^{(n)} (M_q^{(n)} + \partial M_q^{(n)} - S_1) - R_1^{(n)} (M_q^{(n)} - S_1).$$

Distribuyendo  $r_1^{(n)} R_1^{(n)}$

$$\partial (R_1^{(n)} (M_q^{(n)} - S_1)) = r_1^{(n)} R_1^{(n)} (M_q^{(n)} - S_1) + r_1^{(n)} R_1^{(n)} \partial M_q^{(n)} - R_1^{(n)} (M_q^{(n)} - S_1)$$

Sacando factor común  $R_1^{(n)} (M_q^{(n)} - S_1)$

$$\partial (R_1^{(n)} (M_q^{(n)} - S_1)) = (r_1^{(n)} - I) R_1^{(n)} (M_q^{(n)} - S_1) + r_1^{(n)} R_1^{(n)} \partial M_q^{(n)}$$

Despreciando el término  $r_1^{(n)} R_1^{(n)} \partial M_q^{(n)}$  (por tratarse de una rotación diferencial de un incremento diferencial) y aproximándolo con – simplemente –  $r_1^{(n)} R_1^{(n)} \partial M_q^{(n)}$ , la ecuación anterior es equivalente a

$$\partial (R_1^{(n)} (M_q^{(n)} - S_1)) \approx (r_1^{(n)} - I) R_1^{(n)} (M_q^{(n)} - S_1) + R_1^{(n)} \partial M_q^{(n)}$$

Nuevamente partiendo de la ecuación (A1.3) se suplanta  $R_1^{(n)}(M_q^{(n)} - S_1)$  por su equivalente, resultando en

$$\partial(R_1^{(n)}(M_q^{(n)} - S_1)) \approx \frac{1}{\lambda_{1,q}^{(n)}}(r_1^{(n)} - I) \begin{bmatrix} x_{1,q}^{(n)} \\ y_{1,q}^{(n)} \\ -p \end{bmatrix} + R_1^{(n)} \partial M_q^{(n)}.$$

La rotación diferencial  $r$  se expresa según la ecuación (A2.19) para pequeñas rotaciones. Entonces

$$\partial(R_1^{(n)}(M_q^{(n)} - S_1)) \approx \frac{1}{\lambda_{1,q}^{(n)}}(I + \begin{bmatrix} 0 & -\gamma^{(n)} & \beta^{(n)} \\ -\gamma^{(n)} & 0 & -\alpha_1^{(n)} \\ -\beta^{(n)} & \alpha_1^{(n)} & 0 \end{bmatrix} - I) \begin{bmatrix} x_{1,q}^{(n)} \\ y_{1,q}^{(n)} \\ -p \end{bmatrix} + R_1^{(n)} \partial M_q^{(n)}.$$

Aquí se puede simplificar las matrices identidad  $I$  de signo contrario.

$$\partial(R_1^{(n)}(M_q^{(n)} - S_1)) \approx \frac{1}{\lambda_{1,q}^{(n)}} \begin{bmatrix} 0 & -\gamma^{(n)} & \beta^{(n)} \\ -\gamma^{(n)} & 0 & -\alpha_1^{(n)} \\ -\beta^{(n)} & \alpha_1^{(n)} & 0 \end{bmatrix} \begin{bmatrix} x_{1,q}^{(n)} \\ y_{1,q}^{(n)} \\ -p \end{bmatrix} + R_1^{(n)} \partial M_q^{(n)}$$

Realizar el producto matriz por el vector da

$$\partial(R_1^{(n)}(M_q^{(n)} - S_1)) \approx \frac{1}{\lambda_{1,q}^{(n)}} \begin{bmatrix} -\gamma_1^{(n)} y_{1,q}^{(n)} - \beta_1^{(n)} p \\ -\gamma_1^{(n)} x_{1,q}^{(n)} + \alpha_1^{(n)} p \\ -\alpha_1^{(n)} x_{1,q}^{(n)} - \beta_1^{(n)} x_{1,q}^{(n)} \end{bmatrix} + R_1^{(n)} \partial M_q^{(n)}. \quad (\text{A3.5})$$

Incluyéndolo en las ecuaciones (A3.4 a) y (A3.4 b), como sigue

$$\partial x_{1,q}^{(n)} = \lambda_{1,q}^{(n)} \left( 1, 0, \frac{x_{1,q}^{(n)}}{p} \right) \left[ \frac{1}{\lambda_{1,q}^{(n)}} \begin{bmatrix} -\gamma_1^{(n)} y_{1,q}^{(n)} - \beta_1^{(n)} p \\ -\gamma_1^{(n)} x_{1,q}^{(n)} + \alpha_1^{(n)} p \\ -\alpha_1^{(n)} x_{1,q}^{(n)} - \beta_1^{(n)} x_{1,q}^{(n)} \end{bmatrix} + R_1^{(n)} \partial M_q^{(n)} \right]$$

Distribuyendo se obtiene

$$\partial x_{1,q}^{(n)} = \left( 1, 0, \frac{x_{1,q}^{(n)}}{p} \right) \begin{bmatrix} -\gamma_1^{(n)} y_{1,q}^{(n)} - \beta_1^{(n)} p \\ -\gamma_1^{(n)} x_{1,q}^{(n)} + \alpha_1^{(n)} p \\ -\alpha_1^{(n)} x_{1,q}^{(n)} - \beta_1^{(n)} x_{1,q}^{(n)} \end{bmatrix} + \lambda_{1,q}^{(n)} \left( 1, 0, \frac{x_{1,q}^{(n)}}{p} \right) R_1^{(n)} \partial M_q^{(n)}$$

Realizando el producto vectorial

$$\partial x_{1,q}^{(n)} \approx -\gamma_1^{(n)} y_{1,q}^{(n)} \beta_1^{(n)} p + \frac{x_{1,q}^{(n)}}{p} (\alpha_1^{(n)} y_{1,q}^{(n)} - \beta_1^{(n)} x_{1,q}^{(n)}) + \lambda_{1,q}^{(n)} \left[ 1, 0, \frac{x_{1,q}^{(n)}}{p} \right] R_1^{(n)} \partial M_q^{(n)}$$

Finalmente la expresión para el  $x$  observado y el  $x$  calculado resulta

$$x_{1,q} - x_{1,q}^{(n)} \approx \alpha_1^{(n)} \frac{x_{1,q}^{(n)} y_{1,q}^{(n)}}{p} - \beta_1^{(n)} \left( p + \frac{x_{1,q}^{(n)2}}{p} \right) - \gamma_1^{(n)} y_{1,q}^{(n)} + \lambda_{1,q}^{(n)} \left[ 1, 0, \frac{x_{1,q}^{(n)}}{p} \right] R_1^{(n)} \partial M_q^{(n)}$$

$$y_{1,q} - y_{1,q}^{(n)} \approx \alpha_1^{(n)} \left( p + \frac{y_{1,q}^{(n)}}{p} \right) - \beta_1^{(n)} \frac{x_{1,q}^{(n)} y_{1,q}^{(n)}}{p} - \gamma_1^{(n)} y_{1,p}^{(n)} + \lambda_{1,q}^{(n)} \left[ 0, 1, \frac{x_{1,q}^{(n)}}{p} \right] R_1^{(n)} \partial M_q^{(n)}.$$

$$x_{2,q} - x_{2,q}^{(n)} \approx \alpha_2^{(n)} \frac{x_{2,q}^{(n)} y_{2,q}^{(n)}}{p} - \beta_2^{(n)} \left( p + \frac{x_{2,q}^{(n)2}}{p} \right) - \gamma_2^{(n)} y_{2,q}^{(n)} + \lambda_{2,q}^{(n)} \left[ 1, 0, \frac{x_{2,q}^{(n)}}{p} \right] R_2^{(n)} \partial M_q^{(n)}$$

En términos generales lo que se ha logrado es poner  $q$  observaciones de pares homólogos en función de los diferenciales de las rotaciones observadas y estimadas, junto con los diferenciales observado y estimado de puntos en la superficie del objeto. Esto resulta en un vector de incógnitas como el que se muestra a continuación

$$(\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2, \partial M_{x1}, \partial M_{y1}, \partial M_{z1}, \dots, \partial M_{xq}, \partial M_{yq}, \partial M_{zq})^t$$

¿Cuántos puntos son necesarios para plantear tantas ecuaciones como incógnitas?. De las ecuaciones anteriores, surge que cada par de puntos determina cuatro ecuaciones y nueve incógnitas.

No obstante por tratarse de la orientación relativa, puede asumirse  $\alpha_1$  como 0. En definitiva se está fijando esa dirección y trasladando la rotación relativa a las restantes cinco componentes, dado que se trata de un fenómeno a cinco grados de libertad. Entonces dados dos pares de puntos  $s$  y  $t$ , sabemos que cada punto determinará cuatro ecuaciones, pero las del par  $s$  llevarán de incógnitas a  $\partial M$ , mientras que las ecuaciones de  $t$  llevarán a  $\partial M_t$ . En cambio las cinco incógnitas que hacen a  $r_1, r_2$  son comunes a las ocho ecuaciones.

Se trata entonces de encontrar un número  $q$  de pares de puntos de base tal que el número de ecuaciones (cuatro por cada par) con el de incógnitas (cinco fijas más tres por cada par). Esto es, un  $p$  tal que

$$4q = 5 + 3q$$

$$4q - 3q = 5$$

$$q = 5$$

Por lo tanto son necesarios cinco pares de puntos homólogos para asegurar solución única. Con esto el sistema de ecuaciones a resolver resulta de veinte ecuaciones por veinte incógnitas.

De esta manera se ha transformado el problema de orientación relativa en la búsqueda de un vector que resuelva el sistema de ecuaciones planteado. Se ha liberalizado el problema.

Resolver el sistema es el objetivo de cada iteración, tras lo cual las ecuaciones (A3.1) y (A3.2) nos indican cómo ajustar las estimaciones para la siguiente iteración.

De contar con más observaciones que incógnitas puede aproximarse la solución por cuadrados mínimos.

# Estereopsis Automática

## Panorama del Estado del Arte\*

### B.1. Introducción

El problema esencial de la estereopsis digital es la identificación automática de puntos homólogos, problema también denominado apareo estereoscópico (stereo matching) o problema de la correspondencia (correspondence problem) de imágenes estereoscópicas.

Esto consiste en la identificación automática de las proyecciones, en imágenes perspectivas, de los puntos singulares o característicos de una escena, objeto o terreno. A estos puntos perspectivos, en general pares de puntos, se los denomina puntos homólogos. Cuando se trata de imágenes cuyas líneas respetan la geometría epipolar, los valores de paralejo o disparidad son función de la diferencia de posición de estos puntos sobre las imágenes. A partir de estos valores y teniendo en cuenta ciertos parámetros, se obtienen las coordenadas 3D, de puntos de los objetos, con los que genera el modelo numérico de la escena.

Se trabaja con pares estereoscópicos de imágenes digitales, a veces a partir de imágenes fotográficas digitalizadas, para posibilitar su procesamiento por computadora, por lo que el problema de la correspondencia forma parte del área de procesamiento de imágenes digitales.

Cuando se utiliza la expresión puesta en correspondencia se hace alusión a cuatro problemas. Dos corresponden directamente al apareo automático y otros están estrechamente relacionados con el mismo ya que pueden ser de fundamental interés para modelizar particularidades de la visión binocular y contribuir a mejorar la obtención de soluciones:

- imágenes simultáneas, o imágenes no simultáneas con escena idéntica en las tomas;
- imágenes no simultáneas y con modificaciones (movimientos) de la escena en cada toma;
- puesta en correspondencia entre un modelo numérico previamente almacenado y una imagen de ese mismo objeto;
- respuesta perceptual estereoscópica, consiste en comparar un modelo numérico previamente almacenado luego de algún tipo de elaboración, con el modelo numérico resultado de la

\* José Zelasco Alejandra V. Donadio, Daniel E. Mendez.  
Esteropsis automática. Panorama del Estado del Arte.

restitución automática de un par estereoscópico.

El objetivo de este anexo es presentar el estado del arte del problema del apareo automático, particularmente en el caso de imágenes simultáneas considerando el segundo caso (no simultánea y con movimiento en la escena) solo tangencialmente y cuando tiene relación con el primero. Se desarrolla en cinco secciones: aspectos del problema y presentación de dificultades, estrategias y soluciones, clasificación de algoritmos y tendencias.

## B.2. Aspectos Del Problema

Seguidamente, se enuncian y detallan distintos aspectos vinculados con este problema y la concepción de los algoritmos apropiados.

Los temas que se abordan son: el análisis de distintas dificultades en la identificación de puntos homólogos; el tipo de primitiva o característica que se utiliza para el apareo y las restricciones utilizadas por los algoritmos con el objetivo de limitar el espacio de búsqueda.

### B.2.1. Presentación de Dificultades

La identificación automática de puntos homólogos presenta diversos problemas. Según su naturaleza pueden ser clasificados en los debidos al sistema de adquisición de las tomas de vista y en los que provienen de la particularidad de la aplicación. Ambos tipos de dificultades, se manifiestan simultáneamente.

-Debido a las tomas de vista y a las particularidades de la aplicación:

**Variaciones Fotométricas:** La luz reflejada o emitida desde la escena y registrada por el sistema de toma de vistas depende de la posición relativa de los sensores, así como también del ruido consecuencia, en general, de la no linealidad de los sensores. Por el primer motivo, cuando el sensor se desplaza a una nueva posición, o cuando se utilizan dos sensores situados en puntos de vista diferentes, o en medio subacuado, las intensidades de las superficies correspondientes pueden variar. Por otra parte, a causa del ruido, se pueden producir variaciones fotométricas y enmascarar la firma espectral de ciertos objetos.

**Sombras y Reflejos:** Debido a los efectos de la iluminación, puede surgir sombras en las escenas. estas sombras, al igual que los reflejos, aparecen en las imágenes como nuevos terminales u objetos, dificultándose su detección. Tanto las sombras como los reflejos, se pueden presentar en una u otra imagen, causando aparentes occlusiones o manifestándose en ambas. No correspondiendo al mismo punto de la escena su diferencia de ubicación no representa una disparidad real. Además, estas circunstancias también producen variaciones fotométricas en los niveles de intensidad de dos elementos homólogos.

-Debido a las características de la aplicación:

**Oclusiones:** Ocurren cuando una porción de una escena es visible en una imagen pero no en la otra. Se pueden distinguir dos tipos de occlusiones. El trivial y más frecuente es el que se produce en los bordes de las imágenes. El otro, depende de la existencia de discontinuidades acentuadas que provocan obstrucciones para alguno de los puntos de vista.

**Texturas Repetitivas:** En algunas aplicaciones pueden aparecer ocurrencias repetidas. en estos casos, hay múltiples correspondencias posibles y no es sencillo resolver am-

bigüedades. Un ejemplo de este tipo de texturas es el caso de una pared de ladrillos a la vista. El problema se agrava aún más en presencia de variaciones fotométricas u oclusiones que pueden inducir a apareos aparentemente correctos en un área local, pero incorrectos globalmente. Pocos algoritmos de apareo cuentan con estrategias que apunten a resolver este problema, debido a que no es frecuente, pero se considera en ciertas aplicaciones.

**Falta de Textura:** En las escenas del mundo real, la mayoría de los objetos y superficies son texturadas. Las texturas brindan patrones con alta probabilidad del apareado como, por ejemplo, segmentos que son la base de algunos algoritmos. Sin embargo, parte de una escena puede carecer de texturas, lo que no facilita establecer apareos en ese área.

**Superficie y objetos particulares:** Existen diversos tipos de materiales que constituyen los objetos de la escena y que provocan dificultades en el apareo. Entre estos se encuentran los objetos transparentes o translúcidos, más o menos brillantes, que producen distintos efectos en las imágenes. Además, en imágenes en tonos de gris, pueden existir objetos que no puedan ser distinguidos de un fondo, siendo forma y fondo de distinto color pero de una misma intensidad de gris.

Todos estos inconvenientes han inducido a desarrollar algoritmos que resulten apropiados según las aplicaciones. La solución general a la estereopsis automática sigue siendo un problema abierto.

### B.2.2. Primitivas

La comparación de los niveles de intensidad entre la imagen izquierda y la derecha de un par estéreo puede ser un indicador débil de exactitud en las correspondencias debido a varios factores de los mencionados en el párrafo anterior (e.g.: ruidos, distorsiones, oclusiones, sombras, etc.) que producen que los elementos de la imagen alteren sus niveles de gris.

Un gran número de algoritmos tratan de encontrar las características relevantes de una escena, para describir las imágenes. Estas características reciben, según los autores, indistintamente el nombre de primitiva token, objeto, terminal o simplemente característica. Sin embargo, la noción de terminal se puede formalizar por medio del concepto de objeto, como todo subconjunto de puntos conexos de una imagen. A cada objeto se asocia un punto o elemento de imagen, eventualmente por línea, llamado posición, y un conjunto de atributos llamados parámetros descriptivos. Para construir un objeto se puede recurrir a una definición específica o crearse por medio de leyes de composición internas o externas según corresponda. Si las propiedades que caracterizan los terminales se ponen acertadamente de manifiesto conducen a facilitar un apareo correcto. Entre estas propiedades deseables se distinguen el poder descriptivo o contenido semántico y el alto nivel de invariancia o estabilidad a pequeñas alteraciones debidas a las diferentes posiciones de los puntos de vista. Ambas propiedades facilitan la resolución de ambigüedades. Otras, también interesantes, mencionadas en la literatura, son: la generalidad para describir la mayor parte de la información útil de una escena; la disponibilidad para su extracción de forma eficiente y confiable; y la apareabilidad que, en principio, sería consecuencia de las anteriores. No hay que olvidar que deben existir mecanismos de aplicación de restricciones, sobre estas primitivas, para facilitar el apareo.

No siempre las primitivas satisfacen todas estas propiedades deseables, lo frecuente es que las primitivas fáciles de obtener son difíciles de aparear y viceversa.

La elección de la primitiva apropiada depende de ciertas restricciones, como exigencias de precisión, densidad de los mapas de disparidad, etc., impuestas por la naturaleza de la aplicación. No obstante, su elección está también condicionada por las siguientes particularidades:

**Dimensionalidad:** establece que las características puntuales son menos afectadas por la distorsión de perspectiva, que las características de tipo bordes que pueden ser apareadas de manera más confiable.

**Tamaño** (frecuencia espacial): resulta más fácil el pareo de características prominentes.

**Contraste:** se desean primitivas de alto contraste debido a que las de bajo contraste son difíciles de desambiguar.

**Densidad de ocurrencia:** las primitivas esparcidas son más fáciles de aparear, dada su menor ocurrencia, pero producen mapas de disparidad dispersos.

Por otra parte, las primitivas pueden ser clasificadas en distintos niveles teniendo en cuenta la complejidad computacional de su extracción y posterior representación. Se puede establecer básicamente tres niveles: las de nivel más bajo, son las más fáciles de obtener y están conformadas por la intensidad de pixeles y los elementos de contorno; el nivel intermedio esta constituido por contornos, segmentos de líneas, curvas, esquinas, regiones; y el nivel superior corresponde a las formas estructuradas, geométricas o no.

A continuación se enuncian algunas de las primitivas adoptadas con mayor frecuencia:

- **pixeles:** la intensidad de los pixeles se utiliza en los algoritmos que se basan en áreas. Este es un enfoque simple y se fundamenta en la comparación de los niveles de intensidad entre las imágenes izquierda y derecha. Estos métodos deben tener especial cuidado en presencia de ruido, distorsiones y otros factores, que deben ser afrontados.

- **regiones:** la utilización de regiones puede ser buena para representar superficies, pero la segmentación de regiones resulta un problema difícil.

- **segmentos** o líneas: son interesantes para ciertas escenas, pero no resultan apropiados para aquellas que contengan formas complejas, como los escenarios naturales. Por otra parte, la extracción de los segmentos puede sufrir divisiones o uniones espurias.

-**contornos:** los elementos de contorno y los contornos son las primitivas más utilizadas. Estos se manifiestan como transiciones en los niveles de intensidad de la imagen y describen eventos como por ejemplo, los límites de objetos o regiones en un espacio tridimensional. Estos bordes pueden ser extraídos de manera eficiente y apareados confiablemente, utilizando las restricciones apropiadas disponibles.

Es evidente que existe un compromiso entre la confiabilidad que ofrece una primitiva para ser apareada y la complejidad del proceso de extracción que requiere.

En la práctica, las primitivas mas utilizadas incluyen la intensidad de los pixeles o funciones de la intensidad como el gradiente de la imagen, el Laplaciano, el cambio de signo de la segunda derivada (zero crossing) para la detección y localización de elementos de bordes y los segmentos que se conforman mediante su vinculación.

### B.2.3. Restricciones

Tratar todas las correspondencias posibles para resolver el problema del apareo de puntos homólogos está prácticamente fuera de las posibilidades de las computadoras a causa de la explosión combinatoria del problema. Como en todo proceso de búsqueda, cuando el espacio de búsqueda se incrementa, aumenta la probabilidad de error, así como también el costo computacional.

En consecuencia, las soluciones emplean diversas restricciones. Desafortunadamente, las restricciones son usualmente necesarias, pero no necesariamente suficientes para encontrar la solución óptima.

La mayoría de algoritmos de apareo utilizan restricciones para limitar el número de candidatos correspondientes a una característica de una imagen en la otra. Las técnicas de apareo deben ser capaces de utilizar las propiedades de las características y emplear restricciones extras para evaluar los mejores apareos posibles, es decir, que sean consistentes tanto en forma local como global.

Se utilizan dos tipos de restricciones, restricciones locales y restricciones globales. Las restricciones locales, tienen el objeto de limitar o reducir la región de búsqueda de una primitiva correspondiente. Una vez circunscrito este espacio, se analizan las propiedades de las primitivas y solamente serán consideradas como candidatas aquellas que posean atributos similares.

Por otra parte, los apareos candidatos están sujetos a chequeos de consistencia, mediante el empleo de restricciones globales. Este tipo de restricciones consideran las posibles correspondencias en su conjunto y determinan las mejores. Se cumplen de esta manera dos propósitos: la selección del candidato a aparear y la eliminación de apareos falsos.

No existe un orden en particular en la aplicación de las restricciones, ya sean locales o globales, se aplican generalmente en forma simultánea. Además, dependiendo del algoritmo, estas restricciones son implícitas o explícitas.

Por otro lado, es usual clasificar las restricciones según la naturaleza de las mismas. Adoptando este criterio, se puede clasificar como: geométricas, impuestas por el sistema de toma de las vistas; provenientes de la posición relativa de los objetos que están siendo visualizados y debidas al entorno en el que es razonable aplicar el método. Por ejemplo, para tener resultados de cierto nivel de precisión y en ciertos casos la profundidad máxima no debe superar siete veces la distancia entre los puntos de vista, ni la mínima ser inferior a este valor.

Existen además restricciones que pueden provenir de un modelo pre-registrado de la escena o inspecciones previas de las imágenes, efectuadas según una estrategia de control jerárquica o grueso a fino, como se describe en Estrategias de Control

La utilización de las restricciones tiene, entonces, dos finalidades: reducir el espacio de búsqueda, lo que conduce a una fuerte disminución del costo computacional del algoritmo, y efectuar verificaciones de consistencia global para reforzar la confiabilidad de los apareos establecidos.

Se describen a continuación las principales restricciones utilizadas siguiendo la clasificación de restricciones locales y globales.

### a) Restricciones Locales

Las restricciones locales incluyen los principios de epipolaridad, similaridad y lapso de disparidad.

**Epipolaridad:** los puntos correspondientes deben estar presentes en ambas imágenes sobre una misma línea epipolar, esto restringe drásticamente el espacio de búsqueda, reduciéndolo a un problema unidimensional. Las líneas epipoles son rectas que resultan como intersección de los planos de las imágenes con el plano determinado por los dos puntos de vista y el punto objeto. En consecuencia las respectivas perspectivas de dicho punto se encuentran sobre líneas epipoles de las imágenes.

Para optimizar el uso de esta propiedad como restricción es necesario que las líneas de las imágenes (líneas de pixeles) coincidan con las líneas epipoles. Esto requiere que la toma haya respetado la geometría del caso normal, imágenes en el mismo plano y líneas de imagen paralelas a la recta que pasa por los puntos de vista, o que las imágenes hayan sido conducidas a esta situación mediante un proceso de roto-rectificación.

Todas las soluciones estudiadas incluyen la restricción de epipolaridad, y en su mayoría trabajan con imágenes que han sido rotorectificadas. Los procedimientos que son utilizados para lograr la Rotorectificación de las imágenes están fuera del alcance de la presente discusión.

**Similaridad:** es el principio básico para resolver el problema de la correspondencia, los puntos correspondientes tienen características similares en las dos imágenes. Por esta razón, la similaridad es utilizada sin excepción en la búsqueda de los candidatos a aparear y la selección del apareo definitivo.

Sin embargo existen algunos factores que no hacen posible el uso, sin precauciones, de esta restricción. Estos son principalmente, la presencia de ruido, sombra y reflejos, distorsión de la perspectiva pudiendo producir deformaciones en las imágenes, etc. En consecuencia, puede suceder que puntos homólogos pierdan similaridad o que propiedades parecidas caractericen puntos que no son homólogos, por lo que el uso estricto del criterio de similaridad no es aconsejable. Estos problemas son frecuentes y están presentes en la mayoría de las aplicaciones.

Además de estos factores pueden producirse otras situaciones no deseables, cuya ocurrencia es poco frecuentes ya que se presentan sólo en ciertas aplicaciones. Entre estos se mencionan la posibilidad de movimientos en la escena observada, o la presencia de obstáculos no previstos, especialmente cuando se utiliza un solo sensor que debe ser desplazado de un punto de vista a otro. Ejemplos son la presencia de movimientos de vehículos en imágenes aéreas, la aparición de nubes en una imagen satelital o movimiento de flora o fauna en imágenes submarinas.

**Lapso de disparidad:** limitaciones en la disparidad máxima y mínima que surge del conocimiento de la escena observada o de la aplicación de una estrategia de control. Esto es más restrictivo que el criterio general de límites de disparidad debido a las características del método mencionado con anterioridad.

### b) Restricciones Globales

Se encuentra entre estas la unicidad, continuidad, orden, límite del gradiente de disparidad y otras restricciones denominadas particulares.

**Unicidad:** establece que las correspondencias deben ser únicas, a cada elemento de la imagen le corresponde a lo sumo uno en la otra y consecuentemente un único valor de disparidad.

Esta restricción no puede ser estrictamente aplicada tomando los puntos de la imagen como terminales. En efecto, un pixel tiene un tamaño físico y le corresponde una porción de una superficie. No es difícil de imaginar una situación, en la cual, una porción inclinada de la superficie de la escena constituya un cierto número de pixel es en una imagen pero posiblemente un número diferente, mayor o menor en la otra.

**Continuidad de la Superficie:** para cierto tipo de imagen los valores de la disparidad deben variar suavemente a lo largo de la misma. Los cambios de disparidad están localizados frecuentemente en los límites de los objetos de la escena donde se manifiestan discontinuidades en los niveles de intensidad de la imagen.

**Continuidad Figural:** la presencia de una cadena de puntos de borde conectados representa probablemente la proyección de una curva continua en un espacio 3D. Por esta razón, debe existir una continuidad en los valores de disparidad a lo largo de dicha cadena de puntos. Esta restricción es empleada, generalmente, para efectuar verificación de consistencia y robustecer los apareos. Sin embargo, debe tomarse con cuidado en aquellos casos en que la cadena de puntos que constituye un contorno no representen límites de diferentes superficies en la escena.

**Orden:** establecer que el orden posicional relativo de una característica a lo largo de una línea epipolar en una imagen, se preserva en la otra. Una característica presente en la imagen izquierda debe ocurrir en el mismo orden, de izquierda a derecha, sobre la otra imagen a lo largo de las líneas epipoles, salvo si hubiera objetos interpuestos. Esta restricción, entonces, no sirve en caso de discontinuidades bruscas, como por ejemplo en escenas de interiores.

En el caso de los bordes o contornos presentes en el plano de la imagen, un orden de izquierda a derecha sobre una imagen no podrá corresponder con un orden de derecha a izquierda sobre la otra imagen lo que restringe considerablemente el espacio de búsqueda.

**Límites en el Gradiente de Disparidad:** la diferencia de profundidad de objetos contiguos debe, en ciertas aplicaciones, conservarse dentro de ciertos límites.

**Particulares:** este tipo de restricción surgen del conocimiento de las características de la escena, es decir, se puede predecir o se conoce a priori lo observado. Ocurre que se tenga la certeza de que se encuentran dentro de la imagen un cierto número de patrones que son perfectamente identificables. Esta restricción permite generar ciertas reglas básicas y constituye un primer paso en la aplicación de modelos de respuesta perceptual.

### B.3. Estrategias y Soluciones Habituales al Problema del Apareo

Para robustecer las soluciones empleadas en el proceso de apareo y reducir ambigüedades, se han desarrollado estrategias de control. Luego de presentar estas estrategias se describen los métodos clásicos empleados en el proceso de búsqueda de puntos homólogos.

### B.3.1. Estrategias de control

En principio general es el de obtener distintos niveles de descripción de las imágenes y aplicar algoritmos de apareo desde los niveles más sencillos a los más complejos. Los resultados obtenidos de los niveles de menor detalle son utilizados para guiar las búsquedas de correspondencias de los niveles posteriores. Como consecuencia, el empleo de estas estrategias brinda restricciones adicionales, provenientes de inspecciones previas de las imágenes, que se combinan con las restricciones restantes para reforzar el apareo. Estas estrategias responden a estudios de la fusión estereoscópica o estereopsis de la visión binocular humana.

Existen al menos dos formas de instrumentar este tipo de estrategias: las del tipo grueso a fino (coarse to fine) y las estrategias del tipo estructurales, jerárquicas o relacionales.

La diferencia existente entre estos tipos de estrategias, reside en la forma de obtener los distintos niveles de descripción de las imágenes. En las estrategias grueso a fino, se logra mediante la realización de procesos de filtrado sobre las imágenes con los que se atenuan las diferencias de niveles. Mientras que en las otras, se obtiene focalizándose en estructuras de distinto nivel.

#### a) Estrategia de control Grueso a Fino

En un análisis a fino, se trabaja con las imágenes que han sido sometidas a un cambio de resolución. La información obtenida en los niveles inferiores o más gruesos es utilizada para guiar la búsqueda de los niveles más altos o finos.

En esta aproximación, se inicia el apareo en los niveles más gruesos, donde la densidad de características es menor debido al cambio de escala. En estos niveles, se tiene un espacio de búsqueda reducido, que conduce a un apareo más simple. Las correspondencias establecidas se utilizan para guiar los apareos de los niveles más finos y en consecuencia con mayor cantidad de detalles.

El proceso de cambio de resolución puede introducir desplazamientos en la ubicación de las primitivas con respecto a su localización real en la imagen original. Por lo tanto, los apareos no necesariamente son exactos, situación que debe ser tenida en cuenta en los niveles superiores.

Además, los apareos incorrectos de los niveles inferiores pueden prolongarse, conduciendo a errores en los niveles subsecuentes. Sin embargo, el número de errores introducidos por el empleo de esta estrategia es menos significativo que los errores que se ocasionan sin el empleo de la misma, es decir, realizando el apareo sobre los niveles originales de las imágenes.

Esta estrategia no sólo es utilizada por diversos autores, sino que es comercialmente empleada con significativos resultados. Es el caso de la firma ISTAR, del grupo de empresas francesas que desarrollan aplicaciones usando imágenes del satélite SPOT, que se ocupa de la realización de modelos numéricos de terreno a partir de imágenes estereoscópicas.

#### b) Estrategias de control Jerárquicas y Estructurales

Los algoritmos que aplican estas estrategias utilizan primitivas de alto contenido

semántico para representar la información de las imágenes.

En los sistemas jerárquicos, el apareo tiene lugar en sucesivos niveles de descripción de las imágenes. Los distintos niveles de descripción son obtenidos mediante la abstracción de estructuras de alto nivel, sin efectuar un cambio de escala, como sucede en las estrategias grueso a fino.

El objetivo de estas estrategias es el de utilizar los resultados de los niveles inferiores, para limitar y robustecer el apareo de los niveles superiores, como toda estrategia de control. En este caso, se emplean propiedades vinculantes entre los resultados de los niveles inferiores y los superiores para resolver ambigüedades de apareo.

### B.3.2. Técnicas de búsqueda de correspondencias

Existen distintas técnicas para determinar las correspondencias de elementos homólogos. Cada uno de estos métodos, se diferencian según la forma de comparar las primitivas de las imágenes y la elección de los apareos para obtener la función de correspondencia. Se distinguen principalmente cuatro tipos de técnicas, las basadas en correlación matemática, las de programación dinámica, las de relajación, y las de predicción y verificación. A continuación, se describen cada una de ellas.

#### a) Algoritmos basados en Correlación

Los métodos de correlación están entre las primeras técnicas utilizadas en visión estereoscópica por computadora.

Estas técnicas emplean como primitivas los pixeles y como características sus niveles de intensidad. Utilizan principalmente las restricciones de similaridad, geometría epipolar y además, asumen que las profundidades son localmente constantes en la vecindad de un pixel, lo que constituye sólo una aproximación.

Consiste en aparear patrones, formados por porciones o ventanas de una imagen que tienen niveles de gris arbitrarios, con porciones o ventanas equivalentes en la otra imagen. Debido a los efectos mencionados en la sección 2. Presentación de Dificultades nunca existe una copia exacta de una ventana patrón, por lo se debe maximizar el grado de apareo entre dichas ventanas. Para esto se utilizan técnicas de correlación cruzada (cross-correlational).

Los cómputos de correlación son obtenidos comparando una ventana de tamaño fijo en la primera imagen con una ventana que se desplaza en la segunda imagen, generando un conjunto de valores para distintas magnitudes de disparidad. Debido a que generalmente un patrón puede aparear con varias ventanas sobre la otra imagen, se debe, de alguna forma, resolver esta ambigüedad y seleccionar el apareo definitivo, proceso que se denomina habitualmente validación de los apareos.

Las principales dificultades que presentan estos métodos son, la determinación del tamaño de ventana óptimo, la elección del criterio de correlación más confiable, y la selección de un método de validación que resulte apropiado.

En la actualidad estos métodos han sido instrumentados sobre arquitecturas de hardware paralelo con resultados altamente satisfactorios.

### b) Algoritmos basados en Relajación

Los algoritmos basados en relajación son iterativos y localmente paralelos.

En un primer paso, se construye un conjunto inicial con los pareos posibles, asociado a cada primitiva de la imagen izquierda una de la derecha. Este conjunto se organiza como una colección de nodos $\{n_i\}$ , un nodo para cada primitiva de la imagen izquierda. A cada uno de los nodos se asocia un vector  $P$ , que contiene los parámetros o valores de las propiedades de la primitiva en la imagen izquierda, y un conjunto de etiquetas, identificatorias de las terminales de la imagen derecha que son posibles homólogos de la primitiva de la imagen izquierda. Si la primitiva no tiene correspondiente se colocará una etiqueta que indicará esta situación. A cada etiqueta  $e$ , se asocia un número  $p_i(e)$ , que se interpreta como una estimación de probabilidad de que la primitiva identificada con la etiqueta  $e$  sea efectivamente la correspondiente de la primitiva en consideración. Se tiene que  $p_i(1) [0,1]$  y que  $p_i(1)=1$ .

La estimación de la probabilidad inicial se actualizará en cada iteración, teniendo en cuenta la coherencia de la vecindad. Si el número de primitivas, compatibles con el apareamiento 1 dentro de la vecindad, es alto, entonces  $p_i(1)$  aumenta, sino  $p_i(1)$  disminuye.

Después de un número suficiente de iteraciones, un nodo se considera apareable si tiene una etiqueta con la probabilidad elevada, de lo contrario no.

Se destaca que lo esencial de esta técnica es la definición de la regla para la actualización de las probabilidades en cada iteración.

Para reducir ambigüedades en los apareamientos iniciales se puede recurrir a restricciones tales como ángulo de esquinas, magnitud de gradiente, longitud de segmentos, máximo de disparidad, etc. Ejemplos de este tipo de algoritmos pueden encontrarse en

### c) Algoritmos basados en Programación Dinámica

La programación dinámica es una técnica que posibilita la optimización eficiente de una función de costos. El método consiste en dividir un problema general en varios subproblemas más pequeños, resolver cada uno de ellos y finalmente obtener la solución buscada.

El problema del apareo de primitivas, se resuelve por medio de un modelo que consiste en la búsqueda del camino óptimo en un grafo, minimizando una función de costo.

Para ilustrar el método, supóngase que la cantidad máxima de primitivas de la imagen izquierda y derecha es  $N$  y  $M$ , respectivamente. Un camino desde un nodo ( $i = 1$ ) al  $Q = N$  es una asociación de las primitivas de la primera y segunda imagen.

El principio de la división en diferentes pasos, dentro de la programación dinámica, se aplica de la siguiente manera: en cada etapa  $i$ , se asocia el nodo  $(i, j)$  con un camino que garantiza un costo mínimo acumulado hasta la etapa  $i$ . El costo acumulado del camino que termina en el nodo  $(i, j)$  se define como:

$$C(i, j) = \min_{1 < i < k} (f(C(i-1, k), c(i-1, k; i, j)))$$

Donde  $c$  es el costo local de unir el nodo  $(i-1, k)$  con el nodo  $(i, j)$  y  $f(C, c)$  es una función apropiada que combina  $C$  y  $c$ . El problema del apareo se transforma, entonces, en el problema de encontrar el camino de costo óptimal.

La aplicación de las restricciones de orden y unicidad, entre otras, hace posible la obtención de buenos resultados.

El principal elemento a tener en cuenta en esta técnica es la definición de la función de costo de correspondencia entre primitivas.

El algoritmo de programación dinámica normalmente utilizado para instrumentar soluciones al problema del apareo estereoscópico es el algoritmo de Viterbi .

#### **d) Algoritmos basados en Predicción y Verificación**

Estos algoritmos utilizan generalmente primitivas de alto nivel. Inicialmente, se realiza una descripción de las imágenes basadas en estas primitivas. Posteriormente se obtiene el apareo de las primitivas homólogas en dos etapas: de predicción local y de verificación global.

En la etapa predicción local los apareos supuestos se realizan entre las primitivas relacionadas a un plano epipolar común, donde la disparidad está comprendida entre dos valores límites. Además presentan similaridades estrictas en sus parámetros descriptivos.

En la etapa de verificación global, las hipótesis previamente realizadas se propagan en forma recursiva, calculándose nuevos apareos entre las primitivas vecinas, a partir de las previamente apareadas. Las restricciones que se imponen para el apareo de terminales vecinos son en general las siguientes: los terminales deben manifestarse en un cierto plano epipolar común, su disparidad debe ser aproximadamente igual a la disparidad calculada por los terminales apareados previamente, y se debe verificar la similaridad de los parámetros descriptivos aunque menos estrictas que la etapa anterior.

Finalmente, la cantidad de primitivas apareadas en el transcurso de la propagación de la hipótesis se utiliza como una función discriminante para resolver ambigüedades de los apareos y validar las correspondencias definitivas.

### **B.4. Clasificación: Algoritmos Basados en Áreas y en Características**

Normalmente los algoritmos de prueba en correspondencia se pueden agrupar en dos grandes categorías, según las primitivas que utilizan en el apareo: algoritmos basados en áreas y algoritmos basados en características. Esta clasificación es adoptada por varios autores.

Cuando las primitivas de apareo son de bajo nivel y densas, tales como la intensidad de los pixeles, y la estrategia de apareo se aplica localmente en cada pixel de la imagen, se denominan algoritmos basados en áreas.

Por el contrario, si se selecciona un conjunto de primitivas abstractas, como segmentos o contornos, etc. y se aplica el apareo sobre este conjunto, se está en presencia de un algoritmo basado en características.

A continuación, se describen las particularidades de las categorías, enunciando sus principales ventajas y desventajas y una discusión sobre el empleo de una u otra de acuerdo a la aplicación.

#### B.4.1. Basados en Áreas

Los sistemas basados en áreas realizan una correlación de los niveles de gris de porciones o ventanas de las imágenes, asumiendo que ellas presentan alguna similaridad. Utilizan principalmente técnicas de búsqueda basadas en correlación, así como también técnicas de predicción y verificación.

En estos casos, la única primitiva utilizada en la intensidad de los pixeles, representada por el nivel de gris de la imagen.

Estos sistemas ofrecen la ventaja de generar funciones discretas de correspondencia densas. En caso contrario los valores pueden ser interpolados si fuese necesario.

Los algoritmos basados en áreas se han aplicado satisfactoriamente en el análisis de imágenes satelitarias y áreas de terrenos, donde la superficie varía en forma suave y continua.

A causa de la severa asunción de continuidad, no son apropiados para superficies con variaciones bruscas de profundidad. Además, son muy sensibles a las variaciones de intensidad, contrastes e iluminación.

Por otra parte, debido a que estas técnicas resuelven las ambigüedades basándose puramente en información local, la mayoría de los algoritmos que las utilizan no incorporan los criterios de consistencia global.

Un problema adicional de los métodos basados en áreas que aplican un esquema de ventana de correlación es determinar el tamaño de ventana óptimo. Existen desarrollos como el presentado en que utilizan tamaño de ventana variable de acuerdo a las características locales de la imagen.

Aunque el costo computacional de estos algoritmos es elevado, a causa de los numerosos cálculos involucrados en la correlación, gracias a las arquitecturas masivamente paralelas disponibles, es posible reducir considerablemente el tiempo de ejecución de los mismos.

#### B.4.2. Basados en Características

Los sistemas basados en características, en primer lugar identifican primitivas en las imágenes para luego aparearlas. Por este motivo, requieren que las imágenes posean superficies ricas en texturas, para permitir la extracción de terminales.

Estos sistemas adoptan diversas técnicas de búsqueda de apareos, entre las cuales se destacan correlación, programación dinámica, métodos de relajación, etc.

Las primitivas que prevalecen son los puntos de bordes, los segmentos de bordes, esquinas y otras estructuras más complejas, como contornos cerrados y diferentes descriptores.

Cuando estos algoritmos emplean más de un tipo de primitiva, para recolectar información de la imagen, reciben el nombre de algoritmos multiprimitiva.

Los métodos basados en características se distinguen por su rapidez, ya que solamente un subconjunto de pixeles de la imagen se tiene en cuenta para el apareo. Además, son más estables a las variaciones fotométricas, porque los parámetros de las primitivas son

menos sensibles a las mismas que la intensidad de los píxeles.

También, las primitivas pueden ser localizadas en las imágenes con gran precisión, del orden de subpixel. Por consiguiente, se obtiene un alto grado de precisión en la puesta en correspondencia.

Sin embargo, estos sistemas suelen fallar si las primitivas no pueden ser extraídas, de las imágenes, de manera confiable. Asimismo, falta de ciertas precauciones, pueden conducir a errores si existen objetos ocultos o resultar no apropiados para regiones que carecen de características.

Estos algoritmos producen mapas de disparidad raros si las características se encuentran muy espaciadas o distribuidas irregularmente, por lo que deben, frecuentemente, ser interpolados, adicionando un problema.

Contrariamente, en escenas complejas, con alta densidad de características, el número de primitivas a tener en cuenta puede ser demasiado elevado, entonces, el costo computacional del algoritmo se ve incrementado considerablemente, pudiéndose tomar prohibitivo.

Existen trabajos que modelizan e instrumentan un sistema de selección de las primitivas y aprendizaje, esto permite adoptar el conjunto de características más apropiado, es decir, que resulten de mayor poder descriptivo según las circunstancias, de acuerdo a las particularidades que presentan las imágenes.

En general, se incorporan restricciones globales, tales como unicidad y continuidad figural, para auxiliar el proceso de apareo, realizando verificaciones de consistencia global.

#### B.4.3. Algoritmos Basados en Áreas Vs. Basados en Características

Debido a las principales ventajas y desventajas recientemente planteadas, se observa que la elección de un algoritmo basado en áreas o en características, depende del tipo de aplicación en juego. Como regla general se tiene:

- Si las escenas observadas presentan superficies simples, poco texturadas y se puede asumir una hipótesis de continuidad, se utilizan algoritmos basados en áreas; este es el caso de fotografías aéreas de escenas de edificaciones aisladas y terrenos abiertos.

- Si por el contrario, las imágenes son más complejas y permiten la extracción confiable de un número considerable de primitivas para su representación, se prefiere el uso de algoritmos basados en características; como por ejemplo fotografías aéreas de zonas urbanizadas o imágenes capturadas por sensores montados sobre un robot dentro de un edificio. Estas últimas son conocidas como escenas "interiores" o "in-doors". En la actualidad, se han desarrollado sistemas que combinan ambas técnicas.

Estos sistemas realizan apareos basados en áreas y basados en características, para posteriormente, combinar ambos resultados esperando obtener una función de disparidad óptima

Cabe señalar, que en la mayoría de los casos, tanto los algoritmos basados en áreas como los algoritmos basados en características, incorporan la restricción de geometría

epipolar de líneas de imagen para simplificar el apareo, reduciendo el espacio de búsqueda a un problema unidimensional, como se describió en la sección, Restricciones Locales.

### B.5. Tendencias

Existen dos líneas de trabajo para resolver el problema del apareo automático. Por un lado, algunos trabajos que se orientan al desarrollo de una solución general independiente de la aplicación, basados en estudios realizados sobre modelos psicofisiológicos del sistema de visión humano. En este sentido, muchas soluciones trabajan con primitivas de bajo nivel, que se detectan sin necesidad de un análisis sofisticado. Estos trabajos se fundamentan en investigaciones realizadas sobre la estereopsis humana, proceso de fusionar en el cerebro las imágenes captadas por ambos ojos. Estas investigaciones sugerirían que la estereopsis ocurre en forma temprana en el proceso de visión, antes de analizar en forma detallada cada imagen en forma monocular. Los experimentos realizados por Bela Julesz sobre estereogramas aleatorios y las apariciones de autoestereogramas sustentarían esta teoría. Dentro de la línea de estudios que modelizan la visión existen criterios opuestos. En efecto, hay quienes sostienen que el proceso de percepción estereoscópica releva de una actividad inteligente, por la que el ser humano proyecta la forma tridimensional que está motivado a percibir y si bien admiten un automatismo de bajo nivel, lo que justifica lo que sucede con los estereogramas aleatorios, consideran que los mecanismos de la respuesta perceptual mejoran, robustecen y resuelven ambigüedades de apareo.

Por el otro lado, existen líneas de trabajo que intentan desarrollar soluciones para aplicaciones particulares. En estas, el conocimiento de la escena observada, de la presencia patrones específicos, ya conocidos, o del número de ocurrencias de un objeto, puede proveer ciertas reglas que facilitan la solución de los apareos. En algunos casos resultan de estudios psicofisiológicos de la visión binocular por lo que son también un argumento que soporta el modelo de mecanismos de la respuesta perceptual.

## **PUBLICACIONES**

---

**STEREOSCOPIC VISION  
MINIMUM SET OF HOMOLOGOUS POINTS FOR  
ROTORECTIFICATION\***

*José F. Zelasco*<sup>1,2,3</sup>

jfz@fecic.edu.ar

*Teresa Arcomano*<sup>2</sup>

sacomputacion@sinectis.com.ar

*Judith Donayo*<sup>2</sup>

juddonayo@informaticosmix.com

1 Laboratorio de Visión Estéreo – Departamento de Ingeniería Mecánica y Naval – FI-UBA

2 FCEyN-UBA

3 INCA – FCE-UNCPBA

#### **ABSTRACT**

This paper has been developed in the frame of a project of stereo vision for AUV provided with two optical sensors, with several lines of work. One involves recalculating underwater scenes that have been taken from two different locations of the vehicle. This improves the precision, because of the increased distance between the points of view. In successive takes, the lines of images don't coincide with the epipolar lines. With epipolar lines of pixels the problem has been successfully solved as the result of a line of work. In our case, the parameters of relative rotation and translation of two successive images are approximately known. It is, then, necessary to improve the parameter precision to carrying out the rotorectification of the images. This has been also resolved as the result of another line of work. However this solution requires a certain number of pairs of homologous points.

The objective of this paper is to provide those necessary homologous points. For this we calculate the position of the epipolar lines in both successive images. We also find the approximate position of the homologous point of a given one. These elements limit the searching space. An algorithm that carries out compensation by least squares, to search the homologous points in successive images that don't satisfy the epipolar condition of their lines, has been used after being modified. As a comparison we have also developed another algorithm with the same target.

**KEYWORDS:** *stereo matching, epipolar condition, stereoscopic vision, two-view matching, homologous points, stereo image.*

---

\* Artículo presentado en el IASTED International Conference., Pittsburg, Pensilvania - USA, Mayo 15-17, 2000.

***COMPUTER VISION IN AUVS:  
AUTOMATIC ROTO-RECTIFICATION OF  
STEREO IMAGES \****

*José F. Zelasco*<sup>1 2 3</sup>

jfz@fecic.edu.ar

*Teresa Arcomano*<sup>2</sup>

sacomputacion@sinectis.com.ar

*Diego Amin Dagum*<sup>2</sup>

diegum@bigfoot.com

*Judith Donayo*<sup>2</sup>

juddonayo@informaticosmix.com

1 Laboratorio de Visión Estéreo – Departamento de Ingeniería Mecánica y Naval – FI-UBA

2 FCEyN-UBA

3 INCA – FCE-UNCPBA

***ABSTRACT***

This work has been developed in the frame of a project of stereo vision for autonomous underwater vehicles (AUV) provided with two optic sensors, simulating the system of human binocular vision. The project gave an origin to several work lines. To obtain a numeric model of an underwater scene of more precision than the one that would be obtained starting from a couple of simultaneous images, we should re-calculate starting from two images taken in consecutive way. The increase of the distance among the points of view would improve this precision. Normally, in consecutive exposures, the lines of bitmaps do not coincide with the epipolar lines as in case of simultaneous images from calibrated sensors. This complicates the search of homologous points (problem successfully solved as a result of another work line). Parameters of relative rotation and translation of two consecutive images are approximately known. To know them with more accuracy shall allow us to carry out the image roto-rectification and to use the results of the mentioned works. Getting knowledge of a certain number of couples of homologous points, (not less of five pairs), we can calculate with enough precision the relative rotation parameters in order to proceed with image roto-rectification, obtaining a new image couple where epipolar lines match with bitmap lines. This involves three major steps. The first one, concerning with the obtainment of relative orientation among components of stereo couple. That relative orientation will help to find an optimal plane to minimize deformation when projecting images onto. Finally, it follows the projection itself of the couple over optimal plane found. That's how this paper helps to achieve epipolar stereo images. Then, it's possible to apply the 3-D reconstruction process. In conclusion, given any stereo pair of images, the whole process of 3-D numeric model obtainment can be automated.

Key Words: Roto-Rectification, Stereoscopic Vision, Stereo Images, Homologous Points

---

\* Artículo publicado en el OCEANS2000, Providence, Rhode Island -USA, Septiembre 11- 14, 2000.

## **OBTENCIÓN AUTOMÁTICA DEL MODELO NUMÉRICO DE TERRENO EN FOTOGRAFETRÍA AÉREA\***

**José F. Zelasco**<sup>1 2 3</sup>

[jfz@fecic.edu.ar](mailto:jfz@fecic.edu.ar)

**Teresa Arcomano**<sup>2</sup>

[sacomputacion@sinectis.com.ar](mailto:sacomputacion@sinectis.com.ar)

**Diego Amin Dagum**<sup>2</sup>

[diegum@bigfoot.com](mailto:diegum@bigfoot.com)

**Alejandra V. Donadio**<sup>4</sup>

[adonadio@usa.net](mailto:adonadio@usa.net)

**Judith Donayo**<sup>2</sup>

[juddonayo@informaticosmix.com](mailto:juddonayo@informaticosmix.com)

**Daniel E. Mendez**<sup>4</sup>

[mendezde@yahoo.com.ar](mailto:mendezde@yahoo.com.ar)

**Samuel Pomerantz**<sup>5</sup>

[psamue@bigfoot.com](mailto:psamue@bigfoot.com)

1 Laboratorio de Visión Estéreo – Departamento de Ingeniería Mecánica y Naval – FI-UBA

2 FCEyN-UBA

3 INCA – FCE-UNCPBA

4 FCE-UNCPBA

5 FI-UBA

### **RESUMEN**

Cuando las líneas de píxeles o líneas de imagen coinciden con las líneas epipolares del par estéreo, el apareo automático se simplifica ya que la búsqueda de puntos homólogos, se efectúa esencialmente recorriendo líneas de imagen.

Las imágenes aéreas no cumplen esta condición por lo que se requiere la roto rectificación de pares de imágenes. Con esta operación se logra que las líneas de imagen coincidan con las líneas epipolares.

Para lograr la roto rectificación del par de imágenes, se deben conocer un número mínimo de 5 o 6 pares de puntos. Su búsqueda automática no es simple ya que las líneas de imagen no coinciden con las líneas epipolares. Se limita el espacio de búsqueda, calculando, en la segunda imagen de cada par de una corrida de imágenes aéreas, la línea epipolar correspondiente al punto de la primera imagen del par.

Por otra parte, con el par de imágenes anteriores a una imagen dada, se puede conocer las coordenadas objeto de algunos puntos ubicados en el 20 % común al triple estereoscópico. Esto limita aún más el entorno de búsqueda.

Cuando no se dispone de esta información, se reduce el entorno de búsqueda sobre la línea epipolar, estimando un  $\pm \Delta h$  sobre y por debajo de la altura media del vuelo H, dato que en general es conocido al estudiar el proyecto.

Con estos elementos, el conocimiento aproximado de los parámetros de orientación externa etc., y contando con puntos de control de la corrida, se obtiene, usando algoritmos de apareo automático, el modelo numérico de terreno.

\* Artículo publicado en el IX Simposio Latinoamericano de Percepción Remota, Puerto Iguazú - Misiones- Argentina - Noviembre 6- 10, 2000

## Referencias Bibliográficas

- [ 1] J.F.Zelasco, A.Donadio, D.Méndez. **Stereo Matching: Feature and Area Based Solutions. Analysis of Results.** *Proceedings of the IASTED Computer Graphics and Images.* Halifax, Canada, June, 1998).
- [ 2] J.F.Zelasco, D.Dagum. **Relative Orientation and roto-rectification of Stereoscopic Underwater Images.;** *Conference Proceedings of the Oceans'98 IEEE/OAS* vol. 3, Nice, France, September 1998.
- [ 3] R.Etienne. **Reconstrucción du relief par segmentation et mise en correspondance d'objets de couples d'images stéréoscopiques.** *Ecole Nationale de Telecommunications.* Paris 1986.
- [ 4] C.Schmid , R.Mohr. **Matching by Local Invariants,** *INRIA*, Nº 2644.
- [ 5] O.Faugeras, B.Hotz., H.Matthieu. **Real time correlation- based stereo: algorithim, implementación and applications.** *INRIA* Nº 2769.
- [ 6] D.J.Braunegg. **A system that recognizes world Locations with stereo vision.** *IEEE.* Vol 9, Nº 3, pag 303-308. 1993.
- [ 7] O.Faugeras, M.Buffa, Z.Zhang. **A stereo vision-based navigation system for mobile robot.** *INRIA* Nº 1895, 1993
- [ 8] O.Faugeras, S.Laveau, L.Robert, G.Csurka, C.Zeller. **3-D Reconstruction of Urban Scenes from Sequences of Images,** *INRIA* Nº 2572, Julio 1995.
- [ 9] P.Gros, O.Bournez, E.Boyer. **Using geometric quasi-invariants to match and model images of line segments,** *INRIA* Nº 2608, Julio 1995.
- [10] J.F.Zelasco, A.V.Donadio, D.E.Mendez. **Estereoscopía: Plataforma de Prueba y Validación de algoritmos de Puesta en correspondencia.** *ICIE95.*
- [11] J.F.Zelasco, A.V.Donadio, D.E.Mendez. **Estereopsis automática. Panorama del Estado del Arte.** *Anales de la Sociedad Científica Argentina.*1998. Volumen 228-Nº1-ISSN 0037-8437- Pag. 37/52.
- [12] A.V.Donadio, D.E.Méndez. **Visión 3D. Tratamiento de Imágenes Esteroscópicas. Restitución automática. Evaluación de resultados,** *Tesis de Grado. UNCBA, FCE,* 1997.
- [13] D.G.Forney. **The Vitervi algorithm.** *IEEE.* Vol. 14, Nº 3, pag 337-345, Marzo 1992.
- [14] J.F.Zelasco, D.Dagum. **Stereoscopic Underwater Images Relative Orientation and roto-rectification.;** *Signal and Image Processing . IASTED,* October1999, Nassau, Bahamas.
- [15] F.Devernay. **A non-Maxima Suppression Method for Edge Detection with Sub-Pixel Accuracy,** *INRIA* Nº 2724 Noviembre 1995.

- [16] J.Weng, N.Ahuja, T.S.Huang. Matching Two Perspective Views. *IEEE, Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, Nº 8, agosto 1992.
- [17] P.Molina Martinez. **Aplicación y Desarrollo de Algoritmos de Procesamiento de Imágenes Para Determinar el Movimiento de un Robot Móvil.** *Tesis de Grado. Universidad Politécnica de Valencia.*
- [18] A.K.Jain. **Fundamentals of Digital Image Processing.** *Prentice Hall International*, 1989.
- [19] J.S.Lim. **Two Dimensional Signal and Image Processing.** *Prentice Hall International*, 1998.
- [20] G.Strang. **Algebra Lineal y sus Aplicaciones.** *Fondo Educativo Interamericano*, 1982.
- [21] C.Rothwell. **The Importance of Reasoning About Occlusions During Hypothesis Verification in Object Recognition.** *INRIA* Nº 2673, October 1995
- [22] J.F.Zelasco, J.Donayo, T.Arcomano. **Stereoscopic Vision - Minimum Set of Homologous Points for Rotorectification.** *IASTED*. May , 2000, Pittsburg, Pensilvania - USA
- [23] J.F.Zelasco, J.Donayo, T.Arcomano, D. Dagum. **Computer Vision in AUVs: Automatic Roto-Rectification of Stereo Images.** *OCEANS2000* - Providence, Rodhe Island - Septiembre, 2000
- [24] J.F.Zelasco. **Notas del Curso “Imágenes Estereoscópicas. Visión en robótica”**, 1999.

## **Bibliografía de Consulta**

- J.Facon. **Processamento e Analise de Imagens**, *VI EBAI*. Julio 1993.
- T.Kanade, M.Okutomi, **A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment.** *IEEE*, vol 16 Nº 9, pág. 920-932, 1994
- T. Papadopoulo. **Motion Analysis of 3D Rigid Curves from Monocular Image Sequence.** *INRIA* Nº 2779
- B. Ramaekers. **Détermination de la position angulaire instantanée d'un visage dans une séquence en mouvement.** *Université Catholique de Louvain*. Junio 1997
- Z.D.Lan, R.Mohr. **Robust Matching by Partial Correlation.** *INRIA* Nº 2643, Agosto 1995.
- B. Espiau. **Asservissement visuel et commande de la distance focale.** *INRIA* Nº 2613, Julio 1995.
- J.P.Tarel, N.Boujema. **Une approche floue du recalage 3D: généralité et robustesse.** *INRIA* Nº 2716, Noviembre 1995.